AEI–2011–34
MPP–2011–47

Complete N–Point Superstring Disk Amplitude I. Pure Spinor Computation

Carlos R. Mafra[a,b,1], Oliver Schlotterer[b,c,2], and Stephan Stieberger[b,c,3]

[a] *Max–Planck–Institut für Gravitationsphysik*
Albert–Einstein–Institut, 14476 Potsdam, Germany

[b] *Kavli Institute for Theoretical Physics*
University of California, Santa Barbara, CA 93106, USA

[c] *Max–Planck–Institut für Physik*
Werner–Heisenberg–Institut, 80805 München, Germany

In this paper the pure spinor formalism is used to obtain a compact expression for the superstring N–point disk amplitude. The color ordered string amplitude is given by a sum over $(N-3)!$ super Yang–Mills subamplitudes multiplied by multiple Gaussian hypergeometric functions. In order to obtain this result, the cohomology structure of the pure spinor superspace is exploited to generalize the Berends–Giele method of computing super Yang–Mills amplitudes. The method was briefly presented in [1], and this paper elaborates on the details and contains higher-rank examples of building blocks and associated cohomology objects. But the main achievement of this work is to identify these field-theory structures in the pure spinor computation of the superstring amplitude. In particular, the associated set of basis worldsheet integrals is constructively obtained here and thoroughly investigated together with the structure and properties of the amplitude in [2].

[1] e-mail: crmafra@aei.mpg.de

[2] e-mail: olivers@mppmu.mpg.de

[3] e-mail: stephan.stieberger@mpp.mpg.de

Contents

1 Introduction 1

2 The pure spinor formalism 3

2.1. Massless vertex operators and SYM superfields 4

2.2. Tree-level scattering amplitudes 5

2.3. Component expansions of amplitudes: a simple example 6

3 BRST building blocks 7

3.1. OPE residues of vertex operators 7

3.2. Definition of BRST building blocks $T_{123\ldots p}$ 9

3.3. Diagrammatic interpretation of $T_{123\ldots p}$ building blocks 12

3.4. BRST symmetries of building blocks 14

3.5. Explicit construction of $T_{12\ldots p}$ 16

4 Supersymmetric Berends–Giele recursions 19

4.1. Construction of Berends–Giele currents $M_{123\ldots p}$ 19

4.2. Symmetry properties of $M_{12\ldots p}$ 22

4.3. The N–point field-theory tree amplitude 24

4.4. BRST integration by parts and cyclic symmetry 25

4.5. Factorization in cyclically symmetric form 26

5 The superstring tree amplitude in pure spinor superspace 27

5.1. The CFT correlator 28

5.2. A closed formula for $M_{12\ldots p}$ from the superstring 30

5.3. Trading $T_{12\ldots p}$ for $M_{12\ldots p}$ 30

5.4. Worldsheet integration by parts 31

5.5. The complete N–point superstring disk amplitude 33

A The explicit construction of T_{12345} 36

B The solutions for $M_{i_1 i_2 \ldots i_n}$ in terms of BRST building blocks 39

C The cubic graphs of $M_{123\ldots n}$ 42

1. Introduction

The computation of tree-level superstring scattering amplitudes is an important problem since the birth of string theory (see e.g. [3]). But despite being already four decades old, explicit results for tree amplitudes with more than four external legs [4] have only recently been completed using the Ramond–Neveu–Schwarz (RNS) formalism at five points [5], at six points [6] and partially up to seven points [7]. In addition to conceptual issues about higher-point worldsheet integrals, the huge amount of algebraic manipulations required to complete these calculations has proven to be a major obstacle to further developments. When written in terms of ten-dimensional momenta and polarizations, the amplitudes simply become too big.

However, since the year 2000 a new formalism for the superstring which can be used to compute manifestly super-Poincaré invariant scattering amplitudes in superspace is available [8]. A general proof that the disk amplitudes in the pure spinor formalism for an arbitrary number of bosonic and for up to four fermionic external state agree with the standard RNS prescription was given in [9]; and the supersymmetric four-, five- and six-point tree amplitudes have been explicitly computed in [10-13].

In this paper the general problem will be solved; i.e. the complete solution for *all* N–point superstring color-ordered disk amplitude $\mathcal{A}_N \equiv \mathcal{A}(1,2,\ldots,N)$ is given by

$$\mathcal{A}_N = \int\limits_{z_i<z_{i+1}} \prod_{i<j} |z_{ij}|^{-s_{ij}} \Big[\prod_{k=2}^{N-2} \sum_{m=1}^{k-1} \frac{s_{mk}}{z_{mk}} \, \mathcal{A}_{YM}(1,2,\ldots,N) + \mathcal{P}(2,\ldots,N-2) \Big], \qquad (1.1)$$

where $\mathcal{A}_{YM}(1,2,\ldots,N)$ is the color-ordered N–point super Yang–Mills subamplitude in ten dimensions, $\mathcal{P}(2,\ldots,N-2)$ means the summation over all $(N-3)!$ permutations of the labels $(2,\ldots,N-2)$ inside the brackets, and the color ordering of the superstring subamplitude is defined by the integration region $\int\limits_{z_i<z_{i+1}} \equiv \prod_{j=2}^{N-2} \int_{z_{j-1}}^{1} dz_j$.

It is straightforward to obtain subamplitudes associated with different color orderings $(1,2,\ldots,N) \mapsto (1_\sigma, 2_\sigma, \ldots, (N-1)_\sigma, N)$ for $\sigma \in S_{N-1}$ and $i_\sigma \equiv \sigma(i)$ from (1.1). The worldsheet integrand with its $(N-3)!$ kinematic $\mathcal{A}_{YM}$ packages stay the same, only the integration region has to be adapted to

$$I_\sigma \equiv \{z_i \in \mathbf{R},\ 0 = z_{1_\sigma} \leq z_{2_\sigma} \leq \ldots \leq z_{(N-2)_\sigma} \leq z_{(N-1)_\sigma} = 1\},$$

according to the $\sigma \in S_{N-1}$ permutation in question,

$$\mathcal{A}(1_\sigma, 2_\sigma, \ldots, (N-1)_\sigma, N) = \int_{I_\sigma} \prod_{l=2}^{N-2} dz_{l_\sigma} \prod_{i<j} |z_{ij}|^{-s_{ij}} \times \Big[\prod_{k=2}^{N-2} \sum_{m=1}^{k-1} \frac{s_{mk}}{z_{mk}} \, \mathcal{A}_{YM}(1, 2, \ldots, N) + \mathcal{P}(2, \ldots, N-2) \Big]. \tag{1.2}$$

By taking the $\alpha' \to 0$ field-theory limit of (1.2) (in particular of the integrals involved using the methods presented in [2]), it follows that all color-ordered field theory amplitudes can be written in terms of the $(N-3)!$ dimensional basis $\{\mathcal{A}_{YM}(1, 2_\sigma, \ldots, (N-2)_\sigma, N-1, N) \mid \sigma \in S_{N-3}\}$, a result which was proposed in [14] and later proved in [15,16] using monodromy relations in string theory. Furthermore, plugging in the explicit field-theory limits of the integrals appearing in (1.1) (using the method described in [2]), one derives the BCJ relations among different color-ordered subamplitudes discussed in [14].

This paper is organized as follows. In section 2 a brief review of the pure spinor formalism is given; with special emphasis to the elements necessary for the scattering amplitude computations in the following sections. In section 3 the BRST building blocks which encode the information of the pure spinor CFT correlator will be defined and their BRST properties studied at length. In particular, a diagrammatic method which associates arbitrary cubic graphs to certain building block combinations is fully presented (partial results have already been shown in [1]). In section 4 a pure spinor generalization of the recursive method of Berends–Giele [17] to compute super Yang–Mills in ten-dimensions is developed which extends the previous results of [1]. In section 5 the general N–point CFT correlator of the superstring amplitude involved in the pure spinor prescription is obtained in a compact form using the BRST cohomology objects of the previous sections. Finally, using a mixture of pure spinor superspace manipulations together with total derivative relations for the superstring integrals, the superstring N–point amplitude is rewritten in terms of the field-theory subamplitudes as in the result (1.1) presented above. In the appendix A, the calculations involving the explicit derivation of the building block T_{12345} in terms of super Yang–Mills superfields (which were omitted from the main text due to its lenghty nature) are presented in full detail. In appendix B, the explicit expressions for the pure spinor Berends–Giele currents $M_{123\ldots p}$ are written down in terms of BRST building blocks for up to and including $M_{1234567}$. Finally, in appendix C the cubic graphs which were used to find the expressions of appendix B are depicted up to M_{123456} (the 132 graphs used to derive $M_{1234567}$ would occupy too much space and were omitted).

2. The pure spinor formalism

In the pure spinor formalism [8], the worldsheet action for the type IIB superstring is

$$S = \frac{1}{2\pi}\int d^2z\ \left(\frac{1}{2}\partial X^m\overline{\partial}X_m + p_\alpha\overline{\partial}\theta^\alpha + \overline{p}_\alpha\partial\theta^\alpha - \omega_\alpha\overline{\partial}\lambda^\alpha - \overline{\omega}_\alpha\partial\lambda^\alpha\right), \tag{2.1}$$

where $[X^m(z,\overline{z}), \theta^\alpha(z), p_\alpha(z); \overline{\theta}^\alpha(\overline{z}), \overline{p}_\alpha(\overline{z})]$ and $[\lambda^\alpha(z), \omega_\alpha(z); \overline{\lambda}^\alpha(\overline{z}), \overline{\omega}_\alpha(\overline{z})]$ are the Green-Schwarz-Siegel matter variables [18,19] and the Berkovits ghosts. The bosonic pure spinor λ^α satisfies

$$\lambda^\alpha\gamma^m_{\alpha\beta}\lambda^\beta = 0, \qquad m = 0,\dots,9 \quad \alpha,\beta = 1,\dots,16 \tag{2.2}$$

where $\gamma^m_{\alpha\beta}$ are the symmetric 16×16 Pauli matrices in $D = 10$. The right-moving fields have opposite chirality for the type IIA, for the heterotic superstring they are the same as in the RNS formalism, and for the open superstring the boundary conditions relate the two sectors. This paper only considers the open superstring, so the right-moving fields will be ignored.

The supersymmetric momentum and Green-Schwarz constraint are given by

$$\Pi^m(z) = \partial X^m + \frac{1}{2}(\theta\gamma^m\partial\theta), \quad d_\alpha(z) = p_\alpha - \frac{1}{2}(\gamma^m\theta)_\alpha\partial X_m - \frac{1}{8}(\gamma^m\theta)_\alpha(\theta\gamma_m\partial\theta), \tag{2.3}$$

while the ghost contribution to the Lorentz currents is denoted by $N^{mn}(z) = \frac{1}{2}(\lambda\gamma^{mn}w)$. Furthermore, the energy-momentum tensor T with vanishing central charge and the ghost-number current J are given by

$$T(z) = -\frac{1}{2}\Pi^m\Pi_m - d_\alpha\partial\theta^\alpha + \omega_\alpha\partial\lambda^\alpha, \qquad J = \omega_\alpha\lambda^\alpha. \tag{2.4}$$

Finally, the physical spectrum is obtained from the cohomology of the BRST charge [8]

$$Q = \oint \lambda^\alpha(z)d_\alpha(z). \tag{2.5}$$

One can show that these operators satisfy the following relations [8,19,20]

$$\begin{aligned}
&d_\alpha(z)d_\beta(w) \to -\frac{\gamma^m_{\alpha\beta}\Pi_m}{z-w}, \quad \Pi^m(z)\Pi^n(w) \to -\frac{\eta^{mn}}{(z-w)^2}, \quad d_\alpha(z)\,\theta^\beta(w) \to \frac{\delta^\beta_\alpha}{(z-w)}\\
&N^{mn}(z)N_{pq}(w) \to \frac{4}{z-w}N^{[m}_{[p}\delta^{n]}_{q]} - \frac{6}{(z-w)^2}\delta^n_{[p}\delta^m_{q]}, \quad N^{mn}(z)\lambda^\alpha(w) \to -\frac{1}{2}\frac{(\lambda\gamma^{mn})^\alpha}{z-w}\\
&d_\alpha(z)\Pi^m(w) \to \frac{(\gamma^m\partial\theta)_\alpha}{z-w}, \quad \Pi^m(z)X^n(w) \to -\frac{\eta^{mn}}{z-w}, \quad J(z)\lambda^\alpha(w) \to \frac{\lambda^\alpha}{z-w}
\end{aligned} \tag{2.6}$$

where the antisymmetrization bracket $[\ldots]$ encompassing N indices is defined to contain an overall factor of $1/N!$. Furthermore, if $f(X,\theta)$ is a superfield containing only the zero modes of θ and $D_\alpha = \partial_\alpha + \frac{1}{2}(\gamma^m\theta)_\alpha\partial_m$ is the supersymmetric covariant derivative,

$$d_\alpha(z)f(X(w),\theta(w)) \to \frac{D_\alpha f(X(w),\theta(w))}{z-w}, \quad \Pi^m(z)f(X(w),\theta(w)) \to -\frac{k^m f(X(w),\theta(w))}{z-w}.$$

Hence, the action of the BRST operator on superfields is $Qf = \lambda^\alpha D_\alpha f$. It is easy to show using the OPEs of (2.6) and the pure spinor constraint (2.2) that the BRST charge indeed satisfies $Q^2 = 0$. So, the pure spinor formalism can be covariantly quantized, is manifestly space-time supersymmetric and contains no worldsheet spinor fields; avoiding from the outset the issues which make the computation of scattering amplitudes with the RNS and GS formalisms a difficult task.

Throughout this paper $k_m^{12\ldots n}$ stands for $k_m^1 + k_m^2 + \cdots + k_m^n$, the dimensionless (generalized) Mandelstam invariants are given by

$$s_{12\ldots n} = \alpha'(k^1 + k^2 + \cdots + k^n)^2, \tag{2.7}$$

and whenever an α' is not explicitly written down the convention $2\alpha' = 1$ has been used.

2.1. Massless vertex operators and SYM superfields

For the open superstring, the vertex operators for the massless states in unintegrated and integrated forms are given by

$$V^i = \lambda^\alpha A^i_\alpha(x,\theta), \qquad U^i = \partial\theta^\alpha A^i_\alpha + \Pi^m A^i_m + d_\alpha W^\alpha_i + \frac{1}{2}\mathcal{F}^i_{mn}N^{mn}, \tag{2.8}$$

where i denotes the label of the string whose massless modes are described by the ten-dimensional super Yang–Mills (SYM) superfields $[A_\alpha, A_m, W^\alpha, \mathcal{F}_{mn}]$ satisfying [20,21]

$$\begin{aligned} D_\alpha A_\beta + D_\beta A_\alpha &= \gamma^m_{\alpha\beta}A_m, \qquad & D_\alpha A_m &= (\gamma_m W)_\alpha + k_m A_\alpha \\ D_\alpha \mathcal{F}_{mn} &= 2k_{[m}(\gamma_{n]}W)_\alpha, \qquad & D_\alpha W^\beta &= \frac{1}{4}(\gamma^{mn})_\alpha{}^\beta \mathcal{F}_{mn}. \end{aligned} \tag{2.9}$$

Their θ–expansions can be computed using the gauge $\theta^\alpha A_\alpha = 0$ [10,22],

$$\begin{aligned} A_\alpha(x,\theta) &= \frac{1}{2}a_m(\gamma^m\theta)_\alpha - \frac{1}{3}(\xi\gamma_m\theta)(\gamma^m\theta)_\alpha - \frac{1}{32}F_{mn}(\gamma_p\theta)_\alpha(\theta\gamma^{mnp}\theta) + \cdots \\ A_m(x,\theta) &= a_m - (\xi\gamma_m\theta) - \frac{1}{8}(\theta\gamma_m\gamma^{pq}\theta)F_{pq} + \frac{1}{12}(\theta\gamma_m\gamma^{pq}\theta)(\partial_p\xi\gamma_q\theta) + \cdots \\ W^\alpha(x,\theta) &= \xi^\alpha - \frac{1}{4}(\gamma^{mn}\theta)^\alpha F_{mn} + \frac{1}{4}(\gamma^{mn}\theta)^\alpha(\partial_m\xi\gamma_n\theta) + \frac{1}{48}(\gamma^{mn}\theta)^\alpha(\theta\gamma_n\gamma^{pq}\theta)\partial_m F_{pq} + \cdots \\ \mathcal{F}_{mn}(x,\theta) &= F_{mn} - 2(\partial_{[m}\xi\gamma_{n]}\theta) + \frac{1}{4}(\theta\gamma_{[m}\gamma^{pq}\theta)\partial_{n]}F_{pq} + \frac{1}{6}\partial_{[m}(\theta\gamma_{n]}{}^{pq}\theta)(\xi\gamma_q\theta)\partial_p + \cdots \end{aligned} \tag{2.10}$$

where $a_m(X) = e_m e^{ik\cdot X}$, $\xi^\alpha(X) = \chi^\alpha e^{ik\cdot X}$ are the bosonic and fermionic polarizations and $F_{mn} = 2\partial_{[m}a_{n]}$ is the field-strength. Using the OPEs (2.6) and equations of motion (2.9) one can show that

$$(\lambda\gamma^m W^i)(z_i)U^j(z_j) \to \frac{1}{z_j - z_i}\big[(\lambda\gamma^n W^j)\mathcal{F}^i_{mn} - (\lambda\gamma^m W^i)(k^i\cdot A^j) + Q(W^i\gamma^m W^j)\big], \quad (2.11)$$

which will be frequently used in the computations below.

As shown by Howe in 1991 [23], the use of a pure spinor field simplifies the description of ten-dimensional super Yang–Mills, and this is naturally incorporated in the pure spinor formalism. For example, it can be shown that $QV = 0$ is equivalent to putting the SYM superfields on-shell and it also implies that the BRST variation of the integrated vertex U is given by $QU = \partial V$ [20], and many simplifications occur due to this compact description. In fact, it has recently been shown how the cohomology of pure spinor superspace [24,25] is enough to fix all N–point scattering amplitudes of $D = 10$ SYM [26,1]. So unless otherwise stated, all superfield manipulations in the next sections are done on-shell; where both $QV = 0$ and $QU = \partial V$ are satisfied.

2.2. Tree-level scattering amplitudes

The prescription to compute a tree-level open-string scattering amplitude with the pure spinor formalism is given by [8] (see also [9]),

$$\mathcal{A}_N = \left\langle V^1(0)\, V^{(N-1)}(1)\, V^N(\infty) \int dz_2\, U^2(z_2)\, \cdots \int dz_{(N-2)}\, U^{(N-2)}\left(z_{(N-2)}\right) \right\rangle, \quad (2.12)$$

where V^i and U^i are the massless vertex operators of (2.8) and the $SL(2,R)$ invariance of the disk worldsheet has already been used to fix three vertex positions to the convenient values $(z_1, z_{N-1}, z_N) = (0, 1, \infty)$. The pure spinor bracket $\langle \ldots \rangle$ appearing in (2.12) denotes a zero-mode integration prescription for the variables λ^α and θ^α, which are the only ones among $[d_\alpha, \Pi^m, N^{mn}, \theta^\alpha, \partial\theta^\alpha, \lambda^\alpha, w_\alpha]$ to contain zero modes on the disk because they have conformal weight zero [27]. Furthermore, the integration regions of (2.12) encode the different color orderings of the external states. For example, the ordering $A_N(1, 2, 3, \ldots, N)$ is computed when the integration region is $0 = z_1 \le z_2 \le \cdots \le z_{N-2} \le z_{N-1} = 1$.

After integrating out the conformal weight-one variables $[d_\alpha, \Pi^m, N^{mn}, \partial\theta^\alpha]$ from the tree-level amplitude (2.12) using the OPEs of (2.6) and evaluating the world-sheet integrals,

one is left with a generic pure spinor superspace expression containing the zero modes of λ^α and θ^α

$$\mathcal{A}_N = \langle \lambda^\alpha \lambda^\beta \lambda^\gamma f_{\alpha\beta\gamma}^{i_1\dots i_n}(\theta,\alpha')\rangle. \tag{2.13}$$

In (2.13), $f_{\alpha\beta\gamma}^{i_1\dots i_n}(\theta,\alpha')$ is both a composite superfield in the labels $[i_1,\dots,i_n]$ of the external states and a function of the string scale α' satisfying $\lambda^\alpha\lambda^\beta\lambda^\gamma\lambda^\delta D_\delta f_{\alpha\beta\gamma}^{i_1\dots i_n}(\theta,\alpha')=0$. Its specific form in terms of the super Yang–Mills superfields $[A^i_\alpha, A^i_m, W^\alpha_i, \mathcal{F}^i_{mn}]$ follows from the OPE contractions discussed above while its functional dependence on α' is determined by the momentum expansion of n-point hypergeometric integrals [5,6,7]. As explained in [8], the zero-mode integration of $\langle\dots\rangle$ selects from the $\theta-$expansion of the enclosed superfields the unique element in the cohomology of the pure spinor BRST operator at ghost-number three; $(\lambda\gamma^m\theta)(\lambda\gamma^n\theta)(\lambda\gamma^p\theta)(\theta\gamma_{mnp}\theta)$. Its tree-level normalization can be chosen as

$$\langle(\lambda\gamma^m\theta)(\lambda\gamma^n\theta)(\lambda\gamma^p\theta)(\theta\gamma_{mnp}\theta)\rangle = 1, \tag{2.14}$$

and although (2.14) involves only five θ^α out of sixteen, it can be shown to be supersymmetric [8]. Furthermore, given the fact that there is only one scalar in the decomposition of $(\lambda^3\theta^5)$ it is possible to compute any correlator using symmetry arguments and the normalization condition (2.14) [28,29].

2.3. Component expansions of amplitudes: a simple example

Given a pure spinor superspace expression like in (2.13) it is straightforward to perform the θ-expansion of the SYM superfields and select the terms according to (2.14) to obtain the supersymmetric result of the scattering amplitude in terms of the more familiar gluon and gluino polarizations $[e^i_m, \chi^\alpha_i]$ and their momenta k^m_i. For example, let us obtain the 3-gluon scattering from the component expansion of the 3-point amplitude [8],

$$\mathcal{A}_3 = \langle(\lambda A^1)(\lambda A^2)(\lambda A^3)\rangle. \tag{2.15}$$

Plugging in the θ-expansions (2.10) and selecting the terms with a total of five θ's which contain only gluon fields results in,

$$\mathcal{A}_3 = -\frac{1}{64}\left(k^3_m e^1_r e^2_s e^3_n - k^2_m e^1_r e^2_n e^3_s + k^1_m e^1_n e^2_r e^3_s\right)\langle(\lambda\gamma^r\theta)(\lambda\gamma^s\theta)(\lambda\gamma_p\theta)(\theta\gamma^{pmn}\theta)\rangle. \tag{2.16}$$

In the appendix of [30] one finds a catalog of the most common pure spinor correlators and, in particular, $\langle(\lambda\gamma^r\theta)(\lambda\gamma^s\theta)(\lambda\gamma_p\theta)(\theta\gamma^{pmn}\theta)\rangle = \frac{1}{120}\delta^{rsp}_{pmn} = \frac{1}{45}\delta^{rs}_{mn}$. Therefore the 3-gluon amplitude (2.16) is given by

$$\mathcal{A}_3 = -\frac{1}{2880}\left((e^1\cdot e^2)(k^2\cdot e^3)+(e^1\cdot e^3)(k^1\cdot e^2)+(e^2\cdot e^3)(k^3\cdot e^1)\right). \tag{2.17}$$

Performing the above steps becomes a tedious task when higher-point calculations are involved. Fortunately, this procedure is suitable for an automated handling [31,32].

3. BRST building blocks

Only terms which are in the cohomology of the pure spinor BRST charge (2.5) contribute to the n-point scattering amplitude (2.13). Therefore it will be convenient to foresee the BRST properties of the objects which naturally appear in the tree-level calculation of (2.12). With this intent in mind, in this section the OPEs among the massless vertex operators (2.8) are used to define composite superfields $L_{2131\ldots p1}$ and their BRST properties are studied in detail. It will be found that these superfields transform covariantly under the BRST charge and generically contain BRST-exact parts. A prescription to consistently remove these parts will then be given and that will define the so-called *BRST building blocks*: $T_{123\ldots p}$.

In a later section these building blocks will be used to define other composite superfields $M_{123\ldots p}$ and $E_{123\ldots p}$ with well-defined BRST cohomology properties. They will turn out to be the natural objects with which to write the superstring scattering amplitudes. In the course of doing that, several general structures of the string tree amplitudes will become apparent – like the fact that they can be written using a $(N-3)!$ dimensional basis of integrals as conjectured some years ago in [6].

3.1. OPE residues of vertex operators

Motivated by the computations one needs to perform when computing tree-level higher-point amplitudes [11,12,13] it is convenient to define composite superfields $L_{2131\ldots p1}$ as

$$\lim_{z_2\to z_1} V^1(z_1)U^2(z_2) \to \frac{L_{21}}{z_{21}}, \quad \lim_{z_p\to z_1} L_{2131\ldots(p-1)1}(z_1)U^p(z_p) \to \frac{L_{2131\ldots(p-1)1p1}}{z_{p1}}, \qquad (3.1)$$

which transform covariantly under the action of the pure spinor BRST charge [26]. To see this one uses $QV=0$ and $QU=\partial V$ to obtain

$$QL_{2131\ldots p1} = \lim_{z_p\to z_1} z_{p1}\big[\,\big(QL_{2131\ldots(p-1)1}\big)\,(z_1)U^p(z_p) - L_{2131\ldots(p-1)1}(z_1)\partial V^p(z_p)\big]. \qquad (3.2)$$

The OPE in the first term of (3.2) can be computed using the definition (3.1) recursively while the second term evaluates to $\sum_{j=1}^{p-1} s_{jp}L_{2131\ldots(p-1)1}V_p$; as one can easily show by using $\partial V^i = (\partial\lambda^\alpha)A^i_\alpha + \Pi^m k_m V^i + \partial\theta^\alpha D_\alpha V^i$ and the OPEs of (2.6). Therefore,

$$\begin{aligned} QL_{21} &= s_{12}V_1V_2, \\ QL_{2131} &= (s_{13}+s_{23})L_{21}V_3 + s_{12}(L_{31}V_2 + V_1L_{32}), \end{aligned}$$

$$
\begin{aligned}
QL_{213141} =& (s_{14}+s_{24}+s_{34})L_{2131}V_4 + (s_{13}+s_{23})(L_{21}L_{43}+L_{2141}V_3)\\
&+ s_{12}(L_{3141}V_2 + L_{31}L_{42} + L_{41}L_{32} + V_1L_{3242}),\\
QL_{21314151} =& (s_{15}+s_{25}+s_{35}+s_{45})L_{213141}V_5 + (s_{14}+s_{24}+s_{34})(L_{213151}V_4 + L_{2131}L_{54})\\
&+ (s_{13}+s_{23})(L_{214151}V_3 + L_{2141}L_{53} + L_{2151}L_{43} + L_{21}L_{4353})\\
&+ s_{12}(L_{314151}V_2 + V_1L_{324252} + L_{3141}L_{52} + L_{3151}L_{42} + L_{4151}L_{32}\\
&\quad + L_{31}L_{4252} + L_{41}L_{3252} + L_{51}L_{3242}),
\end{aligned}
\tag{3.3}
$$

while $QL_{2131\ldots p1}$ for $p \geq 6$ can be also be easily obtained (the general BRST variation of a object related to $L_{2131\ldots p1}$ will be written down in the next subsection).

The expressions for $L_{2131\ldots p1}$ in terms of SYM superfields can be obtained using the OPEs of (2.6) in the definition (3.1). For example,

$$
L_{21} \equiv \lim_{z_2\to z_1} z_{21}V^1(z_1)U^2(z_2) = -A^1_m(\lambda\gamma^m W^2) - V^1(k^1\cdot A^2) + Q(A^1W^2). \tag{3.4}
$$

Similar calculations yield the expressions for $L_{2131\ldots p1}$ and one can show that (discarding BRST-exact quantities for reasons to be explained in later sections) they are given by:

$$
\begin{aligned}
L_{21} =& - A^1_m(\lambda\gamma^m W^2) - V^1(k^1\cdot A^2),\\
L_{2131} =& - L_{21}(k^{12}\cdot A^3) - \big[\big(L_{31} + V^1(k^1\cdot A^3)\big)(k^1\cdot A^2) - (1\leftrightarrow 2)\big]\\
&- (\lambda\gamma^m W^3)\big((W^1\gamma_m W^2) - k^2_m(A^1\cdot A^2)\big)\\
L_{213141} =& - L_{2131}(k^{123}\cdot A^4) - (L_{2141} + L_{21}(k^{12}\cdot A^4))(k^{12}\cdot A^3)\\
&- \Big[(L_{3141} + L_{31}(k^{13}\cdot A^4))(k^1\cdot A^2) + (L_{41} + V^1(k^1\cdot A^4))(k^1\cdot A^3)(k^1\cdot A^2)\\
&- \frac{1}{4}(\lambda\gamma^m W^4)(W^2\gamma^{pq}\gamma_m W^3)\mathcal{F}^1_{pq} - (1\leftrightarrow 2)\Big]\\
&+ (\lambda\gamma^m W^4)\big((W^1\gamma^n W^2) - k^n_2(A^1\cdot A^2)\big)\mathcal{F}^3_{mn}\\
L_{21314151} =& - L_{213141}(k^{1234}\cdot A^5) - \big(L_{213151} + L_{2131}(k^{123}\cdot A^5)\big)(k^{123}\cdot A^4)\\
&- \big[L_{214151} + L_{2141}(k^{124}\cdot A^5) + \big(L_{2151} + L_{21}(k^{12}\cdot A^5)\big)(k^{12}\cdot A^4)\big](k^{12}\cdot A^3)\\
&- \Big[\big[L_{314151} + L_{3141}(k^{134}\cdot A^5) + (L_{3151} + L_{31}(k^{13}\cdot A^5))(k^{13}\cdot A^4)\\
&+ \big(L_{4151} + L_{41}(k^{14}\cdot A^5) + (L_{51} + V^1(k^1\cdot A^5))(k^1\cdot A^4)\big)(k^1\cdot A^3)\big](k^1\cdot A^2)\\
&+ (\lambda\gamma^m W^5)\big[\frac{1}{4}(W^1\gamma^{pq}\gamma^n W^3)\mathcal{F}^2_{pq}\mathcal{F}^4_{mn} + \frac{1}{16}(W^4\gamma^m\gamma^{pq}\gamma^{rs}W^1)\mathcal{F}^2_{rs}\mathcal{F}^3_{pq}\big] - (1\leftrightarrow 2)\Big]\\
&+ (\lambda\gamma^m W^5)\Big[(W^1\gamma^n W^2)(\mathcal{F}^4_{mp}\mathcal{F}^3_{np} - (W^3\gamma_m W^4)k^3_n - \frac{1}{2}(W^4\gamma_m\gamma_n\gamma^p W^3)k^{12}_p)\\
&- \frac{1}{2}(W^3\gamma^{pq}\gamma_m W^4)\mathcal{F}^1_{pa}\mathcal{F}^2_{qa} + (A^1\cdot A^2)\big(\mathcal{F}^3_{pq}\mathcal{F}^4_{mp}k^2_q + (W^3\gamma_m W^4)(k^2\cdot k^3)\big)\Big],
\end{aligned}
\tag{3.5}
$$

and can be checked to satisfy the BRST identities (3.3).

Due to the recursive definition of $L_{2131\ldots p1}$ care must be taken when discarding BRST-exact terms when evaluating the OPEs for the next $p+1$ step. For example, if the BRST-exact term in L_{21} is kept then it follows that [12]

$$\begin{aligned} L_{2131} &= \big[A^1_m(\lambda\gamma^m W^2) + V^1(k^1\cdot A^2)\big](k^{12}\cdot A^3) \\ &+(\lambda\gamma^m W^3)\big[A^1_m(k^1\cdot A^2) + A^{1\,n}\mathcal{F}^2_{mn} - (W^1\gamma_m W^2)\big] \\ &+s_{12}\big[(A^1W^3)V^2 - (A^2W^3)V^1\big] + (s_{13}+s_{23})(A^1W^2)V^3. \end{aligned} \tag{3.6}$$

Equation (3.3) implies that after discarding $Q(A^iW^j)$ from L_{ji} the last line of (3.6) must be discarded as well, in order for $QL_{2131} = s_{12}(L_{31}V_2 + V_1L_{32}) + (s_{13}+s_{23})L_{21}V_3$ continue to hold. Equivalently, one can consider the expressions in (3.5) as an explicit representation for composite superfields $L_{2131\ldots p1}$ which satisfy the BRST identities of (3.3).

It is worth mentioning that the BRST-exact terms dropped from L_{ji}, L_{jiki} and L_{jikili} were observed to cancel out in the final superspace expressions for the five- and six-point computations of [12,13]. This seems natural in view of the requirement that the overall amplitude should live in the BRST cohomology like its basic ingredients, the vertex operators. This will be the main idea to be exploited in the next subsection.

Furthermore, the energy-momentum tensor and the ghost-number current of (2.4) can be used together with the OPEs of (2.6) to show that the conformal weight h of $L_{2131\ldots p1}$ and its ghost number are given by,

$$h\,(L_{2131\ldots p1}) = (k^1+\cdots+k^n)^2 \neq 0, \qquad \text{ghost } \#(L_{2131\ldots p1}) = +1. \tag{3.7}$$

This will prove essential to argue that the BRST cohomology for composite superfields it generically empty.

3.2. Definition of BRST building blocks $T_{123\ldots p}$

The definition of a rank-q BRST building block $T_{123\ldots q}$ follows from two steps

$$L_{2131\ldots q1} \xrightarrow{(i)} \tilde{T}_{123\ldots q} \xrightarrow{(ii)} T_{123\ldots q} \tag{3.8}$$

which are designed to remove BRST-exact terms in $L_{2131\ldots q1}$ and in $\tilde{T}_{123\ldots q}$ while still preserving the fundamental BRST variation identities (3.3) when the combined redefinition $L_{2131\ldots q1} \longrightarrow T_{123\ldots q}$ is used in both sides of (3.3).

The first step (i) of (3.8) to obtain $\tilde{T}_{123\ldots q1}$ from the composite superfield $L_{2131\ldots q1}$ depends on all the previous redefinitions of $L_{2131\ldots p1}$ with $p < q$ which were made to get the BRST building blocks $T_{123\ldots p}$. Its purpose is to absorb the extra terms (in the left-hand side) when the substitutions $L_{2131\ldots p1} \to T_{123\ldots p}$ are made in the right-hand side of the BRST variation identity for $QL_{2131\ldots q1}$. Therefore the first step (i) ensures that $Q\tilde{T}_{123\ldots q}$ is written in terms of $T_{123\ldots p}$ rather than $L_{2131\ldots p1}$,

$$\begin{aligned} Q\tilde{T}_{123} &= s_{12}(T_{13}V_2 + V_1T_{23}) + (s_{13}+s_{23})T_{12}V_3 \\ Q\tilde{T}_{1234} &= (s_{14}+s_{24}+s_{34})T_{123}V_4 + (s_{13}+s_{23})(T_{12}T_{34}+T_{124}V_3) \\ &\quad + s_{12}(T_{134}V_2 + T_{13}T_{24} + T_{14}T_{23} + V_1T_{234}), \end{aligned} \tag{3.9}$$

and similarly for $\tilde{T}_{123\ldots q}$ with $q \geq 5$.

One can check using (3.9) that there are certain specific combinations of $\tilde{T}$'s which are BRST-closed, like for example $Q(\tilde{T}_{123}+\tilde{T}_{231}+\tilde{T}_{312}) = 0$. Furthermore, it was shown in (3.7) that the composite superfields $L_{2131\ldots p1}$ (and therefore also $\tilde{T}_{123\ldots p}$) have conformal weights $h \neq 0$, so those combinations must also be BRST-exact – because the cohomology of Q at ghost-number $+1$ is non-trivial only at zero conformal weight[4].

So the second step (ii) of (3.8) will involve searching for sums of $\tilde{T}_{123\ldots q}$ which are BRST-closed in order to subtract the corresponding BRST-exact parts from $\tilde{T}_{123\ldots q}$. In principle these sums can be found by a brute-force analysis of the identities in (3.9), but in subsection 3.4 a simple diagrammatic method to find all those sums will be presented. That in turn allows one to obtain the explicit expressions for all $q-1$ BRST-exact parts $R^{(I)}_{123\ldots q}$ of $\tilde{T}_{123\ldots q}$;

$$\sum \tilde{T}_{123\ldots q} = QR^{(I)}_{123\ldots q}, \quad I = 1, 2, 3, \ldots, q-1, \tag{3.10}$$

where the $q-1$ different sums will involve different label permutations of $\tilde{T}_{123\ldots q}$ with $\pm$ signs, see subsection 3.4 for their precise forms.

The prescription to remove the BRST-exact parts from $\tilde{T}_{123\ldots q}$ – which completes the second step (ii) of (3.8) – will be explained in subsection 3.5. After doing that, the previous BRST-closed sums of $\tilde{T}_{123\ldots q}$ become BRST-symmetries of the building blocks $T_{123\ldots q}$, i.e.,

$$\sum T_{123\ldots q} = 0. \tag{3.11}$$

[4] We thank Nathan Berkovits for illuminating discussions on this point.

In summary, the two steps in (3.8) are:

(i) Redefine $L_{2131\ldots q1} \to \tilde{T}_{123\ldots q}$ such that $Q\tilde{T}_{123\ldots q}$ is expressed in terms of building blocks $T_{123\ldots p}$ of lower-level $p < q$.

(ii) Remove the BRST-exact parts of $\tilde{T}_{123\ldots q}$ given by (3.10) such that $T_{123\ldots q}$ satisfies the symmetry properties (3.11).

The composite superfields $T_{123\ldots q}$ defined in this way are the BRST building blocks and obey the following identities,

$$\begin{aligned}
QT_{12} =& \, s_{12}V_1V_2,\\
QT_{123} =& \, (s_{13}+s_{23})T_{12}V_3 + s_{12}(T_{13}V_2 + V_1T_{23}),\\
QT_{1234} =& \, (s_{14}+s_{24}+s_{34})T_{123}V_4 + (s_{13}+s_{23})(T_{12}T_{34} + T_{124}V_3)\\
&+ s_{12}(T_{134}V_2 + T_{13}T_{24} + T_{14}T_{23} + V_1T_{234}),\\
QT_{12345} =& \, (s_{15}+s_{25}+s_{35}+s_{45})T_{1234}V_5 + (s_{14}+s_{24}+s_{34})(T_{1235}V_4 + T_{123}T_{45})\\
&+ (s_{13}+s_{23})(T_{1245}V_3 + T_{124}T_{35} + T_{125}T_{34} + T_{12}T_{345})\\
&+ s_{12}(T_{1345}V_2 + V_1T_{2345} + T_{134}T_{25} + T_{135}T_{24} + T_{145}T_{23}\\
&+ T_{13}T_{245} + T_{14}T_{235} + T_{15}T_{234})
\end{aligned} \tag{3.12}$$

and so forth. The relations (3.12) can be generalized as follows,

$$QT_{12\ldots n} = \sum_{j=2}^{n} \sum_{\alpha\in P(\beta_j)} (s_{1j}+s_{2j}+\cdots+s_{j-1,j})\, T_{12\ldots j-1,\{\alpha\}}\; T_{j,\{\beta_j\setminus\alpha\}}, \tag{3.13}$$

where $\beta_j = \{j+1,\ldots,n\}$, $P(\beta_j)$ is the powerset of β_j and $V_i \equiv T_i$. Furthermore, the first few BRST symmetries of (3.11) are given by

$$\begin{aligned}
0 =& \, T_{12} + T_{21},\\
0 =& \, T_{123} + T_{231} + T_{312},\\
0 =& \, T_{1234} - T_{1243} + T_{3412} - T_{3421},\\
0 =& \, T_{12345} - T_{12354} + T_{12543} - T_{12453} + T_{45321} - T_{45312},
\end{aligned} \tag{3.14}$$

where each higher-order building block $T_{123\ldots q}$ inherits all the lower-order identities in its first $q-1$ labels (this can be seen from the recursive definition of $L_{2131\ldots p1}$ in (3.1)). For example, T_{1234} not only satisfies the third equation of (3.14) but also the previous two in the form of $T_{1234} + T_{2134} = T_{1234} + T_{2314} + T_{3124} = 0$. Using the diagrammatic

method explained below, the following general BRST symmetries for building blocks will be derived,

$$\begin{aligned} p = 2n+1: \quad & T_{12\ldots n+1[n+2[\ldots[2n-1[2n,2n+1]]\ldots]]} - 2T_{2n+1\ldots n+2[n+1[\ldots[3[21]]\ldots]]} = 0 \\ p = 2n: \quad & T_{12\ldots n[n+1[\ldots[2n-2[2n-1,2n]]\ldots]]} + T_{2n\ldots n+1[n[\ldots[3[21]]\ldots]]} = 0. \end{aligned} \tag{3.15}$$

The notation $[i[jk]]$ means consecutive antisymmetrization of pairs of labels starting from the outermost label, e.g. $[i[jk]] = 1/2(i[jk] - [jk]i) = 1/4(ijk - ikj - jki + kji)$

3.3. Diagrammatic interpretation of $T_{123\ldots p}$ building blocks

As discussed in [14], every color-ordered tree-level field theory amplitude can be arranged into a form which manifests the kinematic poles that appear,

$$A_{YM}(1,2,\ldots,N) = \sum_i \frac{n_i}{\prod_{\alpha_i} p_{\alpha_i}^2} \tag{3.16}$$

where the sum is over the set of $(2N-4)!/((N-1)!(N-2)!)$ diagrams with only cubic vertices, n_i represent some kinematic numerator factor and $p_{\alpha_i}^2$ are the propagators of each diagram. Using this representation for the N–point amplitudes it was suggested in [26] that the BRST cohomology of the pure spinor formalism might be enough to fix the ten-dimensional SYM amplitudes, bypassing the need to perform the $\alpha' \to 0$ limit of their corresponding open superstring amplitudes. To that end it is useful to require that the numerator factors n_i have BRST transformations which are proportional to the Mandelstam invariants associated to their poles, $Qn_i = \sum_j p_{\alpha_j}^2 m_j$ for some m_j. This makes sure that each term in Qn_i cancel one of the poles and different terms can be concocted to yield an overall BRST-closed amplitude. So in order for the empirical cohomology method of [26] to work, one needs to have explicit mappings between cubic diagrams and ghost-number *three* pure spinor superspace expressions. Although some lower-order examples were presented in [26], a general solution was still missing. But as it became clear later, it is better to have mappings between cubic diagrams and ghost-number *one* composite superfields; the BRST building blocks. This realization led to the discovery in [1] of a general recursive method to construct expressions in the cohomology of the BRST charge with the correct properties of N–point SYM amplitudes. So in this section we describe in detail the solution of [1] to find the general dictionary between cubic-vertex diagrams and ghost-number one pure spinor building blocks.

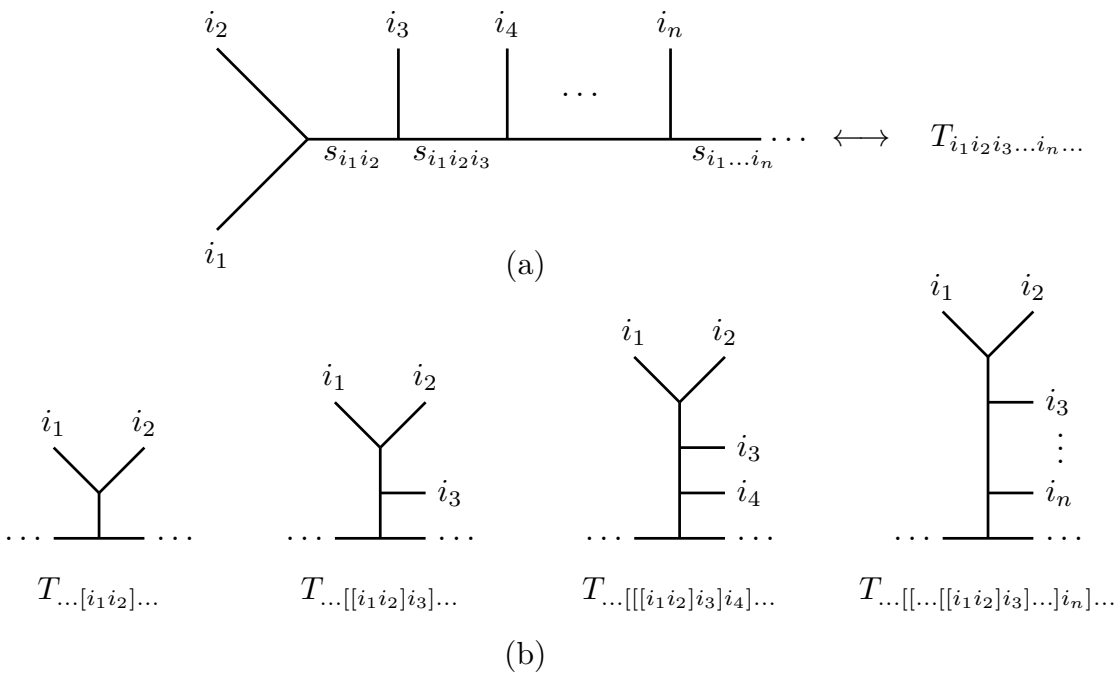


Fig. 1 (a) A tail-end cubic diagram with kinematic poles $\{s_{i_1 i_2}, \ldots, s_{i_1 i_2 \ldots i_n}\}$ corresponds to the building block $T_{i_1 i_2 \ldots i_n \ldots}$. (b) Branches of cubic diagrams and their associated building blocks. The motivation behind this dictionary lies on the fact that all kinematic invariants specified by the cubic graphs are present in the BRST variation of their corresponding building blocks.

The idea to obtain the dictionary is to find the precise sums of building blocks whose BRST variation contains the same set of Mandelstam variables associated to a particular cubic diagram. And this problem can be solved by understanding the patterns present in the BRST variation identities of (3.13).

To see this consider the diagram (a) of Fig. 1 where one leg has been removed and which contains the set of kinematic poles $\{s_{i_1 i_2}, s_{i_1 i_2 i_3}, \ldots, s_{i_1 \ldots i_n}\}$. From equation (3.13) one checks that all terms in the BRST variation of $T_{i_1 i_2 i_3 \ldots i_n \ldots}$ contain at least one of those Mandelstam variables without exception, schematically

$$QT_{i_1 i_2 i_3 \ldots i_n \ldots} \longrightarrow \{s_{i_1 i_2}, s_{i_1 i_2 i_3}, \ldots, s_{i_1 i_2 i_3 \ldots i_n \ldots}\} \tag{3.17}$$

where the trailing dots on the labels of the building block correspond to the amputated part of the diagram. Given this match, we associate the building block of (3.17) to the cubic graph of Fig. 1 (a).

To find the appropriate BRST building blocks which can be associated with the *branches* containing two amputated legs in Fig. 1 (b), note the pattern that certain sums of $T_{123 \ldots p}$

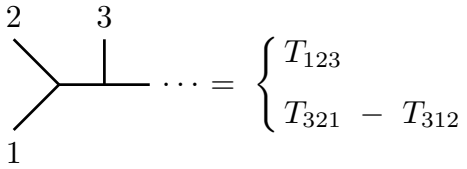


Fig. 2 Two different ways to interpret the same diagram give rise to an identity for T_{ijk}. In the first expression it is viewed as a tail-end graph, while in the second it is interpreted as a branch.

with different label orderings have a different set of Mandelstam invariants in their BRST variation. As seen on (3.17), the BRST variation of $T_{i_1 i_2 \ldots i_n}$ contains all elements of the set $\{s_{i_1 i_2}, s_{i_1 i_2 i_3}, \ldots, s_{i_1 \ldots i_n}\}$ but antisymmetrization in certain labels replaces some elements by others, e.g.

$$\begin{aligned} QT_{i_1 \ldots i_p [jk] r_1 \ldots r_q} &\longrightarrow s_{jk} \text{ instead of } s_{i_1 i_2 \ldots i_p j} \\ QT_{i_1 \ldots i_p [j[kl]] r_1 \ldots r_q} &\longrightarrow s_{kl}, s_{jkl} \text{ instead of } s_{i_1 \ldots i_p j}, s_{i_1 \ldots i_p jk} \\ QT_{i_1 \ldots i_p [j[k[lm]]] r_1 \ldots r_q} &\longrightarrow s_{lm}, s_{klm}, s_{jklm} \text{ instead of } s_{i_1 \ldots i_p j}, s_{i_1 \ldots i_p jk}, s_{i_1 \ldots i_p jkl} \ , \end{aligned} \tag{3.18}$$

where the two sets of dots in the building blocks correspond to the amputated parts of the graphs (b) in Fig. 1. The patterns shown in (3.18) therefore justify the general dictionary given in Fig. 1(b).

3.4. BRST symmetries of building blocks

It is not difficult to use the BRST variations of $\tilde{T}_{123\ldots q}$ in (3.9) to find their BRST-closed sums for small q by trial and error. Since the cohomology at conformal weight $h \neq 0$ is empty, these same BRST-closed combinations of $\tilde{T}$'s are also BRST-exact. As explained in the previous subsection, the removal of the BRST-exact parts of $\tilde{T}_{123\ldots q}$ gives rise to the definition of the building block $T_{123\ldots q}$ and at the same time the BRST-closed sum of $\tilde{T}$'s translates into a symmetry of the associated $T_{12\ldots n}$ (see equation (3.11)). Therefore it is imperative to find the general BRST-closed sums of $\tilde{T}$'s, or equivalently, the general symmetries of T's.

So in this subsection we use the diagrammatic interpretation of building blocks to predict the symmetry properties of $T_{12\ldots n}$ which in turn allow the BRST-exact parts of $\tilde{T}_{123\ldots n}$ to be found (see subsection 3.5).

As a first example, consider the diagram of Fig. 2. In the first expression the diagram is interpreted as a tail-end graph like the one depicted in (a) of Fig. 1 and is associated with

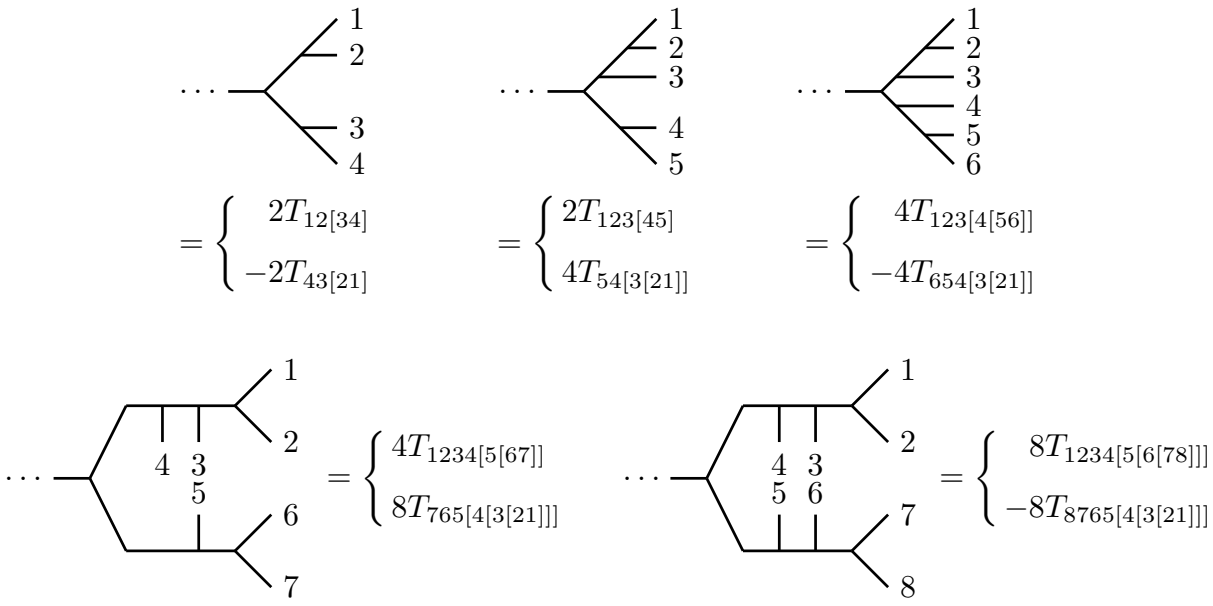


Fig. 3 Diagrammatic derivation of the BRST symmetries of higher-order building blocks. The top (bottom) line corresponds to the building block association which follow from reading the diagram in a counter-clockwise (clockwise) direction.

the building block T_{123}. However, in the second expression the diagram is viewed as a branch like the first graph of (b) in Fig. 1, where one of the "missing" legs now contains the label 3 and it is therefore associated with $2T_{3[21]} = T_{321} - T_{312}$. The fact that both interpretations have to agree implies the symmetry identity (3.14) for T_{ijk},

$$0 = T_{123} - T_{321} + T_{312} = T_{123} + T_{231} + T_{312}.$$

The relative sign between the two viewpoints is fixed by the fact that diagram associated with $T_{12\ldots n}$ catch a $(-1)^{n-1}$ sign under inversion $(1, 2, 3, \ldots, n-1, n) \leftrightarrow (n, n-1, \ldots, 1)$. Hence, we have to make sure that the sign of $T_{123\ldots n}$ relative to $T_{n,n-1,\ldots 21}$ is $(-1)^n$ in (3.11), e.g. $T_{123} + (-1)^3 T_{321} + \cdots = 0$.

This same idea can be used to obtain the BRST symmetries for higher-order building blocks. For example, the symmetries of $T_{123\ldots n}$ for $n = 4$, 5, 6, 7, 8 are obtained from the diagrams of Fig. 6,

$$\begin{aligned} 0 &= 2T_{12[34]} + 2T_{43[21]}, \\ 0 &= 2T_{123[45]} - 4T_{54[3[21]]}, \\ 0 &= 4T_{123[4[56]]} + 4T_{654[3[21]]}, \\ 0 &= 4T_{1234[5[67]]} - 8T_{765[4[3[21]]]}, \\ 0 &= 8T_{1234[5[6[78]]]} + 8T_{8765[4[3[21]]]}. \end{aligned} \tag{3.19}$$

Using the BRST variations (3.12) we checked up to $T_{12345678}$ that these relations are indeed BRST-closed and obtained their explicit BRST-exact parts for up to $\tilde{T}_{12345}$. The latter was made using the explicit expressions of $\tilde{T}_{123\ldots p}$ in terms of super Yang–Mills superfields to find the explicit solutions $R^{(n)}_{123\ldots p}$ of equation (3.10), and that will be presented in the next section.

To write down the generalization of (3.19) to higher $p > 8$, let us distinguish between odd and even ranks for ease of notation:

$$\begin{aligned} p = 2n+1: &\quad T_{12\ldots n+1[n+2[\ldots[2n-1[2n,2n+1]]\ldots]]} - 2T_{2n+1\ldots n+2[n+1[\ldots[3[21]]\ldots]]} = 0 \\ p = 2n: &\quad T_{12\ldots n[n+1[\ldots[2n-2[2n-1,2n]]\ldots]]} + T_{2n\ldots n+1[n[\ldots[3[21]]\ldots]]} = 0 \end{aligned} \tag{3.20}$$

The relations for $p = 2n+1$ and $p = 2n$ involve $3 \cdot 2^{n-1}$ and 2^n terms, respectively.

We should emphasize again that the lower rank identities for $T_{12\ldots q}$ carry over to $T_{12\ldots p}$ with $p > q$. The last labels $q+1, \ldots, p$ are then simply left untouched, e.g. $0 = T_{(12)345} = T_{[123]45} = T_{12[34]5} + T_{43[21]5}$ at rank $p = 5$. By applying the $p-1$ symmetries available at rank p, one can successively move a particular label to the first position, i.e. express $T_{i_1 i_2 \ldots i_p}$ as a combination of $T_{1 j_1 j_2 \ldots j_{p-1}}$. Hence, there are $(p-1)!$ independent rank-p building blocks $T_{i_1 i_2 \ldots i_p}$.

3.5. Explicit construction of $T_{12\ldots p}$

The definition of the first BRST building block T_{12} requires only the step (ii) in (3.8), as there are no lower-order redefinitions to take into account in the first step (i); that is $\tilde{T}_{12} \equiv L_{21}$. From the BRST variation of $\tilde{T}_{12}$ in (3.3) together with the equations of motion (2.9) one sees that its symmetric part is BRST-closed; $Q(\tilde{T}_{21}+\tilde{T}_{12}) = s_{12}(V_1V_2+V_2V_1) = 0$, and also BRST-exact [26]

$$\tilde{T}_{21} + \tilde{T}_{12} = -Q(A^1 \cdot A^2) \equiv -QD_{12}. \tag{3.21}$$

As discussed in (3.11), the definition of the BRST building block T_{12} must be made to satisfy $T_{12} + T_{21} = 0$. This is accomplished by

$$T_{12} = \tilde{T}_{[21]} = \tilde{T}_{21} + \frac{1}{2} QD_{12}. \tag{3.22}$$

The definition of the building block T_{123} now proceeds using both steps of (3.8). The first redefinition $L_{2131} \overset{(i)}{\to} \tilde{T}_{123}$ is found by substituting $L_{ji} = \tilde{T}_{ij} = T_{ij} - \frac{1}{2}QD_{ij}$ in the right-hand side of QL_{2131} in (3.3), which leads to:

$$\begin{aligned} &Q\left(L_{2131} + \frac{1}{2}s_{12}\big[D_{13}V_2 - D_{23}V_1\big] + \frac{1}{2}(s_{13}+s_{23})D_{12}V_3\right) \\ &\qquad = s_{12}(T_{13}V_2 + V_1T_{23}) + (s_{13}+s_{23})T_{12}V_3. \end{aligned}$$

Therefore by defining

$$\tilde{T}_{123} = L_{2131} + \frac{1}{2}s_{12}\big[D_{13}V_2 - D_{23}V_1\big] + \frac{1}{2}(s_{13}+s_{23})D_{12}V_3, \tag{3.23}$$

one obtains the desired identity $Q\tilde{T}_{123} = s_{12}(T_{13}V_2 + V_1T_{23}) + (s_{13}+s_{23})T_{12}V_3$.

Two BRST-closed combinations of $\tilde{T}_{ijk}$ are easily identified,

$$Q(\tilde{T}_{123} + \tilde{T}_{213}) = 0, \quad Q(\tilde{T}_{123} + \tilde{T}_{312} + \tilde{T}_{231}) = 0, \tag{3.24}$$

and one can show using SYM equations of motion (2.9) that they originate as the BRST variation of ghost number zero superfields $R^{(1)}_{123}$, $R^{(2)}_{123}$ [13,1]

$$\tilde{T}_{123} + \tilde{T}_{213} = QR^{(1)}_{123}, \quad \tilde{T}_{123} + \tilde{T}_{312} + \tilde{T}_{231} = QR^{(2)}_{123}, \tag{3.25}$$

where $R^{(1)}_{123} = D_{12}(k^{12}\cdot A^3)$, $\quad R^{(2)}_{123} = D_{12}(k^2\cdot A^3) + \text{cyclic}(123)$. The BRST building block T_{123} is obtained by removing these BRST-exact pieces

$$T_{123} = \tilde{T}_{123} - QS^{(1)}_{123}, \quad S^{(1)}_{123} = \frac{1}{2}R^{(1)}_{123} + \frac{1}{3}R^{(2)}_{[12]3}, \tag{3.26}$$

which implies the following BRST symmetries for T_{ijk}:

$$T_{123} + T_{213} = T_{123} + T_{312} + T_{231} = 0. \tag{3.27}$$

The definition of T_{1234} is done similarly and uses the information from the lower-order redefinitions of L_{21} and L_{2131}. First one rewrites L_{ji} and L_{jiki} in terms of T_{ij} and T_{ijk} in the RHS of the identity for QL_{213141} given in (3.3). After some algebra one finds

$$\begin{aligned}\tilde{T}_{1234} = \; & L_{213141} - \frac{1}{4}\big[(s_{13}+s_{23})D_{12}QD_{34} + s_{12}\,(D_{13}QD_{24} + D_{14}QD_{23})\,\big] \\ & + \frac{1}{2}\big[(s_{13}+s_{23})\,(D_{12}T_{34} - D_{34}T_{12}) + s_{12}\,(D_{13}T_{24} + D_{14}T_{23} - D_{23}T_{14} - D_{24}T_{13})\,\big] \\ & - (s_{14}+s_{24}+s_{34})S^{(1)}_{123}V_4 - (s_{13}+s_{23})S^{(1)}_{124}V_3 + s_{12}(S^{(1)}_{234}V_1 - S^{(1)}_{134}V_2)\end{aligned} \tag{3.28}$$

which satisfies the required property of

$$\begin{aligned}Q\tilde{T}_{1234} = \; & s_{12}(T_{134}V_2 + T_{13}T_{24} + T_{14}T_{23} + V_1T_{234}) \\ & + (s_{13}+s_{23})(T_{12}T_{34} + T_{124}V_3) + (s_{14}+s_{24}+s_{34})T_{123}V_4.\end{aligned} \tag{3.29}$$

Using (3.29) it is easy to check that the lower-order identities of $\tilde{T}_{123}$ given by (3.24) are inherited by the first three labels of $\tilde{T}_{1234}$ and that there is one additional BRST identity involving the fourth label,

$$Q\left(\tilde{T}_{1234}+\tilde{T}_{2134}\right)=Q\left(\tilde{T}_{1234}+\tilde{T}_{3124}+\tilde{T}_{2314}\right)=Q\left(\tilde{T}_{1234}-\tilde{T}_{1243}+\tilde{T}_{3412}-\tilde{T}_{3421}\right)=0,$$

in accord with the discussions of section 3.4. Using the SYM equations of motion in a long sequence of calculations shows that these combinations are indeed BRST-exact,

$$\begin{aligned}\tilde{T}_{1234}+\tilde{T}_{2134}&=QR^{(1)}_{1234}\\ \tilde{T}_{1234}+\tilde{T}_{3124}+\tilde{T}_{2314}&=QR^{(2)}_{1234}\\ \tilde{T}_{1234}-\tilde{T}_{1243}+\tilde{T}_{3412}-\tilde{T}_{3421}&=QR^{(3)}_{1234},\end{aligned}\tag{3.30}$$

where

$$\begin{aligned}R^{(1)}_{1234}=&-R^{(1)}_{123}(k^{123}\cdot A^4)-\frac{1}{4}s_{12}\big[D_{13}D_{24}+D_{14}D_{23}\big],\\ R^{(2)}_{1234}=&-R^{(2)}_{123}(k^{123}\cdot A^4)-\frac{1}{4}\big[s_{12}D_{23}D_{14}+s_{23}D_{24}D_{13}+s_{13}D_{34}D_{12}\big],\\ R^{(3)}_{1234}=&(k^1\cdot A^2)\big[D_{14}(k^4\cdot A^3)-D_{13}(k^3\cdot A^4)\big]-(k^2\cdot A^1)\big[D_{24}(k^4\cdot A^3)-D_{23}(k^3\cdot A^4)\big]\\ &+\frac{1}{4}D_{12}D_{34}(s_{14}+s_{23}-s_{13}-s_{24})+D_{12}\big[(k^4\cdot A^3)(k^2\cdot A^4)-(k^3\cdot A^4)(k^2\cdot A^3)\big]\\ &+D_{34}\big[(k^2\cdot A^1)(k^4\cdot A^2)-(k^1\cdot A^2)(k^4\cdot A^1)\big]+(W^1\gamma^mW^2)(W^3\gamma_mW^4).\end{aligned}\tag{3.31}$$

Removing these BRST-exact parts leads to the rank-four BRST building block—which is accomplished with the second redefinition $\tilde{T}_{1234}\xrightarrow{(ii)}T_{1234}$,

$$T_{1234}=\tilde{T}_{1234}-QS^{(2)}_{1234},\tag{3.32}$$

where $S^{(2)}_{1234}$ is defined recursively by

$$\begin{aligned}S^{(2)}_{1234}&=\frac{3}{4}S^{(1)}_{1234}+\frac{1}{4}(S^{(1)}_{1243}-S^{(1)}_{3412}+S^{(1)}_{3421})+\frac{1}{4}R^{(3)}_{1234}\\ S^{(1)}_{1234}&=\frac{1}{2}R^{(1)}_{1234}+\frac{1}{3}R^{(2)}_{[12]34}.\end{aligned}\tag{3.33}$$

To see that (3.32) and (3.33) imply the BRST symmetries of

$$T_{1234}+T_{2134}=T_{1234}+T_{3124}+T_{2314}=T_{1234}-T_{1243}+T_{3412}-T_{3421}=0,\tag{3.34}$$

it suffices to check that the following identities hold,

$$\begin{aligned} S^{(2)}_{1234} + S^{(2)}_{2134} &= R^{(1)}_{1234} \\ S^{(2)}_{1234} + S^{(2)}_{3124} + S^{(2)}_{2314} &= R^{(2)}_{1234} \\ S^{(2)}_{1234} - S^{(2)}_{1243} + S^{(2)}_{3412} - S^{(2)}_{3421} &= R^{(3)}_{1234}. \end{aligned} \tag{3.35}$$

Following this same procedure for $L_{21314151}$ is straightforward but somewhat tedious, therefore the calculations leading to the explicit superfield expression for the building block T_{12345} will be deferred to the Appendix A.

As will be explained in subsection 4.4, the explicit superfield expressions for T_{ij}, T_{ijk}, T_{ijkl} and T_{ijklm} allows one to obtain the expansions of any superstring or field-theory amplitudes up to $N = 11$ legs in terms of momenta and polarization [31].

4. Supersymmetric Berends–Giele recursions

In subsection 3.3 we have given a superfield representation in terms of $T_{i_1\ldots i_p}$ for each color ordered diagram made of cubic vertices with p on-shell external leg and one off-shell leg. In this section, we combine these diagrams to $p+1$ point field theory amplitudes with one off-shell leg. These objects were firstly considered in [17] in order to derive recursion relations for gluon scattering at tree-level and were referred to as "currents". The pure spinor supersymmetric analogue of the p-point Berends–Giele current J_p will be referred to as $M_{12\ldots p}$.

These $M_{12\ldots p}$ allow for a compact representation of the ten-dimensional N–point SYM amplitude $\mathcal{A}_{YM}(1,\ldots,N)$ which nicely exhibits its factorization channels. The recursive nature of the Berends–Giele currents is inherited by the amplitudes and leads to the recursive method to compute higher-point SYM amplitudes presented below.

4.1. Construction of Berends–Giele currents $M_{123\ldots p}$

The Berends–Giele currents $M_{123\ldots p}$ are written in terms of building blocks $T_{123\ldots p}$ and Mandelstam invariants $\{s_{12}, s_{123}, \ldots, s_{123\ldots p}\}$ and follow from the recursive definition

$$\begin{aligned} E_{123\ldots p} &\equiv \sum_{j=1}^{p-1} M_{12\ldots j} M_{j+1\ldots p} \\ QM_{123\ldots p} &\equiv E_{123\ldots p}, \end{aligned} \tag{4.1}$$

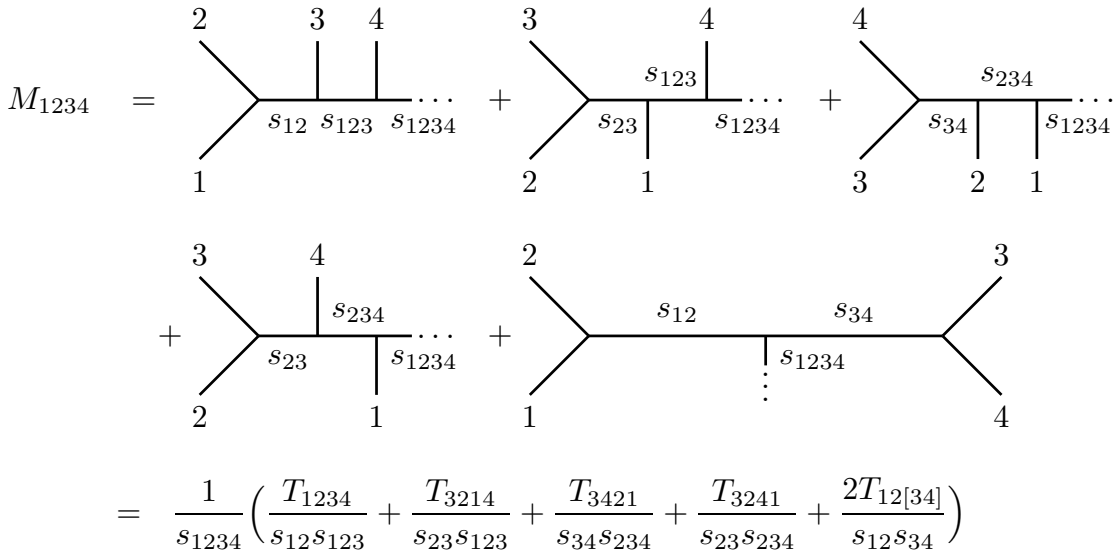
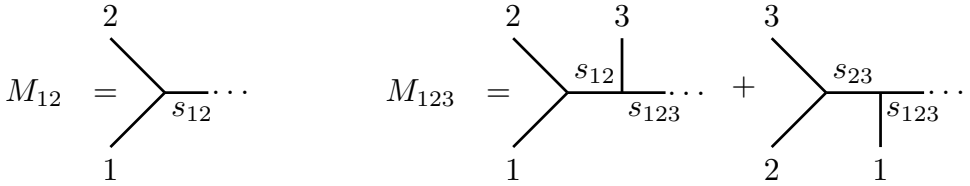


Fig. 4 Diagrammatic construction of the Berends–Giele current M_{1234} in terms of the cubic graphs of the five-point amplitude with one leg off-shell.

where $M_1 = V_1$. Although the defining system (4.1) is purely algebraic, it can be conveniently solved with the recourse of a diagrammatic interpretation for $M_{123\ldots p}$. To see this, the current $M_{123\ldots p}$ is first associated to the sum of $(2p-2)!/(p!(p-1)!)$ cubic graphs which enter the $p+1$ amplitude where the leg $p+1$ is put off-shell. Using the dictionary of subsection 3.3 each one of these cubic graphs can be written in terms of building blocks $T_{123\ldots p}$ and their relative signs are fixed by requiring the system (4.1) to be satisfied. For example, using the cubic graphs for the three- and four-point amplitudes the currents M_{12} and M_{123} are interpreted as



while M_{1234} is associated to the graphs of the color-ordered five-point amplitude shown in Fig. 4. Under the dictionary of subsection 3.3 these graphs correspond to the following expressions in terms of building blocks

$$M_{12} = \frac{T_{12}}{s_{12}}\,, \qquad M_{123} = \frac{1}{s_{123}}\left(\frac{T_{123}}{s_{12}} + \frac{T_{321}}{s_{23}}\right), \tag{4.2}$$
$$M_{1234} = \frac{1}{s_{1234}}\left(\frac{T_{1234}}{s_{12}s_{123}} + \frac{T_{3214}}{s_{23}s_{123}} + \frac{T_{3421}}{s_{34}s_{234}} + \frac{T_{3241}}{s_{23}s_{234}} + \frac{2T_{12[34]}}{s_{12}s_{34}}\right),$$

where their signs can be fixed by requiring that they form a solution of (4.1). To see this one uses the BRST variations (3.13) to obtain

$$\begin{aligned} QM_{12} &= V_1V_2 = M_1M_2, \\ QM_{123} &= \frac{V_1T_{23}}{s_{23}} + \frac{T_{12}V_3}{s_{12}} = M_1M_{23} + M_{12}M_3, \\ QM_{1234} &= \frac{V_1}{s_{234}}\Big(\frac{T_{234}}{s_{23}} + \frac{T_{432}}{s_{34}}\Big) + \frac{T_{12}T_{34}}{s_{12}s_{34}} + \Big(\frac{T_{123}}{s_{12}} + \frac{T_{321}}{s_{23}}\Big)\frac{V_4}{s_{123}} \\ &= M_1M_{234} + M_{12}M_{34} + M_{123}M_4 \end{aligned} \tag{4.3}$$

and therefore the expressions for M_{12}, M_{123} and M_{1234} given above form a solution of the system (4.1) up to this order. Using this method it is straightforward to obtain higher-point currents, and the explicit expressions of currents up to $M_{1234567}$ will be given in the Appendix B.

Therefore by using the diagrammatic interpretation of $M_{123\ldots p}$ in terms of the $p+1$ amplitude with one leg off-shell one is able to efficiently construct any higher-order current in terms of building blocks. However, in the later section 5.2 we will derive a formula for $M_{123\ldots p}$ in terms of the field-theory limit $\alpha' \to 0$ of hypergeometric integrals occurring in a $p+2$ point string theory amplitude. This allows for a direct computation of $M_{12\ldots p}$, therefore bypassing the need to draw the cubic diagrams of the $(p{+}1)$-point SYM amplitude to find their corresponding building blocks.

Note that (4.1) can be written as

$$QM_{12\ldots p} = \sum_{j=1}^{p-1} M_{12\ldots j}M_{j+1\ldots p} \tag{4.4}$$

and therefore one can interpret the action of Q as cutting $M_{12\ldots p}$ in each way compatible with the color ordering, see Fig. 5. Furthermore, equation (4.4) is the supersymmetric pure spinor analogue of the recursive construction of the Berends–Giele gluon currents in [17], whose schematic form is

$$J_n \sim \frac{1}{s_{12\ldots n}}\left(\sum_{m=1}^{n-1} J_m, J_{n-m} + \sum_{m=1}^{n-2}\sum_{k=m+1}^{n-1} J_m J_{k-m} J_{n-k}\right). \tag{4.5}$$

The cubic term in the lower order currents represents the four gluon vertex in the QCD action. It does not enter into our supersymmetric version (4.4) which encompasses diagrams with cubic vertices only. After multiplying the external propagator $1/s_{12\ldots n}$ to the left hand side of (4.5) one could symbolically reproduce (4.4) by identifying $s_{12\ldots n} \equiv Q$.

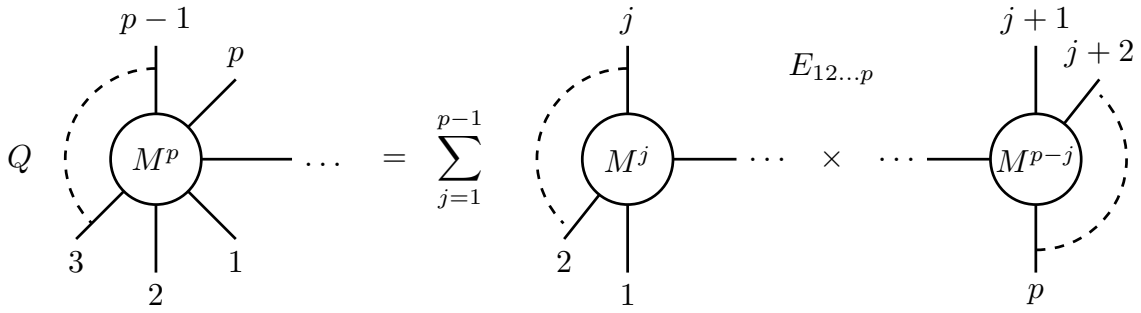


Fig. 5 Decomposition of $M_{12\ldots p}$ into its factorization channels under the action of the pure spinor BRST charge; $QM_{12\ldots p} = \sum_{j=1}^{p-1} M_{12\ldots j} M_{j+1\ldots p}$.

4.2. Symmetry properties of $M_{12\ldots p}$

As a further motivation for identifying $M_{12\ldots p}$ with supersymmetric Berends–Giele currents, we discuss their symmetry properties in this subsection. First of all, M_{12} trivially satisfies $M_{12} + M_{21} = 0$ because the building block T_{ij} is antisymmetric. Similar identities hold for M_{123}

$$M_{123} + M_{231} + M_{312} = 0, \qquad M_{123} - M_{321} = 0, \tag{4.6}$$

as one can easily check by plugging in the expression for M_{ijk} given in (4.2).

At higher $n \geq 4$, this generalizes as follows:

$$\begin{aligned} M_{12\ldots n} &= (-1)^{n-1} M_{n\ldots 21} \\ \sum_{\sigma\in\text{cyclic}} M_{\sigma(1,2,\ldots,n)} &= 0. \end{aligned} \tag{4.7}$$

The proof of these identities is most conveniently carried out on the level of the corresponding $E_{12\ldots n} = QM_{12\ldots n} = \sum_{p=1}^{n-1} M_{12\ldots p} M_{p+1\ldots n}$. Since all the BRST closed components of the $M_{12\ldots n}$ have been removed by construction of its $T_{12\ldots n}$ constituents, the BRST variation $E_{12\ldots n}$ contains all information on the symmetry properties of its $M_{12\ldots n}$ ancestor. The reflection identity can be easily checked by induction, and the vanishing cyclic sum follows from

$$\begin{aligned} \sum_{\sigma\in\text{cyclic}} E_{\sigma(1,2,\ldots,n)} &= \sum_{\sigma\in\text{cyclic}} \sum_{p=1}^{n-1} M_{\sigma(1,2,\ldots,p)} M_{\sigma(p+1,\ldots,n)} \\ &= \sum_{\sigma\in\text{cyclic}} \sum_{p=1}^{n-1} \frac{1}{2} (M_{\sigma(1,2,\ldots,p)} M_{\sigma(p+1,\ldots,n)} + M_{\sigma(p+1,\ldots,n)} M_{\sigma(1,2,\ldots,p)}) = 0 \end{aligned} \tag{4.8}$$

where the last step exploits the overall cyclic sum to shift all labels of the second term by p and that the $M_{12\ldots p}$ anticommute.

The properties (4.7) are shared by the n-gluon Berends–Giele currents J_n of [17] and can be naturally explained by the construction of currents $M_{123\ldots n}$ as $(n+1)$-point amplitudes with one off-shell leg. Inspired by this explanation, we explicitly checked using the expressions of Appendix B that $M_{12\ldots n}$ for $n \leq 7$ also satisfy an additional relation – obtained by removing the $(n+1)$-th leg from the $(n+1)$-point Kleiss-Kuijf identity [33]:

$$M_{\{\beta\},1,\{\alpha\}} = (-1)^{n_\beta} \sum_{\sigma \in \mathrm{OP}(\{\alpha\},\{\beta^T\})} M_{1,\{\sigma\}}. \tag{4.9}$$

The summation range $\mathrm{OP}(\{\alpha\},\{\beta^T\})$ denotes the set of all the permutations of $\{\alpha\} \bigcup \{\beta^T\}$ that maintain the order of the individual elements of both sets $\{\alpha\}$ and $\{\beta^T\}$. The notation $\{\beta^T\}$ represents the set $\{\beta\}$ with reversed ordering of its n_β elements. The Kleiss-Kuijf identity is well known to reduce the number of independent color ordered $n+1$ point amplitudes down to $(n-1)!$.

The specialization of (4.9) to sets $\{\beta\}$ with one element only, say $\{\beta\} = \{n\}$, reproduces the second property of (4.7). However, this so-called dual Ward identity or photon decoupling identity by itself is not sufficient for a reduction to $(n-1)!$ independent $M_{i_1 i_2 \ldots i_n}$ at $n \geq 6$ [33]. Since there are only $(n-1)!$ independent $T_{i_1 i_2 \ldots i_n}$ which constitute the $M_{i_1 i_2 \ldots i_n}$, also the latter must have a basis of no more than $(n-1)!$ elements. This suggests the Kleiss-Kuijf identity (4.9) to hold beyond our checks for $n \leq 7$.

The reflection- and Kleiss-Kuijf identity for the $M_{12\ldots n}$ are inherited from their associated $n+1$ point amplitudes with one leg off-shell. The off-shellness of one leg is no obstruction for the aforementioned identities to hold because they do not involve any kinematic factors. However, the field theory version of the monodromy relations [15,16]

$$\begin{aligned} s_{12}\,\mathcal{A}_{YM}(2,1,3,\ldots,N) + (s_{12}+s_{13})\mathcal{A}_{YM}(2,3,1,\ldots,N) + \cdots \\ + (s_{12}+\cdots+s_{1,N-1})\mathcal{A}_{YM}(2,3,\ldots,N-1,1,N) = 0 \end{aligned} \tag{4.10}$$

rely on having on-shell momenta, so the $M_{12\ldots n}$ do not obey any analogue of (4.10) and cannot be reduced to $(n-2)!$ independent permutations.

4.3. The N–point field-theory tree amplitude

The expressions found for $QM_{12\ldots p} = E_{12\ldots p}$ might look familiar from lower order field theory amplitudes such as

$$\begin{aligned} \mathcal{A}_{YM}(1,2,3) &= \langle V_1 V_2 V_3 \rangle = \langle E_{12} V_3 \rangle \\ \mathcal{A}_{YM}(1,2,3,4) &= \langle \left(\frac{V_1 T_{23}}{s_{23}} + \frac{T_{12} V_3}{s_{12}} \right) V_4 \rangle = \langle E_{123} V_4 \rangle \end{aligned} \tag{4.11}$$

From $QV = 0$, one might naively expect that the three-point amplitude would be BRST-exact, $\mathcal{A}(1,2,3) = \langle Q(T_{12}V_3/s_{12}) \rangle$, and thus doomed to vanish. However, all Mandelstam invariants s_{ij} vanish in the momentum phase space of three massless particles – therefore writing $V_1 V_2 = Q(T_{12}/s_{12})$ is not allowed and BRST triviality of the amplitude is avoided.

More generally, the prefactor $M_{12\ldots p} \sim 1/s_{12\ldots p}$ in the p point current is incompatible with putting the external state with $k_{p+1} = -\sum_{i=1}^{p} k_i$ on-shell $k_{p+1}^2 = 0$. Since N particle kinematics forbids the existence of $M_{12\ldots N-1}$, the corresponding $E_{12\ldots N-1}$ is not BRST exact. Hence, the following expression for the N–point field theory amplitude is in the cohomology of the pure spinor BRST charge[5][1]

$$\mathcal{A}_{YM}(1,2,\ldots,N) = \langle E_{12\ldots N-1} V_N \rangle = \sum_{j=1}^{N-2} \langle M_{12\ldots j} M_{j+1\ldots N-1} V_N \rangle. \tag{4.12}$$

The diagrammatic representation of $\sum_{j=1}^{p-1} M_{12\ldots j} M_{j+1\ldots p}$ in Fig. 5 can be uplifted to the on-shell $N = p+1$ point amplitude $\mathcal{A}_{YM}(1,\ldots,N)$ where an additional cubic vertex connects the N^{th} leg with the two currents of rank j and $N-1-j$, respectively.
The N–point formula (4.12) is analogous to the Berends–Giele formula for the color ordered N gluon amplitude of [17]. The latter is written as a product of a rank $N-1$ current J_{N-1} and another J_1 for the N^{th} leg, multiplied by the Mandelstam factor $s_{12\ldots N-1}$ to cancel the divergent propagator; $\mathcal{A}_{YM} = s_{12\ldots N-1} J(1,\ldots,N-1)\, J(N)$. In our case, the somewhat artificial object $s_{12\ldots N-1} J_{N-1}$ is replaced by $E_{12\ldots N-1}$, which could be written as $QM_{12\ldots N-1}$ in a larger momentum phase space. Therefore this parallel also suggests the schematic identification $s_{12\ldots N-1} \to Q$ mentioned after (4.5).

[5] It is interesting to note that the cohomology formula (4.12) together with the property of $E_{n(n-1)\ldots 1} = (-1)^{n-1} E_{12\ldots n}$ (which follows from (4.7)) imply that if the amplitude satisfies the reflection property of $\mathcal{A}(n,n-1,\ldots,1) = (-1)^n \mathcal{A}(1,2,\ldots,n)$ then it is also *cyclically* symmetric, $\mathcal{A}(2,3,\ldots,n,1) = \mathcal{A}(1,2,\ldots,n)$.

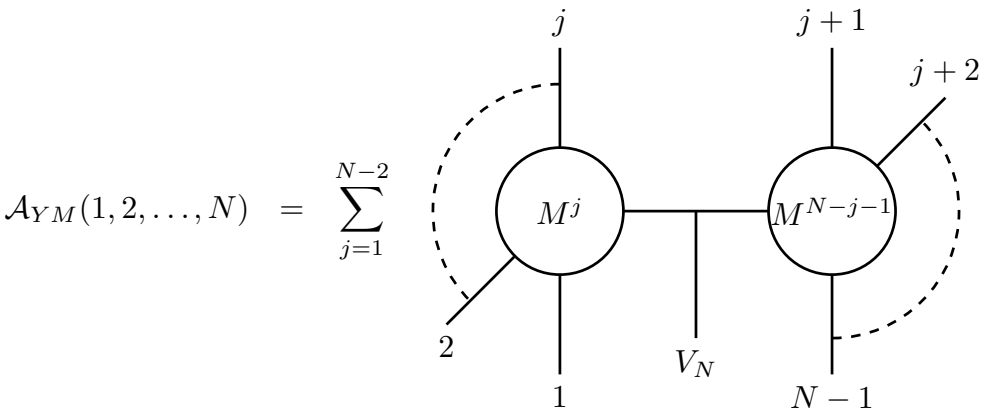


Fig. 6 Berends–Giele decomposition of $\mathcal{A}_{YM}$ according to the pure spinor cohomology formula (4.12).

4.4. BRST integration by parts and cyclic symmetry

The strength of our presentation (4.12) of the N–point field theory amplitude is the manifestation of its factorization properties. But singling out a particular leg V_N obscures the cyclic symmetry required for color stripped amplitudes. The essential tool to restore manifest cyclicity is BRST integration by parts,

$$\langle M_{i_1\ldots i_p} E_{j_1\ldots j_q}\rangle = \langle E_{i_1\ldots i_p} M_{j_1\ldots j_q}\rangle. \tag{4.13}$$

Using the definition of $E_{123\ldots p}$ in (4.1) it follows that,

$$E_{12\ldots N-1}V_N = E_{23\ldots N}V_1 + \sum_{j=2}^{N-2}\big(M_{12\ldots j}E_{j+1\ldots N} - E_{12\ldots j}M_{j+1\ldots N}\big), \tag{4.14}$$

therefore $\langle E_{12\ldots N-1}V_N\rangle = \langle E_{23\ldots N}V_1\rangle$ and the N–point subamplitude (4.12) is cyclically invariant. However, to obtain a formula with *manifest* cyclic symmetry one needs to explicitly use BRST integration by parts in (4.12). And as a byproduct of that, the maximum rank of the Berends–Giele currents needed for the N–point amplitude is reduced. To see this, note that the term containing the maximum rank of $M_{i_1\ldots i_p}$ appearing in the N–point amplitude (4.12) is $p = N-2$ and has the form $\langle M_{i_1\ldots i_{N-2}}V_{i_{N-1}}V_N\rangle$, therefore the use of (4.14) leads to

$$\langle M_{i_1\ldots i_{N-2}}V_{i_{N-1}}V_N\rangle = \langle M_{i_1\ldots i_{N-2}}QM_{i_{N-1}N}\rangle = \langle E_{i_1\ldots i_{N-2}}M_{i_{N-1}N}\rangle, \tag{4.15}$$

so the BRST integration reduced the maximum rank to $p = N-3$ (because $E_{12\ldots(N-2)}$ contains at most $M_{12\ldots N-3}$). It turns out that the cohomology formula (4.12) allows enough

BRST integration by parts as to reduce the maximum rank of the currents to $p = [N/2]$, leading to manifestly cyclic-symmetric amplitudes

$$\begin{aligned}
\mathcal{A}_{YM}(1,2,\ldots,5) &= \langle M_{12}V_3M_{45}\rangle \ + \ \text{cyclic}(12345)\\
\mathcal{A}_{YM}(1,2,\ldots,6) &= \frac{1}{3}\langle M_{12}M_{34}M_{56}\rangle + \frac{1}{2}\langle M_{123}E_{456}\rangle \ + \ \text{cyclic}(123456)\\
\mathcal{A}_{YM}(1,2,\ldots,7) &= \langle M_{123}M_{45}M_{67}\rangle + \langle V_1M_{234}M_{567}\rangle \ + \ \text{cyclic}(1234567)\\
\mathcal{A}_{YM}(1,2,\ldots,8) &= \langle M_{123}M_{456}M_{78}\rangle + \frac{1}{2}\langle M_{1234}E_{5678}\rangle \ + \ \text{cyclic}(12345678)
\end{aligned} \tag{4.16}$$

The fractional prefactors $\frac{1}{2}$ or $\frac{1}{3}$ compensate for the fact that cyclic orbits for particularly symmetric superfield kinematics are shorter than the number N of legs. At $N = 6$, for instance, $M_{12}M_{34}M_{56}$ has just one distinct cyclic image $M_{23}M_{45}M_{61}$, hence the full cyclic(123456) overcounts the occurring diagrams by a factor of three.

4.5. Factorization in cyclically symmetric form

In this subsection, we introduce a cyclically symmetric presentation of SYM amplitudes where their factorization into two Berends–Giele currents becomes even more obvious.

One can check by evaluating the BRST variations that the amplitudes in (4.16) can be equivalently written as

$$\begin{aligned}
\mathcal{A}_{YM}(1,2,\ldots,4) &= \frac{1}{2}\langle M_{12}QM_{34}\rangle + \text{cyclic}(1234)\\
\mathcal{A}_{YM}(1,2,\ldots,5) &= \frac{1}{4}\Big(\langle M_{12}QM_{345}\rangle + \langle M_{123}QM_{45}\rangle\Big) + \text{cyclic}(12345)\\
\mathcal{A}_{YM}(1,2,\ldots,6) &= \frac{1}{6}\Big(\langle M_{12}QM_{3456}\rangle + \langle M_{123}QM_{456}\rangle + \langle M_{1234}QM_{56}\rangle\Big) + \text{cyclic}(123456)\\
\mathcal{A}_{YM}(1,2,\ldots,7) &= \frac{1}{8}\Big(\langle M_{12}QM_{34567}\rangle + \langle M_{123}QM_{4567}\rangle + \langle M_{1234}QM_{567}\rangle\\
&\qquad + \langle M_{12345}QM_{67}\rangle\Big) \ + \ \text{cyclic}(1234567)\\
\mathcal{A}_{YM}(1,2,\ldots,8) &= \frac{1}{10}\Big(\langle M_{12}QM_{345678}\rangle + \langle M_{123}QM_{45678}\rangle + \langle M_{1234}QM_{5678}\rangle\\
&\qquad + \langle M_{12345}QM_{678}\rangle + \langle M_{123456}QM_{78}\rangle\Big) + \text{cyclic}(12345678)
\end{aligned} \tag{4.17}$$

Note that some terms in the formulæ are naively overcounted by a factor of 2 because the cyclic orbits of $\langle M_{12\ldots j}QM_{j+1\ldots N}\rangle$ and $\langle M_{12\ldots N-j}QM_{N-j+1\ldots N}\rangle$ are the same. The

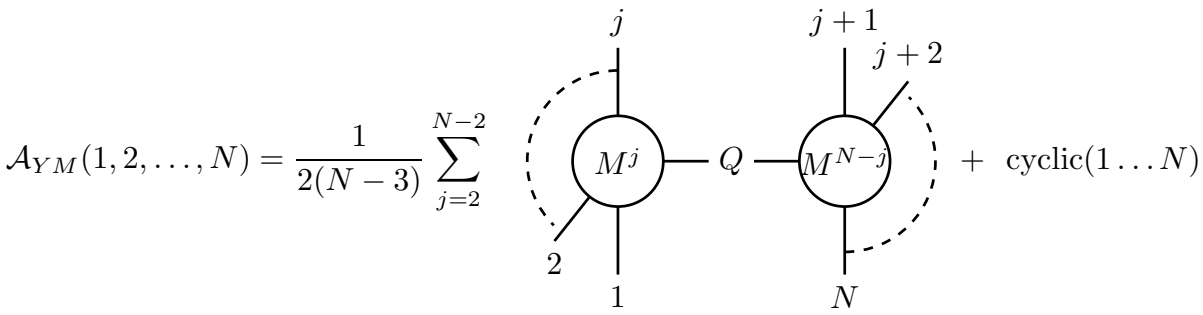


Fig. 7 Cyclic factorization of the N–point field-theory amplitude $\mathcal{A}_{YM}(1,2,\ldots,N)$ into different Berends–Giele partitions according to equation (4.18).

purpose of including both of them is to obtain a uniform overall coefficient in (4.17) and to simplify the transition to the general N–point formula,

$$\mathcal{A}_{YM}(1,2,\ldots,N) = \frac{1}{2(N-3)} \sum_{j=2}^{N-2} \langle M_{12\ldots j} Q M_{j+1\ldots N} \rangle \;+\; \text{cyclic}(1\ldots N) \tag{4.18}$$

whose graphical representation is shown in Fig. 7. We have explicitly checked up to $N = 10$ points that the formula (4.18) exactly reproduces the expression $\mathcal{A}_{YM} = \langle E_{12\ldots N-1} V_N \rangle$ of [1], including prefactors.

The factorization formula (4.18) can also be interpreted as coming from the factorization channels of two amplitudes with one leg x off-shell each with the form $\langle E_{12\ldots j} V_x \rangle$ and $\langle V_x E_{j+1\ldots N} \rangle$ that are connected by a pure spinor propagator which effectively replaces[6] $V_x\, V_x \to \frac{1}{Q}$, resulting in

$$\begin{aligned} \mathcal{A}_{YM}(1,2,\ldots,N) &= \frac{1}{2(N-3)} \sum_{j=2}^{N-2} \langle E_{12\ldots j} \frac{1}{Q}\, E_{j+1\ldots N} \rangle + \text{cyclic}(1\ldots N) \\ &= \frac{1}{2(N-3)} \sum_{j=2}^{N-2} \langle M_{12\ldots j} Q M_{j+1\ldots N} \rangle + \text{cyclic}(1\ldots N) \end{aligned}$$

which reproduces the formula (4.18).

5. The superstring tree amplitude in pure spinor superspace

This section derives our central result (5.22) for the superstring N point tree amplitude of the massless gauge multiplet. The BRST building blocks $T_{12\ldots p}$ and their combinations

[6] CM thanks Nathan Berkovits for suggesting back in 2006 how one could view an operation like $V_x\, V_x \to \frac{1}{Q}$ as possibly being related to a massless propagator in pure spinor superspace.

to form supersymmetric Berends–Giele currents $M_{12\ldots p}$ turn out to be very efficient bookkeeping devices to handle the kinematic structures of a superstring amplitude in a universal way, i.e. for any number N of external legs.

According to the tree level prescription (2.12), the task in computing superstring amplitudes in the canonical color ordering $(1, 2, \ldots, N)$ is to evaluate the CFT correlator

$$\prod_{j=2}^{N-2} \int dz_j \langle V^1(0) V^{(N-1)}(1) V^N(\infty) U^2(z_2) U^3(z_3) \ldots U^{(N-2)}(z_{(N-2)}) \rangle \tag{5.1}$$

integrated over $z_1 = 0 \leq z_2 \leq \cdots \leq z_{N-2} \leq z_{N-1} = 1$. We will first of all give a representation of (5.1) in terms of $(N-2)!$ different z_i polynomials in the integrand. Then, performing manipulations on the level of both the building blocks and the associated integrals reduces the number of distinct integrals to $(N-3)!$ each of which multiplies a full-fledged SYM amplitude (4.12) in a color ordering specific to the integral.

5.1. The CFT correlator

Since the conformal $h = 1$ primaries $[\partial\theta^\alpha, \Pi^m, d_\alpha, N^{mn}]$ within the integrated vertex do not have zero modes at tree level, the correlator (5.1) can be computed by summing all their OPE singularities. Generically, this gives rise to a set of $(N-2)!$ worldsheet functions where all the z_{ij} appear as single poles, and additionally to a set of double pole integrands $\sim z_{ij}^{-2}$. It has been observed in [13] that the role of the double pole integrals is to correct the numerators of the $(N-2)!$ single pole integrals such that any OPE residue $L_{jiki\ldots li}$ is transformed to the associated BRST building block $T_{ijk\ldots l}$. This is the consequence of a subtle interplay between the integrals along the lines of subsection 5.4, in particular the tachyon poles due to double pole integrals are cancelled by the superfield kinematics in a highly nontrivial way.

A bit of care is needed to reduce the single pole residue among two integrated vertices $U^i(z_i)U^j(z_j)$ to the more basic $L_{jiki\ldots li}$ superfields which appear when $U^j U^k \ldots U^l$ successively approach an unintegrated vertex V^i. The required manipulations are based on the independence of correlation functions on the order of integrating out the $h = 1$ fields [12]. The relations up to the six point case can be found in [12,13],

$$\begin{aligned} V^1(z_1)U^2(z_2)U^3(z_3) &\sim \frac{L_{3121} - L_{2131}}{z_{23}z_{31}} =: \frac{2L_{[31,21]}}{z_{23}z_{31}} \\ V^1(z_1)U^2(z_2)U^3(z_3)U^4(z_4) &\sim \frac{L_{413121} - L_{412131} + L_{213141} - L_{312141}}{z_{23}z_{34}z_{41}} =: \frac{4L_{[41,[31,21]]}}{z_{23}z_{34}z_{41}}, \end{aligned} \tag{5.2}$$

we are picking out one particular residue here when the arguments approach each other in the order $z_2 \to z_3 \to z_1$ and $z_2 \to z_3 \to z_4 \to z_1$, respectively. This order is reflected in the specific z_{ij} in the denominator.

Higher order analogues of (5.2) involve nested antisymmetrizations:

$$\begin{aligned} V^1(z_1)U^2(z_2)U^3(z_3)U^4(z_4)U^5(z_5) &\sim \frac{8L_{[51,[41,[31,21]]]}}{z_{23}z_{34}z_{45}z_{51}} \\ V^1(z_1)U^2(z_2)U^3(z_3)\cdots U^p(z_p) &\sim \frac{2^{p-2}L_{[p1,[(p-1)1,[\ldots,[41,[31,21]]\ldots]]]}}{z_{23}z_{34}\cdots z_{p-1,p}z_{p1}} \end{aligned} \tag{5.3}$$

When all the single pole numerators are reduced to $L_{jiki\ldots li}$ and the double pole corrections are absorbed into $L_{jiki\ldots li} \mapsto T_{ijk\ldots l}$, the integrated correlator (5.1) assumes a manifestly symmetric form in the labels $2, 3, \ldots, N-2$ of the U^j vertices

$$\begin{aligned} &\prod_{j=2}^{N-2}\int dz_j \langle V^1(0)V^{(N-1)}(1)V^N(\infty)\, U^2(z_2)\, U^3(z_3)\cdots U^{(N-2)}(z_{(N-2)})\rangle \qquad (5.4) \\ &= \prod_{j=2}^{N-2}\int dz_j \prod_{i<j}|z_{ij}|^{-s_{ij}} \sum_{p=1}^{N-2}\Big\langle \frac{T_{12\ldots p}\, T_{N-1,N-2,\ldots,p+1}V_N}{(z_{12}z_{23}\cdots z_{p-1,p})(z_{N-1,N-2}z_{N-2,N-3}\cdots z_{p+2,p+1})} \\ &\qquad\qquad + \ \mathcal{P}(2,3,\ldots,N-2)\Big\rangle, \end{aligned}$$

where $\mathcal{P}(2,3,\ldots,N-2)$ denotes a symmetric sum over the $(N-3)!$ permutations of the labels $(2,3,\ldots,N-2)$. The z_{ij} polynomials associated with a specific BRST building block $T_{ij_1j_2\ldots j_p}$ follow an intriguing pattern (where the first label i belongs to an unintegrated vertex V^1 or V^{N-1} and the remaining ones to the integrated vertices $j_k \in \{2,3,\ldots,N-2\}$):

$$T_{ij_1j_2\ldots j_p} \quad \leftrightarrow \quad \frac{1}{z_{ij_1}z_{j_1j_2}z_{j_2j_3}\cdots z_{j_{p-1},j_p}} \tag{5.5}$$

Since there are $(N-3)!$ permutations of the $(2,3,\ldots,N-2)$ labels and the p sum collects $(N-2)$ distinct permutation orbits, (5.4) yields an expression for the N–point superstring amplitude (2.12) in terms of $(N-2)!$ kinematic numerators and hypergeometric integrals,

$$\begin{aligned} \mathcal{A}_N \ &\equiv \mathcal{A}(1,2,\ldots,N) \ = \ \prod_{j=2}^{N-2}\int dz_j \prod_{i<j}|z_{ij}|^{-s_{ij}} \\ &\sum_{p=1}^{N-2}\Big\langle \frac{T_{12\ldots p}T_{N-1,N-2,\ldots,p+1}V_N}{(z_{12}z_{23}\cdots z_{p-1,p})(z_{N-1,N-2}\cdots z_{p+2,p+1})} \ + \ \mathcal{P}(2,\ldots,N-2)\Big\rangle. \end{aligned} \tag{5.6}$$

The cases $N=5$ and $N=6$ of (5.6) reproduce the formulæ obtained in [13,26] and (5.6) has also been used in [34] to obtain (via the field-theory limit $\alpha' \to 0$) local expressions for all $(2N-5)!!$ kinematic numerators entering the field-theory N–point amplitude which manifestly satisfy all BCJ numerator identities [14].

5.2. A closed formula for $M_{12\ldots p}$ from the superstring

In this subsection we will show that the result (5.6) for the N–point superstring amplitude allows to extract a closed formula for the Berends–Giele current $M_{12\ldots p}$. The p sum in (5.6) partitions the legs $2,3,\ldots,N-2$ into two groups – one of them gets connected to leg 1, the other to leg $N-1$. The same structure is also present in the cohomology formula (4.12) for the field-theory amplitude; $\mathcal{A}^N_{YM}=\sum_{p=1}^{N-2}\langle M_{12\ldots p}M_{p+1\ldots N-1}V_N\rangle$.

Since the kinematic factors within individual terms of the p sum are linearly independent, we can directly compare the $p=N-2$ term on both sides of $\mathcal{A}_N \overset{\alpha'\to 0}{\longrightarrow} \mathcal{A}^N_{YM}$ – with the string- and field-theory amplitudes given respectively by (5.6) and (4.12):

$$\mathcal{A}_N=(2\alpha')^{N-3}\prod_{j=2}^{N-2}\int dz_j\prod_{i<j}|z_{ij}|^{-2\alpha' s_{ij}}\left\langle\frac{T_{12\ldots N-2}V_{N-1}V_N}{z_{12}z_{23}\cdots z_{N-3,N-2}}\;+\;\mathcal{P}(2,\ldots,N-2)+\cdots\right\rangle$$

$$\overset{\alpha'\to 0}{\longrightarrow}\;\langle M_{12\ldots N-2}V_{N-1}V_N\rangle\;+\;\cdots \tag{5.7}$$

This yields a closed-formula solution for the rank $p=N-2$ current $M_{12\ldots p}$,

$$M_{12\ldots p}=\lim_{\alpha'\to 0}(2\alpha')^{p-1}\prod_{j=2}^{p}\int_{z_{j-1}}^{1}dz_j\prod_{i<j}^{p+1}|z_{ij}|^{-2\alpha' s_{ij}}\left(\frac{T_{12\ldots p}}{z_{12}z_{23}\cdots z_{p-1,p}}+\mathcal{P}(2,3,\ldots,p)\right), \tag{5.8}$$

where $z_1=0$ and $z_{p+1}=1$ as customary for a $(p+2)-$point amplitude. For example, using the momentum expansion of the five-point superstring integrals [5] and the BRST symmetry $T_{123}+T_{231}+T_{312}=0$ of (3.14) the following M_{123} is generated

$$\begin{aligned}M_{123}&=\lim_{\alpha'\to 0}(2\alpha')^2\int_0^1 dz_2\int_{z_2}^1 dz_3\prod_{i<j}^{4}|z_{ij}|^{-2\alpha' s_{ij}}\left(\frac{T_{123}}{z_{12}z_{23}}+\frac{T_{132}}{z_{13}z_{32}}\right)\\&=\frac{T_{123}}{s_{12}s_{123}}+\frac{T_{123}}{s_{23}s_{123}}-\frac{T_{132}}{s_{23}s_{123}}=\frac{T_{123}}{s_{12}s_{123}}+\frac{T_{321}}{s_{23}s_{123}},\end{aligned} \tag{5.9}$$

which is easily shown to satisfy $QM_{123}=E_{123}$. Similarly, we checked that the formula (5.8) correctly generates solutions of (4.4) up to and including $M_{1234567}$.

5.3. Trading $T_{12\ldots p}$ for $M_{12\ldots p}$

As will be shown in the next subsections, in order to simplify even further the expression (5.6) of the superstring N–point amplitude it will be convenient to trade the BRST building blocks $T_{12\ldots p}$ for the Berends–Giele currents $M_{12\ldots p}$.

This exchange will be possible because of the particular pattern (5.5) of z_{ij} dependence along with the $T_{12\ldots p}$. The lowest order example of $T \leftrightarrow M$ conversion is a triviality $\frac{T_{12}}{z_{12}} = \frac{s_{12}}{z_{12}} M_{12}$, but already the simplest generalization is a result of partial fraction relations and the symmetry properties of T_{ijk}:

$$\frac{T_{123}}{z_{12}z_{23}} \; + \; \mathcal{P}(2,3) = \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) M_{123} \; + \; \mathcal{P}(2,3). \tag{5.10}$$

Similar identities have been checked at $p = 4$ and $p = 5$ level:

$$\begin{aligned}
\frac{T_{1234}}{z_{12}z_{23}z_{34}} \; + \; \mathcal{P}(2,3,4) &= \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \left(\frac{s_{14}}{z_{14}} + \frac{s_{24}}{z_{24}} + \frac{s_{34}}{z_{34}} \right) M_{1234} + \; \mathcal{P}(2,3,4) \\
\frac{T_{12345}}{z_{12}z_{23}z_{34}z_{45}} \; + \; \mathcal{P}(2,3,4,5) &= \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \left(\frac{s_{14}}{z_{14}} + \frac{s_{24}}{z_{24}} + \frac{s_{34}}{z_{34}} \right) \\
&\times \left(\frac{s_{15}}{z_{15}} + \frac{s_{25}}{z_{25}} + \frac{s_{35}}{z_{35}} + \frac{s_{45}}{z_{45}} \right) M_{12345} \; + \; \mathcal{P}(2,3,4,5).
\end{aligned} \tag{5.11}$$

These identities heavily rely on the interplay of different terms in the permutation sum and on the symmetry properties (3.20) of the BRST building blocks which leave no more than $(p-1)!$ independent permutations of $T_{i_1\ldots i_p}$ at level p.

The natural n point generalization of (5.10) and (5.11) reads as follows:

$$\begin{aligned}
\frac{T_{12\ldots p}}{z_{12}z_{23}\cdots z_{p-1,p}} \; + \; \mathcal{P}(2,\ldots,p) &= \prod_{k=2}^{p} \sum_{m=1}^{k-1} \frac{s_{mk}}{z_{mk}} M_{12\ldots p} \; + \; \mathcal{P}(2,\ldots,p) \\
\frac{T_{N-1,N-2,\ldots,p+1}}{z_{N-1,N-2}\cdots z_{p+2,p+1}} \; + \; \mathcal{P}(2,\ldots,p) &= \prod_{k=p+1}^{N-2} \sum_{n=k+1}^{N-1} \frac{s_{nk}}{z_{nk}} M_{N-1,N-2,\ldots,p+1} \\
+ \; \mathcal{P}(2,\ldots,p) &= \prod_{k=p+1}^{N-2} \sum_{n=k+1}^{N-1} \frac{s_{kn}}{z_{kn}} M_{p+1,p+2,\ldots,N-1} \; + \; \mathcal{P}(2,\ldots,p),
\end{aligned} \tag{5.12}$$

where in the last line the rank $N-1-p$ Berends–Giele current with leg $N-1$ involved was reflected via (4.7); $M_{N-1,\ldots,p+1} = (-1)^{N-p-2} M_{p+1,\ldots,N-1}$.

5.4. Worldsheet integration by parts

This subsection focuses on the integrals rather than the kinematic factors in the superstring amplitude. The chain of $\frac{s_{mk}}{z_{mk}}$ sums which appears as a result of (5.12) when all the $T_{12\ldots p}$ are converted to $M_{12\ldots p}$ is particularly suitable to perform integration by parts with respect to z_j variables. Further details on the structure and manipulations of the integrals can be found in [2].

The key idea is the vanishing of boundary terms in the worldsheet integrals:

$$\int dz_j \cdots \int dz_{N-2} \, \frac{\partial}{\partial z_k} \frac{\prod_{i<j} |z_{ij}|^{-s_{ij}}}{z_{i_1 j_1} \cdots z_{i_{N-4} j_{N-4}}} = 0. \tag{5.13}$$

This identity provides relations between the integrals in an N–point superstring amplitude with $N-3$ powers of z_{ij} in the denominator. They become particularly easy if the differentiation variable z_k does not appear in the denominator (i.e. if $k \notin \{i_l, j_l\}$) because $\frac{\partial}{\partial z_k}$ only hits the $\prod_{m \neq k} |z_{mk}|^{-s_{mk}}$ factor in that case:

$$\int dz_2 \cdots \int dz_{N-2} \frac{\prod_{i<j} |z_{ij}|^{-s_{ij}}}{z_{i_1 j_1} \cdots z_{i_{N-4} j_{N-4}}} \sum_{\substack{m=1 \\ m \neq k}}^{N-1} \frac{s_{mk}}{z_{mk}} = 0. \tag{5.14}$$

This can be directly applied to the integrands on the right hand side of (5.10), (5.11) and (5.12), namely:

$$\begin{aligned}
&\prod_{j=2}^{3} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) = \prod_{j=2}^{3} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \frac{s_{34}}{z_{34}} \\
&\prod_{j=2}^{4} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \left(\frac{s_{14}}{z_{14}} + \frac{s_{24}}{z_{24}} + \frac{s_{34}}{z_{34}} \right) \\
&\qquad = \prod_{j=2}^{4} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \frac{s_{45}}{z_{45}} \begin{cases} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \\ \left(\frac{s_{34}}{z_{34}} + \frac{s_{35}}{z_{35}} \right) \end{cases} \\
&\prod_{j=2}^{5} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \left(\frac{s_{14}}{z_{14}} + \frac{s_{24}}{z_{24}} + \frac{s_{34}}{z_{34}} \right) \left(\frac{s_{15}}{z_{15}} + \frac{s_{25}}{z_{25}} + \frac{s_{35}}{z_{35}} + \frac{s_{45}}{z_{45}} \right) \\
&\qquad = \prod_{j=2}^{5} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \frac{s_{56}}{z_{56}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \left(\frac{s_{45}}{z_{45}} + \frac{s_{46}}{z_{46}} \right)
\end{aligned} \tag{5.15}$$

In the general N point case, it is most economic to leave the first $[N/2]-1$ factors of $\sum_{m=1}^{k-1} \frac{s_{mk}}{z_{mk}}$ as they are, and to integrate the remaining $[(N-3)/2]$ such factors by parts:

$$\begin{aligned}
&\prod_{j=2}^{N-2} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \cdots \left(\frac{s_{1,N-2}}{z_{1,N-2}} + \cdots + \frac{s_{N-1,N-2}}{z_{N-1,N-2}} \right) \\
&= \prod_{j=2}^{N-2} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \cdots \left(\frac{s_{1,[N/2]}}{z_{1,[N/2]}} + \cdots + \frac{s_{[N/2]-1,[N/2]}}{z_{[N/2]-1,[N/2]}} \right)
\end{aligned}$$

$$\times \left(\frac{s_{[N/2]+1,[N/2]+2}}{z_{[N/2]+1,[N/2]+2}} + \cdots + \frac{s_{[N/2]+1,N-1}}{z_{[N/2]+1,N-1}} \right) \cdots \left(\frac{s_{N-3,N-2}}{z_{N-3,N-2}} + \frac{s_{N-3,N-1}}{z_{N-3,N-1}} \right) \frac{s_{N-2,N-1}}{z_{N-2,N-1}}$$
$$= \prod_{j=2}^{N-2} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \left(\prod_{k=2}^{[N/2]} \sum_{m=1}^{k-1} \frac{s_{mk}}{z_{mk}} \right) \left(\prod_{k=[N/2]+1}^{N-2} \sum_{n=k+1}^{N-1} \frac{s_{kn}}{z_{kn}} \right) \tag{5.16}$$

In contrast to the $T_{12\ldots p} \to M_{12\ldots p}$ reshuffling identities from the previous subsection, (5.15) and (5.16) are valid before summing over permutations of $(2, 3, \ldots, N-2)$.

5.5. The complete N–point superstring disk amplitude

This subsection completes the derivation of the striking result (5.22) for the superstring N–point amplitude $\mathcal{A}_N \equiv \mathcal{A}(1, 2, \ldots, N)$ by combining the results of the previous subsections. Let us first look at the four-, five- and six-point examples to get a better feeling of the mechanisms at work.

After using $T_{ij} = s_{ij} M_{ij}$, the total derivative relation $\frac{s_{23}}{z_{23}} \mapsto \frac{s_{12}}{z_{12}}$ as well as $E_{123} = M_{12} V_3 + V_1 M_{23}$, the four-point open string disk amplitude is easily seen to be

$$\begin{aligned}
\mathcal{A}_4 &= \int dz_2 \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\langle \frac{T_{12} V_3 V_4}{z_{12}} + \frac{V_1 T_{32} V_4}{z_{32}} \right\rangle \\
&= \int dz_2 \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\langle \frac{s_{12}}{z_{12}} M_{12} V_3 V_4 + \frac{s_{23}}{z_{23}} V_1 M_{23} V_4 \right\rangle \\
&= \int dz_2 \prod_{i<j} |z_{ij}|^{-s_{ij}} \ \frac{s_{12}}{z_{12}} \left\langle (M_{12} V_3 + V_1 M_{23}) V_4 \right\rangle \\
&= \int dz_2 \prod_{i<j} |z_{ij}|^{-s_{ij}} \ \frac{s_{12}}{z_{21}} \mathcal{A}_{\rm YM}(1, 2, 3, 4).
\end{aligned}$$

Similarly, the five-point superstring amplitude (5.6) contains six different integrands and kinematic terms. After applying (5.10), the T_{ij} and T_{ijk} conspire to give M_{ij} and M_{ijk} with modified integrals, then we use integration by parts according to (5.15) on the way to the third equality of (5.18). Remarkably, many of the initially $(N-2)! = 6$ distinct integrals now coincide: The three kinematic terms $M_{123} V_4 V_5$, $M_{12} M_{34} V_5$ and $V_1 M_{234} V_5$ are multiplied by the same integral after partial integration, the same is true for the $(2 \leftrightarrow 3)$ permutation. That is why we can identify color ordered field-theory amplitudes (4.12) in the last line:

$$\mathcal{A}_5 = \int dz_2 dz_3 \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\langle \frac{T_{123} V_4 V_5}{z_{12} z_{23}} + \frac{T_{12} T_{43} V_5}{z_{12} z_{43}} + \frac{V_1 T_{432} V_5}{z_{43} z_{32}} + (2 \leftrightarrow 3) \right\rangle$$

$$= \int dz_2 dz_3 \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\langle \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) M_{123} V_4 V_5 + \frac{s_{12} s_{34}}{z_{12} z_{34}} M_{12} M_{34} V_5 \right.$$
$$\left. + \frac{s_{43}}{z_{43}} \left(\frac{s_{42}}{z_{42}} + \frac{s_{32}}{z_{32}} \right) V_1 M_{432} V_5 + (2 \leftrightarrow 3) \right\rangle \tag{5.18}$$
$$= \int dz_2 dz_3 \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\{ \frac{s_{12} s_{34}}{z_{12} z_{34}} \langle M_{123} V_4 V_5 + M_{12} M_{34} V_5 + V_1 M_{234} V_5 \rangle + (2 \leftrightarrow 3) \right\}$$
$$= \int dz_2 dz_3 \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\{ \frac{s_{12} s_{34}}{z_{12} z_{34}} \mathcal{A}_{\rm YM}(1,2,3,4,5) + \frac{s_{13} s_{24}}{z_{13} z_{24}} \mathcal{A}_{YM}(1,3,2,4,5) \right\}$$

Simplifying the six-point amplitudes $\mathcal{A}_6$ follows similar steps. In this case, (5.11) takes care of the conversion of T_{ijkl} into M_{ijkl}, then integration by parts makes the four integrals within a given $(2,3,4)$ permutation coincide:

$$\mathcal{A}_6 = \prod_{j=2}^{4} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\langle \frac{T_{1234} V_5 V_6}{z_{12} z_{23} z_{34}} + \frac{T_{123} T_{54} V_6}{z_{12} z_{23} z_{54}} + \frac{T_{12} T_{543} V_6}{z_{12} z_{54} z_{43}} + \frac{V_1 T_{5432} V_6}{z_{54} z_{43} z_{32}} + \mathcal{P}(2,3,4) \right\rangle$$
$$= \prod_{j=2}^{4} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\langle \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \left(\frac{s_{14}}{z_{14}} + \frac{s_{24}}{z_{24}} + \frac{s_{34}}{z_{34}} \right) M_{1234} V_5 V_6 \right.$$
$$+ \frac{s_{12}}{z_{12}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \frac{s_{45}}{z_{45}} M_{123} M_{45} V_6 + \frac{s_{12}}{z_{12}} \frac{s_{45}}{z_{45}} \left(\frac{s_{34}}{z_{34}} + \frac{s_{35}}{z_{35}} \right) M_{12} M_{543} V_6$$
$$\left. + \frac{s_{45}}{z_{45}} \left(\frac{s_{34}}{z_{34}} + \frac{s_{35}}{z_{35}} \right) \left(\frac{s_{52}}{z_{52}} + \frac{s_{42}}{z_{42}} + \frac{s_{32}}{z_{32}} \right) V_1 M_{5432} V_6 \ + \ \mathcal{P}(2,3,4) \right\rangle$$
$$= \prod_{j=2}^{4} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\{ \frac{s_{12} s_{45}}{z_{12} z_{45}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \langle M_{1234} V_5 V_6 + M_{123} M_{45} V_6 \right.$$
$$\left. + M_{12} M_{345} V_6 + V_1 M_{2345} V_6 \rangle \ + \ \mathcal{P}(2,3,4) \right\} \tag{5.19}$$
$$= \prod_{j=2}^{4} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\{ \frac{s_{12} s_{45}}{z_{12} z_{45}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \mathcal{A}_{\rm YM}(1,2,3,4,5,6) \ + \ \mathcal{P}(2,3,4) \right\}$$

The identities (5.11) and (5.15) are sufficient to also reduce the superstring seven-point amplitude $\mathcal{A}_7$ to its field-theory constituents:

$$\mathcal{A}_7 = \prod_{j=2}^{5} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\langle \frac{T_{12345} V_6 V_7}{z_{12} z_{23} z_{34} z_{45}} + \frac{T_{1234} T_{65} V_7}{z_{12} z_{23} z_{34} z_{65}} + \frac{T_{123} T_{654} V_7}{z_{12} z_{23} z_{65} z_{54}} \right.$$
$$\left. + \frac{T_{12} T_{6543} V_7}{z_{12} z_{65} z_{54} z_{43}} + \frac{V_1 T_{65432} V_7}{z_{65} z_{54} z_{43} z_{32}} \ + \ \mathcal{P}(2,3,4,5) \right\rangle$$
$$= \prod_{j=2}^{5} \int dz_j \prod_{i<j} |z_{ij}|^{-s_{ij}} \left\{ \frac{s_{12} s_{56}}{z_{12} z_{56}} \left(\frac{s_{13}}{z_{13}} + \frac{s_{23}}{z_{23}} \right) \left(\frac{s_{45}}{z_{45}} + \frac{s_{46}}{z_{46}} \right) \mathcal{A}_{\rm YM}(1,2,3,4,5,6,7) \right.$$

$$+\ \mathcal{P}(2,3,4,5)\Big\}. \tag{5.20}$$

The N–point generalization is based on introducing currents $M_{i_1 i_2 \ldots i_p}$ via (5.12) followed by integration by parts using (5.16). The latter makes the integral independent on p such that the z_{ij} can be placed outside the p sum and SYM amplitudes emerge from the kinematics.

$$\begin{aligned}
\mathcal{A}_N &= \prod_{j=2}^{N-2}\int dz_j \prod_{i<j}|z_{ij}|^{-s_{ij}}\Big\langle \sum_{p=1}^{N-2}\frac{T_{12\ldots p}\,T_{N-1,N-2,\ldots,p+1}V_N}{(z_{12}z_{23}\cdots z_{p-1,p})(z_{N-1,N-2}\cdots z_{p+2,p+1})} \\
&\qquad\qquad +\ \mathcal{P}(2,3,\ldots,N-2)\Big\rangle \\
&= \prod_{j=2}^{N-2}\int dz_j \prod_{i<j}|z_{ij}|^{-s_{ij}}\Big\langle \sum_{p=1}^{N-2}\left(\prod_{k=2}^{p}\sum_{m=1}^{k-1}\frac{s_{mk}}{z_{mk}}M_{12\ldots p}\right) \\
&\qquad\qquad \times\left(\prod_{k=p+1}^{N-2}\sum_{n=k+1}^{N-1}\frac{s_{kn}}{z_{kn}}M_{p+1,\ldots,N-2,N-1}\right)V_N\ +\ \mathcal{P}(2,3,\ldots,N-2)\Big\rangle \\
&= \prod_{j=2}^{N-2}\int dz_j \prod_{i<j}|z_{ij}|^{-s_{ij}}\Big\{\left(\prod_{k=2}^{[N/2]}\sum_{m=1}^{k-1}\frac{s_{mk}}{z_{mk}}\right)\left(\prod_{k=[N/2]+1}^{N-2}\sum_{n=k+1}^{N-1}\frac{s_{kn}}{z_{kn}}\right) \\
&\qquad\qquad \times\sum_{p=1}^{N-2}\langle M_{12\ldots p}M_{p+1\ldots N-2,N-1}V_N\rangle\ +\ \mathcal{P}(2,3,\ldots,N-2)\Big\} \\
&= \prod_{j=2}^{N-2}\int dz_j \prod_{i<j}|z_{ij}|^{-s_{ij}}\Big\{\left(\prod_{k=2}^{[N/2]}\sum_{m=1}^{k-1}\frac{s_{mk}}{z_{mk}}\right)\left(\prod_{k=[N/2]+1}^{N-2}\sum_{n=k+1}^{N-1}\frac{s_{kn}}{z_{kn}}\right) \\
&\qquad\qquad \times\mathcal{A}_{YM}(1,2,3,\ldots,N-1,N)\ +\ \mathcal{P}(2,3,\ldots,N-2)\Big\}.
\end{aligned} \tag{5.21}$$

Equivalently, by undoing the total derivative relation used in (5.21) the full N–point superstring amplitude becomes

$$\mathcal{A}_N = \int\limits_{z_i<z_{i+1}}\prod_{i<j}|z_{ij}|^{-s_{ij}}\Big[\prod_{k=2}^{N-2}\sum_{m=1}^{k-1}\frac{s_{mk}}{z_{mk}}\,\mathcal{A}_{YM}(1,2,\ldots,N)+\mathcal{P}(2,\ldots,N-2)\Big], \tag{5.22}$$

where the integration region $\int_{z_i<z_{i+1}} \equiv \prod_{j=2}^{N-2}\int_{z_{j-1}}^{1} dz_j$ is responsible for dictating which color-ordered string subamplitude is being computed. Therefore the end result of all these pure spinor superspace manipulations is that the N–point superstring disk amplitude is written in terms of the explicit sum of $(N-3)!$ basis of field-theory amplitudes multiplied

by an equal number of hypergeometric integrals, as mentioned in the Introduction and further elaborated in [2].

Acknowledgements: We thank Dimitrios Tsimpis for his contributions at earlier stages of this project and for useful discussions. We also thank Nathan Berkovits for pointing out the general argument about the BRST-exactness of the sums in subsection 3.2. Furthermore, we thank the Kavli Institute for Theoretical Physics in Santa Barbara for hospitality and partial financial support. C.M. also thanks the Werner-Heisenberg-Institut in München for hospitality and partial financial support during preparation of this work and acknowledges support by the Deutsch-Israelische Projektkooperation (DIP H52). O.S. is indebted to UCLA and in particular to Martin Ammon for hospitality during preparation of this work. St.St. is grateful to UCLA, Caltech, and SLAC for hospitality and financial support during completion of this work. This research was supported in part by the National Science Foundation under Grant No. NSF PHY0551164.

Appendix A. The explicit construction of T_{12345}

In order to find the appropriate redefinition of $L_{21314151}$ leading to $\tilde{T}_{12345}$ one simply uses the known redefinitions of $[L_{21}, L_{2131}, L_{213141}] \to [T_{12}, T_{123}, T_{1234}]$ in the right-hand side of (3.3). Even though it is not obvious, all terms from these lower-order redefinitions group together into a BRST-exact combination which can be moved to the left-hand side of (3.3). Doing that finally leads to the definition of $\tilde{T}_{12345}$, given by

$$\begin{aligned}
\tilde{T}_{12345} = {} & L_{21314151} \\
& - \frac{1}{4}(s_{13}+s_{23})\big[D_{12}D_{34}V_5(s_{35}+s_{45}) + D_{12}D_{35}V_4 s_{34} - D_{12}D_{45}V_3 s_{34}\big] \\
& - \frac{1}{4}s_{12}\Big[D_{13}D_{24}V_5(s_{25}+s_{45}) + D_{14}D_{23}V_5(s_{25}+s_{35}) + D_{15}D_{23}V_4(s_{24}+s_{34}) \\
& + s_{24}(D_{13}D_{25}V_4 - D_{13}D_{45}V_2) + s_{13}(D_{34}D_{25}V_1 + D_{35}D_{24}V_1) \\
& + s_{23}(D_{14}D_{25}V_3 - D_{14}D_{35}V_2 + D_{15}D_{24}V_3 - D_{15}D_{34}V_2) + s_{14}D_{45}D_{23}V_1\Big] \\
& - (s_{15}+s_{25}+s_{35}+s_{45})S^{(2)}_{1234}V_5 - (s_{14}+s_{24}+s_{34})\left(S^{(1)}_{123}L_{54} + S^{(2)}_{1235}V_4\right) \\
& - (s_{13}+s_{23})\left(S^{(1)}_{124}L_{53} + S^{(1)}_{125}L_{43} - S^{(1)}_{345}L_{21} + S^{(2)}_{1245}V_3\right) \\
& - s_{12}\Big[S^{(1)}_{134}L_{52} + S^{(1)}_{135}L_{42} + S^{(1)}_{145}L_{32} + S^{(2)}_{1345}V_2 - (1\leftrightarrow 2)\Big] \\
& - \frac{1}{2}\Big[T_{123}D_{45}(s_{14}+s_{24}+s_{34}) + (T_{125}D_{34} - T_{345}D_{12} + T_{124}D_{35})(s_{13}+s_{23}) \\
& + s_{12}(T_{134}D_{25} + T_{135}D_{24} + T_{145}D_{23} - (1\leftrightarrow 2))\Big] \qquad \text{(A.1)}
\end{aligned}$$

which, by construction, is guaranteed to satisfy

$$\begin{aligned} Q\tilde{T}_{12345} = &+ (s_{15}+s_{25}+s_{35}+s_{45})T_{1234}V_5 + (s_{14}+s_{24}+s_{34})(T_{1235}V_4 + T_{123}T_{45}) \\ &+ (s_{13}+s_{23})(T_{1245}V_3 + T_{124}T_{35} + T_{125}T_{34} + T_{12}T_{345}) \\ &+ s_{12}(T_{1345}V_2 + V_1T_{2345} + T_{134}T_{25} + T_{135}T_{24} + T_{145}T_{23} \\ &+ T_{13}T_{245} + T_{14}T_{235} + T_{15}T_{234}). \end{aligned} \tag{A.2}$$

One can also show that[7]

$$\begin{aligned} \tilde{T}_{12345} + \tilde{T}_{21345} =& QR^{(1)}_{12345}, \\ \tilde{T}_{12345} + \tilde{T}_{23145} + \tilde{T}_{31245} =& QR^{(2)}_{12345}, \\ \tilde{T}_{12345} - \tilde{T}_{12435} + \tilde{T}_{34125} - \tilde{T}_{34215} =& QR^{(3)}_{12345}, \\ \tilde{T}_{12345} - \tilde{T}_{12354} + \tilde{T}_{45123} - \tilde{T}_{45213} - \tilde{T}_{45312} + \tilde{T}_{45321} =& QR^{(4)}_{12345} \end{aligned} \tag{A.3}$$

where the BRST-exact parts are given by

$$\begin{aligned} R^{(1)}_{12345} =& D_{12}(k^{12}\cdot A^3)(k^{123}\cdot A^4)(k^{1234}\cdot A^5) + \frac{1}{6}(s_{13}+s_{23})D_{12}\Big[D_{45}\left((k^4\cdot A^3)-(k^5\cdot A^3)\right) \\ &+ D_{35}\left((k^5\cdot A^4)-(k^3\cdot A^4)\right) - 2D_{34}\left((k^3\cdot A^5)+2(k^4\cdot A^5)\right)\Big], \\ R^{(2)}_{12345} =& D_{12}(k^2\cdot A^3)(k^{123}\cdot A^4)(k^{1234}\cdot A^5) + \frac{1}{6}\Big[s_{12}D_{13}\left(D_{45}((k^4\cdot A^2)-(k^5\cdot A^2))\right. \\ &\left. + D_{25}((k^5\cdot A^4)-(k^2\cdot A^4)) - 2D_{24}((k^2\cdot A^5)+2(k^4\cdot A^5))\right) + \text{cyclic}(123)\Big], \\ R^{(3)}_{12345} =& -(W^1\gamma^m W^2)(W^3\gamma^m W^4)(k^{1234}\cdot A^5) \\ &+ \left[D_{12}(k^3\cdot A^4)(k^2\cdot A^3)(k^{1234}\cdot A^5) + \frac{1}{3}(s_{24}-2s_{23})D_{34}D_{12}(k^4\cdot A^5) - (3\leftrightarrow 4)\right] \\ &+ \frac{1}{6}(s_{14}+s_{24})\left[D_{25}D_{34}((k^2\cdot A^1)-(k^5\cdot A^1)) + D_{15}D_{34}((k^5\cdot A^2)-(k^1\cdot A^2))\right] \\ &+ \frac{1}{6}(s_{23}+s_{24})\left[D_{45}D_{12}((k^4\cdot A^3)-(k^5\cdot A^3)) + D_{35}D_{12}((k^5\cdot A^4)-(k^3\cdot A^4))\right] \\ &+ \Big[\left(D_{13}(k^1\cdot A^2)(k^3\cdot A^4) + D_{24}(k^2\cdot A^1)(k^4\cdot A^3) + D_{34}(k^1\cdot A^2)(k^4\cdot A^1)\right)(k^{1234}\cdot A^5) \\ &+ \frac{1}{3}(s_{24}-2s_{14})D_{34}D_{12}(k^2\cdot A^5) - (1\leftrightarrow 2)\Big] \\ R^{(4)}_{12345} =& (W^1\gamma^m W^2)\left[(W^4\gamma^n W^5)\mathcal{F}^3_{mn} - (W^4\gamma^m W^5)(k^{12}\cdot A^3)\right] \\ &+ \Big[(W^1\gamma^m W^2)(W^3\gamma^m W^5)(k^5\cdot A^4) + \frac{1}{4}(W^1\gamma^m W^2)(W^5\gamma^{np}\gamma^m W^3)\mathcal{F}^4_{np} \end{aligned}$$

[7] The tedious algebra was handled using FORM [32].

$$
\begin{aligned}
&+ D_{12}(k^2\cdot A^3)(k^{23}\cdot A^4)(k^4\cdot A^5) + D_{12}(k^1\cdot A^3)(k^2\cdot A^4)(k^4\cdot A^5)\\
&+ \frac{1}{6}D_{12}D_{35}(k^3\cdot A^4)s_{23} + \frac{5}{6}D_{12}D_{35}(k^5\cdot A^4)s_{23} + \frac{1}{3}D_{12}D_{45}(k^4\cdot A^3)s_{23}\\
&+ D_{14}(k^1\cdot A^2)(k^{12}\cdot A^3)(k^4\cdot A^5) + D_{25}(k^2\cdot A^1)(k^{12}\cdot A^3)(k^5\cdot A^4)\\
&+ D_{34}(k^2\cdot A^1)(k^3\cdot A^2)(k^4\cdot A^5) + D_{35}(k^3\cdot A^1)(k^1\cdot A^2)(k^5\cdot A^4) - (4\leftrightarrow 5)\Big]\\
&+ \Big[(W^2\gamma^m W^3)(W^4\gamma^m W^5)(k^2\cdot A^1) + \frac{1}{4}(W^4\gamma^m W^5)(W^1\gamma^{np}\gamma^m W^3)\mathcal{F}^2_{np}\\
&+ D_{13}(k^1\cdot A^2)(k^3\cdot A^4)(k^4\cdot A^5) - D_{13}(k^5\cdot A^4)(k^1\cdot A^2)(k^3\cdot A^5)\\
&+ D_{45}(k^2\cdot A^1)(k^3\cdot A^2)(k^5\cdot A^3) + D_{45}(k^5\cdot A^1)(k^1\cdot A^2)(k^{12}\cdot A^3)\\
&+ \frac{1}{3}D_{12}D_{45}(k^2\cdot A^3)(-2s_{15}+s_{25}+s_{35}) + \frac{1}{6}D_{13}D_{45}(k^3\cdot A^2)(s_{15}+s_{25}+s_{35})\\
&- \frac{1}{6}D_{13}D_{45}(k^1\cdot A^2)(s_{15}+s_{25}-5s_{35}) - (1\leftrightarrow 2)\Big] \qquad \text{(A.4)}
\end{aligned}
$$

Removing these BRST-exact parts is accomplished by the second redefinition $\tilde{T}_{12345} \longrightarrow T_{12345}$, leading to the rank-five BRST building block

$$T_{12345} = \tilde{T}_{12345} - QS^{(3)}_{12345}, \qquad \text{(A.5)}$$

where the expression for $S^{(3)}_{12345}$ can be written recursively as

$$
\begin{aligned}
S^{(3)}_{12345} =& \frac{4}{5}S^{(2)}_{12345} + \frac{1}{5}\left(S^{(2)}_{12354} - S^{(2)}_{45123} + S^{(2)}_{45213} + S^{(2)}_{45312} - S^{(2)}_{45321}\right) + \frac{1}{5}R^{(4)}_{12345},\\
S^{(2)}_{12345} =& \frac{3}{4}S^{(1)}_{12345} + \frac{1}{4}(S^{(1)}_{12435} - S^{(1)}_{34125} + S^{(1)}_{34215}) + \frac{1}{4}R^{(3)}_{12345}, \qquad \text{(A.6)}\\
S^{(1)}_{12345} =& \frac{1}{2}R^{(1)}_{12345} + \frac{1}{3}R^{(2)}_{[12]345}.
\end{aligned}
$$

To see that (A.5) and (A.6) imply all the BRST-symmetries of T_{12345}

$$
\begin{aligned}
0 &= T_{12345} + T_{21345}\\
0 &= T_{12345} + T_{31245} + T_{23145}\\
0 &= T_{12345} - T_{12435} + T_{34125} - T_{34215}\\
0 &= T_{12345} - T_{12354} + T_{45123} - T_{45213} - T_{45312} + T_{45321} \qquad \text{(A.7)}
\end{aligned}
$$

it suffices to check that the following identities hold,

$$
\begin{aligned}
S^{(3)}_{12345} + S^{(3)}_{21345} =& R^{(1)}_{12345}\\
S^{(3)}_{12345} + S^{(3)}_{31245} + S^{(3)}_{23145} =& R^{(2)}_{12345} \qquad \text{(A.8)}\\
S^{(3)}_{12345} - S^{(3)}_{12435} + S^{(3)}_{34125} - S^{(3)}_{34215} =& R^{(3)}_{12345}\\
S^{(3)}_{12345} - S^{(3)}_{12354} + S^{(3)}_{45123} - S^{(3)}_{45213} - S^{(3)}_{45312} + S^{(3)}_{45321} =& R^{(4)}_{12345}.
\end{aligned}
$$

Having the explicit superfield expressions for the building blocks up to T_{12345} allows all component amplitudes up to $N = 11$ to be evaluated.

Appendix B. The solutions for $M_{i_1 i_2 \ldots i_n}$ in terms of BRST building blocks

From the relation between $M_{123\ldots n}$ and the cubic diagrams of the $(n+1)$-point amplitude discussed in subsection 4.1, it follows that the solutions for M_{123}, M_{1234}, M_{12345}, M_{123456} and $M_{1234567}$ which satisfy (4.1) contain 2, 5, 14, 42 and 132 different kinematic pole configurations, which are represented by the cubic-graph expansion of the tree amplitudes. Their explicit expressions can then be read off from the dictionary between those cubic graphs and the BRST building blocks; as discussed in subsection 3.3. Furthermore, using the antisymmetry on the first two labels of $T_{ijk\ldots}$, one can always choose an ordering such that all terms in$M_{123\ldots n}$ have a positive coefficient, leading to:

$$M_{12} = \frac{T_{12}}{s_{12}}\,, \tag{B.1}$$

$$M_{123} = \frac{1}{s_{123}}\left(\frac{T_{123}}{s_{12}} + \frac{T_{321}}{s_{23}}\right), \tag{B.2}$$

$$M_{1234} = \frac{1}{s_{1234}}\left(\frac{T_{1234}}{s_{12}s_{123}} + \frac{T_{3214}}{s_{23}s_{123}} + \frac{T_{3241}}{s_{23}s_{234}} + \frac{T_{3421}}{s_{34}s_{234}} + \frac{2T_{12[34]}}{s_{12}s_{34}}\right), \tag{B.3}$$

$$\begin{aligned} M_{12345} = \frac{1}{s_{12345}}\Bigg[&\frac{1}{s_{1234}}\left(\frac{T_{12345}}{s_{12}s_{123}} + \frac{T_{32145}}{s_{23}s_{123}} + \frac{T_{32415}}{s_{23}s_{234}} + \frac{T_{34215}}{s_{34}s_{234}} + \frac{2T_{12[34]5}}{s_{12}s_{34}}\right) \\ &+ \frac{1}{s_{2345}}\left(\frac{T_{34251}}{s_{34}s_{234}} + \frac{T_{32451}}{s_{23}s_{234}} + \frac{T_{34521}}{s_{34}s_{345}} + \frac{T_{54321}}{s_{45}s_{345}} + \frac{2T_{45[23]1}}{s_{23}s_{45}}\right) \\ &+ \frac{2T_{123[45]}}{s_{12}s_{123}s_{45}} + \frac{2T_{321[45]}}{s_{23}s_{123}s_{45}} + \frac{2T_{453[12]}}{s_{45}s_{345}s_{12}} + \frac{2T_{435[12]}}{s_{34}s_{345}s_{12}}\Bigg], \end{aligned} \tag{B.4}$$

$$\begin{aligned} M_{123456} = \frac{1}{s_{123456}}\Bigg[&\frac{4T_{12[34][56]}}{s_{12}s_{34}s_{56}s_{1234}} + \frac{4T_{34[56][21]}}{s_{12}s_{34}s_{56}s_{3456}} + \frac{4T_{123[[45]6]}}{s_{12}s_{45}s_{123}s_{456}} + \frac{4T_{123[4[56]]}}{s_{12}s_{56}s_{123}s_{456}} \\ &+ \frac{4T_{231[[54]6]}}{s_{23}s_{45}s_{123}s_{456}} + \frac{4T_{231[4[65]]}}{s_{23}s_{56}s_{123}s_{456}} + \frac{2T_{345[21]6}}{s_{12}s_{34}s_{345}s_{12345}} + \frac{2T_{3456[21]}}{s_{12}s_{34}s_{345}s_{3456}} \\ &+ \frac{2T_{12[34]56}}{s_{12}s_{34}s_{1234}s_{12345}} + \frac{2T_{123[45]6}}{s_{12}s_{45}s_{123}s_{12345}} + \frac{2T_{543[21]6}}{s_{12}s_{45}s_{345}s_{12345}} + \frac{2T_{5436[21]}}{s_{12}s_{45}s_{345}s_{3456}} \\ &+ \frac{2T_{4563[12]}}{s_{12}s_{45}s_{456}s_{3456}} + \frac{2T_{1234[56]}}{s_{12}s_{56}s_{123}s_{1234}} + \frac{2T_{5643[21]}}{s_{12}s_{56}s_{456}s_{3456}} + \frac{2T_{231[54]6}}{s_{23}s_{45}s_{123}s_{12345}} \\ &+ \frac{2T_{456[23]1}}{s_{23}s_{45}s_{456}s_{23456}} + \frac{2T_{34[56]21}}{s_{34}s_{56}s_{23456}s_{3456}} + \frac{2T_{23[54]16}}{s_{23}s_{45}s_{12345}s_{2345}} + \frac{2T_{23[54]61}}{s_{23}s_{45}s_{2345}s_{23456}} \\ &+ \frac{2T_{2314[65]}}{s_{23}s_{56}s_{123}s_{1234}} + \frac{2T_{2341[65]}}{s_{23}s_{56}s_{234}s_{1234}} + \frac{2T_{234[65]1}}{s_{23}s_{56}s_{234}s_{23456}} + \frac{2T_{564[32]1}}{s_{23}s_{56}s_{456}s_{23456}} \\ &+ \frac{2T_{3421[56]}}{s_{34}s_{56}s_{234}s_{1234}} + \frac{2T_{342[56]1}}{s_{34}s_{56}s_{234}s_{23456}} + \frac{T_{321456}}{s_{23}s_{123}s_{1234}s_{12345}} + \frac{T_{324156}}{s_{23}s_{234}s_{1234}s_{12345}} \end{aligned}$$

$$+\frac{T_{324516}}{s_{23}s_{234}s_{12345}s_{2345}}+\frac{T_{324561}}{s_{23}s_{234}s_{2345}s_{23456}}+\frac{T_{342156}}{s_{34}s_{234}s_{1234}s_{12345}}+\frac{T_{342516}}{s_{34}s_{234}s_{12345}s_{2345}}$$
$$+\frac{T_{342561}}{s_{34}s_{234}s_{2345}s_{23456}}+\frac{T_{345216}}{s_{34}s_{345}s_{12345}s_{2345}}+\frac{T_{345261}}{s_{34}s_{345}s_{2345}s_{23456}}+\frac{T_{345621}}{s_{34}s_{345}s_{23456}s_{3456}}$$
$$+\frac{T_{543216}}{s_{45}s_{345}s_{12345}s_{2345}}+\frac{T_{543261}}{s_{45}s_{345}s_{2345}s_{23456}}+\frac{T_{123456}}{s_{12}s_{123}s_{1234}s_{12345}}+\frac{T_{543621}}{s_{45}s_{345}s_{23456}s_{3456}}$$
$$+\frac{T_{546321}}{s_{45}s_{456}s_{23456}s_{3456}}+\frac{T_{564321}}{s_{56}s_{456}s_{23456}s_{3456}}\Big], \tag{B.5}$$

$s_{1234567}M_{1234567} =$

$$+\frac{8T_{12[34][[56]7]}}{s_{12}s_{34}s_{56}s_{567}s_{1234}}+\frac{8T_{54[67][1[23]]}}{s_{23}s_{45}s_{67}s_{123}s_{4567}}+\frac{8T_{12[34][5[67]]}}{s_{12}s_{34}s_{67}s_{567}s_{1234}}+\frac{8T_{45[67][3[12]]}}{s_{12}s_{45}s_{67}s_{123}s_{4567}}$$
$$+\frac{4T_{567[34][12]}}{s_{12}s_{34}s_{56}s_{567}s_{34567}}+\frac{4T_{12[34][56]7}}{s_{12}s_{34}s_{56}s_{1234}s_{123456}}+\frac{4T_{43[56][12]7}}{s_{12}s_{34}s_{56}s_{123456}s_{3456}}+\frac{4T_{43[56]7[12]}}{s_{12}s_{34}s_{56}s_{3456}s_{34567}}$$
$$+\frac{4T_{435[12][67]}}{s_{12}s_{34}s_{67}s_{345}s_{12345}}+\frac{4T_{435[67][12]}}{s_{12}s_{34}s_{67}s_{345}s_{34567}}+\frac{4T_{765[34][12]}}{s_{12}s_{34}s_{67}s_{567}s_{34567}}+\frac{4T_{12[34]5[67]}}{s_{12}s_{34}s_{67}s_{1234}s_{12345}}$$
$$+\frac{4T_{123[45][67]}}{s_{12}s_{45}s_{67}s_{123}s_{12345}}+\frac{4T_{453[12][67]}}{s_{12}s_{45}s_{67}s_{345}s_{12345}}+\frac{4T_{453[67][12]}}{s_{12}s_{45}s_{67}s_{345}s_{34567}}+\frac{4T_{45[67]3[12]}}{s_{12}s_{45}s_{67}s_{34567}s_{4567}}$$
$$+\frac{4T_{123[[45]6]7}}{s_{12}s_{45}s_{123}s_{456}s_{123456}}+\frac{4T_{5467[[12]3]}}{s_{12}s_{45}s_{123}s_{456}s_{4567}}+\frac{4T_{7654[[12]3]}}{s_{12}s_{67}s_{123}s_{567}s_{4567}}+\frac{4T_{1234[5[67]]}}{s_{12}s_{67}s_{123}s_{567}s_{1234}}$$
$$+\frac{4T_{321[45][67]}}{s_{23}s_{45}s_{67}s_{123}s_{12345}}+\frac{4T_{32[45]1[67]}}{s_{23}s_{45}s_{67}s_{12345}s_{2345}}+\frac{4T_{32[45][67]1}}{s_{23}s_{45}s_{67}s_{2345}s_{234567}}+\frac{4T_{5647[[12]3]}}{s_{12}s_{56}s_{123}s_{456}s_{4567}}$$
$$+\frac{4T_{1234[[56]7]}}{s_{12}s_{56}s_{123}s_{567}s_{1234}}+\frac{4T_{5674[[12]3]}}{s_{12}s_{56}s_{123}s_{567}s_{4567}}+\frac{4T_{123[4[56]]7}}{s_{12}s_{56}s_{123}s_{456}s_{123456}}+\frac{4T_{5647[1[23]]}}{s_{23}s_{56}s_{123}s_{456}s_{4567}}$$
$$+\frac{4T_{3214[[56]7]}}{s_{23}s_{56}s_{123}s_{567}s_{1234}}+\frac{4T_{5674[1[23]]}}{s_{23}s_{56}s_{123}s_{567}s_{4567}}+\frac{4T_{321[4[56]]7}}{s_{23}s_{56}s_{123}s_{456}s_{123456}}+\frac{4T_{3214[5[67]]}}{s_{23}s_{67}s_{123}s_{567}s_{1234}}$$
$$+\frac{4T_{7654[1[23]]}}{s_{23}s_{67}s_{123}s_{567}s_{4567}}+\frac{4T_{3241[5[67]]}}{s_{23}s_{67}s_{234}s_{567}s_{1234}}+\frac{4T_{324[5[67]]1}}{s_{23}s_{67}s_{234}s_{567}s_{234567}}+\frac{4T_{45[67][23]1}}{s_{23}s_{45}s_{67}s_{234567}s_{4567}}$$
$$+\frac{4T_{321[[45]6]7}}{s_{23}s_{45}s_{123}s_{456}s_{123456}}+\frac{4T_{5467[1[23]]}}{s_{23}s_{45}s_{123}s_{456}s_{4567}}+\frac{4T_{3241[[56]7]}}{s_{23}s_{56}s_{234}s_{567}s_{1234}}+\frac{4T_{324[[56]7]1}}{s_{23}s_{56}s_{234}s_{567}s_{234567}}$$
$$+\frac{4T_{3421[[56]7]}}{s_{34}s_{56}s_{234}s_{567}s_{1234}}+\frac{4T_{342[[56]7]1}}{s_{34}s_{56}s_{234}s_{567}s_{234567}}+\frac{4T_{3421[5[67]]}}{s_{34}s_{67}s_{234}s_{567}s_{1234}}+\frac{4T_{342[5[67]]1}}{s_{34}s_{67}s_{234}s_{567}s_{234567}}$$
$$+\frac{2T_{435[12]67}}{s_{12}s_{34}s_{345}s_{12345}s_{123456}}+\frac{2T_{4356127}}{s_{12}s_{34}s_{345}s_{123456}s_{3456}}+\frac{2T_{43567[12]}}{s_{12}s_{34}s_{345}s_{3456}s_{34567}}$$
$$+\frac{2T_{12[34]567}}{s_{12}s_{34}s_{1234}s_{12345}s_{123456}}+\frac{2T_{453[12]67}}{s_{12}s_{45}s_{345}s_{12345}s_{123456}}+\frac{2T_{4536[12]7}}{s_{12}s_{45}s_{345}s_{123456}s_{3456}}$$
$$+\frac{2T_{45367[12]}}{s_{12}s_{45}s_{345}s_{3456}s_{34567}}+\frac{2T_{4563[12]7}}{s_{12}s_{45}s_{456}s_{123456}s_{3456}}+\frac{2T_{45637[12]}}{s_{12}s_{45}s_{456}s_{3456}s_{34567}}$$

$$+\frac{2T_{45673[12]}}{s_{12}s_{45}s_{456}s_{34567}s_{4567}}+\frac{2T_{1234[56]7}}{s_{12}s_{56}s_{123}s_{1234}s_{123456}}+\frac{2T_{123[45]67}}{s_{12}s_{45}s_{123}s_{12345}s_{123456}}$$
$$+\frac{2T_{6543[12]7}}{s_{12}s_{56}s_{456}s_{123456}s_{3456}}+\frac{2T_{65437[12]}}{s_{12}s_{56}s_{456}s_{3456}s_{34567}}+\frac{2T_{65473[12]}}{s_{12}s_{56}s_{456}s_{34567}s_{4567}}$$
$$+\frac{2T_{65743[12]}}{s_{12}s_{56}s_{567}s_{34567}s_{4567}}+\frac{2T_{12345[67]}}{s_{12}s_{67}s_{123}s_{1234}s_{12345}}+\frac{2T_{67543[12]}}{s_{12}s_{67}s_{567}s_{34567}s_{4567}}$$
$$+\frac{2T_{321[45]67}}{s_{23}s_{45}s_{123}s_{12345}s_{123456}}+\frac{2T_{456[23]17}}{s_{23}s_{45}s_{456}s_{123456}s_{23456}}+\frac{2T_{456[23]71}}{s_{23}s_{45}s_{456}s_{23456}s_{234567}}$$
$$+\frac{2T_{4567[23]1}}{s_{23}s_{45}s_{456}s_{234567}s_{4567}}+\frac{2T_{32[45]167}}{s_{23}s_{45}s_{12345}s_{123456}s_{2345}}+\frac{2T_{3241[56]7}}{s_{23}s_{56}s_{234}s_{1234}s_{123456}}$$
$$+\frac{2T_{32[45]617}}{s_{23}s_{45}s_{123456}s_{2345}s_{23456}}+\frac{2T_{32[45]671}}{s_{23}s_{45}s_{2345}s_{23456}s_{234567}}+\frac{2T_{3214[56]7}}{s_{23}s_{56}s_{123}s_{1234}s_{123456}}$$
$$+\frac{2T_{324[56]17}}{s_{23}s_{56}s_{234}s_{123456}s_{23456}}+\frac{2T_{324[56]71}}{s_{23}s_{56}s_{234}s_{23456}s_{234567}}+\frac{2T_{654[23]17}}{s_{23}s_{56}s_{456}s_{123456}s_{23456}}$$
$$+\frac{2T_{654[23]71}}{s_{23}s_{56}s_{456}s_{23456}s_{234567}}+\frac{2T_{6547[23]1}}{s_{23}s_{56}s_{456}s_{234567}s_{4567}}+\frac{2T_{6574[23]1}}{s_{23}s_{56}s_{567}s_{234567}s_{4567}}$$
$$+\frac{2T_{32145[67]}}{s_{23}s_{67}s_{123}s_{1234}s_{12345}}+\frac{2T_{32415[67]}}{s_{23}s_{67}s_{234}s_{1234}s_{12345}}+\frac{2T_{32451[67]}}{s_{23}s_{67}s_{234}s_{12345}s_{2345}}$$
$$+\frac{2T_{3245[67]1}}{s_{23}s_{67}s_{234}s_{2345}s_{234567}}+\frac{2T_{6754[23]1}}{s_{23}s_{67}s_{567}s_{234567}s_{4567}}+\frac{2T_{3421[56]7}}{s_{34}s_{56}s_{234}s_{1234}s_{123456}}$$
$$+\frac{2T_{342[56]17}}{s_{34}s_{56}s_{234}s_{123456}s_{23456}}+\frac{2T_{342[56]71}}{s_{34}s_{56}s_{234}s_{23456}s_{234567}}+\frac{2T_{657[34]21}}{s_{34}s_{56}s_{567}s_{234567}s_{34567}}$$
$$+\frac{2T_{34[56]217}}{s_{34}s_{56}s_{123456}s_{23456}s_{3456}}+\frac{2T_{34[56]271}}{s_{34}s_{56}s_{23456}s_{234567}s_{3456}}+\frac{2T_{34[56]721}}{s_{34}s_{56}s_{234567}s_{3456}s_{34567}}$$
$$+\frac{2T_{34215[67]}}{s_{34}s_{67}s_{234}s_{1234}s_{12345}}+\frac{2T_{34251[67]}}{s_{34}s_{67}s_{234}s_{12345}s_{2345}}+\frac{2T_{3425[67]1}}{s_{34}s_{67}s_{234}s_{2345}s_{234567}}$$
$$+\frac{2T_{5432[67]1}}{s_{45}s_{67}s_{345}s_{2345}s_{234567}}+\frac{2T_{543[67]21}}{s_{45}s_{67}s_{345}s_{234567}s_{34567}}+\frac{2T_{54[67]321}}{s_{45}s_{67}s_{234567}s_{34567}s_{4567}}$$
$$+\frac{2T_{34521[67]}}{s_{34}s_{67}s_{345}s_{12345}s_{2345}}+\frac{2T_{3452[67]1}}{s_{34}s_{67}s_{345}s_{2345}s_{234567}}+\frac{2T_{345[67]21}}{s_{34}s_{67}s_{345}s_{234567}s_{34567}}$$
$$+\frac{2T_{675[34]21}}{s_{34}s_{67}s_{567}s_{234567}s_{34567}}+\frac{2T_{54321[67]}}{s_{45}s_{67}s_{345}s_{12345}s_{2345}}+\frac{T_{1234567}}{s_{12}s_{123}s_{1234}s_{12345}s_{123456}}$$
$$+\frac{T_{3214567}}{s_{23}s_{123}s_{1234}s_{12345}s_{123456}}+\frac{T_{3241567}}{s_{23}s_{234}s_{1234}s_{12345}s_{123456}}+\frac{T_{3245167}}{s_{23}s_{234}s_{12345}s_{123456}s_{2345}}$$
$$+\frac{T_{3245617}}{s_{23}s_{234}s_{123456}s_{2345}s_{23456}}+\frac{T_{3245671}}{s_{23}s_{234}s_{2345}s_{23456}s_{234567}}+\frac{T_{3456217}}{s_{34}s_{345}s_{123456}s_{23456}s_{3456}}$$

$$
\begin{aligned}
&+\frac{T_{3421567}}{s_{34}s_{234}s_{1234}s_{12345}s_{123456}}+\frac{T_{3425167}}{s_{34}s_{234}s_{12345}s_{123456}s_{2345}}+\frac{T_{3452617}}{s_{34}s_{345}s_{123456}s_{2345}s_{23456}}\\
&+\frac{T_{3425617)}}{s_{34}s_{234}s_{123456}s_{2345}s_{23456}}+\frac{T_{3425671)}}{s_{34}s_{234}s_{2345}s_{23456}s_{234567}}+\frac{T_{3452167}}{s_{34}s_{345}s_{12345}s_{123456}s_{2345}}\\
&+\frac{T_{3452671}}{s_{34}s_{345}s_{2345}s_{23456}s_{234567}}+\frac{T_{3456271}}{s_{34}s_{345}s_{23456}s_{234567}s_{3456}}+\frac{T_{3456721}}{s_{34}s_{345}s_{234567}s_{3456}s_{34567}}\\
&+\frac{T_{5432167}}{s_{45}s_{345}s_{12345}s_{123456}s_{2345}}+\frac{T_{5432617}}{s_{45}s_{345}s_{123456}s_{2345}s_{23456}}+\frac{T_{5436217}}{s_{45}s_{345}s_{123456}s_{23456}s_{3456}}\\
&+\frac{T_{5432671}}{s_{45}s_{345}s_{2345}s_{23456}s_{234567}}+\frac{T_{5436271}}{s_{45}s_{345}s_{23456}s_{234567}s_{3456}}+\frac{T_{5436721}}{s_{45}s_{345}s_{234567}s_{3456}s_{34567}}\\
&+\frac{T_{5463217}}{s_{45}s_{456}s_{123456}s_{23456}s_{3456}}+\frac{T_{5463271}}{s_{45}s_{456}s_{23456}s_{234567}s_{3456}}+\frac{T_{5463721}}{s_{45}s_{456}s_{234567}s_{3456}s_{34567}}\\
&+\frac{T_{5467321}}{s_{45}s_{456}s_{234567}s_{34567}s_{4567}}+\frac{T_{5643217}}{s_{56}s_{456}s_{123456}s_{23456}s_{3456}}+\frac{T_{5643271}}{s_{56}s_{456}s_{23456}s_{234567}s_{3456}}\\
&+\frac{T_{5643721}}{s_{56}s_{456}s_{234567}s_{3456}s_{34567}}+\frac{T_{5647321}}{s_{56}s_{456}s_{234567}s_{34567}s_{4567}}+\frac{T_{5674321}}{s_{56}s_{567}s_{234567}s_{34567}s_{4567}}\\
&+\frac{T_{7654321}}{s_{67}s_{567}s_{234567}s_{34567}s_{4567}}. \qquad \text{(B.6)}
\end{aligned}
$$

Appendix C. The cubic graphs of $M_{123\ldots n}$

As discussed in section 4.1, the expressions for $M_{123\ldots n}$ of Appendix B were found using the dictionary between the cubic diagrams of the $(n+1)-$point amplitude with one leg off-shell and BRST building blocks. The graphs which compose the expressions for $M_{123},\ldots,M_{123456}$ are given below (the 132 graphs used in the derivation of $M_{1234567}$ would occupy to much space and were omitted).

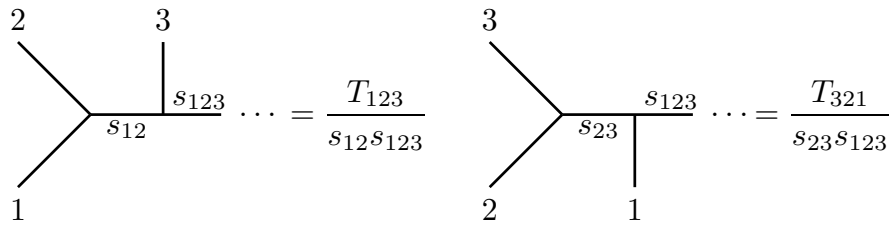


Fig. 8 The two cubic diagrams which constitute M_{123}.

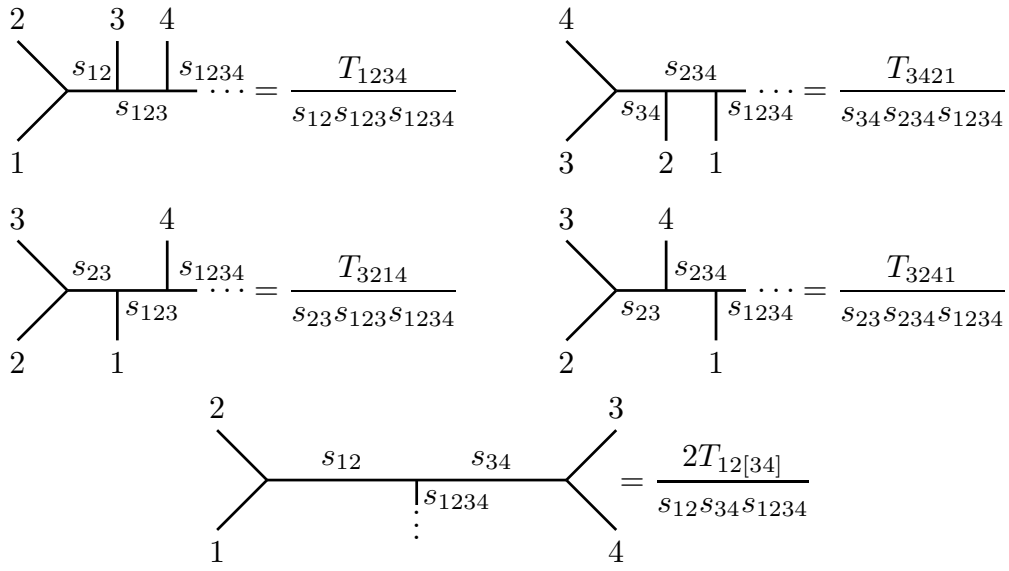
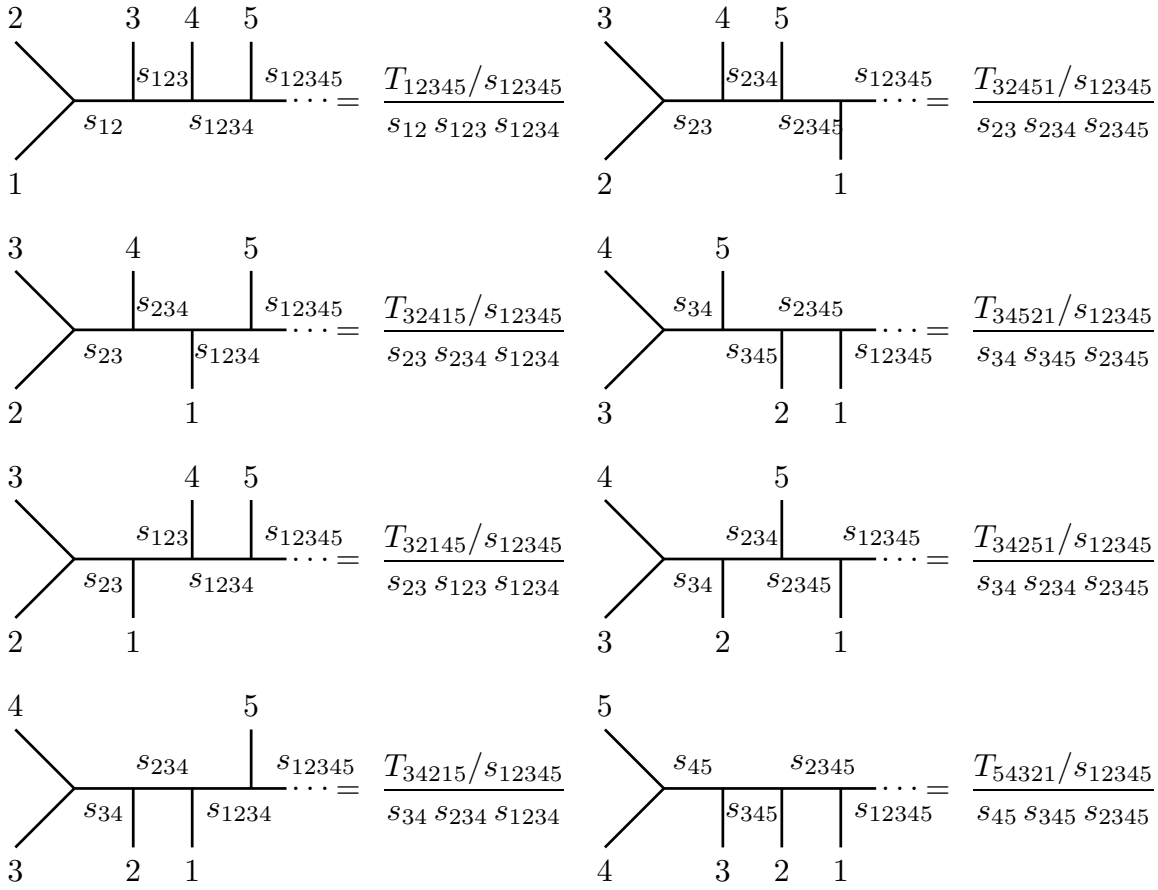


Fig. 9 The five cubic diagrams which constitute M_{1234}. The signs match the corresponding terms given in the formula (B.3).



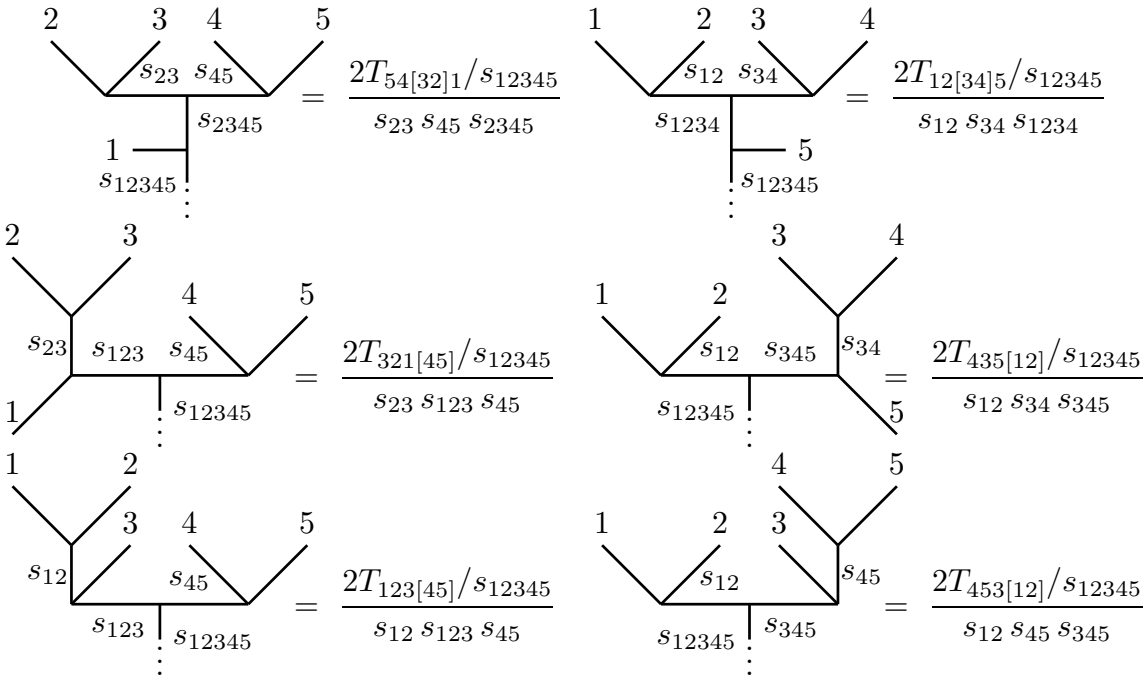
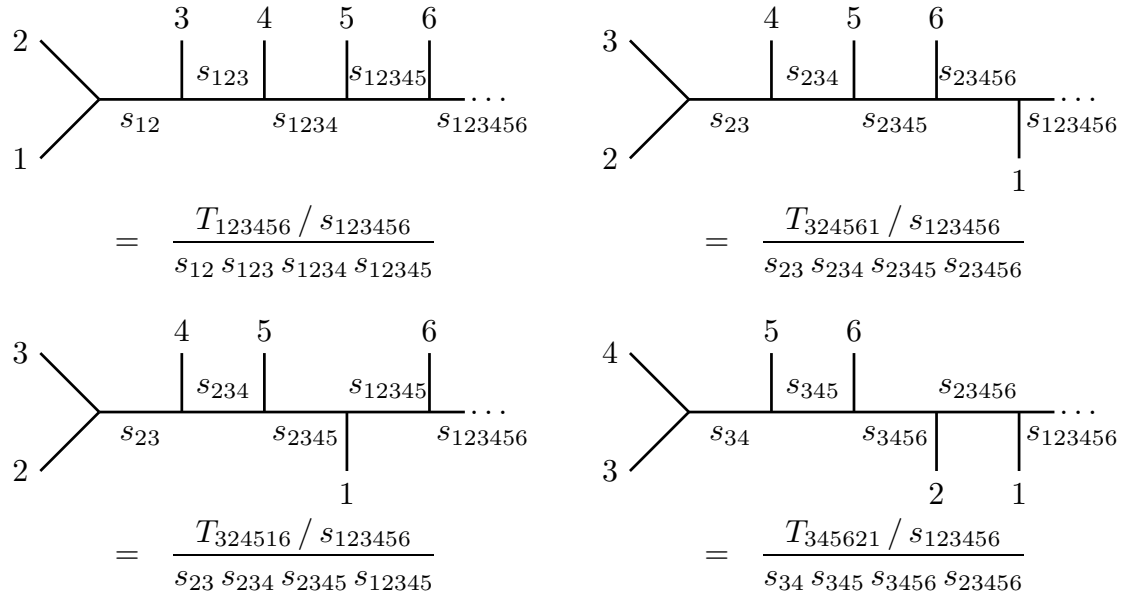
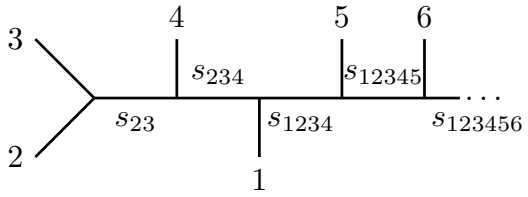
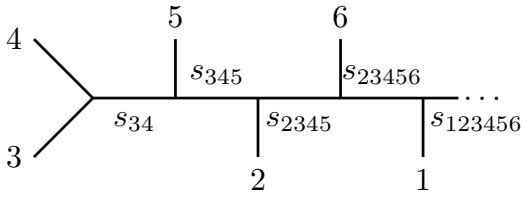


Fig. 10 The 14 cubic diagrams which constitute M_{12345}. The signs of their corresponding formulæ are in one-to-one agreement with the terms in expression for M_{12345} given by (B.4), which is reproduced by summing all 14 graphs displayed here.

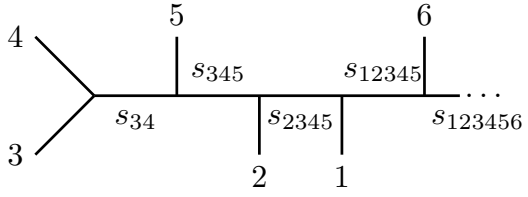




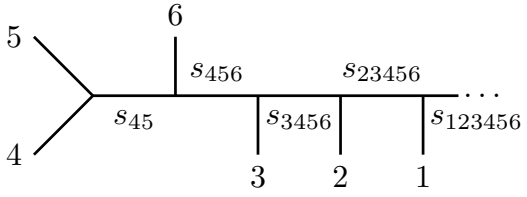
$$= \frac{T_{324156} \,/\, s_{123456}}{s_{23}\, s_{234}\, s_{1234}\, s_{12345}}$$



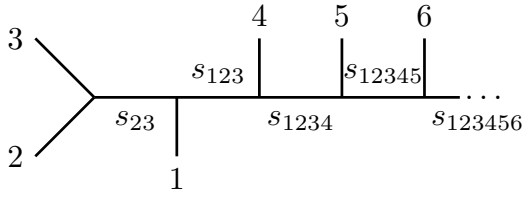
$$= \frac{T_{345261} \,/\, s_{123456}}{s_{34}\, s_{345}\, s_{2345}\, s_{23456}}$$



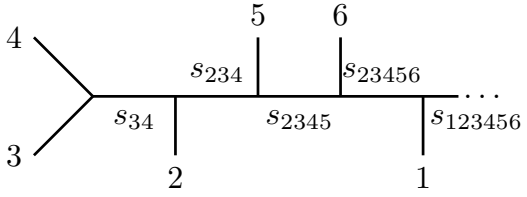
$$= \frac{T_{345216} \,/\, s_{123456}}{s_{34}\, s_{345}\, s_{2345}\, s_{12345}}$$



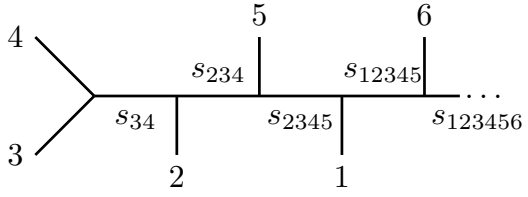
$$= \frac{T_{546321} \,/\, s_{123456}}{s_{45}\, s_{456}\, s_{3456}\, s_{23456}}$$



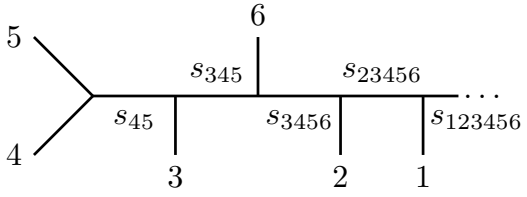
$$= \frac{T_{321456} \,/\, s_{123456}}{s_{23}\, s_{123}\, s_{1234}\, s_{12345}}$$



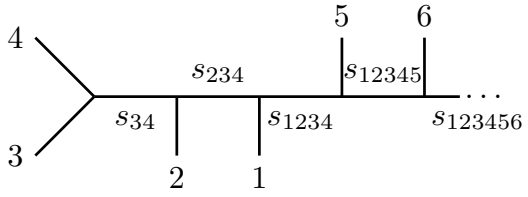
$$= \frac{T_{342561} \,/\, s_{123456}}{s_{34}\, s_{234}\, s_{2345}\, s_{23456}}$$



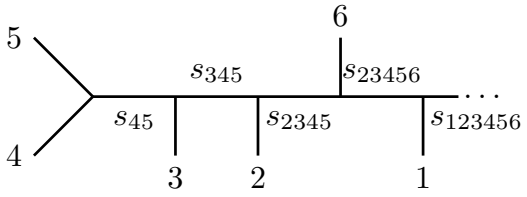
$$= \frac{T_{342516} \,/\, s_{123456}}{s_{34}\, s_{234}\, s_{2345}\, s_{12345}}$$



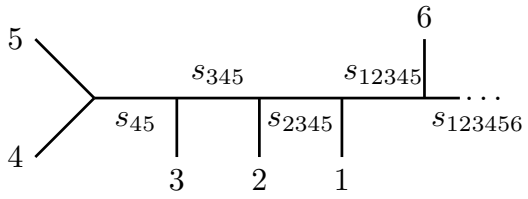
$$= \frac{T_{543621} \,/\, s_{123456}}{s_{45}\, s_{345}\, s_{3456}\, s_{23456}}$$



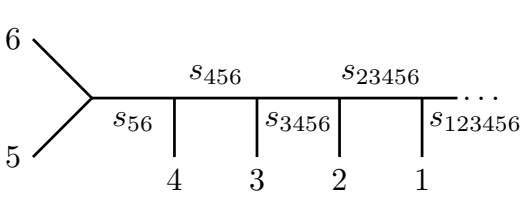
$$= \frac{T_{342156} \,/\, s_{123456}}{s_{34}\, s_{234}\, s_{1234}\, s_{12345}}$$



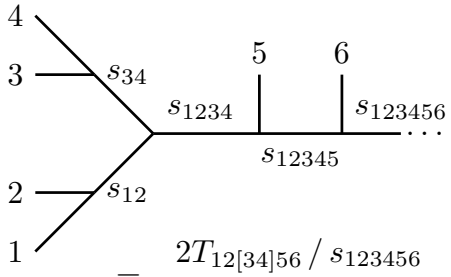
$$= \frac{T_{543261} \,/\, s_{123456}}{s_{45}\, s_{345}\, s_{2345}\, s_{23456}}$$



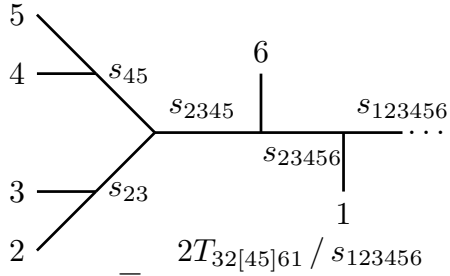
$$= \frac{T_{543216} \,/\, s_{123456}}{s_{45}\, s_{345}\, s_{2345}\, s_{12345}}$$



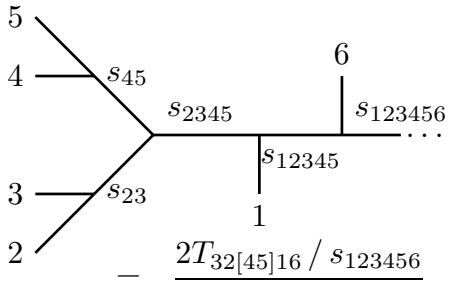
$$= \frac{T_{564321} \,/\, s_{123456}}{s_{56}\, s_{456}\, s_{3456}\, s_{23456}}$$



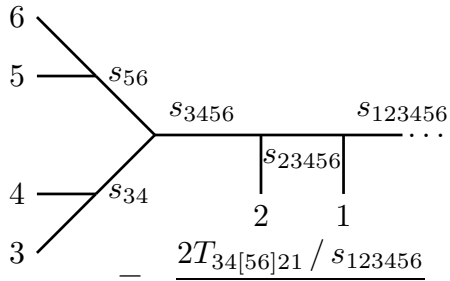
$$= \frac{2T_{12[34]56} \,/\, s_{123456}}{s_{12}\, s_{34}\, s_{1234}\, s_{12345}}$$



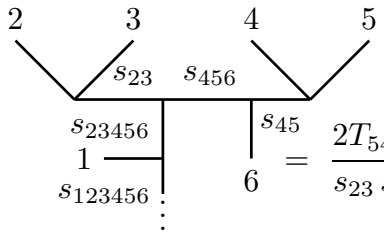
$$= \frac{2T_{32[45]61} \,/\, s_{123456}}{s_{23}\, s_{45}\, s_{2345}\, s_{23456}}$$



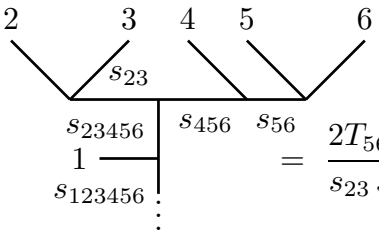
$$= \frac{2T_{32[45]16} \,/\, s_{123456}}{s_{23}\, s_{45}\, s_{2345}\, s_{12345}}$$



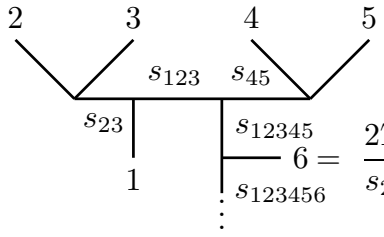
$$= \frac{2T_{34[56]21} \,/\, s_{123456}}{s_{34}\, s_{56}\, s_{3456}\, s_{23456}}$$



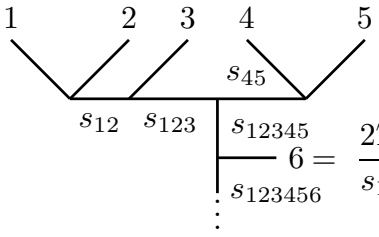
$$= \frac{2T_{546[32]1}/s_{123456}}{s_{23}\, s_{45}\, s_{456}\, s_{23456}}$$



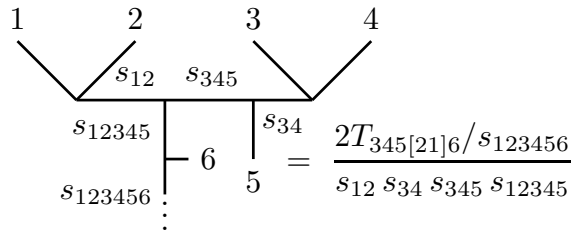
$$= \frac{2T_{564[32]1}/s_{123456}}{s_{23}\, s_{56}\, s_{456}\, s_{23456}}$$



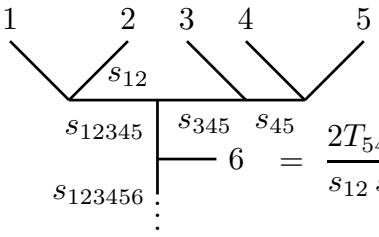
$$= \frac{2T_{321[45]6}/s_{123456}}{s_{23}\, s_{45}\, s_{123}\, s_{12345}}$$



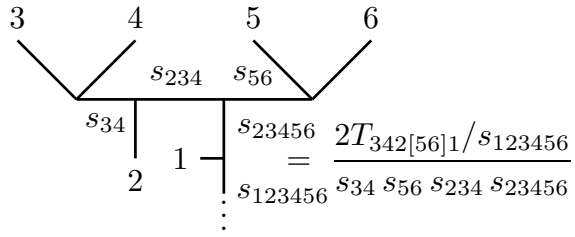
$$= \frac{2T_{123[45]6}/s_{123456}}{s_{12}\, s_{45}\, s_{123}\, s_{12345}}$$



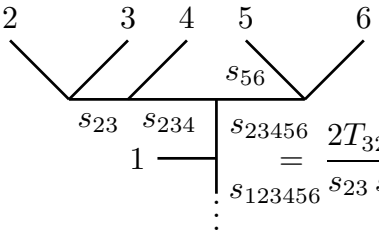
$$= \frac{2T_{345[21]6}/s_{123456}}{s_{12}\,s_{34}\,s_{345}\,s_{12345}}$$



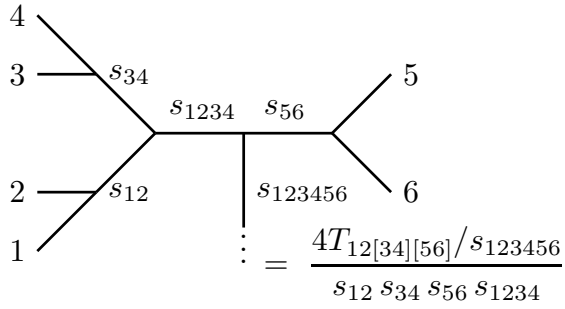
$$= \frac{2T_{543[21]6}/s_{123456}}{s_{12}\,s_{45}\,s_{345}\,s_{12345}}$$



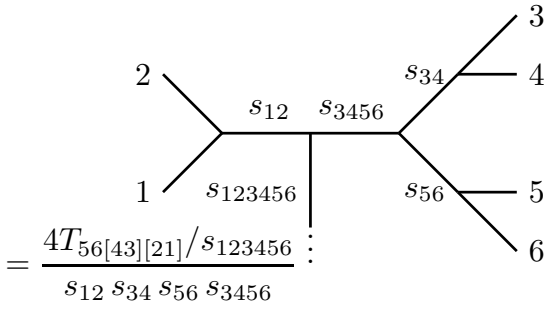
$$= \frac{2T_{342[56]1}/s_{123456}}{s_{34}\,s_{56}\,s_{234}\,s_{23456}}$$



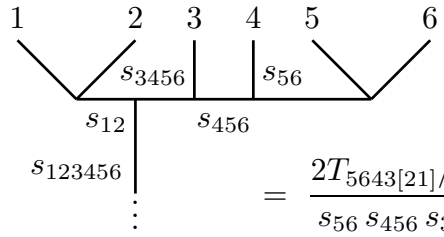
$$= \frac{2T_{324[56]1}/s_{123456}}{s_{23}\,s_{56}\,s_{234}\,s_{23456}}$$



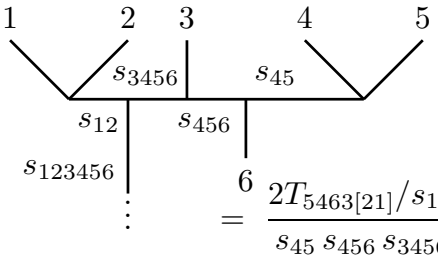
$$= \frac{4T_{12[34][56]}/s_{123456}}{s_{12}\,s_{34}\,s_{56}\,s_{1234}}$$



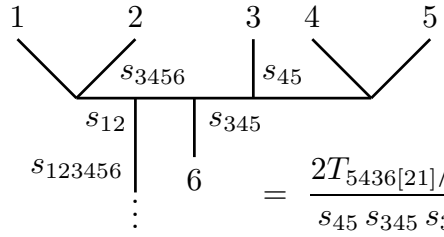
$$= \frac{4T_{56[43][21]}/s_{123456}}{s_{12}\,s_{34}\,s_{56}\,s_{3456}}$$



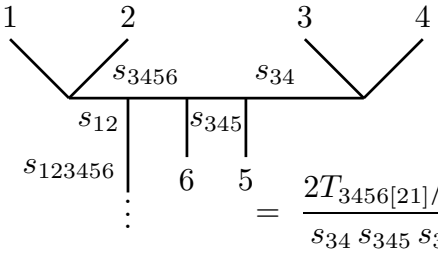
$$= \frac{2T_{5643[21]}/s_{123456}}{s_{56}\,s_{456}\,s_{3456}\,s_{12}}$$



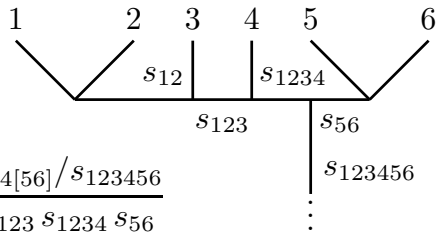
$$= \frac{2T_{5463[21]}/s_{123456}}{s_{45}\,s_{456}\,s_{3456}\,s_{12}}$$



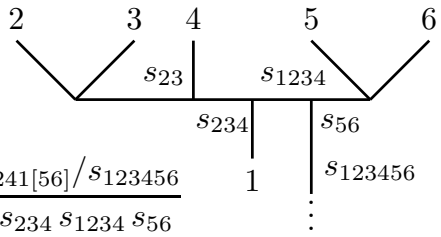
$$= \frac{2T_{5436[21]}/s_{123456}}{s_{45}\,s_{345}\,s_{3456}\,s_{12}}$$



$$= \frac{2T_{3456[21]}/s_{123456}}{s_{34}\,s_{345}\,s_{3456}\,s_{12}}$$



$$= \frac{2T_{1234[56]}/s_{123456}}{s_{12}\,s_{123}\,s_{1234}\,s_{56}}$$



$$= \frac{2T_{3241[56]}/s_{123456}}{s_{23}\,s_{234}\,s_{1234}\,s_{56}}$$

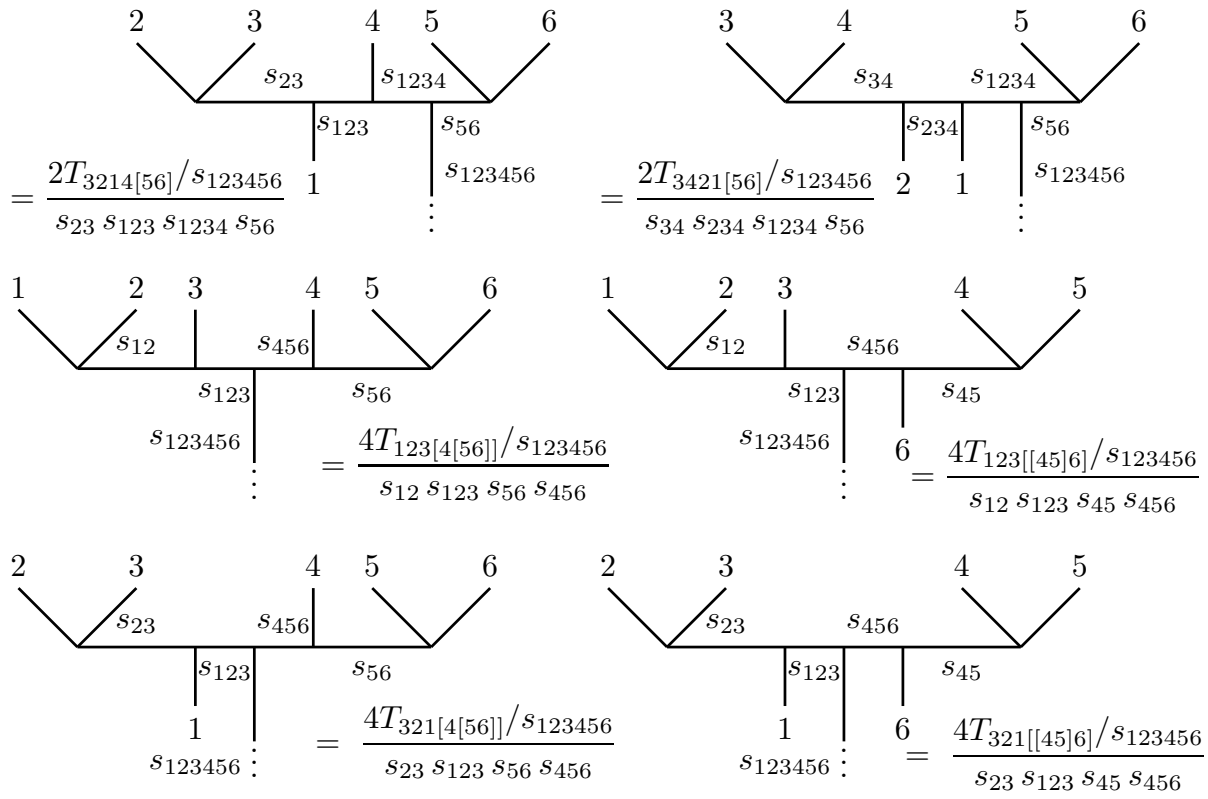


Fig. 11 The 42 cubic diagrams which constitute M_{123456}. The signs of their corresponding formulæ are in one-to-one agreement with the terms in expression for M_{123456} given by (B.5), which is reproduced by summing all 42 graphs displayed here.

References

[1] C.R. Mafra, O. Schlotterer, S. Stieberger and D. Tsimpis, "A recursive formula for N-point SYM tree amplitudes," arXiv:1012.3981 [hep-th], to appear in Phys. Rev. D.

[2] C.R. Mafra, O. Schlotterer and S. Stieberger, "Complete N–Point Superstring Disk Amplitude. II. Amplitude and Hypergeometric Function Structure," AEI–2011–35, MPP-2011–65.

[3] M.B. Green, J.H. Schwarz, E. Witten, "Superstring Theory. Vol. 1: Introduction," Cambridge, Uk: Univ. Pr. (1987) 469 P. (Cambridge Monographs On Mathematical Physics).

[4] M.B. Green, J.H. Schwarz, "Supersymmetrical Dual String Theory. 2. Vertices and Trees," Nucl. Phys. **B198**, 252-268 (1982);
J.H. Schwarz, "Superstring Theory," Phys. Rept. **89**, 223-322 (1982);
A.A. Tseytlin, "Vector Field Effective Action in the Open Superstring Theory," Nucl. Phys. **B276**, 391 (1986).

[5] R. Medina, F.T. Brandt and F.R. Machado, "The open superstring 5-point amplitude revisited," JHEP **0207**, 071 (2002) [arXiv:hep-th/0208121] ;
L.A. Barreiro and R. Medina, "5-field terms in the open superstring effective action," JHEP **0503**, 055 (2005) [arXiv:hep-th/0503182].

[6] D. Oprisa and S. Stieberger, "Six gluon open superstring disk amplitude, multiple hypergeometric series and Euler-Zagier sums," arXiv:hep-th/0509042 ;
S. Stieberger, T.R. Taylor, "Amplitude for N-Gluon Superstring Scattering," Phys. Rev. Lett. **97**, 211601 (2006). [hep-th/0607184]; "Multi-Gluon Scattering in Open Superstring Theory," Phys. Rev. **D74**, 126007 (2006). [hep-th/0609175]; "Complete Six-Gluon Disk Amplitude in Superstring Theory," Nucl. Phys. **B801**, 128-152 (2008). [arXiv:0711.4354 [hep-th]].

[7] S. Stieberger and T.R. Taylor, "Supersymmetry Relations and MHV Amplitudes in Superstring Theory," Nucl. Phys. B **793**, 83 (2008) [arXiv:0708.0574 [hep-th]].

[8] N. Berkovits, "Super-Poincare covariant quantization of the superstring," JHEP **0004**, 018 (2000) [arXiv:hep-th/0001035].

[9] N. Berkovits, B.C. Vallilo, "Consistency of superPoincare covariant superstring tree amplitudes," JHEP **0007**, 015 (2000). [hep-th/0004171].

[10] G. Policastro and D. Tsimpis, "R^4, purified," Class. Quant. Grav. **23**, 4753 (2006) [arXiv:hep-th/0603165].

[11] C.R. Mafra, "Pure Spinor Superspace Identities for Massless Four-point Kinematic Factors," JHEP **0804**, 093 (2008). [arXiv:0801.0580 [hep-th]].

[12] C.R. Mafra, "Simplifying the Tree-level Superstring Massless Five-point Amplitude," JHEP **1001**, 007 (2010) [arXiv:0909.5206 [hep-th]].

[13] C.R. Mafra, O. Schlotterer, S. Stieberger and D. Tsimpis, "Six Open String Disk Amplitude in Pure Spinor Superspace," Nucl. Phys. B **846**, 359 (2011) [arXiv:1011.0994 [hep-th]].

[14] Z. Bern, J.J.M. Carrasco, H. Johansson, "New Relations for Gauge-Theory Amplitudes," Phys. Rev. **D78**, 085011 (2008). [arXiv:0805.3993 [hep-ph]].

[15] N.E.J. Bjerrum-Bohr, P.H. Damgaard, P. Vanhove, "Minimal Basis for Gauge Theory Amplitudes," Phys. Rev. Lett. **103**, 161602 (2009). [arXiv:0907.1425 [hep-th]].

[16] S. Stieberger, "Open & Closed vs. Pure Open String Disk Amplitudes," [arXiv:0907.2211 [hep-th]].

[17] F.A. Berends, W.T. Giele, "Recursive Calculations for Processes with n Gluons," Nucl. Phys. **B306**, 759 (1988).

[18] M.B. Green, J.H. Schwarz, "Covariant Description of Superstrings," Phys. Lett. **B136**, 367-370 (1984).

[19] W. Siegel, "Classical Superstring Mechanics," Nucl. Phys. **B263**, 93 (1986).

[20] N. Berkovits, "ICTP lectures on covariant quantization of the superstring," [hep-th/0209059].

[21] E. Witten, "Twistor-Like Transform In Ten-Dimensions," Nucl. Phys. B **266**, 245 (1986).

[22] J.P. Harnad and S. Shnider, "Constraints And Field Equations For Ten-Dimensional Superyang-Mills Theory," Commun. Math. Phys. **106**, 183 (1986) ;
P.A. Grassi and L. Tamassia, "Vertex operators for closed superstrings," JHEP **0407**, 071 (2004) [arXiv:hep-th/0405072].

[23] P.S. Howe, "Pure spinors lines in superspace and ten-dimensional supersymmetric theories," Phys. Lett. **B258**, 141-144 (1991).

[24] N. Berkovits, "Explaining Pure Spinor Superspace," [hep-th/0612021].

[25] N. Berkovits, "Cohomology in the pure spinor formalism for the superstring," JHEP **0009**, 046 (2000). [hep-th/0006003].

[26] C.R. Mafra, "Towards Field Theory Amplitudes From the Cohomology of Pure Spinor Superspace," JHEP **1011**, 096 (2010) [arXiv:1007.3639 [hep-th]].

[27] E. D'Hoker, D.H. Phong, "The Geometry of String Perturbation Theory," Rev. Mod. Phys. **60**, 917 (1988).

[28] N. Berkovits, C.R. Mafra, "Equivalence of two-loop superstring amplitudes in the pure spinor and RNS formalisms," Phys. Rev. Lett. **96**, 011602 (2006). [hep-th/0509234].

[29] C. Stahn, "Fermionic superstring loop amplitudes in the pure spinor formalism," JHEP **0705**, 034 (2007). [arXiv:0704.0015 [hep-th]].

[30] N. Berkovits and C.R. Mafra, "Some superstring amplitude computations with the non-minimal pure spinor formalism," JHEP **0611**, 079 (2006) [arXiv:hep-th/0607187].

[31] C.R. Mafra, "PSS: A FORM Program to Evaluate Pure Spinor Superspace Expressions," [arXiv:1007.4999 [hep-th]].

[32] J.A.M. Vermaseren, "New features of FORM," arXiv:math-ph/0010025. ;
M. Tentyukov and J.A.M. Vermaseren, "The multithreaded version of FORM," arXiv:hep-ph/0702279.

[33] R. Kleiss, H. Kuijf, "Multi - Gluon Cross-sections And Five Jet Production At Hadron Colliders," Nucl. Phys. **B312**, 616 (1989).

[34] C.R. Mafra, O. Schlotterer, S. Stieberger, "Explicit BCJ Numerators from Pure Spinors," [arXiv:1104.5224 [hep-th]].