

Master thesis  
**Amplitudes from string theory and the  
CHY formalism**

advisors: Emil Bjerrum-Bohr & Poul Henrik Damgaard

Christian Baadsgaard Jepsen

*Niels Bohr International Academy,  
The University of Copenhagen*

August 12, 2015



## Abstract

This thesis examines and compares string theory and the CHY formalism as methods of computing tree-level amplitudes. While the two formalisms operate with different mathematical languages, the mechanisms by which scattering amplitudes can be calculated from them bear such a striking resemblance that either formalism can be studied to gain insight to the other. Based on the divergent behaviour of string theory integrals as the Regge slope tends to zero, a simple combinatorial algorithm is derived for computing the field theory limit of a large class of string theory integrals. Prompted by this result, an analogue derivation of a complementary algorithm is carried out in a CHY setting of complex contour integrals enclosing the solutions to a set of rational equations. By employing this algorithm, an identification is made between individual scalar Feynman diagrams and a certain class of integrands, and a method is provided by which all other integrands can be reduced to those of this class, whereby the algorithm becomes a sufficient tool for calculating gluon and graviton amplitudes. Lastly, in string theory, the algorithmic rules are used to derive a closed-form expression for the colour-ordered  $\phi^p$  tree-amplitude.

*The diagrams comprehend the profoundest secrets of the universe; and the power of exciting the various motions of the universe depends on their explanation;- the power to effect transmutation depends on the understanding of the diagrams of Changes. — Confucius, commentary on the I Ching<sup>[2]</sup>*

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>String theory integration rules</b>	<b>5</b>
2.1	Motivation and set-up . . . . .	5
2.2	Deriving the rules . . . . .	7
2.3	Discussion and examples . . . . .	11
<b>3</b>	<b>CHY integration rules</b>	<b>13</b>
3.1	Motivation and set-up . . . . .	13
3.2	Deriving the rules . . . . .	15
3.3	Discussion and examples . . . . .	19
<b>4</b>	<b>Reduction of higher-order poles</b>	<b>23</b>
4.1	Even $n$ reductions . . . . .	24
4.2	Odd $n$ reductions . . . . .	25
4.3	Numerators and templateless diagrams . . . . .	25
<b>5</b>	<b>Cycle integrals</b>	<b>27</b>
5.1	The first string theory–CHY duality . . . . .	27
5.2	Polygon decomposition: Dual Feynman diagrams . . . . .	28
5.3	Brute force phi-to-the- $p^{\text{th}}$ . . . . .	33
<b>6</b>	<b><math>\phi^p</math> from string theory</b>	<b>35</b>
6.1	$\phi^4$ and the second string theory–CHY duality . . . . .	35
6.2	$\phi^p$ rules . . . . .	37
6.3	The $\phi^p$ procedure . . . . .	38
6.4	$\phi^3$ . . . . .	42
6.5	$\phi^4$ . . . . .	42
6.6	$\phi^5$ . . . . .	46
6.7	$\phi^p$ for $p > 4$ . . . . .	50
6.8	Discussion . . . . .	51
<b>7</b>	<b>Möbius invariance and massive particles</b>	<b>54</b>
<b>8</b>	<b>Conclusion</b>	<b>57</b>
<b>9</b>	<b>Acknowledgement</b>	<b>58</b>
	<b>Appendices</b>	<b>59</b>
<b>A</b>	<b>Table of CHY integrals with higher order poles</b>	<b>59</b>
A.1	$n = 4$ . . . . .	59
A.2	$n = 5$ . . . . .	59
A.3	$n = 6$ . . . . .	59
<b>B</b>	<b>Expanding diverging integrals beyond leading order</b>	<b>62</b>

# 1 Introduction

As high-precision experimental particle physics continues to advance, the community of theoretical particle physicists are, in their predictions, continually faced with the challenge of keeping pace with the accuracy that is required by the experiments. To meet this end, several methods have been developed to compute field theory amplitudes, which encode the essential information of scattering processes, and in studying the amplitudes, structures have been uncovered that help provide a deeper understanding of the underlying theory. One of the most recent developments in this process, is the entirely novel way of calculating scattering amplitudes that was discovered by Cachazo, He, and Yuan, henceforth abbreviated CHY.

In 2013 CHY observed the existence of a certain set of equations that relate the Mandelstam variables of  $n$  massless particles to the  $n$ -punctured Riemann sphere. These equations, which they dubbed *the scattering equations*, exhibit several remarkable properties: they are Möbius invariant, general kinematic invariants can be directly extracted from them, and their solutions satisfy KLT orthogonality. For this reason CHY conjectured that the Yang-Mills and gravity S-matrices could be expressed in terms of the scattering equations [3]. CHY quickly followed up this result by providing explicit formulas for the tree-level amplitudes of gluons and gravitons [4] and shortly after also supplied a formulation of cubic scalar theory in terms of the scattering equations [5]. Subsequent developments by CHY include the extension of the formalism to Einstein-Yang-Mills and Einstein-Yang-Mills-scalar theory [6] and to Einstein-Maxwell, Yang-Mills-scalar, quartic scalar, Born-Infeld, and Dirac-Born-Infeld theory as well as to the non-linear sigma model [7]. By using the BCFW recursion relations, Dolan and Goddard have been able to provide rigorous proofs of the CHY formulas for the cubic scalar and Yang-Mills amplitudes [8]. While the original CHY formulas applied only to massless particles, it has since been realized that a simple procedure exists for adding masses to the particles [8],[9].

Despite the compactness of the CHY formulas, they do not in themselves provide a practical method of calculating amplitudes. The formulas express amplitudes as integrals fully localized, by delta functions, on the solutions to the scattering equations. By solving the scattering equations, plugging the solutions into the correct rational function, and summing over all solutions, one obtains the amplitude. But the number of solutions to the scattering equations increases factorially with the number of particles, and the complexity of the individual solutions escalates drastically, making the task of solving the equations tremendously difficult – even though the final answer arrived at after summing over the solutions is usually very simple. This state of affairs begs the question of whether CHY type integrals can in general be evaluated without actually solving the scattering equations, and recent papers [10],[11] have shown that this is indeed the case. Finding easier ways of calculating CHY integrals and investigating whether the CHY formalism can potentially provide a more powerful means of calculating amplitudes than conventional methods is therefore an active field of research and is the main subject of this thesis.

The approach adopted here in order to pursue this course of inquiry is to seek inspiration from string theory. By taking the infinite-tension limit of string theory, one can derive ordinary field theory. When calculating amplitudes by taking this limit of string theory, one encounters integrals that diverge in the limit but which are multiplied with a pre-factor (the Regge slope) that tends to zero such that the product converges on the correct finite value [12]. Therefore, in taking the field theory limit, the parts of the integration domain

that do not diverge are killed off by the pre-factor, so that in the end only the diverging parts of the integration domain, which in many cases consist of only a finite number of individual points, contributes to the final result. This process is clearly reminiscent of the CHY procedure where amplitudes are calculated from integrals that are fully localized to a discrete set of points, and in 2013 it was discovered that the scattering equations could be reproduced by taking a chiral infinite-tension limit of string theory [13], [14]. Further motivation for using a string theory approach to study the scattering equations can be found in [15], where the picture changing formalism is used to re-express the superstring amplitude in a form where it can be converted into an equivalent CHY integral by a slight modification of the measure of integration.

This thesis presents the results of work carried out in collaboration with Emil Bjerrum-Bohr, Poul Henrik Damgaard, and Jacob Bourjaily, and is organized as follows: Section 2 considers the general form of string theory integrals and derives a set of rules for evaluating them. Section 3 develops a similar set of rules for generic integrals appearing in the CHY formalism. More specifically, it will be shown that a certain subset of CHY integrals evaluate to a sum of  $\phi^3$  Feynman diagrams, and a procedure will be stated for determining which integrals are of this type and for evaluating them. Section 4 demonstrates how to calculate the CHY integrals that are not equal to sums of Feynman diagrams, and which are therefore not immediately encompassed by the integration rules of section 3, but which nonetheless are part of the CHY expressions for gluons and gravitons. By comparing the string theory and CHY integration rules of sections 2 and 3, a connection between the two formalisms is established in section 5, where it is also shown how individual Feynman diagrams tie into either formalism through a dual type of diagrams. Section 6 employs the string theory integration rules to generalize the Koba-Nielsen formula by constructing a closed-form expression for the colour-ordered tree-level  $\phi^p$  amplitude. In order not to complicate matters with extraneous considerations, attention is restricted to massless particles throughout sections 2 to 6, but in section 7 this deficiency is remedied as it is shown how to extend the results to massive particles.

Most of the results in this thesis can also be found in two papers on arXiv: [16], [17].

## 2 String theory integration rules

### 2.1 Motivation and set-up

While the question of whether string theory is fundamentally a true description of nature remains very much unresolved, the theory in any event deserves careful study for the insight it brings into quantum field theory. One application where string theory has proved particularly fruitful is in the calculation of scattering amplitudes. Historically, one of the first discoveries along this line of research was the Veneziano amplitude and its extension to the Koba-Nielsen formula, from which the the colour-ordered  $\phi^3$  amplitude can be calculated by letting the Regge slope  $\alpha'$  tend to zero:<sup>1</sup>

$$\mathcal{A}_n^{\phi^3} = \lim_{\alpha' \rightarrow 0} (\alpha')^{n-3} \int d\mu_n \Lambda_n(\alpha', k, z) \frac{1}{\prod_{i=1}^n (z_i - z_{i+1})}, \quad z_{n+1} = z_1. \quad (1)$$

---

<sup>1</sup>Throughout the thesis, I omit pre-factors such as colour and coupling constants and  $2^{n-3}$ , which have no bearing on the enquiry at hand.

Here  $\Lambda_n(\alpha', k, z)$  denotes the Koba-Nielsen factor, given by

$$\Lambda_n(\alpha', k, z) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n (z_i - z_j)^{\alpha' s_{ij}}, \quad s_{ij} = k_i \cdot k_j, \quad (2)$$

and  $d\mu_n$  is the string theory integration measure, which is defined as

$$d\mu_n = \delta(z_A - z_A^0) \delta(z_B - z_B^0) \delta(z_C - z_C^0) \quad (3)$$

$$(z_A - z_B)(z_A - z_C)(z_B - z_C) \prod_{i=2}^n \theta(z_{i-1} - z_i) \prod_{i=1}^n dz_i,$$

where  $\theta(z)$  is the Heaviside function. Because of gauge freedom, one can freely choose  $z_A$ ,  $z_B$ , and  $z_C$  without changing the result of the integration.

Another example of a string theory amplitude is the tree amplitude of  $n$  massless vector particles from the bosonic string:<sup>[18]</sup>

$$A_n^{\text{b.s.}} = (\alpha')^{(n-4)/2} \int d\mu_n \Lambda_n(k, z) H_n^{\text{b.s.}}(k, \epsilon, z) \quad (4)$$

where  $H_n^{\text{b.s.}}(k, \epsilon, z)$  is given by the part of

$$\exp \sum_{i \neq j} \left[ \frac{1}{2} \frac{\epsilon_i \cdot \epsilon_j}{(z_i - z_j)^2} - \sqrt{\alpha'} \frac{k_i \cdot \epsilon_j}{z_i - z_j} \right]. \quad (5)$$

that is linear in each polarization vector  $\epsilon_i$ .

There is also a tree-amplitude formula for  $n$  massless vector particles for the super string:<sup>[18]</sup>

$$A_n^{\text{s.s.}} = (\alpha')^{(n-4)/2} \int d\mu_n \int \left( \prod d\theta_i \right) \prod_{i < j} (z_i - z_j - \theta_i \theta_j)^{\alpha' s_{ij}} H_n^{\text{s.s.}}(k, \epsilon, z, \theta), \quad (6)$$

where the integrand  $H_n^{\text{s.s.}}$  is defined in terms of the Grassmann variables  $\theta_i$  and  $\phi_i$  as follows:

$$H_n^{\text{s.s.}} = \int \left( \prod_i d\phi_i \right) \prod_{i < j} \exp \left[ \sqrt{\alpha'} \frac{(\theta_i - \theta_j)(\phi_i \epsilon_i \cdot k_j + \phi_j \epsilon_j \cdot k_i)}{z_i - z_j} - \frac{\phi_i \phi_j \epsilon_i \cdot \epsilon_j}{z_i - z_j} - \frac{\theta_i \theta_j \phi_i \phi_j \epsilon_i \cdot \epsilon_j}{(z_i - z_j)^2} \right]. \quad (7)$$

There are many other string theory amplitudes, but these few examples will serve to illustrate a general property of string tree-amplitudes, namely that they can all be expanded into terms of the following form:

$$I_n[H] = \int d\mu_n \Lambda_n(\alpha', k, z) H(z), \quad (8)$$

where  $H(z)$  is a product of differences lifted to some powers:

$$H(z) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n (z_i - z_j)^{c_{ij}}. \quad (9)$$

Another common property of the string theory tree-amplitudes is that they exhibit  $SL(2, \mathbb{C})$ , or Möbius, invariance – that is to say, invariance under the following transformation of the integration variables:

$$z_i \rightarrow \frac{Az_i + B}{Cz_i + D}, \quad \text{with } AD - BC = 1, \quad \text{for } i = 1, \dots, n. \quad (10)$$

This invariance is what ensures the freedom in selecting  $z_A, z_B, z_C, z_A^0, z_B^0$ , and  $z_C^0$  in the measure (3). To retain this property we impose on the exponents  $c_{ij}$  in (9) the requirement that for each  $i = 1, \dots, n$

$$\sum_{\substack{j=1 \\ j \neq i}}^n c_{ij} = -2, \quad \text{where } c_{ij} = c_{ji}. \quad (11)$$

## 2.2 Deriving the rules

Seeing how integrals of the form (8) constitute the building blocks of string theory tree-amplitudes, it is a worthwhile endeavour – which we will presently engage in – to examine how these types of integrals behave as  $\alpha'$  tends to zero. For concreteness, it will be convenient to work in a specific gauge, and the one we will adopt is the following:

$$z_A = z_1, \quad z_A^0 = \infty, \quad z_B = z_2, \quad z_B^0 = 1, \quad z_C = z_n, \quad z_C^0 = 0. \quad (12)$$

We now set out to find a general method for evaluating  $I_n[H]$  to leading order in  $\alpha'$ . The cases of most interest are those where  $I_n[H]$  diverges in the  $\alpha' \rightarrow 0$  limit. Such divergences can come about when some of the variables  $z_i, i \in \{2, \dots, n\}$  tend to the same value since  $H(z)$  contains differences  $(z_i - z_j)$  raised to negative powers. Because of the integration domain, only consecutive variables can tend to the same value.

Consider, therefore, the variables  $z_i$  to  $z_{i+m}$  and define

$$y_j = z_i - z_j, \quad \text{for } j = i, i + 1, \dots, i + m. \quad (13)$$

In that case

$$\int_0^{z_i} dz_{i+1} \int_0^{z_{i+1}} dz_{i+2} \dots \int_0^{z_{i+m-1}} dz_{i+m} = \int_0^{z_i} dy_{i+m} \int_0^{y_{i+m}} dy_{i+m-1} \dots \int_0^{y_{i+2}} dy_{i+1}. \quad (14)$$

Now define

$$\epsilon = y_{i+m}, \quad \tilde{y}_j = \frac{y_j}{y_{i+m}} \quad \text{for } j = i, i + 1, \dots, i + m. \quad (15)$$

Then

$$\begin{aligned} & \int_0^{z_i} dy_{i+m} \int_0^{y_{i+m}} dy_{i+m-1} \dots \int_0^{y_{i+2}} dy_{i+1} = \\ & \int_0^{z_i} d\epsilon \epsilon^{m-1} \int_0^1 d\tilde{y}_{i+m-1} \int_0^{\tilde{y}_{i+m-1}} d\tilde{y}_{i+m-2} \dots \int_0^{\tilde{y}_{i+2}} d\tilde{y}_{i+1}. \end{aligned} \quad (16)$$

Now if  $j, k \in \{i, i + 1, \dots, i + m\}$ , then

$$(z_j - z_k)^c = (y_k - y_j)^c = \epsilon^c (\tilde{y}_k - \tilde{y}_j)^c.$$

Consequently, in changing from variables  $z_i, z_{i+1}, \dots, z_{i+m-1}, z_{i+m}$  to variables  $z_i, \tilde{y}_{i+1}, \dots, \tilde{y}_{i+m-1}, \epsilon$ , from the measure  $d\mu_n$  we pick up a factor of

$$\epsilon^{m-1}, \quad (17)$$

from  $H(z)$  we pick up a factor of

$$\prod_{i \leq j < l \leq i+m} \epsilon^{c_{jl}}, \quad (18)$$

and from  $\Lambda_n$  we pick up a factor of

$$\prod_{i \leq j < l \leq i+m} \epsilon^{\alpha' s_{jl}}. \quad (19)$$

We see then, that if we make the definitions

$$C = \sum_{i \leq j < l \leq i+m} c_{jl}, \quad P = \sum_{i \leq j < l \leq i+m} s_{jl}, \quad (20)$$

then on changing from variable  $z_i, z_{i+1}, \dots, z_{i+m-1}, z_{i+m}$  to variables  $z_i, \tilde{y}_{i+1}, \dots, \tilde{y}_{i+m-1}, \epsilon$ , the integral over  $\epsilon$  will be given by

$$\int_0^1 d\epsilon \epsilon^{m-1+C+\alpha'P} (1 + \mathcal{O}(\epsilon)) = B[m + C + \alpha'P, 1] + \int_0^1 d\epsilon \epsilon^{m-1+C+\alpha'P} \mathcal{O}(\epsilon), \quad (21)$$

where  $B[x, y]$  is the beta function. The  $\mathcal{O}(\epsilon)$  term comes from factors  $(z_l - z_j)^{c_{lj}}$  where  $l \notin \{i, i+1, \dots, i+m\}$  but  $j \in \{i, i+1, \dots, i+m\}$ . For in that case

$$\begin{aligned} (z_l - z_j)^{c_{lj}} &= (z_l - z_i + y_j)^{c_{lj}} = (z_l - z_i + \epsilon \tilde{y}_j)^{c_{lj}} = (z_l - z_i)^{c_{lj}} \left(1 + \frac{\epsilon \tilde{y}_j}{z_l - z_i}\right)^{c_{lj}} \\ &= (z_l - z_i)^{c_{lj}} (1 + \mathcal{O}(\epsilon)). \end{aligned} \quad (22)$$

The same argument applies to the factors in  $\Lambda_n(\alpha', p, z)$ .

The beta function  $B[x, 1]$  has a simple pole at  $x = 0$  while it remains finite for all other values of  $x$ . We see, then, that the condition for the integral to have a divergence in the  $\alpha' \rightarrow 0$  limit as variables  $z_i$  to  $z_{i+m}$  tend to the same value is that  $m + C = 0$ . We can formally state this divergence criteria thus:

*In order for a string theory integral  $I_n[H]$  to diverge in the  $\alpha' \rightarrow 0$  limit in the part of the integration domain where consecutive variables  $z_j$  with  $j \in \tau = \{i, i+1, \dots, i+m\}$  tend to the same value, the sum of the exponents to the differences  $(z_j - z_l)$  with  $j, l \in \tau$  that appear in  $H(z)$  must equal  $m$ , ie. one minus the number of elements in  $\tau$ .<sup>2</sup>*

When the divergence criteria is satisfied, the integral over  $\epsilon$  yields a propagator times  $(\alpha')^{-1}$  plus a term that is finite in the  $\alpha' \rightarrow 0$  limit:

$$\int_0^1 d\epsilon \epsilon^{-1+\alpha'P} (1 + \mathcal{O}(\epsilon)) = \frac{1}{\alpha'P} + \mathcal{O}((\alpha')^0). \quad (23)$$

---

<sup>2</sup>It is tacitly understood, here and in the following sections, that the subset set has at least two and less than  $n - 1$  members.



Retaining only the leading order term in  $\alpha'$ , we can replace factors of the form  $(z_l - z_j)^{c_{lj}}$  with  $(z_l - z_i)^{c_{lj}}$  when  $l \notin \{i, i+1, \dots, i+m\}$  while  $j \in \{i, i+1, \dots, i+m\}$ . And similarly we can replace  $(z_l - z_j)^{\alpha' s_{lj}}$  factors. The full integral therefore factors into two parts:

$$\int_0^1 dz_3 \int_0^{z_3} dz_4 \dots \int_0^{z_{i-1}} dz_i \int_0^{z_i} dz_{i+m+1} \int_0^{z_{i+m+1}} dz_{i+m+2} \dots \int_0^{z_{n-2}} dz_{n-1} \Lambda_n^1(\alpha', k, z) H_1(z) \\ \times \int_0^1 d\tilde{y}_{i+m-1} \int_0^{\tilde{y}_{i+m-1}} d\tilde{y}_{i+m-2} \dots \int_0^{\tilde{y}_{i+2}} d\tilde{y}_{i+1} \Lambda_n^2(\alpha', k, \tilde{y}) H_2(\tilde{y}), \quad (24)$$

where we define

$$H_1(z) = \lim_{z_{i+1}, z_{i+2}, \dots, z_{i+m} \rightarrow z_i} \frac{H(z)}{\prod_{l=i}^{i+m-1} \prod_{j=l+1}^{i+m} (z_l - z_j)^{c_{lj}}}, \quad (25)$$

$$\Lambda_n^1(\alpha', k, z) = \lim_{z_{i+1}, z_{i+2}, \dots, z_{i+m} \rightarrow z_i} \frac{\Lambda(\alpha', p, z)}{\prod_{l=i}^{i+m-1} \prod_{j=l+1}^{i+m} (z_l - z_j)^{\alpha' s_{lj}}}, \quad (26)$$

$$H_2(\tilde{y}) = \prod_{l=i}^{i+m-1} \prod_{j=l+1}^{i+m} (\tilde{y}_l - \tilde{y}_j)^{c_{lj}}, \text{ and} \quad (27)$$

$$\Lambda_n^2(\alpha', k, \tilde{y}) = \prod_{l=i}^{i+m-1} \prod_{j=l+1}^{i+m} (\tilde{y}_l - \tilde{y}_j)^{\alpha' s_{lj}}, \quad (28)$$

We note that while  $\tilde{y}_j$  is a variable for  $i < j < i+m$ ,  $\tilde{y}_i = 1$  and  $\tilde{y}_{i+m} = 0$ , as can be seen from the above definitions.

Each of the two parts that the integral factors into in equation (24) are of exactly the same functional form as the original integral. The same analysis as above will therefore yield the same divergence criteria for each of the two parts of the integral, which can in turn contain divergences and separate into smaller sub-parts and pick up additional factors of propagators and  $(\alpha')^{-1}$ . But if by  $\tau' = \{p, p+1, \dots, q-1, q\}$  we denote the indices to a set consecutive variables that gives rise to an additional divergence in the  $\alpha' \rightarrow 0$  limit as they tend to the same value, then it is not possible that

- $p < i \leq q \leq i+m$ , or
- $i \leq p \leq i+m < q$ .

On the contrary, it must hold true that either

1.  $p < q < i$ , or
2.  $i+m < p < q$ , or
3.  $p < i < i+m < q$ , or
4.  $i \leq p < q \leq i+m$ .

The reason is that each divergence splits the variables into two separate groups so that subsequent divergences can only come about due to variables in the same group tending to the same value.

In the first and second cases:  $\tau \cap \tau' = \emptyset$ . In the third case:  $\tau \subset \tau'$ . In the fourth case:  $\tau' \subset \tau$ . Hence,  $\tau$  and  $\tau'$  are either *nested* or *disjoint*.

Iterating the above analysis in evaluating  $I_n[H]$ , we pick up propagators and inverse  $\alpha'$  factors as variables indexed by  $\tau_1$  or  $\tau_2$  or ... or  $\tau_r$  tend to the same value, with any two  $\tau_l$  and  $\tau_k$  with  $l, k \in \{1, 2, \dots, r\}$ , being either nested or disjoint. The procedure can be carried out until one is left with an integral that does not diverge in the  $\alpha' \rightarrow 0$  limit but evaluates to some constant  $K$ . The resultant expression after iterating this procedure will then be

$$\frac{K}{(\alpha')^r} \prod_{i=1}^r \frac{1}{\left(\frac{1}{2} \sum_{j \in \tau_i} k_j\right)^2}, \quad (29)$$

Now, it is possible for  $I_n[H]$  to diverge in the  $\alpha' \rightarrow 0$  limit when variables  $z_i$  to  $z_j$  tend to the same value and to also diverge when variables  $z_k$  to  $z_l$  tend to the same value, where  $i < k < j < l$ . In that case the divergences occur in different parts of the integration domain of the full integral. Since the integral over the full domain is equal to the sum of the integrals over any partitioning of the domain, one must sum over divergences that occur in different parts of the domain. That is to say, propagators carrying external legs indexed by sets of numbers that are neither nested nor disjoint must be included in different terms of the final answer.

Our analysis has hitherto been carried out in a specific gauge. But since  $I_n[H]$  is invariant with respect to the choice of gauge, we may formulate the integration rules we have arrived at in a gauge-invariant form. In so doing we must consider a subset  $\tau \subset \mathbb{Z}_n$  of consecutive variables equivalent to its complement. For, by momentum conservation,

$$\frac{1}{\left(\sum_{i \in \tau} k_i\right)^2} = \frac{1}{\left(\sum_{i \in \tau^c} k_i\right)^2}. \quad (30)$$

Möbius invariance ensures that the divergence criteria can equivalently be applied to a subset  $\tau$  of variables and its complement  $\tau^c$ . For, let

$$C_\tau = \sum_{\substack{i, j \in \tau \\ i < j}} c_{i, j} \quad \text{and} \quad C_{\tau^c} = \sum_{\substack{i, j \in \tau^c \\ i < j}} c_{i, j} \quad \text{and} \quad c_{i, j} = c_{j, i}. \quad (31)$$

Then, using equation 11, we find that:

$$-2|\tau| = \sum_{i \in \tau} \sum_{j=1}^n c_{i, j} = 2C_\tau + \sum_{i \in \tau} \sum_{j \in \tau^c} c_{i, j}. \quad (32)$$

And similarly we have that

$$-2|\tau^c| = 2C_{\tau^c} + \sum_{i \in \tau^c} \sum_{j \in \tau} c_{i, j}. \quad (33)$$

From these two equations we see that

$$C_\tau + |\tau| = C_{\tau^c} + |\tau^c|, \quad (34)$$

from which the following bi-implication follows:

$$C_\tau = 1 - |\tau| \Leftrightarrow C_{\tau^c} = 1 - |\tau^c|. \quad (35)$$

We are now ready to state the string theory integration rules:

- Identify all subsets  $\tau \subset \mathbb{Z}$  that contain only consecutive numbers, where 1 and  $n$  are considered consecutive, and that have the property that  $C_\tau = 1 - |\tau|$ . Complementary subsets are to be considered equivalent:  $\tau \simeq \tau^c$ .

Any two such subsets  $\tau$  and  $\tau'$  shall be considered compatible if  $\tau \in \tau'$ , or  $\tau' \in \tau$ , or  $\tau^c \in \tau'$ , or  $\tau \in (\tau')^c$ .

- Out of all these subsets, form all possible collections of subsets  $T = \{\tau_1, \tau_2, \dots, \tau_r\}$  such that any two subsets in the same collection are compatible.

We shall say that a collection  $T$  is maximal if for any other collection  $T'$  we have that  $|T| \geq |T'|$ .

- To leading order in  $\alpha'$ ,  $I_n[H]$  will then be given by

$$\frac{1}{(\alpha')^r} \sum_{\text{maximal collections } T} K_T \prod_{\tau \in T} \frac{1}{\left(\sum_{i \in \tau} k_i\right)^2}, \quad (36)$$

where  $r$  is the number of elements in the maximal collections and  $K_T$  are constants that can be found by carrying out the residual, non-divergent, integrations.

### 2.3 Discussion and examples

As we have already seen, string theory integrals are often accompanied by a pre-factor of  $\alpha'$  raised to some power. On taking the infinite-tension limit, such pre-factors kill off all integrals that do not have a sufficient divergence in the  $\alpha' \rightarrow 0$  limit. When the pre-factor is  $(\alpha')^{n-3}$ , the only integrals that remain after taking the limit are those with maximal collections  $T$  that contain  $r = n - 3$  compatible subsets so that there are no residual integrations.

As to the cases where it is necessary to perform residual integrations, one would expect  $K_T$  to be some purely numerical factor. For after carrying out the diverging integrations, one is left with an integral that remains finite as  $\alpha'$  tends to zero, and so one should think that it would be permissible in this integral to immediately set  $\alpha'$  to zero so that the Koba-Nielsen factor  $\Lambda_n(\alpha', k, z)$  becomes equal to unity, removing the dependency on the  $k_i$  and leaving a purely numerical integral. Consider for example this integrand:

$$H(z) = \frac{1}{(z_1 - z_3)(z_1 - z_4)(z_2 - z_4)(z_2 - z_5)(z_3 - z_5)}. \quad (37)$$

After gauge-fixing we can write the integral thus:

$$I_5[H] = \int_0^1 dz_3 \int_0^{z_3} dz_4 \frac{(1 - z_3)^{\alpha' s_{23}} (1 - z_4)^{\alpha' s_{24}} (z_3 - z_4)^{\alpha' s_{34}} z_3^{\alpha' s_{35}} z_4^{\alpha' s_{45}}}{(1 - z_4) z_3}. \quad (38)$$

One would be tempted to take the  $\alpha' \rightarrow 0$  limit of  $I_5[H]$  by just setting  $\alpha'$  to zero:

$$\lim_{\alpha' \rightarrow 0} I_5[H] = \int_0^1 dz_3 \int_0^{z_3} dz_4 \frac{1}{(1 - z_4) z_3} = \frac{\pi^2}{6}. \quad (39)$$

The correct result is indeed produced. But this does not always happen. For consider this integrand:

$$H(z) = \frac{1}{(z_1 - z_2)^2 (z_3 - z_4)^2}. \quad (40)$$

On taking the  $\alpha' \rightarrow 0$  limit of the integral we now find that

$$\lim_{\alpha' \rightarrow 0} I_4[H] = \lim_{\alpha' \rightarrow 0} \int_0^1 dz_3 \frac{(1 - z_3)^{\alpha' s_{23}} z_3^{s_{34}}}{z_3^2} = \lim_{\alpha' \rightarrow 0} B[1 + \alpha' s_{23}, \alpha' s_{34} - 1] = -\frac{s_{23} + s_{34}}{s_{34}}. \quad (41)$$

We see that now  $K_T$  does depend on the momenta, and if  $s_{23}$  and  $s_{34}$  are both positive, the result is negative despite the fact that the integrand is positive in the whole integration domain. Conventional intuition fails in this case because  $I_n[H]$  is not a well-defined Riemann integral due to the quadratic divergence near the  $z_3 = 0$  region of the integration domain. Analytical extension is required. We may observe that in general the residual integration constants  $K_T$  will depend on the kinematic invariants  $s_{ij}$  when the integrals require analytical extension. Integrals of the form

$$\int_0^1 dx \frac{1}{x^\beta} \quad (42)$$

are finite when  $\beta < 1$ , diverge when  $\beta = 1$ , and are finite after analytical extension when  $\beta > 1$ . Hence, the divergence criterion that was formulated above provides the demarcation between the integrals that require analytical extension and those that do not. We can therefore formulate another rule:

*If and only if, for a given string theory integral  $I_n[H]$ , there is no subset  $\tau \subset \{1, 2, \dots, n\}$  of consecutive numbers such that  $C_\tau < 1 - |\tau|$ , then an analytical extension is not necessary in order to take the  $\alpha' \rightarrow 0$  limit of  $I_n[H]$ , and the residual integration constants  $K_T$  will all be purely numerical constants.<sup>3</sup>*

While the computation of the constants  $K_T$  can pose an exceedingly difficult challenge, the integration rules trivialize the computation of the  $\alpha' \rightarrow 0$  limit of string theory integrals that are maximally divergent.

For example, consider the integral  $I_6[H]$  with

$$H(z) = \frac{1}{(z_1 - z_2)(z_1 - z_5)(z_2 - z_4)(z_3 - z_4)(z_3 - z_6)(z_5 - z_6)}. \quad (43)$$

Selecting from the pairs of equivalent subsets the subset not containing  $z_1$ , we have that the subsets of variables that will yield a divergence are the following:

$$\begin{aligned} \{3, 4\} &: \text{two variables, one factor connecting them in the numerator,} \\ \{5, 6\} &: \text{two variables, one factor connecting them in the numerator,} \\ \{2, 3, 4\} &: \text{three variables, two factors connecting them in the numerator,} \\ \{3, 4, 5, 6\} &: \text{four variables, three factors connecting them in the numerator.} \end{aligned} \quad (44)$$

These subsets are all compatible with each other except that  $\{2, 3, 4\}$  and  $\{3, 4, 5, 6\}$  are incompatible. We can therefore form two maximal sets of compatible subsets:

- 1)  $\{3, 4\}, \{5, 6\}, \{2, 3, 4\}$ ,
- 2)  $\{3, 4\}, \{5, 6\}, \{3, 4, 5, 6\}$ .

---

<sup>3</sup>Strictly speaking, analytical extension may also be needed when  $C_\tau = 1 - |\tau|$  depending on the sign of the Mandelstam variables, since, as a Riemann integral,  $\int_0^1 x^{\epsilon-1}$  is only well-defined for  $\epsilon > 0$ . But this has no bearing on the evaluation of  $K_T$ .

Since there are no residual integrations, all  $K_T$  are equal to one, and we find that, to leading order in  $\alpha'$ , the integral is given as follows:

$$I_6[H] = \frac{1}{s_{34} s_{56}} \left( \frac{1}{s_{234}} + \frac{1}{s_{3456}} \right) \frac{1}{(\alpha')^3}. \quad (45)$$

### 3 CHY integration rules

#### 3.1 Motivation and set-up

We now set aside string theory for a moment and instead turn our attention to the CHY formalism for computing the tree-amplitude of  $n$  massless particles. The scattering equations, which constitute the cornerstone of the formalism, are a simple set of rational equations:

$$0 = S_i(z) = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{s_{ij}}{z_i - z_j}, \quad i = 1, \dots, n. \quad (46)$$

Because of the identities

$$0 = \sum_{i=1}^n S_i, \quad 0 = \sum_{i=1}^n S_i z_i, \quad 0 = \sum_{i=1}^n S_i z_i^2, \quad (47)$$

which hold true for any values of the variables  $z_i$ , only  $n - 3$  of the scattering equations are independent. Consequently, (46) constitutes an under-determined system of equations unless three of the variables  $z_i$  are assigned fixed values. When this is done, the equations have  $(n - 3)!$  solutions<sup>[19]</sup>.

In terms of the scattering equations, CHY have formulated the following integration measure:

$$d\Omega_n = \frac{(z_r - z_s)(z_s - z_t)(z_t - z_r)}{dz_r dz_s dz_t} (z_{r'} - z_{s'})(z_{s'} - z_{t'})(z_{t'} - z_{r'}) \prod_{i=1}^n dz_i \prod_{\substack{i=1 \\ i \neq r', s', t'}}^n \delta(S_i). \quad (48)$$

By employing this measure, CHY have been able present various tree-amplitudes in a remarkably compact form. Their formulas for cubic scalars<sup>[5]</sup>, gluons, and gravitons<sup>[4]</sup>, are respectively given by

$$\mathcal{A}_n^{\phi^3} = (-1)^n \int d\Omega_n \frac{1}{\prod_{i=1}^n (z_i - z_{i+1})^2}, \quad (49)$$

$$\mathcal{A}_n^{\text{gluons}} = \int d\Omega_n \frac{\text{Pf}'\Psi}{\prod_{i=1}^n (z_i - z_{i+1})}, \quad (50)$$

$$\mathcal{A}_n^{\text{gravitons}} = \int d\Omega_n (\text{Pf}'\Psi)^2, \quad (51)$$

where  $\Psi$  is the following  $2n \times 2n$  anti-symmetric matrix:

$$\Psi = \begin{pmatrix} A & -C^T \\ C & B \end{pmatrix}, \quad (52)$$

where the sub-matrices  $A$ ,  $B$ , and  $C$  are defined thus:

$$A_{ij} = \begin{cases} \frac{k_i \cdot k_j}{z_i - z_j}, & \text{if } i \neq j, \\ 0, & \text{if } i = j, \end{cases} \quad B_{ij} = \begin{cases} \frac{\epsilon_i \cdot \epsilon_j}{z_i - z_j}, & \text{if } i \neq j, \\ 0, & \text{if } i = j, \end{cases} \quad C_{ij} = \begin{cases} \frac{\epsilon_i \cdot k_j}{z_i - z_j}, & \text{if } i \neq j, \\ -\sum_{l \neq i} \frac{\epsilon_i \cdot k_l}{z_i - z_l}, & \text{if } i = j, \end{cases} \quad (53)$$

and by  $\text{Pf}'\Psi$  we denote the reduced Pfaffian of  $\Psi$ , which is defined as

$$\text{Pf}'\Psi \equiv \frac{(-1)^{i+j}}{z_i - z_j} \text{Pf}\Psi_{i,j}. \quad (54)$$

where  $\Psi_{i,j}$  denotes the sub-matrix of  $\Psi$  obtained by removing rows and columns  $i$  and  $j$  with  $i, j \leq n$ . When evaluated on the scattering equations,  $\text{Pf}'\Psi$  is independent of the choice of  $i$  and  $j$ . It should also be noted that equations (49) to (51) hold true irrespective of the choice of  $z_r, z_s, z_t, z_{r'}, z_{s'},$  and  $z_{t'}$  in the measure (48).

A property common to all the CHY formulas is that the integrals can be expanded into terms of the following form:

$$\mathcal{I}_n[\mathcal{G}] = \int d\Omega_n \mathcal{G}(z), \quad (55)$$

where  $\mathcal{G}(z)$  is a product of differences lifted to integer powers:

$$\mathcal{G}(z) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n (z_i - z_j)^{c_{ij}}, \quad \text{with } c_{ij} \in \mathbb{Z}. \quad (56)$$

Note that we have adopted the convention that the differences  $(z_i - z_j)$  in  $\mathcal{G}(z)$  are expressed such that the index of the minuend is always less than that of the subtrahend. Another shared feature of the CHY formulas is Möbius invariance. Because of the integration measure, this entails that the variables in (??) have weight  $-4$ , instead of  $-2$  as in string theory, by which we mean that

$$\sum_{\substack{j=1 \\ j \neq i}}^n c_{ij} = -4, \quad \text{where } c_{ij} = c_{ji}. \quad (57)$$

When working with specific integrands  $\mathcal{G}(z)$ , it is convenient to represent  $\mathcal{I}_n[\mathcal{G}]$  diagrammatically by drawing  $n$  vertices numbered from 1 to  $n$  and then drawing an edge connecting vertices  $i$  and  $j$  for each factor of  $(z_i - z_j)^{-1}$  in  $\mathcal{G}(z)$ , counted with multiplicity. I will refer to such diagrams as CHY diagrams. When  $\mathcal{G}(z)$  has no factors of  $(z_i - z_j)$  in the numerator, Möbius invariances dictates that there are four edges incident to each vertex. In that case the CHY diagram is a so-called 4-regular graph.

For example, the diagram



represents the integral  $\mathcal{I}_n[\mathcal{G}]$  with

$$\mathcal{G}(z) = \frac{1}{(z_1 - z_2)^2(z_1 - z_4)(z_1 - z_5)(z_2 - z_3)(z_2 - z_4)(z_3 - z_4)(z_3 - z_5)^2(z_4 - z_5)}. \quad (59)$$

### 3.2 Deriving the rules

Integrals of the form (55) satisfying (57) we will call CHY integrals. If one can compute such integrals, then, by dressing them up with suitable prefactors and summing up, one can calculate gluon and graviton amplitudes via equations (50) and (51). We will therefore now set about finding a general set of rules for evaluating such integrals without having to resort to the cumbersome undertaking of actually solving the scattering equations. Our approach to this task is greatly inspired by [8].

It will be convenient to work in a specific gauge, and we will make the following choice:

$$z_r = z_{r'} = z_1 = \infty, \quad z_s = z_{s'} = z_2 = 1, \quad z_t = z_{t'} = z_n = 0. \quad (60)$$

And so if we define

$$G(z) = \lim_{z_1 \rightarrow \infty} \frac{\mathcal{G}(z)}{z_1^4}, \quad (61)$$

then we can express  $\mathcal{I}_n[\mathcal{G}]$  as follows:

$$\mathcal{I}_n[\mathcal{G}] = \int \prod_{i=3}^{n-1} dz_i G(z) \prod_{i=3}^{n-1} \delta(S_i). \quad (62)$$

Now, instead of working with delta functions, it is possible to interpret  $\mathcal{I}_n[\mathcal{G}]$  as the residue of a complex contour integral enclosing the solutions to the scattering equations:

$$\frac{1}{(2\pi i)^{n-3}} \mathcal{I}_n[\mathcal{G}] = \frac{1}{(2\pi i)^{n-3}} \oint_{S_3=S_4=\dots=S_{n-1}=0} \prod_{i=3}^{n-1} dz_i \frac{G(z)}{\prod_{i=3}^{n-1} S_i}. \quad (63)$$

The advantage of this formulation is that it enables one to make use of the powerful tools proffered by complex analysis in calculating  $\mathcal{I}_n[\mathcal{G}]$ . Specifically, we can make use of the global residue theorem, which tells us that the residues at the solutions to the scattering equations are equal to minus the sum of all other residues.

The question that then faces us is for which functions  $G(z)$  that the integrand has poles elsewhere than at the solutions to the scattering equations. Such other poles can only come about because of factors of  $(z_i - z_j)$  in the denominator of  $G(z)$ , which cause the integrand to diverge as variables tend to the same values.

We first consider the case when variables  $z_i$  tend to  $z_n = 0$  for  $i \in \tau$ , where  $\tau$  is a subset of  $\mathbb{Z}_n$  containing  $n$  but not 1 and 2. We also assume that  $\tau$  has at least one other member than  $n$ , and we denote this member  $a$ .

We make the following definitions:

$$z_a = \epsilon, \quad z_i = \epsilon x_i \quad \text{for} \quad i \in \tau.$$

Note that  $x_a = 1$  and  $x_n = 0$ .

In terms of the new variables, differences between the old variables can be rewritten as follows:

$$(z_i - z_j) = z_i (1 + \mathcal{O}(\epsilon)) \quad \text{for } i \notin \tau \text{ and } j \in \tau, \quad (64)$$

$$(z_i - z_j) = -z_j (1 + \mathcal{O}(\epsilon)) \quad \text{for } i \in \tau \text{ and } j \notin \tau, \quad (65)$$

$$(z_i - z_j) = \epsilon(x_i - x_j) \quad \text{for } i, j \in \tau. \quad (66)$$

Hence, it is seen that  $G(z)$  factors into a function of variables  $z_i$  with  $i \notin \tau$  and a function of variables  $x_i$  with  $i \in \tau$  :

$$G(z) = \epsilon^g \hat{G}(z) \tilde{G}(x) (1 + \mathcal{O}(\epsilon)) (-1)^{n_{\text{inv}}}, \quad (67)$$

where we have made the following definitions:

$$g = \sum_{\substack{i, j \in \tau \\ i < j}} c_{ij}, \quad (68)$$

$$\hat{G}(z) = \lim_{z_i \rightarrow 0 \forall i} \frac{G(z)}{\prod_{\substack{i, j \in \tau \\ i < j}} (z_i - z_j)^{c_{ij}}} \quad (69)$$

$$\tilde{G}(x) = \prod_{\substack{i, j \in \tau \\ i < j}} (x_i - x_j)^{c_{ij}} \quad (70)$$

And by  $n_{\text{inv}}$  we denote the number of factors  $(z_i - z_j)$  with  $i \in \tau$  and  $j \notin \tau$ , the which factors we will denote "inversion factors".

Having introduced the new variables, we can rewrite the integration measure thus:

$$\prod_{i=3}^{n-1} dz_i = d\epsilon \epsilon^{|\tau|-2} \prod_{i \in \tau^c \setminus \{1,2\}} dz_i \prod_{i \in \tau \setminus \{a,n\}} dx_i, \quad (71)$$

We can expand the  $S_i$  to leading order in  $\epsilon$  as follows:

$$S_i = \frac{1}{\epsilon} (1 + \mathcal{O}(\epsilon)) \tilde{S}_i, \quad \tilde{S}_i = \sum_{\substack{j \in \tau \\ j \neq i}} \frac{S_{ij}}{x_i - x_j}, \quad \text{for } i \in \tau, \quad (72)$$

$$S_i = \hat{S}_i + \mathcal{O}(\epsilon), \quad \hat{S}_i = \sum_{\substack{j \notin \tau \\ j \neq i}} \frac{S_{ij}}{z_i - z_j} + \sum_{j \in \tau} \frac{S_{ij}}{z_i}, \quad \text{for } i \notin \tau. \quad (73)$$

The product of the  $S_i$  can therefore be written in the following manner:

$$\prod_{i=3}^{n-1} S_i = \epsilon^{1-|\tau|} (1 + \mathcal{O}(\epsilon)) \prod_{i \in \tau^c \setminus \{1,2\}} \hat{S}_i \prod_{i \in \tau \setminus \{a,n\}} \tilde{S}_i. \quad (74)$$

Putting these results together, we have that  $\mathcal{I}_n[\mathcal{G}]$  is given by

$$\frac{(-1)^{n_{\text{inv}}}}{(2\pi i)^{n-3}} \oint_{S_3=\dots=S_{n-1}=0} d\epsilon \prod_{i \in \tau^c \setminus \{1,2\}} dz_i \prod_{i \in \tau \setminus \{a,n\}} dx_i \frac{\hat{G}(z) \tilde{G}(x)}{\prod_{i \in \tau^c \setminus \{1,2\}} \hat{S}_i \prod_{i \in \tau \setminus \{a,n\}} \tilde{S}_i} \epsilon^{g+2|\tau|-1} (1 + \mathcal{O}(\epsilon)). \quad (75)$$

From this expression we see that at  $\epsilon = 0$ ,



- there is no pole if  $g > 2 - 2|\tau|$ ,
- there is a simple pole if  $g = 2 - 2|\tau|$ ,
- there is a double pole if  $g = 1 - 2|\tau|$ , and
- there is a triple pole if  $g = -2|\tau|$ .

Poles of higher order than three are possible, but because of Möbius invariance they can only appear when  $\mathcal{G}$  has factors of  $(z_i - z_j)$  in the numerator.

We do not presume to be able to state general integration rules for CHY integrals with higher-order poles, and for this reason we will in the following restrict attention to the CHY integrals for which there are no higher order poles, that is, the CHY integrals for which, irrespective of gauge choice or the ordering of the external legs, it never happens that  $g < 2 - 2|\tau|$ . This condition can equivalently be formulated as follows:

$\mathcal{I}_n[\mathcal{G}]$  has no higher order poles if and only if there is no subset  $\rho$  of  $\mathbb{Z}_n$  such that  $\sum_{\substack{i,j \in \rho \\ i < j}} c_{ij} < 2 - 2|\rho|$ .

In the present case, then, we assume that  $g = 2 - 2|\tau|$ . Now, we observe that

$$\begin{aligned} \sum_{i \in \tau} \tilde{S}_i x_i &= \sum_{\substack{i,j \in \tau \\ j \neq i}} \frac{s_{ij}}{x_i - x_j} x_i = \sum_{\substack{i,j \in \tau \\ j \neq i}} \frac{s_{ij}}{x_i - x_j} \left( \frac{x_i + x_i}{2} \right) = \sum_{\substack{i,j \in \tau \\ j \neq i}} \frac{s_{ij}}{x_i - x_j} \left( \frac{x_i - x_j}{2} \right) \\ &= \sum_{\substack{i,j \in \tau \\ j \neq i}} \frac{s_{ij}}{2} = \sum_{\substack{i,j \in \tau \\ i < j}} s_{ij} = s_\tau. \end{aligned} \quad (76)$$

On applying the global residue theorem, we rewrite  $\mathcal{I}_n[\mathcal{G}]$  as minus the sum of other residues. One of these residues will be the residue at  $\epsilon = 0$ ,  $\tilde{S}_i = 0$  for all  $i \in \tau \setminus \{a, n\}$ . For this residue, we have that

$$s_\tau = \sum_{i \in \tau} \tilde{S}_i x_i = \tilde{S}_a x_a = \tilde{S}_a \quad (77)$$

since  $x_n = 0$ . Therefore this residue times minus one is equal to

$$\begin{aligned} &-\frac{(-1)^{n_{\text{inv}}}}{s_\tau} \frac{1}{(2\pi i)^{n-|\tau|-4}} \oint_{\hat{S}_i=0 \vee i \in \tau^c \setminus \{1,2\}} \prod_{i \in \tau^c \setminus \{1,2\}} dz_i \frac{\hat{G}(z)}{\prod_{i \in \tau^c \setminus \{1,2\}} \hat{S}_i} \times \\ &\frac{1}{(2\pi i)^{|\tau|}} \oint_{\tilde{S}_i=0 \vee i \in \tau \setminus \{a,n\}} \prod_{i \in \tau \setminus \{a,n\}} dx_i \frac{\tilde{G}(x)}{\prod_{i \in \tau \setminus \{a,n\}} \tilde{S}_i}. \end{aligned} \quad (78)$$

From here, one could apply the counting rule to  $\hat{G}(z)$  and  $\tilde{G}(x)$  to look for further poles and iterate this procedure until the integrations had been completely carried out. Because the rescaled variables separate completely from the non-rescaled ones, it is clear that the sets of external legs carried by the propagators will be either nested or disjoint. It is also clear that in order to obtain a non-zero contribution to  $\mathcal{I}_n[\mathcal{G}]$ , we have to be able to iterate this procedure  $n - 3$  times so as to satiate all the integrations. Each time we

apply this procedure and invoke the global residue theorem we pick up a factor of minus one so that in total we pick up a factor of  $(-1)^{n-3}$  in addition to a factor of  $(-1)^{\mathcal{N}_{\text{inv}}}$ , where  $\mathcal{N}_{\text{inv}}$  denotes the total number of inversion factors.

While the factorization (78) was carried out in a specific gauge, gauge invariance ensures us that it holds true generally. That is to say, whenever we have a subset  $\tau \subset \mathbb{Z}_n$  such that

$$\sum_{\substack{i,j \in \tau \\ i < j}} c_{ij} = 2 - 2|\tau|, \quad (79)$$

then  $\mathcal{I}[\mathcal{G}]$  has a residue given by  $s_\tau^{-1}$  times (an integral over variables  $z_i$  with  $i \in \tau$ ) times (an integral over variables  $z_i$  with  $i \notin \tau$ ). Whenever  $\tau$  does not consist of consecutive variables, there will be variables  $z_a, z_b$ , and  $z_c$  with canonical ordering  $z_a \rightarrow z_b \rightarrow z_c$  such that  $z_a, z_c \in \tau$  while  $z_b \notin \tau$ , and in that case one must be mindful of inversion factors.

As in string theory, momentum conservation dictates that complementary subsets must be considered equivalent. The derivation in equations (30) to (35) carries directly over to the CHY formalism to show that one can equally well consider subsets  $\tau$  and  $\tau^c$  in checking for poles. When working in a specific gauge, one encounters only one of the two subsets, namely the one that does not contain the leg that is fixed to infinity.

Here one could ask the question, what if a subset contains the variable fixed at infinity, while its complement contains the legs fixed to 0 and 1. Say for instance that  $\mathcal{G}(z)$  contains a factor of  $(z_2 - z_n)^{-2}$ . Then the counting rule tells us that the subset  $\tau = \{2, n\}$  has a pole where variables  $z_2$  and  $z_n$  tend to the same value. But when working in the gauge (60), variables  $z_2$  and  $z_n$  cannot tend to the same value to produce a propagator  $(s_{2n})^{-1}$ , and neither can we obtain a propagator  $(s_{z_n \setminus \{2, n\}})^{-1}$  since  $s_{z_n \setminus \{2, n\}}$  contains the variable fixed to infinity. The solution to this conundrum is that in the gauge (60) the propagator  $(s_{2n})^{-1}$  arises due to a pole at infinity. Thus we see that the choice of gauge can shift poles to and from infinity since one can equally well chose not to fix the two variables  $z_2$  and  $z_n$ , so that the propagator  $(s_{2n})^{-1}$  will come about from a regular pole. Thereby gauge invariance spares us the trouble of dealing with poles at infinity.

It may be instructive to see an example of how a full residue can be evaluated by iteratively applying the global residue theorem. This can be done graphically with the CHY diagrams introduced in the previous subsection:

$$\begin{aligned}
 & \begin{array}{c} \text{Diagram 1: A square with vertices 1 (top), 2 (left), 3 (bottom), 4 (right). Diagonals 1-3 and 2-4 are present. External legs 5 and 6 are attached to vertices 3 and 4 respectively.} \end{array} \rightarrow \frac{-1}{s_{45}} \begin{array}{c} \text{Diagram 2: A square with vertices 1 (top), 2 (left), 3 (bottom), 4 (right). Diagonals 1-3 and 2-4 are present. External legs 5 and 6 are attached to vertices 3 and 4 respectively.} \end{array} \rightarrow \frac{1}{s_{45} s_{36}} \begin{array}{c} \text{Diagram 3: A square with vertices 1 (top), 2 (left), 3 (bottom), 4 (right). Diagonals 1-3 and 2-4 are present. External legs 5 and 6 are attached to vertices 3 and 4 respectively.} \end{array} \\
 & \rightarrow \frac{-1}{s_{45} s_{36} s_{3456}} \begin{array}{c} \text{Diagram 4: A triangle with vertices 1 (top), 2 (left), 3 (right). External legs 4, 5, and 6 are attached to vertices 1, 2, and 3 respectively.} \end{array} = \frac{-1}{s_{45} s_{36} s_{3456}} \quad (80)
 \end{aligned}$$

At the first arrow we pick up a minus when applying the global residue theorem. At the second arrow we pick up three factors of minus one: one from each factor of  $(z_3 - z_{4,5})$

and one from applying the global residue theorem. At the third arrow we pick up a minus from the global residue theorem.

To obtain the full result for the diagram on the left-hand side, one must sum over all distinct ways of reducing the diagram to a double-lined triangle.

To wrap up, the rules for evaluating  $\mathcal{I}_n[\mathcal{G}]$  can be stated as follows:

- If there is any subset  $\tau \subset \mathbb{Z}_n$  such that

$$\sum_{\substack{i,j \in \tau \\ i < j}} c_{ij} < 2 - 2|\tau|, \quad (81)$$

then the integrand of  $\mathcal{I}_n[\mathcal{G}]$  has a higher order pole and cannot be written as a product of propagators.

- If there are no higher order poles, then  $\mathcal{I}_n[\mathcal{G}]$  can be evaluated through the following steps:
  - Determine all poles by finding all subsets  $\tau \subset \mathbb{Z}_n$  such that

$$\sum_{\substack{i,j \in \tau \\ i < j}} c_{ij} = 2 - 2|\tau|, \quad (82)$$

and assign to each subset a propagator  $(s_\tau)^{-1}$ . Complementary subsets are considered equivalent.

- There is a residue, which is equal to the product of the propagators of the subsets, for each collection  $T$  of  $n - 3$  subsets that are pairwise compatible in the sense that for any  $\tau, \tau' \in T$  either  $\tau \in \tau'$ , or  $\tau' \in \tau$ , or  $\tau^{\complement} \in \tau'$ , or  $\tau \in (\tau')^{\complement}$ . Add together all the residues to obtain  $\mathcal{I}_n[\mathcal{G}]$  up to an overall sign of  $(-1)^{n+\mathcal{N}_{\text{inv}}+1}$ .

It should be remarked that in stating these integration rules, we have characterized which CHY integrals that evaluate to one or a sum of  $\phi^3$  Feynman diagrams. It is exactly these integrals that can be calculated with the integration rules.

### 3.3 Discussion and examples

To familiarize ourselves with the integration rules, let us apply them to a few examples. But first, it would be convenient to have a simple way of ascertaining whether a given integral can be evaluated with the rules or if it has higher-order poles. For integrals with trivial numerators, the diagrammatic representation as 4-regular graphs has the benefit of allowing for the immediate determination of this issue. For integrals with third-order poles are those whose CHY graph has a subset  $\tau$  of vertices connected to each other with  $2|\tau|$  edges. But, by Möbius invariance, this implies that there can be no more edges incident on the vertices in  $\tau$ , meaning that the vertices in  $\tau$  are not connected to those in  $\tau^{\complement}$ . So  $\mathcal{I}_n[\mathcal{G}]$  has a triple pole if and only if the CHY graph is disconnected (e.g. a graph that has two vertices connected by four edges). Integrals with second-order poles have a subset  $\rho$  of vertices with  $2|\rho| - 2$  edges connecting them with each other. So the vertices in  $\rho$  are connected with those in  $\rho^{\complement}$  by two edges. In other words,  $\mathcal{I}_n[\mathcal{G}]$  has a double pole

if and only if the CHY graph can be disconnected by removing two edges (e.g. a graph with a triple line).

Let us consider, then, some integrals with CHY diagrams that cannot be separated into two graphs by removing two edges, and which can therefore be directly evaluated with the integration rules.

As a first example, we consider the integral  $\mathcal{I}_5[\mathcal{G}]$  with the following CHY diagram:



Up to complementarity, this graph has but two subsets of vertices with  $(2 \text{ times } \# \text{ of vertices} - 2)$  edges connecting them:

$$\begin{aligned} & \{1, 2\}, \text{ or, equivalently, } \{3, 4, 5\} \text{ and} \\ & \{3, 5\}, \text{ or, equivalently, } \{1, 2, 4\}. \end{aligned} \quad (84)$$

These two subsets are compatible. If we adopt our standard gauge-choice,  $(z_3 - z_4)$  is the only inversion factor. So  $\mathcal{N}_{\text{inv}} = 1$ , and the overall sign is minus. We conclude, then, that

$$\mathcal{I}_5[\mathcal{G}] = -\frac{1}{s_{12}s_{35}}. \quad (85)$$

As a second example, consider the integral  $\mathcal{I}_6[\mathcal{G}]$  with the following CHY diagram:



To evaluate the integral we first enumerate the subsets of vertices with net  $(2 \text{ times } \# \text{ of points} - 2)$  edges connecting them:

$$\begin{aligned} & \left. \begin{array}{l} \{1, 2\} \\ \{3, 6\} \\ \{4, 5\} \end{array} \right\} \text{two vertices, two edges} \\ & \left. \begin{array}{l} \{1, 2, 6\} \end{array} \right\} \text{three vertices, four edges} \end{aligned} \quad (87)$$

These subsets are all compatible except that  $\{3, 6\}$  is incompatible with  $\{1, 2, 6\}$ . Up to the sign, then,  $\mathcal{I}_6[\mathcal{G}]$  is given by

$$\frac{1}{s_{12}s_{45}} \left( \frac{1}{s_{36}} + \frac{1}{s_{126}} \right). \quad (88)$$

Neither subsets  $\{1, 2\}$ ,  $\{4, 5\}$ , nor  $\{1, 2, 6\}$ , being consecutive, give rise to inversion factors.  $\{3, 6\}$ , on the other hand, gives rise to two:  $(z_3 - z_4)$  and  $(z_3 - z_5)$ . In either case  $\mathcal{N}_{\text{inv}}$  is even, and so the overall sign is minus.

In both the two examples we have just considered,  $\mathcal{G}(z)$  had no  $(z_i - z_j)$  factors in the numerator. But the integration rules also apply to integrals with non-trivial numerators. Such integrals can be represented diagrammatically by adding to the type of diagram described above a dotted line connecting vertices  $i$  and  $j$  for every factor of  $(z_i - z_j)$  in the numerator. In that case Möbius invariance dictates that the number of normal lines minus the sum of dotted lines incident on each vertex must equal four.

Now let us consider the integral  $\mathcal{I}_6[\mathcal{G}]$  with

$$\mathcal{G}(z) = \frac{(z_2 - z_6)}{(z_1 - z_2)^2(z_1 - z_6)^2(z_2 - z_3)^2(z_2 - z_4)(z_3 - z_4)(z_3 - z_5)(z_4 - z_5)(z_4 - z_6)(z_5 - z_6)^2}, \quad (89)$$

which can be represented by the following diagram:



The subsets of vertices that can contribute a propagator are the following:

$$\left. \begin{array}{l} \{1, 2\} \\ \{1, 6\} \\ \{2, 3\} \\ \{5, 6\} \end{array} \right\} \text{two vertices, two lines} \quad (91)$$

$$\left. \begin{array}{l} \{1, 2, 3\} \\ \{2, 3, 4\} \end{array} \right\} \text{three vertices, four lines}$$

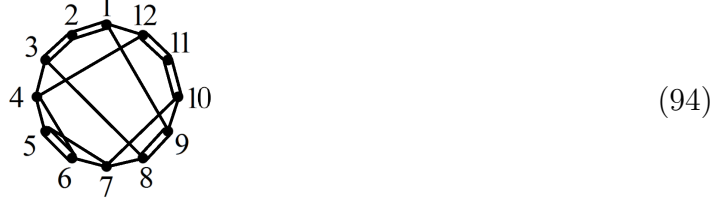
The vertices  $\{1, 2, 6\}$  are connected by four normal lines, but the dotted line counts minus one, so this subset does not make it to the list. Of the subsets that are on the list, we can form the following four maximal groups of compatible subsets:

$$\begin{aligned} & \{1, 2\}, \{5, 6\}, \{1, 2, 3\} \\ & \{1, 6\}, \{2, 3\}, \{2, 3, 4\} \\ & \{2, 3\}, \{5, 6\}, \{1, 2, 3\} \\ & \{2, 3\}, \{5, 6\}, \{2, 3, 4\} \end{aligned} \quad (92)$$

Since all the subsets consist of consecutive numbers, there are no inversion factors:  $\mathcal{N}_{\text{inv}} = 0$ . Hence, the overall sign is minus, and we conclude that the integral is given by:

$$-\left(\frac{1}{s_{12}} + \frac{1}{s_{23}}\right) \frac{1}{s_{56}s_{123}} - \left(\frac{1}{s_{16}} + \frac{1}{s_{56}}\right) \frac{1}{s_{23}s_{234}}. \quad (93)$$

To see that the difficulty in applying the integration rules actually scales quite slowly with the number of external legs, let us, for a final example, consider the 12-point integral with the following diagram:



As usual, we list the subsets of vertices with  $(2 \text{ times } \# \text{ of vertices} - 2)$  edges connecting them:

$$\begin{array}{l}
 \left. \begin{array}{l}
 \{1, 2\} \\
 \{2, 3\} \\
 \{5, 6\} \\
 \{8, 9\} \\
 \{10, 11\} \\
 \{11, 12\}
 \end{array} \right\} \text{two vertices, two edges} \\
 \\
 \left. \begin{array}{l}
 \{1, 2, 3\} \\
 \{4, 5, 6\} \\
 \{5, 6, 7\} \\
 \{10, 11, 12\}
 \end{array} \right\} \text{three vertices, four edges} \\
 \{4, 5, 6, 7\} \quad \left. \right\} \text{four vertices, six edges} \\
 \{1, 2, 3, 8, 9\} \quad \left. \right\} \text{five vertices, eight edges}
 \end{array} \tag{95}$$

All these sets are compatible with each other in the sense that any two sets are either nested or disjoint – except for the following three overlapping sets:

$$\begin{array}{l}
 \{1, 2\} \text{ overlaps with } \{2, 3\} \\
 \{10, 11\} \text{ overlaps with } \{11, 12\}, \text{ and} \\
 \{4, 5, 6\} \text{ overlaps with } \{5, 6, 7\}.
 \end{array}$$

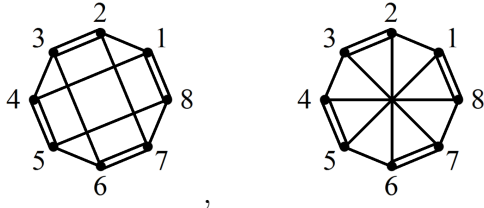
This leaves  $2^3$  different ways of combining nine compatible subsets. Summing over the corresponding products of propagators we get the result:

$$\left( \frac{1}{s_{1,2}} + \frac{1}{s_{2,3}} \right) \frac{1}{s_{5,6}} \frac{1}{s_{8,9}} \left( \frac{1}{s_{10,11}} + \frac{1}{s_{11,12}} \right) \frac{1}{s_{1,2,3}} \left( \frac{1}{s_{4,5,6}} + \frac{1}{s_{5,6,7}} \right) \frac{1}{s_{10,11,12}} \frac{1}{s_{4,5,6,7}} \frac{1}{s_{1,2,3,8,9}}. \tag{96}$$

Only one of the subsets does not consist of consecutive numbers – namely  $\{1, 2, 3, 8, 9\}$ , or, equivalently,  $\{4, 5, 6, 7, 10, 11, 12\}$  – so only this set can give rise to inversion factors. We find that  $(z_7 - z_8)$  is the only inversion factor, so  $\mathcal{N}_{\text{inv}} = 1$  and the overall sign is plus.

The integrals without higher-order poles, which can be directly computed from the integration rules, can always be interpreted as a sum of  $\phi^3$  Feynman diagrams, as is also

seen to be the case for the above examples. But the mapping to Feynman diagrams is not injective. For instance, both of the CHY integrals with the diagrams


(97)

evaluate to the following expression:

$$\frac{1}{s_{18} s_{23} s_{45} s_{67} s_{1458}}. \quad (98)$$

## 4 Reduction of higher-order poles

We have now seen how to compute CHY integrals that only have simple poles. But in order to calculate gluon and gravity amplitudes in the CHY formalism one must also be able to deal with higher-order poles. This can be seen from equations (50) and (51). For example, in expanding out  $\text{Pf}'\Psi$  in the gluon formula, there will be terms that pick up one factor of  $(z_1 - z_2)^{-1}$  from the  $A$  sub-matrix and another from the  $B$  sub-matrix, which together with the factor from  $\prod_{i=1}^n (z_i - z_{i+1})^{-1}$  totals a factor of  $(z_1 - z_2)^{-3}$ . In other words, there will be terms whose CHY diagrams have a triple line. And as shown above, such integrals have a second-order pole. However, in expanding out the gluon formula, there will not be terms with triple poles since the factor of  $\prod_{i=1}^n (z_i - z_{i+1})^{-1}$  in the integrand implies that the CHY graph is connected. But in the case of gravity, there will be terms with double and triple poles.

The approach we will adopt here to tackle such higher-order poles is to make use of the fact that every CHY integral  $\mathcal{I}[\mathcal{G}]$  is equal to the sum of a fixed expression evaluated at the various solutions to the scattering equations. This entails that any identity valid on solutions to the scattering equations can be used to rewrite  $\mathcal{G}$ . One could, for instance, use the scattering equations themselves and replace any factor of  $(z_1 - z_2)^{-1}$  in the integrand with

$$-\frac{1}{s_{12}} \sum_{i=3}^n \frac{s_{1i}}{z_1 - z_i}. \quad (99)$$

But with this particular substitution one would end up with  $n - 2$  terms that would not be Möbius invariant individually, though their sum would. We would therefore not have a procedure for computing the individual terms

What is needed are identities relating different Möbius invariant integrals. But the facts that the Pfaffian of the matrix  $\Psi$ , defined in equation (52), is identically zero and that, on solutions to the scattering equations, the reduced Pfaffian  $\text{Pf}'\Psi$  is invariant with respect to which two rows and columns  $i, j \leq n$  that are removed from  $\Psi$  provide us with just such identities. Indeed  $\text{Pf}'\Psi$  can be thought of as a generating function for a host of CHY integral identities since there are  $\frac{1}{2}n(n - 1)$  different choices of  $i$  and  $j$ , and any physically meaningful values can be assigned to the dot products  $k_a \cdot \epsilon_b$  and  $\epsilon_a \cdot \epsilon_b$ .

Depending on the integral to be evaluated, some identities might be substantially more practical than others, but in general, the less terms in an equation, the more useful it is. Let us therefore consider some of the simplest identities. As these take on somewhat different appearances depending on the parity of the number of external particles, it is convenient to treat the two cases separately.

### 4.1 Even $n$ reductions

A subset of the CHY integral identities that can be extracted from  $\Psi$ , are those that only involve the sub-matrix  $A$ ,

$$A_{ij} = \begin{cases} \frac{k_i \cdot k_j}{z_i - z_j}, & \text{if } i \neq j, \\ 0, & \text{if } i = j, \end{cases} \quad (100)$$

Because the rows and columns are not independent, the Pfaffian of  $A$  is always zero. For even  $n$ , this fact provides a non-trivial identity relating  $(n - 1)!!$  CHY diagrams. But there is another Pfaffian identity that is more practical because it relates only  $2(n - 3)!!$  CHY diagrams, and this identity is the invariance of the reduced Pfaffian: on solutions to the scattering equations, the value of  $\text{Pf}'A$  is independent of the choice of  $i$  and  $j$ .

Diagrammatically, the invariance of the reduced Pfaffian can be interpreted as follows: If we draw a 3-regular graph with  $n$  vertices (we will call this graph the 'template'), then we need to superimpose a 1-regular graph in order to have a graphical representation of a CHY integral. The Pfaffian of  $A$  can be regarded as the sum (with sign) over all possible 1-regular graphs given  $n$  vertices, where each term is multiplied with a Mandelstam variable  $s_{ab}$  if the 1-regular graph has an edge connecting vertices  $a$  and  $b$ . The reduced Pfaffian is the same sum, except that one edge is fixed and not multiplied with a Mandelstam variable. The invariance of the reduced Pfaffian tells us that if on our 3-regular graph we connect any two vertices  $i$  and  $j$  with an additional edge and sum (with sign) over all ways of connecting the  $n - 2$  remaining vertices to form a 4-regular graph while multiplying each term with the proper Mandelstam variables, then this sum will be independent of the choice of  $i$  and  $j$ .

As an example, consider the following template:

$$\quad \quad \quad (101)$$

By applying to the template the fact that the reduced Pfaffian is the same whether we fix legs 1 and 4 or legs 4 and 5, we obtain the following identity:

$$s_{23}s_{56} \quad - s_{25}s_{36} \quad + s_{26}s_{35} \quad = \quad (102)$$



$$s_{12}s_{36} \begin{array}{c} 2 \quad 1 \\ \diagup \quad \diagdown \\ 3 \quad 6 \\ \diagdown \quad \diagup \\ 4 \quad 5 \end{array} - s_{13}s_{26} \begin{array}{c} 2 \quad 1 \\ \diagup \quad \diagdown \\ 3 \quad 6 \\ \diagdown \quad \diagup \\ 4 \quad 5 \end{array} + s_{16}s_{23} \begin{array}{c} 2 \quad 1 \\ \diagup \quad \diagdown \\ 3 \quad 6 \\ \diagdown \quad \diagup \\ 4 \quad 5 \end{array} .$$

This equation relates a CHY integral with a second-order pole to five integrals that only have simple poles.

## 4.2 Odd $n$ reductions

When  $n$  is odd,  $\text{Pf} A$  and  $\text{Pf}' A$  vanish trivially, and so these identities are of no use. But if by  $A_i$  we denote the sub-matrix of  $A$  obtained by removing row and column  $i$  from  $A$ , then we also have the non-trivial identity that  $\text{Pf} A_i = 0$  for all  $i$  when evaluated on solutions to the scattering equations.

The vanishing of  $\text{Pf} A_i$  provides an identity that relates  $(n-2)!!$  CHY diagrams. Graphically, the identity can be represented as follows:

As a template we draw a graph with  $n$  vertices and  $\frac{1}{2}(3n+1)$  edges such that there are three edges incident to each vertex except for one vertex to which four edges are incident. If we then sum (with sign) over all the ways of drawing the remaining  $\frac{1}{2}(n-1)$  edges so that we obtain a 4-regular graph, while multiplying each term with the Mandelstam variable  $s_{ab}$  if an edge connecting vertices  $a$  and  $b$  is added to the template, then the result will be zero.

As an example, consider the following template:

$$\begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad 4 \end{array} . \quad (103)$$

Performing the weighted sum over the different ways of completing the diagram, we get zero:

$$0 = s_{23}s_{45} \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad 4 \end{array} - s_{24}s_{35} \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad 4 \end{array} + s_{25}s_{34} \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad 4 \end{array} . \quad (104)$$

## 4.3 Numerators and templateless diagrams

In some cases the Pfaffian identities permit one to immediately express an unknown CHY diagram in terms of some that can be evaluated with the integration rules. But in general it will be necessary to invoke several Pfaffian identities involving several unknown CHY integrals and then solve a system of equations. One will also have to start with simple diagrams and then work towards more complicated ones. A diagram with four edges connecting the same two vertices can only appear in Pfaffian identities where the template

has two vertices connected with three edges, so the identities involving quadruple-line diagrams will always involve only diagrams that cannot be calculated with the integration rules. Hence, one will have to first consider other Pfaffian identities relating the triple-line diagrams to yet simpler diagrams.

For simplicity, we have up to this point restricted focus to the special case when  $H(z)$  has a trivial numerator. But CHY integrals whose integrands have non-trivial numerators can also be re-expressed via Pfaffian identities, and in fact it is necessary to calculate CHY integrals with cross ratios in order to compute gluon or graviton amplitudes by expanding out  $\text{Pf}'\Psi$ . This is because of the diagonal entries of the matrix  $C$ :

$$C_{ii} = - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\epsilon_i \cdot k_j}{z_i - z_j}. \quad (105)$$

If, in a CHY integral  $\mathcal{I}[\mathcal{G}]$ , one were to expand out the terms of  $C_{ii}$  in the above form, the individual terms would not be Möbius invariant, only the sum would. But  $C_{ii}$  can be re-written in a manifestly Möbius invariant form thus:

$$C_{ii} = - \frac{\epsilon_i \cdot k_1}{z_i - z_1} - \sum_{\substack{j=2 \\ j \neq i}}^n \frac{\epsilon_i \cdot k_j}{z_i - z_j} = \sum_{\substack{j=2 \\ j \neq i}}^n \frac{\epsilon_i \cdot k_j}{z_i - z_1} - \sum_{\substack{j=2 \\ j \neq i}}^n \frac{\epsilon_i \cdot k_j}{z_i - z_j} = \sum_{\substack{j=2 \\ j \neq i}}^n \frac{\epsilon_i \cdot k_j (z_1 - z_j)}{(z_i - z_1)(z_i - z_j)}. \quad (106)$$

Of course there is nothing special about leg number 1, and in re-writing  $C_{ii}$  one could just as easily have used momentum conservation to remove any other momentum except that indexed by  $i$ .

While CHY integrals with cross-ratios can also be expressed in terms of other CHY integrals by selecting a template with a dotted line, there are CHY integrals for which no template exists. These are the integrals represented by a 4-regular CHY graph that has no 3-regular sub-graph. An example is this one



Such diagrams do not appear when expanding out the CHY formula for gluon amplitudes, but they do appear in the graviton case. While Pfaffian identities from the sub-matrix  $A$  cannot be used to reduce these diagrams, other identities can be extracted from  $\text{Pf}'\Psi$  to serve this purpose. Setting to zero all dot products  $\epsilon_i \cdot \epsilon_j$  except  $\epsilon_1 \cdot \epsilon_2$  and  $\epsilon_4 \cdot \epsilon_5$  and all dot products  $\epsilon_i \cdot k_j$  except  $\epsilon_3 \cdot k_1$ ,  $\epsilon_3 \cdot k_2$ ,  $\epsilon_6 \cdot k_4$ , and  $\epsilon_6 \cdot k_5$ , the identity between the reduced Pfaffian obtained by removing rows and columns 1 and 3 and the reduced

Pfaffian obtained by removing rows and columns 3 and 6 takes on the following form:

$$s_{456} \text{ (diagram)} = -s_{14} \text{ (diagram)} - s_{15} \text{ (diagram)} - s_{24} \text{ (diagram)} - s_{25} \text{ (diagram)}. \quad (108)$$

## 5 Cycle integrals

### 5.1 The first string theory–CHY duality

There is a class of CHY integrands  $\mathcal{G}(z)$  that deserves special attention. This is the class of integrands that can be written as the product of two so-called Hamiltonian cycles, that is, the integrands for which

$$\mathcal{G}(z) = \text{Cycle}(z) \text{Cycle}_2(z), \quad (109)$$

where the cycles are of the form

$$\prod_{i=1}^n \frac{1}{z_{\sigma(i)} - z_{\sigma(i+1)}}, \quad \sigma \in S_n. \quad (110)$$

Without loss of generality, we can assume that the ordering of  $\text{Cycle}_2$  is the canonical ordering of the external legs so that

$$\text{Cycle}_2(z) = \prod_{i=1}^n \frac{-1}{z_i - z_{i+1}}, \quad z_{n+1} = z_1. \quad (111)$$

Now given any subset  $\tau$  of the external points or vertices, the number of edges connecting them with each other can be at most  $|\tau| - 1$  per cycle. Hence, the number of edges connecting vertices in  $\tau$  with themselves can be at most  $2|\tau| - 2$ . In other words: double-cycle integrals do not have higher-order poles. They can always be evaluated with the integration rules.

We next observe that in order for a subset  $\tau$  to have  $2|\tau| - 2$  edges within itself so that it can contribute a propagator to result of the integration,  $|\tau| - 1$  edges must come from  $\text{Cycle}_2$ . In other words: The propagators always carry consecutive legs. This fact has also been proved by CHY in the scattering equation formalism, where they have shown that any two-cycle CHY integral evaluates to the sum of all Feynman diagrams that are compatible with the orderings of both cycles.<sup>[5]</sup>

As a third observation, we remark that since the subsets  $\tau$  that can contribute a propagator are always consecutive, there are no inversion factors:  $\mathcal{N}_{\text{inv}} = 0$ . The overall sign,

therefore, is merely  $(-1)^{n+1}$  provided that differences  $(z_i - z_j)$  are always written such that the minuend has the lowest index.

Next, we note that since we have fixed  $\text{Cycle}_2$ , the double-cycle integral depends only on the first cycle, and for this reason the integration rules simplify. In order for a subset  $\tau$  to be able to contribute a propagator there must be  $|\tau| - 1$  factors of  $(z_i - z_j)^{-1}$  in  $\text{Cycle}(z)$  with  $i, j \in \tau$ . Harking back to section 2, we see that if we set  $H(z) = \text{Cycle}(z)$  in the string theory integral, then the CHY integration rules become identical to the string theory integration rules.

This leads us to what I shall term *the first string theory–CHY duality*:

$$\lim_{\alpha' \rightarrow 0} (\alpha')^{n-3} \int d\mu_n \Lambda_n(\alpha', k, z) \text{Cycle}(z) = (-1)^n \int d\Omega_n \frac{\text{Cycle}(z)}{\prod_{i=1}^n (z_i - z_{i+1})}. \quad (112)$$

We see that, up to an over-all sign, one can interchange string theory and CHY cycle integrations by the following simple procedure:

$$(\alpha')^{n-3} d\mu_n \Lambda_n(\alpha', k, z) \quad \leftrightarrow \quad d\Omega_n \prod_{i=1}^n (z_i - z_{i+1})^{-1}. \quad (113)$$

And so we have arrived at exactly the translation prescription between string theory and CHY integrals that was first put forward in [15].

## 5.2 Polygon decomposition: Dual Feynman diagrams

We have seen that for any cycle we obtain the same result whether we integrate over it with the string theory measure or divide it with the cyclic product of neighbouring differences and integrate over it with the CHY measure. The result of either integration we will denote by  $\mathcal{I}[\text{Cycle}]$ . Now, another reason why the double-cycle special case merits special attention is the fact that there is a one-to-one correspondence between the cycles for which  $\mathcal{I}[\text{Cycle}]$  is non-zero and individual  $\phi^3$  Feynman diagrams that may contain sub-amplitudes and that are planar with respect to the canonical ordering.

In string theory, the question of how to identify various terms in the expansion of integrals with individual Feynman diagrams was studied already in the early days of the Veneziano model<sup>[20]</sup>, while, in the scattering equation formalism, CHY in 2013 pointed out an integrand–Feynman diagram correspondence and provided a beautiful diagrammatic description of it by what they termed a polygon decomposition.<sup>[5]</sup>

The correspondence and the diagrammatic representation, which apply equally well to string theory and the CHY formalism, and which we will be able to prove with the integration rules, can be described thus: Given a tree-level Feynman diagram in  $\phi^3$  scalar field theory, possibly involving  $p$ -point sub-amplitudes, the cycle for which  $\mathcal{I}[\text{Cycle}]$  evaluates to the given Feynman diagram can be determined through the following steps:

1. Draw the Feynman diagram.
2. Dissociate the Feynman diagram into its component vertices.
3. Replace each Feynman-vertex with a polygon with the same number of polygon-vertices as legs in the Feynman-vertex.
4. Connect the polygons at the adjacent polygon-vertices.

From the polygon diagram constructed by this procedure, one can directly read of the cycle: If the diagram has an edge connecting vertices  $i$  and  $j$ , then  $\text{Cycle}(z)$  contains a factor of  $(z_i - z_j)^{-1}$ .

The procedure is perhaps best understood through an example, so I urge the reader to consider figure 1. We note that the Feynman diagram shown at step 1 is actually the sum of  $2 \cdot 2 \cdot 5 = 20$  individual  $\phi^3$  Feynman diagrams since at each of the two four-point vertices and at the five-point vertex one has to sum over all the factorization channels of the respective sub-amplitudes. Specifically, if we introduce a function that is equal to the sum of the factorization channels of the  $p$ -point sub-amplitudes so that it can be expressed in terms of the  $\phi^3$  scattering amplitude as follows:

$$\mathcal{B}_p[1, 2, \dots, p-1] = \mathcal{A}_p^{\phi^3}[k_1, k_2, \dots, k_{p-1}, -\sum_{i=1}^{p-1} k_i], \quad (114)$$

then the Feynman diagram in figure 1 is equal to

$$\frac{\mathcal{B}_4[3, 4, 5] \mathcal{B}_4[3+4+5, 6, 7+8] \mathcal{B}_5[9, 10, 11, 12]}{s_{1,2} s_{7,8} s_{3,4,5} s_{9,10,11,12} s_{3,4,5,6,7,8}} \quad (115)$$

In following the steps of the procedure, one finds that  $\mathcal{I}[\text{Cycle}]$  evaluates to the Feynman diagram under consideration when the cycle is given as follows:

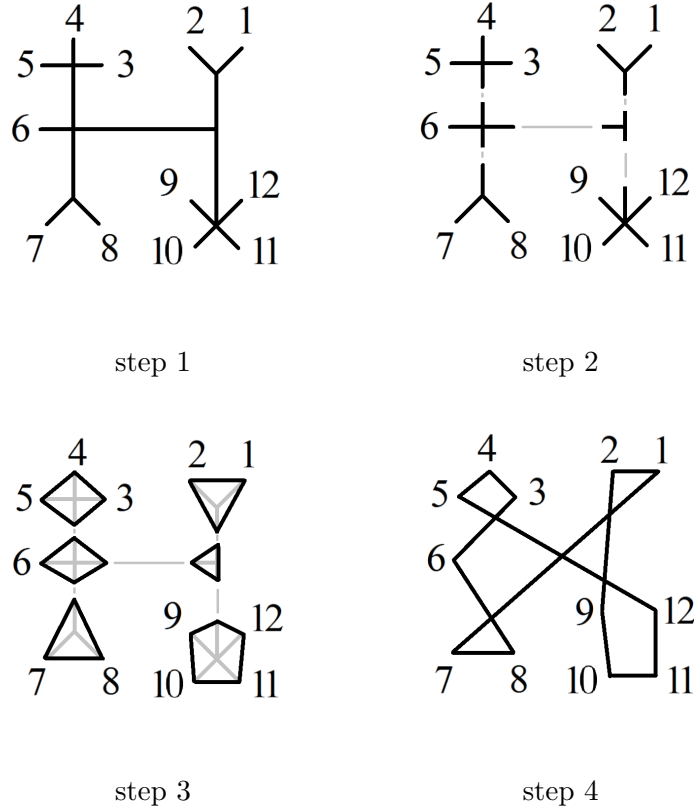
$$\text{Cycle}(z) = \frac{1}{(z_1 - z_2)(z_2 - z_9)(z_9 - z_{10})(z_{10} - z_{11})(z_{11} - z_{12})(z_5 - z_{12})} \times \frac{1}{(z_4 - z_5)(z_3 - z_4)(z_3 - z_6)(z_6 - z_8)(z_7 - z_8)(z_1 - z_7)} \quad (116)$$

While the procedure for finding the cycle corresponding to a Feynman diagram is entirely straightforward, the same is not quite true when going in the reverse direction. This is because the polygons formed by a cycle changes with the relative distances of the external points, and drawing the external points equidistantly on a circle will not always result in internal polygons that exhibit the Feynman diagram that the cycle integral evaluates to. But for non-zero cycle integrals it is always possible to move the points into such positions that the correct polygons appear. There are cycles that cannot be transformed into a set of polygons meeting at the the vertices without also forming internal polygons, such as in the following cases:

$$, \quad (117)$$

For such cycles, the integral always vanishes:  $\mathcal{I}[\text{Cycle}] = 0$ . It holds true in general that for cycles without at least two edges connecting neighbouring vertices,  $\mathcal{I}[\text{Cycle}] = 0$ .

It will be noted that the identification between  $p$ -gons and  $p$ -point sub-amplitudes provides us with a large number of identities between integrals since integrals with sub-amplitudes must be equal to the sum of the integrals corresponding to the individual



**Figure 1:** The procedure for determining which cycle evaluates to a given Feynman diagram.

Feynman diagrams of the expansion of the sub-amplitude. For example, from the Feynman diagram expansion of the  $\phi^3$  5-point amplitude we obtain the following identity:

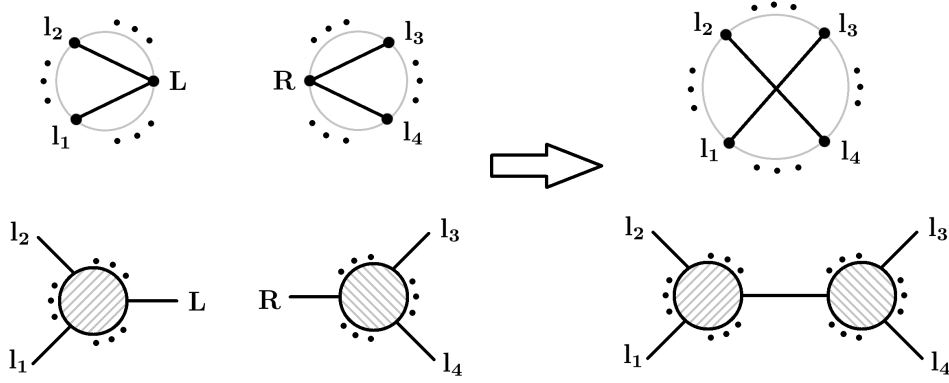
$$\begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} = \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} \quad (118)$$

The identity holds true only post-integration. But we can conclude that subtracting the five integrands on the right-hand side of equation (118) from the left-hand side must result in some Möbius invariant quantity that integrates to zero. And indeed by partial fractioning we find the following integrand-level identity:

$$\begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} = \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} + \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array} - 2 \begin{array}{c} 1 \\ \diagup \quad \diagdown \\ 2 \quad \text{---} \quad 5 \\ \diagdown \quad \diagup \\ 3 \quad \text{---} \quad 4 \end{array}. \quad (119)$$

### 5.2.1 Proof of the polygon decomposition

We will here outline an inductive proof as to why the dual Feynman diagrams obtained by the polygon decomposition do indeed produce the correct cycles. The base case is already firmly established in the literature. In string theory, the base case is the Koba-Nielsen



**Figure 2:** Merging cycles corresponds to merging Feynman diagrams.

formula (1), which tells us that one obtains the tree-level  $\phi^3$  amplitude by integrating the canonical cycle

$$\prod_{i=1}^n \frac{1}{z_i - z_{i+1}} \quad (120)$$

with respect to the string theory measure. In the CHY formalism, the base case is but another way of phrasing the CHY  $\phi^3$  formula (49): The CHY diagram for the  $\phi^3$   $n$ -point amplitude is a double-lined  $n$ -gon. What we need to show, then, is the inductive step, which can be expressed thus:

*Take any two external legs  $L$  and  $R$  of any two Feynman diagrams and connect them. The new Feynman diagram so constructed corresponds to the cycle obtained by connecting the vertices  $L$  and  $R$  of the cycles corresponding to the original two Feynman diagrams.*

This statement can best be conveyed pictorially, and we urge the reader to consider figure 2.

In order to justify the statement, we start by considering two Feynman diagrams, which, by the inductive assumption, can each be written as a cycle integral:

- $\mathcal{I}[\text{Cycle}_L]$ , whose points we shall label with  $1, 2, \dots, n$ , and
- $\mathcal{I}[\text{Cycle}_R]$ , whose points shall be labelled  $n + 1, n + 2, \dots, n + n'$ .

Let  $L, l_1, l_2 \in \{1, 2, \dots, n\}$  be external points of the first Feynman diagram with canonical ordering  $L \rightarrow l_1 \rightarrow l_2 \rightarrow L$ , and let the points  $l_1$  and  $l_2$  be connected to point  $L$  in  $\text{Cycle}_L$ . Let  $R, l_3, l_4 \in \{n + 1, n + 2, \dots, n + n'\}$  be external points of the second Feynman diagram with canonical ordering  $R \rightarrow l_3 \rightarrow l_4 \rightarrow R$ , and let the points  $l_3$ , and  $l_4$  be connected to point  $R$  in the  $\text{Cycle}_R$ . In that case the Feynman diagram constructed by connecting legs  $L$  and  $R$  of the two original Feynman diagrams will be equal to the following cycle integral:

$$\mathcal{I} \left[ \frac{(z_{l_1} - z_L)(z_L - z_{l_2})(z_{l_3} - z_R)(z_R - z_{l_4})}{(z_{l_2} - z_{l_4})(z_{l_1} - z_{l_3})} \text{Cycle}_L \text{Cycle}_R \right] \quad (121)$$

To understand why this is so, we direct the reader's attention to the following observations:

- By merging vertices  $L$  and  $R$  of the two original cycles in the manner shown on figure 2 we do indeed obtain a new Hamiltonian cycle, so that Möbius invariance is retained both in string theory and the CHY formalism. The cycle integral is therefore well-defined and must evaluate to either zero or a Feynman diagram.
- In order to be non-zero, according to the integration rules, the combined cycle must have  $n + n' - 5$  pairwise compatible propagators, which is equal to one plus the sum of the numbers of propagators in the two Feynman diagrams.
- Since both the edges removed from the first cycle were connected to  $L$ , any subset  $q$  of vertices  $\mathbb{Z}_n \setminus \{L\}$  from  $\text{Cycle}_L$  will be connected by the same number of edges. Therefore any subset of vertices that gave a propagator  $\mathcal{I}[\text{Cycle}_L]$  also contributes a propagator to the integral over the combined cycle, and no other subset of vertices from  $\mathcal{I}[\text{Cycle}_L]$  contributes a propagator. The same is true for  $\mathcal{I}[\text{Cycle}_R]$ .
- $\text{Cycle}_L$  contains  $n$  vertices and  $n$  edges. In forming the combined cycle we removed one of these vertices and two of the edges. Therefore, if in the combined cycle we consider the set of vertices  $q = \mathbb{Z}_n \setminus \{L\}$  of external points, there will be  $n-2 = |q|-1$  lines connecting these points with each other. Hence, by the integration rules, this set gives rise to a propagator.
- Finally, we note that there is no subset  $\tau$  of the vertices of the combined cycle that gives rise to a propagator and for which both  $\tau$  and its complement contain vertices from both  $\text{Cycle}_L$  and  $\text{Cycle}_R$ . For if there were such a subset, there would have to be an edge connecting the vertices in  $\tau$  from  $\text{Cycle}_L$  with the vertices in  $\tau$  from  $\text{Cycle}_R$ . And there are only two edges connecting vertices from  $\text{Cycle}_L$  with vertices from  $\text{Cycle}_R$ , viz. the edge between vertices  $l_1$  and  $l_3$  and that between vertices  $l_2$  and  $l_4$ . So  $\tau$  would have to contain either the vertices  $l_1$  and  $l_3$  or vertices  $l_2$  and  $l_4$ , and the same would have to hold true for  $\tau^c$ . But this is impossible, for the vertices in  $\tau$  would have to be consecutive, and so would the ones in  $\tau^c$ , since only subsets of consecutive vertices can give rise to propagators in cycle integrals.

Thus we conclude that the subsets of the vertices of the combined cycle that can give rise to a propagator are exactly those for which the subset itself or its complement either

1. is a subset of vertices that gave a propagator to  $\mathcal{I}[\text{Cycle}_L]$ , or
2. is a subset of vertices that gave a propagator to  $\mathcal{I}[\text{Cycle}_R]$ , or
3. is equal to  $q = \{1, 2, \dots, n\} \setminus \{L\}$ .








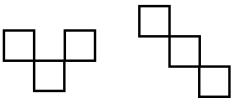

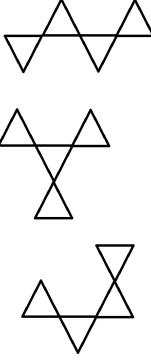
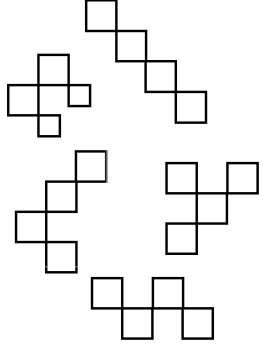
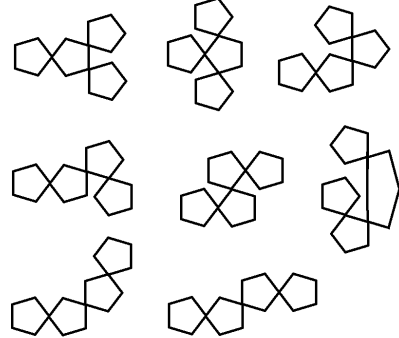
And given two subsets from two different groups, they will either be disjoint or nested. By the integration rules, then, we have that the integral over the combined cycle evaluates to

$$\begin{aligned} & \sum \text{products of } n - 3 \text{ compatible propagators from } \mathcal{I}[\text{Cycle}_L] \times & (122) \\ & \sum \text{products of } n' - 3 \text{ compatible propagators from } \mathcal{I}[\text{Cycle}_R] \times \frac{1}{s_q}. \end{aligned}$$

But this is equal to

$$\mathcal{I}[\text{Cycle}_L] \times \mathcal{I}[\text{Cycle}_R] \times \frac{1}{s_q}, \quad (123)$$



$p$			
$2p - 2$			
$3p - 4$			
$4p - 6$			

**Figure 3:** The first few pure polygon cycle structures

which we recognize as the value of the Feynman diagram obtained by merging legs  $L$  and  $R$  of the original two Feynman diagrams.

It is clear that the procedure shown on figure 2 preserves the polygon structure of the original cycles and simply attaches vertices  $L$  and  $R$ . By repeatedly connecting vertices of various  $p$ -gon cycles corresponding to  $\phi^3$   $p$ -point amplitudes, one can construct any  $\phi^3$  Feynman diagram with any sub-amplitudes. Thus we recursively prove that a cycle can be found for any allowed Feynman diagram by a polygon decomposition.

### 5.3 Brute force phi-to-the- $p^{\text{th}}$

We have seen that the cycles that yield a non-zero result upon integration are those that can be represented as a set of polygons meeting at the vertices. Such cycles can be catalogued according to their polygons: how many they consists of, what type of polygons – solely triangles or squares or pentagons etc. or combinations – they are built from, and how the polygons are attached to each other. On figure 3 we list the first few non-mixed cycle structures.

The structures of the cycles mirror those of Feynman diagrams of scalar theories, pure triangles corresponding to  $\phi^3$ , pure squares to  $\phi^4$ , etc. This similarity can be used to build up  $n$ -point amplitudes for  $\phi^p$  theories in string theory and the CHY formalism by considering all the structures that can be formed by connecting  $p$ -gons together at the vertices – without forming internal polygons – and for each structure counting up all the different ways of canonically ordering the structure, up to reflection and rotation. It will

be necessary, though, to explicitly cancel the  $p$ -point  $\phi^3$  sub-amplitudes.

As an example to illustrate the principle, we can compute the  $\phi^4$  10-point amplitude:

$$\mathcal{A}_{10}^{\phi^4} = \mathcal{A}[\text{diagram 1}] + \mathcal{A}[\text{diagram 2}] + \mathcal{A}[\text{diagram 3}] + \mathcal{A}[\text{diagram 4}] + \mathcal{A}[\text{diagram 5}], \quad (124)$$

where the five terms are respectively given by

$$\begin{aligned} \mathcal{A}[\text{diagram 1}] &= \frac{1}{\mathcal{B}_4[1, 2, 3] \mathcal{B}_4[6, 7, 8] \mathcal{B}_4[1 + 2 + 3, 4, 10] \mathcal{B}_4[5, 6 + 7 + 8, 9]} \quad (125) \\ &+ 4 \text{ rotations} = \frac{1}{s_{123} s_{678} s_{56789}} + \dots \quad (5 \text{ terms}), \end{aligned}$$

$$\begin{aligned} \mathcal{A}[\text{diagram 2}] &= \frac{1}{\mathcal{B}_4[1, 2, 3] \mathcal{B}_4[4, 5, 6] \mathcal{B}_4[8, 9, 10] \mathcal{B}_4[1 + 2 + 3, 4 + 5 + 6, 7]} \quad (126) \\ &+ 9 \text{ rotations} = \frac{1}{s_{123} s_{456} s_{8910}} + \dots \quad (10 \text{ terms}), \end{aligned}$$

$$\begin{aligned} \mathcal{A}[\text{diagram 3}] &= \frac{1}{\mathcal{B}_4[1, 2, 3] \mathcal{B}_4[6, 7, 8] \mathcal{B}_4[1 + 2 + 3, 4, 5] \mathcal{B}_4[6 + 7 + 8, 9, 10]} \quad (127) \\ &+ 9 \text{ rotations and reflections} = \frac{1}{s_{123} s_{678} s_{12345}} + \dots \quad (10 \text{ terms}), \end{aligned}$$

$$\begin{aligned} \mathcal{A}[\text{diagram 4}] &= \frac{1}{\mathcal{B}_4[1, 2, 3] \mathcal{B}_4[7, 8, 9] \mathcal{B}_4[1 + 2 + 3, 4, 10] \mathcal{B}_4[5, 6, 7 + 8 + 9]} \quad (128) \\ &+ 19 \text{ rotations and reflections} = \frac{1}{s_{123} s_{789} s_{56789}} + \dots \quad (20 \text{ terms}), \end{aligned}$$

$$\begin{aligned} \mathcal{A}[\text{diagram 5}] &= \frac{1}{\mathcal{B}_4[5, 6, 7] \mathcal{B}_4[8, 9, 10] \mathcal{B}_4[3, 4, 5 + 6 + 7] \mathcal{B}_4[1, 2, 8 + 9 + 10]} \quad (129) \\ &+ 9 \text{ rotations} = \frac{1}{s_{567} s_{8910} s_{34567}} + \dots \quad (10 \text{ terms}). \end{aligned}$$

The extension of this  $\phi^p$  procedure to mixed vertices is straightforward, although the number of cycle structures increases drastically when there are several types of  $p$ -gons, and so does the number of different copies of the same cycle structure since having different types of polygons reduces the dihedral and chiral symmetry of cycle structures.

## 6 $\phi^p$ from string theory

### 6.1 $\phi^4$ and the second string theory–CHY duality

The  $\phi^p$  procedure reviewed in the last section suffers from the deficiency that it requires one to count up a large number of contributing terms, which each have to be dressed with a non-trivial pre-factor that depends on the external momenta. And so the question arises if there do not exist more compact expressions for the pure scalar amplitudes in string theory and the CHY formalism. In the case of  $\phi^4$  theory, up to two-loop amplitudes have been calculated with string theory by [21],[22],[23], though they have not derived a closed-form tree-level formula. But such a formula does in fact exist, for by a compactification of Yang-Mills theory, it is possible to use the four-vertex of this theory to build up quartic scalar tree-amplitudes. In this manner, as a corollary to their construction of Yang-Mills-scalar theory, CHY have derived the following formula:<sup>[7]</sup>

$$\mathcal{A}_n^{\phi^4} = \int d\Omega_n \frac{\text{Pf}' A}{\prod_{i=1}^n (z_i - z_{i+1})} \sum_{\substack{\text{connected perfect} \\ \text{matchings}}} \frac{1}{PM(z)} \quad (130)$$

Here,  $PM(z)$  denotes a so-called perfect matching, that is, a product of differences  $(z_i - z_j)$  such that each  $z_i$  appears in exactly one factor; diagrammatically, it can be represented with a 1-regular graph. The sum in (130) is over all the perfect matchings represented by a connected 1-regular graph, or, equivalently, the  $PM(z)$  for which, if one selects any proper subset  $\rho \subset \mathbb{Z}_n$  of consecutive numbers, then  $PM(z)$  will contain at least one factor  $(z_i - z_j)$  with  $i \in \rho$  and  $j \notin \rho$ .

Incidentally, it happens that each term where the 1-regular graph of  $PM(z)$  is a  $\phi^4$  Feynman diagram, the term is equal to just that Feynman diagram, while terms for which  $PM(z)$  has a connected graph that is not a  $\phi^4$  Feynman diagram evaluate to zero. This behaviour is not at all obvious since, if one wishes to apply the integration rules, one must for each term in equation (130) expand out  $\text{Pf}' A$  into  $(n-3)!!$  terms, which one can then apply the integration rules to and then sum up to obtain either 0 or a  $\phi^4$  Feynman diagram. But from the perspective of string theory, this behaviour can be understood.

Tree-level Yang-Mills amplitudes can be calculated both from the bosonic string string (4) and from the CHY formalism (50). When we expand out the two formulas in terms with different pre-factors composed of  $\epsilon_i \cdot \epsilon_j$  and  $k_i \cdot \epsilon_j$ , the formulas must agree term by term for terms with pre-factors consisting of only  $\epsilon_i \cdot \epsilon_j$  since the polarization vectors can be chosen independently of each other. Hence, we arrive at *the second string theory–CHY duality*:

$$\lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-4}{2}} \int d\mu_n \frac{\Lambda_n(\alpha', k, z)}{[PM(z)]^2} = \int d\Omega_n \frac{\text{Pf}' A}{PM(z) \prod_{i=1}^n (z_i - z_{i+1})}. \quad (131)$$

By summing the string theory expression over connected perfect matchings, we obtain a  $\phi^4$  formula completely equivalent to equation (130). But now the behaviour of the individual terms follows from the integration rules of section 2 with

$$H(z) = \frac{1}{[PM(z)]^2}. \quad (132)$$

The second string theory–CHY duality is in a certain regard more interesting than the first. For the first duality relates only the most well-behaved string theory and CHY integrals: Because of the  $(\alpha')^{n-3}$  pre-factor in equation (112), the only string theory integrals that are not killed off in the  $\alpha' \rightarrow 0$  limit, are those that have  $n - 3$  compatible divergences, and which therefore require no residual integration, so that the  $K_T$  constants in (36) are all equal to unity. And the CHY integrals in equation (112), being cycle integrals, do not have any higher-order poles.

Now, if, in equation (131),  $PM(z)$  is a connected perfect matching, as in the  $\phi^4$  formula, then these integrals are also of the well-behaved kind: After expanding out  $\text{Pf}'A$ , the CHY integrals have no higher-order poles, and while there are integrations left after taking care of the divergences in the string theory integrals, the residual integrations always trivially produce unity. But the second duality holds true also when  $PM(z)$  is not a perfect matching, and in these cases the CHY integrals have higher-order poles, and the string theory integrals are so strongly divergent that analytical extension is needed to make sense of the  $\alpha' \rightarrow 0$  limit, which entails that the  $K_T$  constants depend on the external momenta.

To see this play out in a concrete example, consider the perfect matching

$$PM(z) = (z_1 - z_2)(z_3 - z_6)(z_4 - z_8)(z_5 - z_7), \quad (133)$$

which can be represented diagrammatically as



In the  $z_1 = \infty$ ,  $z_2 = 1$ ,  $z_8 = 0$  gauge, the left-hand side of equation (131) takes on the form

$$\lim_{\alpha' \rightarrow 0} (\alpha')^2 \int d\mu_n \frac{1}{(z_3 - z_6)^2 z_4^2 (z_5 - z_7)^2} \prod_{j=i+1}^n (z_i - z_j)^{\alpha' s_{ij}}. \quad (135)$$

The integration can be carried out by changing to the rescaled variables  $x_3 = z_3$ ,  $x_i = \frac{z_i}{z_{i-1}}$  for  $i = 4, 5, 6, 7$ . The integral then becomes the following:

$$\begin{aligned} & \int_0^1 dx_7 x_7^{\alpha' s_{78}} (1 - x_7)^{\alpha' s_{67}} \\ & \int_0^1 dx_6 x_6^{\alpha' s_{678} + 1} (1 - x_6)^{\alpha' s_{56}} \frac{(1 - x_6 x_7)^{\alpha' s_{57}}}{(1 - x_6 x_7)^2} \\ & \int_0^1 dx_5 x_5^{\alpha' s_{5678}} (1 - x_5)^{\alpha' s_{45}} (1 - x_5 x_6)^{\alpha' s_{46}} (1 - x_5 x_6 x_7)^{\alpha' s_{47}} \\ & \int_0^1 dx_4 \frac{x_4^{\alpha' s_{45678}}}{x_4} (1 - x_4)^{\alpha' s_{34}} (1 - x_4 x_5)^{\alpha' s_{35}} \frac{(1 - x_4 x_5 x_6)^{\alpha' s_{36}}}{(1 - x_4 x_5 x_6)^2} (1 - x_4 x_5 x_6 x_7)^{\alpha' s_{37}} \\ & \int_0^1 dz_3 \frac{x_3^{\alpha' s_{345678}}}{x_3^2} (1 - x_3)^{\alpha' s_{23}} (1 - x_3 x_4)^{\alpha' s_{24}} (1 - x_3 x_4 x_5)^{\alpha' s_{25}} (1 - x_3 x_4 x_5 x_6)^{\alpha' s_{26}} (1 - x_3 x_4 x_5 x_6)^{\alpha' s_{27}} \end{aligned} \quad (136)$$

In taking the  $\alpha' \rightarrow 0$  limit, the only parts of the integration domain that are not killed off are:

1. the region where  $x_6, x_7 \rightarrow 1$  and  $x_4, x_5 \rightarrow 1$
2. the region where  $x_6, x_7 \rightarrow 1$  and  $x_4 \rightarrow 0$ .

The left-hand side of equation (131) is therefore equal to

$$\lim_{\alpha' \rightarrow 0} \int_0^1 dx_3 \frac{x_3^{\alpha' s_{345678}}}{x_3^2} \left( \frac{(1-x_3)^{\alpha' s_{23}}}{s_{45678}} + \frac{(1-x_3)^{\alpha' (s_{23}+s_{24}+s_{25}+s_{26}+s_{27})}}{s_{34567}} \right) \frac{1}{s_{567}}, \quad (137)$$

which after analytical extension will be found, by expanding out the beta function, to be equal to

$$\frac{1}{s_{567} s_{45678}} \left( 1 + \frac{s_{23}}{s_{345678}} \right) + \frac{1}{s_{567} s_{34567}} \left( 1 + \frac{s_{23} + s_{24} + s_{25} + s_{26} + s_{27}}{s_{345678}} \right). \quad (138)$$

Performing the calculation in the CHY formalism produces exactly the same result.

## 6.2 $\phi^p$ rules

While it is possible to extract the quartic scalar amplitudes from Yang-Mills theory, the theory cannot yield  $\phi^p$  for  $p$  greater than 4, so we must begin afresh if we are to find a compact way of expressing the general  $\phi^p$  tree-amplitude. A major obstacle lies in our way if we are to pursue this goal via the scattering equations. For CHY integrals have the same dimension as  $\phi^3$  Feynman diagrams, so the greater  $p$  is, the greater the dimensional mismatch between CHY integrals and  $\phi^p$  Feynman diagrams. In the CHY  $\phi^4$  formula (130) the correct dimension is obtained because of the Mandelstam variables in  $\text{Pf}'A$ , but it is difficult to see a general way of remedying the mismatch.

In string theory, we are not faced with this difficulty, for here the dimensionality comes from the  $\alpha'$  pre-factor and is determined by the exponent of this factor, which is a parameter we can modify. It appears a string theory  $\phi^p$  formula does indeed lie within our grasp, and we will now direct our steps to the pursuance of this goal.

We start with some observations for  $\phi^p$  theory: The only non-zero  $n$ -point amplitudes are those for which  $n = L(p-2) + 2$  with  $L \in \mathbb{N}$ . For these  $n$ -point amplitudes, each Feynman diagram carries  $L-1$  propagators. The number of external legs carried by each propagator is equal to 1 modulo  $p-2$  and is greater than one and less than  $n-p+2 = (L-1)(p-2) + 2$ . In other words, the number of external legs a propagator can carry is of the form  $l(p-2) + 1$  with  $1 < l < L$ .

Based on these observations, we can formulate the following algorithm for computing the non-zero colour-ordered  $\phi^p$  amplitude  $\mathcal{A}_n^{\phi^p}$ , where  $n = L(p-2) + 2$ :

- Count up all subsets  $\tau \in \mathbb{Z}_n$  of consecutive numbers (1 and  $n$  being considered consecutive) for which  $|\tau| = (p-2)l + 1$  for some integer  $l$  bigger than one and less than  $L$ . Complementary subsets are considered equivalent.
- To each of these subsets  $\tau$  associate a propagator

$$P_\tau = \frac{1}{\sum_{\substack{i,j \in \tau \\ i < j}} s_{ij}}. \quad (139)$$

We recall that two subsets  $\tau$  and  $\tau'$  are said to be compatible if  $\tau \in \tau'$  or  $\tau' \in \tau$  or  $\tau^c \in \tau'$  or  $\tau' \in \tau^c$ .

- Form all possible collections  $T$  of pairwise compatible subsets with  $|T| = L - 1$ .
- We then have that

$$\mathcal{A}_n^{\phi^p} = \sum_T \prod_{\tau \in T} P_\tau. \quad (140)$$

### 6.3 The $\phi^p$ procedure

The question we will now address is whether we can combine the string theory integration rules of section 2 with the algorithm of the previous subsection to find a  $\phi^p$  formula. Specifically, we will investigate whether there exists a function  $H_n^p(z)$  of the form

$$H_n^p(z) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n (z_i - z_j)^{c_{i,j}}, \quad (141)$$

such that the colour-ordered  $\phi^p$  amplitude  $\mathcal{A}_n^{\phi^p}$  can be produced with the following ansatz:

$$\lim_{\alpha' \rightarrow 0} (\alpha')^q K_n^p \int d\mu \Lambda_n(\alpha', k, z) H_n^p(z). \quad (142)$$

Since we wish to calculate partial amplitudes, our expressions must be invariant under cyclic interchanges of the indices of the external legs. Hence the exponents in (141) must satisfy the following for all  $i, j, k$ :

$$c_{i,i+k} = c_{j,j+k}. \quad (143)$$

Invariance to inversion of the ordering of the external legs follows as a direct consequence, as can be seen by choosing  $j = i - k$ :

$$c_{i,i+k} = c_{i-k,i}. \quad (144)$$

Another consequence of the cyclic invariance, is the fact that if we introduce the definition

$$e_j \equiv c_{i,i+j}, \quad (145)$$

then  $H_n^p$  is uniquely characterized by the set  $\{e_j\}$ :

$$H_n^p(z) = \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} (z_i - z_{i+j})^{e_j}. \quad (146)$$

Note that, as a consequence of the inversion invariance, we must have that

$$e_i = e_{n-i}. \quad (147)$$

Hence, only  $\lfloor n/2 \rfloor$  of the members of the set  $\{e_j\}$  are independent.

We will impose on  $H_n^p(z)$  the requirement that the right-hand side of equation (142) is Möbius invariant:

$$\sum_{\substack{j=1 \\ j \neq i}}^n c_{ij} = -2, \quad i = 1, \dots, n, \quad (148)$$

which is equivalent to the following:

$$\sum_{i=1}^{n-1} e_i = -2, \quad i = 1, \dots, n. \quad (149)$$

Now, we recall that in order for  $\mathcal{A}_n^{\phi^p}$  to be non-zero, we must have that  $n = L(p-2) + 2$  with  $L \in \mathbb{N}$ , in which case each Feynman diagram carries  $L-1$  propagators. In order for the dimensions to match,  $q$  in equation (142) must therefore equal  $L-1 = \frac{n-p}{p-2}$ .

If the right-hand side of equation (142) is to yield a finite answer as  $\alpha'$  tends to zero, then the integral must go as  $(\alpha')^{-q}$  in order to cancel the  $(\alpha')^q$  prefactor. This means that the integral must diverge as  $\alpha'$  tends to zero. Because of the form of our ansatz, divergences can only come about when consecutive variables tend to the same value. We consider, then, the case when variables  $z_i$  to  $z_{i+m}$  tend to the same value. The divergence criteria derived in section 2 states that we obtain an  $(\alpha')^{-1}$  divergence when the following condition is met:

$$-m = C = \sum_{i \leq j < l \leq i+m} c_{jl}. \quad (150)$$

By rewriting  $C$  in terms of the  $e_i$ , we obtain an equivalent form of the divergence condition:

$$-m = C = m e_1 + (m-1) e_2 + \dots + 2 e_{m-1} + e_m = \sum_{i=1}^m (m+1-i) e_i.$$

Now, we recall that, in  $\phi^p$  theory, propagators carrying external legs  $i$  to  $i+m$  have the property that  $m$  is of the form  $m = l(p-2)$  with  $l \in \{1, 2, \dots, L-1\}$ . Therefore, if we define

$$\Sigma_m = \sum_{i=1}^m (m+1-i) e_i, \quad (151)$$

then, in order for equation (142) to describe the  $n$ -point amplitude of  $\phi^p$  theory, we must impose on  $H_n^p$  the requirement that for  $m = l(p-2)$  with  $l \in \{1, 2, \dots, L-1\}$ :

$$\Sigma_m = -m, \quad (152)$$

while for  $m = l(p-2) + j$  with  $l \in \{1, 2, \dots, L-1\}$  and  $j \in \{1, 2, \dots, p-3\}$ :

$$\Sigma_m \neq -m. \quad (153)$$

In addition to these equations and inequalities, we should also impose Möbius invariance, that is, equation (149). Incidentally, however, it turns out that, under inversion invariance, Möbius invariance is equivalent to equation (152) with  $m = L(p-2)$ :

$$-L(p-2) = \Sigma_{L(p-2)} = e_{L(p-2)} + 2e_{L(p-2)-1} + \dots + (L(p-2)-1) e_2 + L(p-2) e_1. \quad (154)$$

For since  $n = L(p - 2) + 2$ , we can invert the indices:  $i \rightarrow n - i$  to obtain the following equation:

$$-L(p - 2) = e_2 + 2e_3 + \dots + (L(p - 2) - 1) e_{L(p-2)} + L(p - 2) e_{L(p-2)+1}. \quad (155)$$

Adding together equations (154) and (155) and dividing by  $L(p - 2)$ , we obtain equation (149).

Because of momentum conservation, the propagator carrying the  $m + 1$  external legs indexed by the numbers  $i$  to  $i + m$  is equal to the propagator carrying all the  $n - m - 1$  other external legs. This fact should be reflected in our expression for  $H_n^p$ :  $\Sigma_m$  and  $\Sigma_{n-m-2}$  must be related. And this is indeed the case. For

$$\Sigma_m = e_m + 2e_{m-1} + \dots + m e_1, \quad (156)$$

and

$$\Sigma_{n-m-2} = e_{n-m-2} + 2e_{n-m-3} + \dots + (n - m - 2) e_1. \quad (157)$$

By inversion invariance, equation (157) can also be written as

$$\Sigma_{n-m-2} = e_{m+2} + 2e_{m+3} + \dots + (n - m - 2) e_{n-1} \quad (158)$$

Subtracting equation (158) from (156) yields:

$$\begin{aligned} \Sigma_m - \Sigma_{n-m-2} &= \sum_{i=1}^m (m + 1 - i) e_i + \sum_{i=m+2}^{n-1} (m + 1 - i) e_i = \sum_{i=1}^{n-1} (m + 1 - i) e_i \\ &= \sum_{i=1}^{n-1} (n - 1 - i) e_i - (n - m - 2) \sum_{i=1}^{n-1} e_i. \end{aligned} \quad (159)$$

The first sum on the bottom right-hand side is equal to  $\Sigma_{L(p-2)}$ , which by Möbius invariance (equation (154)) is equal to  $-L(p - 2) = 2 - n$ . The second sum is, again by Möbius invariance (equation (149)), equal to  $-2$ . We conclude that

$$\Sigma_m - \Sigma_{n-m-2} = n - 2m - 2. \quad (160)$$

Hence, the condition  $\Sigma_m = -m$  is equivalent to the condition  $\Sigma_{n-m-2} = -(n - m - 2)$ . In determining  $m$ , we need therefore only impose equalities or inequalities on  $\Sigma_m$  for  $m \in \{1, 2, \dots, \lfloor n/2 \rfloor - 1\}$ , besides Möbius invariance. So we have  $\lfloor n/2 \rfloor$  independent variables  $e_i$ , on which we impose  $\lfloor n/2 \rfloor$  independent conditions that are each either an equation or an inequality. In  $\phi^p$  theory with  $p > 3$  this is an underdetermined system of equations, which leaves us with one degree of freedom<sup>4</sup> for each of the conditions that is an inequality.

But these remaining degrees of freedom must be subjected to yet new conditions. The conditions we have so far imposed ensure that for each term in  $\mathcal{A}_n^{\phi^p}$  there will be a part of the integration domain of the ansatz (142) that blows up in the  $\alpha' \rightarrow 0$  limit such that

---

<sup>4</sup>By degree of freedom, what is really meant here, is a variable that can assume all values except – because of an inequality – one single number.



after integrating out the divergent beta functions, we get exactly the propagators corresponding to the given term of  $\mathcal{A}_n^{\phi^p}$ . But after integrating out the  $q$  divergent beta functions and cancelling the  $(\alpha')^q$  pre-factor, there remain  $n - 3 - q = n - 2 - L = \frac{np - 3n - 2p + 6}{p - 2}$  finite residual integrations that must be carried out for each term. But if our ansatz (142) is to reproduce the  $\phi^p$  tree-amplitude, then the residual integration for each term must evaluate to the same value. That is to say,  $K_T$  in equation (36) must be the same for each  $T$ .

If the remaining degrees of freedom after imposing equations (152) and (153) and Möbius invariance, can be selected such that all  $K_T$  are equal, then we can cancel the values of the residual integration constants with the normalization constant  $C_n^p$  in the ansatz (142) and so obtain a formula for  $\phi^p$ .

To summarize, the procedure for determining an integral representation of the massless, colour-ordered  $\phi^p$   $n$ -point amplitude, where  $n = L(p - 2) + 2$  for some  $L \in \mathbb{N}$ , consists of these steps:

- For  $m = 1, 2, \dots, \lfloor n/2 \rfloor - 1$ , impose that  $\Sigma_m = -m$  when  $m$  is a multiple of  $p - 2$  and impose that  $\Sigma_m \neq -m$  when  $m$  is not a multiple of  $p - 2$ , where  $\Sigma_m$  is as defined in equation (151). Also, impose Möbius invariance in the form of equation (149) or (154), where it is understood that  $e_i = e_{n-i}$ . Solve these equations and inequalities in order to find the permitted values for the set  $\{e_j\}$ .
- Use the remaining degrees of freedom to select  $\{e_j\}$  such that the residual integrations yield the same value throughout.
- Select the value of  $C_n^p$  that cancels the constant owing to the residual integrations in order to normalize the expression.

The amplitude will then be given by equation (142), where  $H_n^p$  is given by equation (146).

### 6.3.1 Conventions

We saw that for each inequality we impose, we obtain one degree of freedom in the allowed values for the set  $\{e_i\}$ . The solution set for the  $e_i$  can therefore be parameterized by a number of parameters that is equal to the number of inequalities. The parameterization that we will adopt consists in defining

$$x_m = \Sigma_m + m - 1 \tag{161}$$

for each  $m$  that is not a multiple of  $p - 2$ . The  $x_m$  can then assume any real value except for minus one. The general solution to the requirements of divergence and Möbius invariance can then be parametrized by parameters  $x_i$  where  $i$  can assume all integer values from 1 to  $\lfloor (n - 2)/2 \rfloor$  for which  $i \not\equiv 0 \pmod{p - 2}$ . The inequalities demand that the  $x_i$  are different from minus one, but otherwise they can assume any values. We will however restrict them to be greater than minus one, so that the residual integrals will not require analytical extension.

For ease of notation it is convenient to also define

$$x_m = \Sigma_m + m - 1$$

for  $m > \lfloor (n-2)/2 \rfloor$ , though these extra  $x_m$  will not be independent from the first ones. Indeed, from equation (151) it follows that

$$x_i = x_{n-2-i}. \quad (162)$$

Another convention is the choice of gauge. When carrying out the residual integrations to determine the normalization constants, we will need to work in a specific gauge. Once again, we make the following choice:

$$z_1 = \infty, \quad z_2 = 1, \quad z_n = 0. \quad (163)$$

## 6.4 $\phi^3$

The divergence requirements:

$$\begin{aligned} -1 &= e_1 \\ -2 &= 2e_1 + e_2 \\ -3 &= 3e_1 + 2e_2 + e_3 \\ &\vdots \\ \lfloor n/2 \rfloor - 1 &= (\lfloor n/2 \rfloor - 1)e_1 + (\lfloor n/2 \rfloor - 2)e_2 + \dots + e_{\lfloor n/2 \rfloor - 1} \end{aligned} \quad (164)$$

Möbius invariance:

$$-2 = \sum_{i=1}^{n-1} e_i \quad (165)$$

The equations have but one solution:

$$e_1 = e_{n-1} = -1 \quad \text{and} \quad e_i = 0 \quad \text{for} \quad 2 \leq i \leq n-2. \quad (166)$$

The beta function integrations satiate all the variables so that no residual integrations remain, and so  $K_n^3$  is always equal to unity. Thus we recover the Koba-Nielsen formula (1).

## 6.5 $\phi^4$

### 6.5.1 Six-point amplitude

The divergence requirements:

$$\begin{aligned} -1 &\neq e_1 \\ -2 &= 2e_1 + e_2 \end{aligned} \quad (167)$$

Möbius invariance:

$$-2 = e_1 + e_2 + e_3 + e_4 + e_5 = 2e_1 + 2e_2 + e_3 \quad (168)$$

The solutions to the conditions can be written as follows:

$$\begin{aligned} e_1 &= x_1 \\ e_2 &= -2x_1 - 2 \\ e_3 &= 2x_1 + 2. \end{aligned} \quad (169)$$

We have, then, in gauge-fixed form, the following ansatz for  $\mathcal{A}_6^{\phi^4}$ :

$$\lim_{\alpha' \rightarrow 0} \alpha' K_6^4 \int_0^1 dz_3 \int_0^{z_3} dz_4 \int_0^{z_4} dz_5 \Lambda_6 \frac{(1-z_3)^{x_1} (z_3-z_4)^{x_1} (z_4-z_5)^{x_1} z_5^{x_1} (1-z_5)^{2x_1+2} z_3^{2x_1+2}}{(1-z_4)^{2x_1+2} (z_3-z_5)^{2x_1+2} z_4^{2x_1+2}}. \quad (170)$$

The integral diverges in three different integration regions as  $\alpha'$  tends to zero: 1) the region where  $z_3$  and  $z_4$  tend to one, 2) the region where  $z_3$ ,  $z_4$ , and  $z_5$  tend to the same value, and 3) the region where  $z_4$  and  $z_5$  tend to zero. In non-gauge-fixed form the ansatz is by construction invariant with respect to a cyclic permutation of the indices, and for this reason the residual integration must compute to the same value in each of these three regions. We will consider the third region and introduce new variables  $y_4$  and  $y_5$  as follows:

$$\begin{aligned} z_4 &= z_3 y_4 \\ z_5 &= z_3 y_4 y_5. \end{aligned} \quad (171)$$

In terms of the new variables we can write the ansatz as follows:

$$\lim_{\alpha' \rightarrow 0} \alpha' K_6^4 \int_0^1 dz_3 \int_0^1 dy_4 \int_0^1 dy_5 \Lambda_6 \left[ \frac{z_3^{x_1} (1-z_3)^{x_1} y_5^{x_1} (1-y_5)^{x_1}}{y_4} + \mathcal{O}((y_4)^0) \right]. \quad (172)$$

From this expression we read off that the normalization is given in terms of beta functions as follows:

$$K_6^4(x) = (\text{B}[1+x_1, 1+x_1])^{-2}. \quad (173)$$

For explicitness, let us consider the case when  $x_1 = 0$ . In that case the normalization constant becomes unity, and we obtain the following expression for the six-point amplitude:

$$\lim_{\alpha' \rightarrow 0} \alpha' \int d\mu_6 \Lambda_6 \frac{(z_1-z_4)^2 (z_2-z_5)^2 (z_3-z_6)^2}{(z_1-z_3)^2 (z_1-z_5)^2 (z_2-z_4)^2 (z_2-z_6)^2 (z_3-z_5)^2 (z_4-z_6)^2}. \quad (174)$$

### 6.5.2 Eight-point amplitude

The divergence requirements:

$$\begin{aligned} -1 &\neq e_1 \\ -2 &= 2e_1 + e_2 \\ -3 &\neq 3e_1 + 2e_2 + e_3 \end{aligned} \quad (175)$$

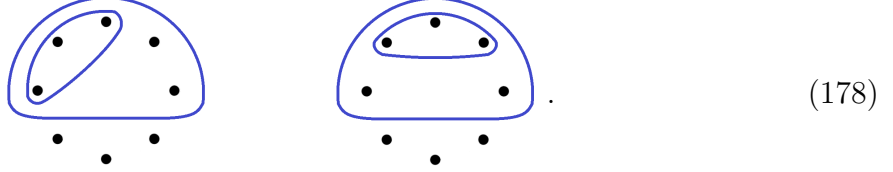
Möbius invariance:

$$-2 = 2e_1 + 2e_2 + 2e_3 + e_4 \quad (176)$$

The solutions to the conditions can be written as follows:

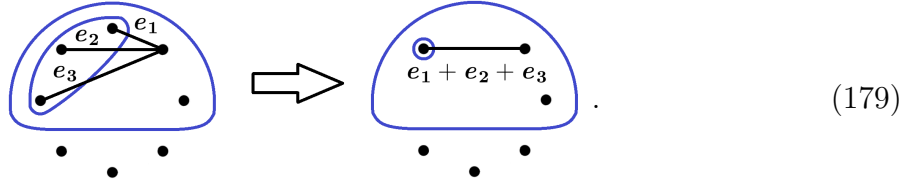
$$\begin{aligned} e_1 &= x_1 \\ e_2 &= -2x_1 - 2 \\ e_3 &= 2 + x_3 + x_1 \\ e_4 &= -2 - 2x_3. \end{aligned} \quad (177)$$

For the given values the  $e_i$ , there will be six diverging regions of the integration domain that produce a term in the  $\alpha' \rightarrow 0$  limit. While the invariance with respect to inversion and cyclic permutations of the indices render most of the residual integrations identical, there are two different types of residual integrations that are not necessarily the same because there are two distinct ways for variables to tend to the same value to produce two compatible divergences that can cancel the  $(\alpha')^2$  pre-factor. These two different types of compatible divergences can be illustrated by what we will call nesting diagrams:



Each point represents a variable  $z_i$ , and each closed curve represents an  $(\alpha')^{-1}$  divergence that comes about because the enclosed points tend to the same value.

When variables tend to the same value, their points in the nesting diagram merge together to one point, and the connections  $e_i$  to this point is equal to the sum of the connections to the individual points that merge together:



Based on this observation we note that the two different types of residual integrals will give the same if the following condition is met:

$$e_1 = e_1 + e_2 + e_3. \quad (180)$$

Given the solutions for the  $e_i$  in equations (177), equation (180) is equivalent to

$$x_1 = x_3. \quad (181)$$

When this condition is satisfied, all residual integrations evaluate to the same value, which can be cancelled by the overall normalization constant:

$$K_8^4(x_1) = (\text{B}[1 + x_1, 1 + x_1])^{-3}. \quad (182)$$

If we set all the  $x_i$  equal to zero, we get the following expression for  $\mathcal{A}_8^{\phi^4}$ :

$$\lim_{\alpha' \rightarrow 0} (\alpha')^2 \int d\mu_8 \Lambda_8 \frac{\prod_{i=1}^5 (z_i - z_{i+3})^2 \prod_{i=1}^3 (z_i - z_{i+5})^2}{\prod_{i=1}^6 (z_i - z_{i+2})^2 \prod_{i=1}^4 (z_i - z_{i+4})^2 \prod_{i=1}^2 (z_i - z_{i+6})^2}. \quad (183)$$

### 6.5.3 $n$ -point amplitude

The divergence and Möbius invariance conditions can be framed thus:

$$\Sigma_i = \begin{cases} x_i - i + 1 & \text{for } i \text{ odd} \\ -i & \text{for } i \text{ even} \end{cases} \quad (184)$$

It follows that

$$\sum_{j=1}^i e_j = \Sigma_i - \Sigma_{i-1} = \begin{cases} x_i & \text{for } i \text{ odd,} \\ -x_{i-1} - 2 & \text{for } i \text{ even.} \end{cases} \quad (185)$$

From these equations we see that we can describe the solutions to the exponents  $e_i$  recursively as follows:

$$e_i = \begin{cases} x_i - \sum_{j=1}^{i-1} e_j & \text{for odd } i, \\ -2 - x_{i-1} - \sum_{j=1}^{i-1} e_j & \text{for even } i. \end{cases} \quad (186)$$

These recursive relations have the following solution:

$$e_i = \begin{cases} x_1, & i = 1 \\ -2 - 2x_{i-1}, & i \text{ even} \\ 2 + x_i + x_{i-2}, & i \text{ even and } i > 1. \end{cases} \quad (187)$$

This solution applies only to the first  $n/2$  exponents  $e_i$ . The remaining exponents can be found from the relation  $e_{n-i} = e_i$ .

The residual integrations will give the same if the following is satisfied:

$$e_1 = \sum_{i=1}^3 e_i = \sum_{i=1}^5 e_i = \sum_{i=1}^7 e_i = \dots \quad (188)$$

In terms of the parameters  $x_i$ , the conditions (188) state that they must all be equal:

$$x_1 = x_3 = x_5 = x_7 = \dots \equiv x \quad (189)$$

The normalization constant becomes the following:

$$K_n^4(x) = (\text{B}[1+x, 1+x])^{-\frac{n-2}{2}}. \quad (190)$$

We can now write down an expression for the non-zero quartic scalar amplitudes:

$$\mathcal{A}_n^{\phi^4} = \lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-4}{2}} K_n^4(x) \int d\mu_n \Lambda_n \frac{\prod_{i=1}^{n-1} (z_i - z_{i+1})^x \prod_{j=1}^{\frac{n-4}{2}} \prod_{i=1}^{n-2j-1} (z_i - z_{i+2j+1})^{2+2x}}{\prod_{j=1}^{\frac{n-2}{2}} \prod_{i=1}^{n-2j} (z_i - z_{i+2j})^{2+2x}}. \quad (191)$$

One particularly simple solution is obtained by setting  $x = 0$ :

$$\mathcal{A}_n^{\phi^4} = \lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-4}{2}} \int d\mu_n \Lambda_n \frac{\prod_{j=1}^{\frac{n-4}{2}} \prod_{i=1}^{n-2j-1} (z_i - z_{i+2j+1})^2}{\prod_{j=1}^{\frac{n-2}{2}} \prod_{i=1}^{n-2j} (z_i - z_{i+2j})^2}. \quad (192)$$

## 6.6 $\phi^5$

### 6.6.1 Eight-point amplitude

The divergence requirements:

$$\begin{aligned}
-1 &\neq e_1 \\
-2 &\neq 2e_1 + e_2 \\
-3 &= 3e_1 + 2e_2 + e_3
\end{aligned} \tag{193}$$

Möbius invariance:

$$-2 = 2e_1 + 2e_2 + 2e_3 + e_4 \tag{194}$$

The solutions to the conditions can be parametrized as follows:

$$\begin{aligned}
e_1 &= x_1 \\
e_2 &= -1 - 2x_1 + x_2 \\
e_3 &= -1 + x_1 - 2x_2 \\
e_4 &= 2 + 2x_2.
\end{aligned} \tag{195}$$

There is only one way a  $(\alpha')^{-1}$  divergence can come about: four variables tending to the same value. The invariance of the ansatz with respect to inversion and cyclic permutations of the indices therefore ensures that all the residual integrations yield the same value. To determine this value we can consider the case where variables  $z_5, z_6, z_7 \rightarrow 0$  in the standard gauge. We introduce new variables in the usual fashion:

$$\begin{aligned}
z_4 &= z_3 y_4 \\
z_5 &= z_3 y_4 y_5 \\
z_6 &= z_3 y_4 y_5 y_6 \\
z_7 &= z_3 y_4 y_5 y_6 y_7
\end{aligned} \tag{196}$$

In terms of the new variables we can write our ansatz for  $\mathcal{A}_8^{\phi^5}$  as follows:

$$\begin{aligned}
&\lim_{\alpha' \rightarrow 0} (\alpha') K_8^5 \int_0^1 dz_3 \int_0^1 dy_4 \int_0^1 dy_5 \int_0^1 dy_6 \int_0^1 dy_7 \Lambda'_8 \times \\
&\left[ \frac{1}{y_5} z_3^{x_1} y_4^{x_2} (1 - z_3)^{x_1} (1 - y_4)^{x_1} (1 - z_3 y_4)^{-1-2x_1+x_2} \times \right. \\
&\quad \left. y_6^{x_2} y_7^{x_1} (1 - y_6)^{x_1} (1 - y_7)^{x_1} (1 - y_6 y_7)^{-1-2x_1+x_2} + \mathcal{O}((y_5)^0) \right].
\end{aligned} \tag{197}$$

From this expression we see that the normalization constant is given as follows,

$$K_8^5 = \left( C[x_1, x_2, x_1, x_1, -1 - 2x_1 + x_2] \right)^{-2} \tag{198}$$

where we have introduced what we can call the Stieberger function:

$$C[a, b, c, d, e] = \int_0^1 dx \int_0^1 dy x^a y^b (1 - x)^c (1 - y)^d (1 - xy)^e. \tag{199}$$

The Stieberger function itself can be rewritten in terms of beta functions and a hypergeometric function:<sup>[24]</sup>

$$C[a, b, c, d, e] = B[1 + a, 1 + c] B[1 + b, 1 + d] {}_3F_2 \left[ \begin{matrix} 1 + a, 1 + b, -e \\ 2 + a + c, 2 + b + d \end{matrix}; 1 \right]. \quad (200)$$

Another way of writing the normalization constant is the following

$$K_8^5 = \left( \int d\mu_4 \prod_{j=1}^3 \prod_{i=1}^{4-j} (z_i - z_{i+j})^{e_j} \right)^{-2}. \quad (201)$$

We note that

$$C[0, 0, 0, 0, -1] = \zeta(2) = \frac{\pi^2}{6}. \quad (202)$$

Therefore, if for explicitness we consider the case when the  $x_i$  are all zero, then we find that  $\mathcal{A}_8^{\phi^5}$  is given by

$$\lim_{\alpha' \rightarrow 0} \alpha' \frac{36}{\pi^4} \int d\mu_8 \frac{\prod_{i=1}^4 (z_i - z_{i+4})^2}{\prod_{i=1}^6 (z_i - z_{i+2}) \prod_{i=1}^5 (z_i - z_{i+3}) \prod_{i=1}^3 (z_i - z_{i+5}) \prod_{i=1}^2 (z_i - z_{i+6})}. \quad (203)$$

### 6.6.2 Eleven-point amplitude

The divergence requirements:

$$\begin{aligned} -1 &\neq e_1 \\ -2 &\neq 2e_1 + e_2 \\ -3 &= 3e_1 + 2e_2 + e_3 \\ -4 &\neq 4e_1 + 3e_2 + 2e_3 + e_4 \end{aligned} \quad (204)$$

Möbius invariance:

$$-2 = 2e_1 + 2e_2 + 2e_3 + 2e_4 + 2e_5 \quad (205)$$

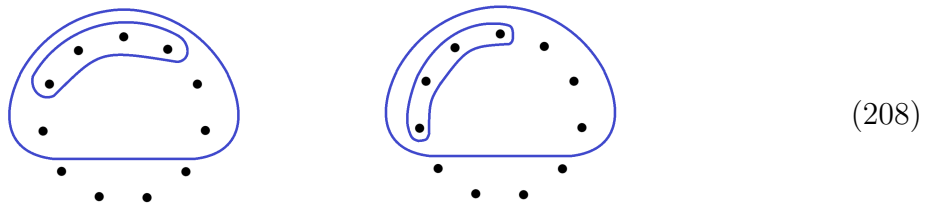
The solutions to the conditions can be parametrized as follows:

$$\begin{aligned} e_1 &= x_1 \\ e_2 &= -1 - 2x_1 + x_2 \\ e_3 &= -1 + x_1 - 2x_2 \end{aligned} \quad (206)$$

$$e_4 = 2 + x_2 + x_4 \quad (207)$$

$$e_5 = -1 - x_4.$$

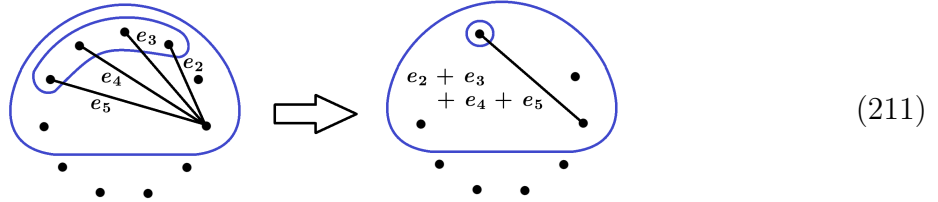
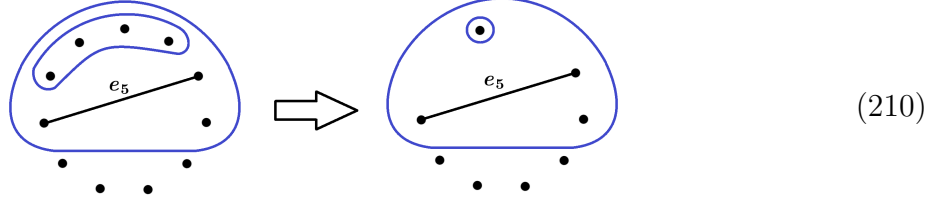
In this case there are two distinct ways two compatible  $(\alpha')^{-1}$  divergences can come about:



The residual integrations will yield the same values if the following conditions are met:

$$\begin{aligned} e_1 &= e_1 + e_2 + e_3 + e_4, \\ e_2 &= e_2 + e_3 + e_4 + e_5, \\ e_2 &= e_5. \end{aligned} \tag{209}$$

These conditions can be understood from the following diagrams:



Given our parametrization of the  $e_i$ , the conditions (209) are equivalent to the following:

$$x_1 = x_2 = x_4. \tag{212}$$

There is one degree of freedom left because the conditions (209) are not independent. The normalization becomes the following:

$$K_{11}^5 = (\mathbb{C}[x_1, x_1, x_1, x_1, -1 - x_1])^{-3}. \tag{213}$$

### 6.6.3 $n$ -point amplitude

The divergence and Möbius invariance conditions can in this case be framed as follows:

$$\Sigma_i = \begin{cases} x_i - i + 1 & \text{for } i \equiv 1 \text{ or } i \equiv 2 \pmod{3}, \\ -i & \text{for } i \equiv 0 \pmod{3}. \end{cases} \tag{214}$$

From these equations it follows that we can describe the solutions to the exponents  $e_i$  recursively as follows:

$$e_i = \Sigma_i - \Sigma_{i-1} - \sum_{j=1}^{i-1} e_j = \begin{cases} x_i - \sum_{j=1}^{i-1} e_j & \text{for } i \equiv 1 \pmod{3}, \\ x_i - x_{i-1} - 1 - \sum_{j=1}^{i-1} e_j & \text{for } i \equiv 2 \pmod{3}, \\ -x_{i-1} - 2 - \sum_{j=1}^{i-1} e_j & \text{for } i \equiv 0 \pmod{3}. \end{cases} \tag{215}$$



These recursive relations have the following solutions:

$$e_i = \begin{cases} x_1, & i = 1, \\ x_i - 2x_{i-1} - 1, & i \equiv 2 \pmod{3}, \\ -2x_{i-1} + x_{i-2} - 1, & i \equiv 0 \pmod{3}, \\ x_i + x_{i-2} + 2, & i \equiv 1 \pmod{3} \text{ and } i > 1. \end{cases} \quad (216)$$

This solution applies only to the first  $\lfloor n/2 \rfloor$  exponents  $e_i$ . The remaining exponents can be found from the relation  $e_{n-i} = e_i$ .

The residual integrations will agree if the following is satisfied:

$$\begin{aligned} e_1 &= \sum_{i=1}^4 e_i = \sum_{i=1}^7 e_i = \sum_{i=1}^{10} e_i = \dots \\ e_2 &= \sum_{i=2}^5 e_i = \sum_{i=2}^8 e_i = \sum_{i=2}^{11} e_i = \dots \\ e_2 &= e_5 = e_8 = e_{11} = \dots \end{aligned} \quad (217)$$

And these conditions are satisfied if all the  $x_i$  are equal:

$$x_1 = x_2 = x_4 = x_5 = x_7 = \dots \equiv x. \quad (218)$$

In that case the full set of exponents  $e_i$  can be described thus:

$$e_i = \begin{cases} x, & i = 1 \text{ or } i = n - 1, \\ -1 - x, & i \equiv 2 \text{ or } i \equiv 0 \pmod{3}, \\ 2 + 2x, & 1 < i < n - 1 \text{ and } i \equiv 1 \pmod{3}. \end{cases} \quad (219)$$

The normalization constant becomes the following:

$$K_n^5(x) = (C[x, x, x, x, -1 - x])^{\frac{n-2}{3}} \quad (220)$$

We arrive at the following expression for the non-zero quintic scalar amplitudes:

$$\mathcal{A}_n^{\phi^5} = \lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-5}{3}} K_n^5(x) \int d\mu_n \Lambda_n \frac{\prod_{i=1}^{n-1} (z_i - z_{i+1})^x \prod_{j=1}^{\frac{n-5}{3}} \prod_{i=1}^{n-3j-1} (z_i - z_{i+3j+1})^{2+2x}}{\prod_{j=1}^{\frac{n-2}{3}} \prod_{i=1}^{n-3j} (z_i - z_{i+3j})^{1+x} \prod_{j=0}^{\frac{n-5}{3}} \prod_{i=1}^{n-3j-2} (z_i - z_{i+3j+2})^{1+x}}. \quad (221)$$

As in the  $\phi^4$  case, the simplest solution is obtained if one sets  $x$  to zero:

$$\mathcal{A}_n^{\phi^5} = \lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-5}{3}} \left( \frac{6}{\pi^2} \right)^{\frac{n-2}{3}} \int d\mu_n \Lambda_n \frac{\prod_{j=1}^{\frac{n-5}{3}} \prod_{i=1}^{n-3j-1} (z_i - z_{i+3j+1})^2}{\prod_{j=1}^{\frac{n-2}{3}} \prod_{i=1}^{n-3j} (z_i - z_{i+3j}) \prod_{j=0}^{\frac{n-5}{3}} \prod_{i=1}^{n-3j-2} (z_i - z_{i+3j+2})}. \quad (222)$$

## 6.7 $\phi^p$ for $p > 4$

The divergence and Möbius invariance conditions can in this case be framed as follows:

$$\Sigma_i = \begin{cases} x_i - i + 1 & \text{for } i \equiv 1, 2, \dots, p-3 \pmod{p-2}, \\ -i & \text{for } i \equiv 0 \pmod{p-2}. \end{cases} \quad (223)$$

From these equations it follows that we can describe the solutions to the exponents  $e_i$  recursively as follows:

$$e_i = \Sigma_i - \Sigma_{i-1} - \sum_{j=1}^{i-1} e_j = \begin{cases} x_i - \sum_{j=1}^{i-1} e_j & \text{for } i \equiv 1 \pmod{p-2}, \\ x_i - x_{i-1} - 1 - \sum_{j=1}^{i-1} e_j & \text{for } i \equiv 2, 3, \dots, p-3 \pmod{p-2}, \\ -x_{i-1} - 2 - \sum_{j=1}^{i-1} e_j & \text{for } i \equiv 0 \pmod{p-2}. \end{cases} \quad (224)$$

These recursive relations have the following solutions:

$$e_i = \begin{cases} x_1, & i = 1 \\ x_i - 2x_{i-1} - 1, & i \equiv 2 \pmod{p-2}, \\ x_i - 2x_{i-1} + x_{i-2}, & i \equiv 3, 4, \dots, p-3 \pmod{p-2}, \\ -2x_{i-1} + x_{i-2} - 1, & i \equiv 0 \pmod{p-2}, \\ x_i + x_{i-2} + 2, & i \equiv 1 \pmod{p-2} \text{ and } i > 1. \end{cases} \quad (225)$$

This solution applies only to the first  $\lfloor n/2 \rfloor$  exponents  $e_i$ . The remaining exponents can be found from the relation  $e_{n-i} = e_i$ .

The residual integrations will agree if the following is satisfied:

$$\begin{aligned} e_1 &= \sum_{i=1}^{p-1} e_i = \sum_{i=1}^{2p-3} e_i = \sum_{i=1}^{3p-5} e_i = \dots \\ e_2 &= \sum_{i=2}^p e_i = \sum_{i=2}^{2p-2} e_i = \sum_{i=2}^{3p-4} e_i = \dots \\ &\vdots \\ e_{p-3} &= \sum_{i=p-3}^{2p-5} e_i = \sum_{i=p-3}^{3p-7} e_i = \sum_{i=p-3}^{4p-9} e_i = \dots \\ e_2 &= e_p = e_{2p-2} = e_{3p-4} = \dots \\ e_3 &= e_{p+1} = e_{2p-1} = e_{3p-3} = \dots \\ &\vdots \\ e_{p-3} &= e_{2p-5} = e_{3p-7} = e_{4p-9} = \dots \end{aligned} \quad (226)$$

And these conditions are satisfied if all the  $x_i$  are equal:

$$x_1 = x_2 = x_4 = x_5 = x_7 = \dots \equiv x. \quad (227)$$

When this is the case, the exponents  $e_i$  for which  $i$  is equivalent to zero or two mod  $p-2$  become identical. Incidentally, when  $i$  is equivalent to zero mod  $p-2$ , the reversed index  $n-i$  is equivalent to two, and vice versa. The other groups of exponents  $e_i$  likewise transform into themselves under reversion of the indices. The full set of exponents  $e_i$  can therefore be described thus:

$$e_i = \begin{cases} x, & i = 1 \text{ or } i = n-1, \\ -1-x, & i \equiv 2 \text{ or } i \equiv 0 \pmod{p-2}, \\ 2+2x, & 1 < i < n-1 \text{ and } i \equiv 1 \pmod{3}, \\ 0, & i \equiv 3, 4, \dots, p-3 \pmod{p-2}. \end{cases} \quad (228)$$

The normalization constant is given by

$$K_n^p(x) = \left( \int d\mu_{p-1} \prod_{j=1}^{p-2} \prod_{i=1}^{p-1-j} (z_i - z_{i+j})^{e_j} \right)^{-\frac{n-2}{p-2}}. \quad (229)$$

We arrive at the following closed-form expression for the colour-ordered  $\phi^p$  amplitude:

$$\lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-p}{p-2}} K_n^p(x) \int d\mu_n \Lambda_n \frac{\prod_{i=1}^{n-1} (z_i - z_{i+1})^x \prod_{j=1}^{\frac{n-2}{p-2}-1} \prod_{i=1}^{n-j(p-2)-1} (z_i - z_{i+j(p-2)+1})^{2+2x}}{\prod_{j=1}^{\frac{n-2}{p-2}} \prod_{i=1}^{n-j(p-2)} (z_i - z_{i+j(p-2)})^{1+x} \prod_{j=0}^{\frac{n-2}{p-2}-1} \prod_{i=1}^{n-j(p-2)-2} (z_i - z_{i+j(p-2)+2})^{1+x}}, \quad (230)$$

If  $x$  is set to zero, the identity simplifies too:

$$\mathcal{A}_n^{\phi^p} = \lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-p}{p-2}} K_n^p(0) \int d\mu_n \Lambda_n \frac{\prod_{j=1}^{\frac{n-2}{p-2}-1} \prod_{i=1}^{n-j(p-2)-1} (z_i - z_{i+j(p-2)+1})^2}{\prod_{j=1}^{\frac{n-2}{p-2}} \prod_{i=1}^{n-j(p-2)} (z_i - z_{i+j(p-2)}) \prod_{j=0}^{\frac{n-2}{p-2}-1} \prod_{i=1}^{n-j(p-2)-2} (z_i - z_{i+j(p-2)+2})}, \quad (231)$$

with the normalization constant given by

$$K_n^p(0) = \left( \prod_{i=1}^{p-3} \left( \int_0^1 dz_i \right) \prod_{i=1}^{p-4} \frac{1}{1 - z_i z_{i+1}} \right)^{-\frac{n-2}{p-2}}. \quad (232)$$

## 6.8 Discussion

In order to characterize the distinct ways that compatible  $(\alpha')^{-1}$  divergences can come about, we found it useful to introduce nesting diagrams. These diagrams are in a one-to-one correspondence with the polygon diagrams of section 5. To illustrate this fact, the first few types of nesting diagrams are tabulated in figure 4, which is the analogue of the bottom half of figure 3.

$3p - 4$			
$4p - 6$			

**Figure 4:** The first few pure polygon cycle structures

A deficiency of the nesting diagrams that one should be mindful of is that, owing to the equivalence of complementary subsets of points, some equivalent diagrams may appear distinct. For example, both diagrams

(233)

correspond to the polygon diagram

(234)

Now, having found a compact expression for the  $\phi^p$  amplitude in terms of a single string theory integrand, an interesting question presents itself: What is the physical significance of this result?

Truth be told, I do not know the answer to this question. But we can make a few observations. Firstly, the  $\phi^p$  expression (230) contains no poles with such a strong divergence that analytical extension is required, and the same was the case with the brute-force  $\phi^p$  procedure of section 5. In string theory, such poles are usually associated with tachyons, so that we may tentatively posit that all scalar amplitudes can be extracted from a tachyon-free theory.

Being tachyon-free is a characteristic of the super string, in contradistinction to the bosonic string. While a direct application of the super string formula for the vector boson amplitude would require analytical extension, it is always possible to remove such tachyonic poles via integration by parts, as discussed in [15]. But we have already seen that the  $\phi^4$  amplitude can be extracted from the bosonic string, in the form of the string theory version of the CHY  $\phi^4$  formula (130):

$$\mathcal{A}_n^{\phi^4} = \lim_{\alpha' \rightarrow 0} (\alpha')^{\frac{n-4}{2}} \int d\mu_n \Lambda_n \sum_{\substack{\text{connected perfect} \\ \text{matchings}}} \frac{1}{[PM(z)]^2}. \quad (235)$$

To probe if the  $\phi^p$  formulas (191), (221), and (230) are bosonic in nature, or whether they bear a closer relation to the super string, we can attempt to expand them beyond leading order in  $\alpha'$  since it is a distinctive trait of the super string amplitudes that the terms linear  $\alpha'$  vanish.

To start with the simplest example, we can consider the  $n = 4$  case of the  $\phi^4$  formula (191), which in gauge-fixed form looks thus:

$$\frac{1}{\text{B}[1+x, 1+x]} \int_0^1 dz (1-z)^{x+\alpha' s_{23}} z^{x+\alpha' s_{34}} = 1 + (s_{23} + s_{34}) \left( \psi(1+x) - \psi(2+2x) \right) \alpha' + \mathcal{O}(\alpha'^2), \quad (236)$$

where  $\psi(x)$  is the digamma function.

Since  $\psi(1+x) - \psi(2+2x)$  ranges from  $-\infty$  to  $-\log(2)$  for  $x \in ]-1, \infty[$ , the term linear in  $\alpha'$  never vanishes, indicating that formula (191) is not related to the super string. But if we set  $x = 0$ , we recover the  $n = 4$  version of the bosonic equation (235) since there is only one connected perfect matching at four points.

Let us proceed unto six points and examine (174), which takes on the following form when we gauge-fix:

$$\alpha' \int_0^1 dz_3 \int_0^{z_3} dz_4 \int_0^{z_4} dz_5 \Lambda'_6 \frac{(1-z_5)^2 z_3^2}{(1-z_4)^2 (z_3-z_5)^2 z_4^2}. \quad (237)$$

Unfortunately, the only method I know of for expanding such integrals beyond leading order is rather involved and would be infeasible if we retained a free parameter  $x_1$ , which is why we now only consider the  $x_1 = 0$  case. The method is explained with an example in appendix B, and applying it to the present integral, we obtain the following expansion:

$$\begin{aligned} \frac{1}{s_{234}} + \frac{1}{s_{345}} + \frac{1}{s_{467}} + \left( \frac{37}{12} + \frac{\pi^2}{9} - \frac{s_{23} + s_{24} + s_{25} + 3s_{34} + 2s_{35} + s_{36} + 3s_{45} + s_{46} + s_{56}}{s_{345}} \right. \\ \left. - \frac{s_{23} + s_{34} + s_{35} + s_{36} + 2s_{45} + s_{46} + 2s_{56}}{s_{456}} \right. \\ \left. - \frac{2s_{23} + s_{24} + s_{25} + 2s_{34} + s_{35} + s_{45} + s_{56}}{s_{234}} \right) \alpha' + \mathcal{O}(\alpha'^2). \end{aligned} \quad (238)$$

Unsurprisingly, we find that the linear  $\alpha'$  term is not zero. Though it is not evident in their present form, the three linear  $\alpha'$  terms that depend on the external momenta are

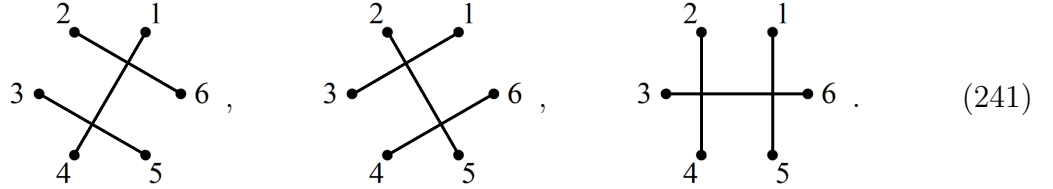
related to each other by cyclic permutations of the indices. This is related to the fact that there are three connected perfect matchings that are related to each other by exactly such cyclic permutations of the indices. For if we expand the  $n = 6$  version of the bosonic  $\phi^4$  formula (235), which in gauge-fixed form looks thus:

$$\alpha' \int_0^1 dz_3 \int_0^{z_3} dz_4 \int_0^{z_4} dz_5 \Lambda'_6 \left( \frac{1}{(1-z_4)^2 z_3^2} + \frac{1}{(z_3-z_5)^2} + \frac{1}{(1-z_5)^2 z_4^2} + \frac{1}{(1-z_5)^2 z_3^2} \right), \quad (239)$$

then we find the following expansion:

$$\begin{aligned} \frac{1}{s_{234}} + \frac{1}{s_{345}} + \frac{1}{s_{467}} + \left( 4 - \frac{s_{23} + s_{24} + s_{25} + 3s_{34} + 2s_{35} + s_{36} + 3s_{45} + s_{46} + s_{56}}{s_{345}} \right. \\ \left. - \frac{s_{23} + s_{34} + s_{35} + s_{36} + 2s_{45} + s_{46} + 2s_{56}}{s_{456}} \right. \\ \left. - \frac{2s_{23} + s_{24} + s_{25} + 2s_{34} + s_{35} + s_{45} + s_{56}}{s_{234}} \right) \alpha' + \mathcal{O}(\alpha'^2), \end{aligned} \quad (240)$$

where each of the last three terms arises from one of the following perfect matchings:



$$\begin{array}{ccc} \begin{array}{c} 2 \quad 1 \\ \diagdown \quad / \\ 3 \quad 6 \\ / \quad \diagdown \\ 4 \quad 5 \end{array}, & \begin{array}{c} 2 \quad 1 \\ / \quad \diagdown \\ 3 \quad 6 \\ \diagdown \quad / \\ 4 \quad 5 \end{array}, & \begin{array}{c} 2 \quad 1 \\ | \quad | \\ 3 \quad 6 \\ | \quad | \\ 4 \quad 5 \end{array}. \end{array} \quad (241)$$

The last perfect matching



$$\begin{array}{c} 2 \quad 1 \\ / \quad \diagdown \\ 3 \quad 6 \\ \diagdown \quad / \\ 4 \quad 5 \end{array} \quad (242)$$

contributes only a linear  $\alpha'$  term that does not depend on the momenta. It is only by a term of that kind that the two expansion (239) and (241) disagree.

It appears, then, that there might be some connection between the bosonic string and the  $\phi^p$  formula derived in this section. In any event there is an uncharted region to explore here.

## 7 Möbius invariance and massive particles

It is a property common to the massless string theory integrands and CHY integrands that they are invariant with respect to Möbius transformations:

$$z_a \rightarrow \frac{Az_a + B}{Cz_a + D}, \quad \text{with } AD - BC = 1, \quad \text{for } a = 1, \dots, n. \quad (243)$$

The transformation has four parameters of which three are independent. Invariance with respect to this type of transformation is what ensures that we can gauge-fix any three

of the variables  $z_i$ . In deriving the integration rules in subsection 2.2 we saw how, concretely, that this invariance ensures the rule of complementarity: that complementary subsets of  $\mathbb{Z}_n$  are equivalent. If, therefore, we wish to extend the integration rules and the  $\phi^p$  formula to massive particles, then it is absolutely necessary that Möbius invariance is preserved.

Under Möbius transformations, differences transform as follows:

$$\begin{aligned} z_a - z_b &\rightarrow \frac{Az_a + B}{Cz_a + D} - \frac{Az_b + B}{Cz_b + D} = \frac{(Az_a + B)(Cz_b + D) - (Az_b + B)(Cz_a + D)}{(Cz_a + D)(Cz_b + D)} \\ &= \frac{(AD - BC)(z_a - z_b)}{(Cz_a + D)(Cz_b + D)} = \frac{z_a - z_b}{(Cz_a + D)(Cz_b + D)}. \end{aligned} \quad (244)$$

The rational functions  $S_i$  that appear in the scattering equations transform as follows:

$$\begin{aligned} S_i &= \sum_{j \neq i} \frac{k_i \cdot k_j}{z_i - z_j} \rightarrow (Cz_i + D) \sum_{j \neq i} \frac{k_i \cdot k_j}{z_i - z_j} (Cz_j + D) = \\ &= (Cz_i + D) \left( DS_i + C \sum_{j \neq i} \frac{k_i \cdot k_j}{z_i - z_j} z_j \right) = \\ &= (Cz_i + D) \left( DS_i - C \sum_{j \neq i} k_i \cdot k_j + Cz_i S_i \right). \end{aligned} \quad (245)$$

We see that when the external particles are massless, so that

$$\sum_{j \neq i} k_i \cdot k_j = \frac{1}{2} \left( \sum_{j \neq i} k_j \right)^2 = \frac{1}{2} k_i^2 = 0, \quad (246)$$

then the scattering functions transform as

$$S_i \rightarrow (Cz_i + D)^2 S_i, \quad (247)$$

so that the scattering equations  $S_i = 0$  are Möbius invariant. But when the particles are massive,

$$\sum_{j \neq i} k_i \cdot k_j \neq 0, \quad (248)$$

and the scattering equations are not Möbius invariant. If, however, as observed by Naculich<sup>[9]</sup>, the scattering equations are modified thus

$$S_i = \sum_{j \neq i} \frac{k_i \cdot k_j}{z_i - z_j} \rightarrow \sum_{j \neq i} \frac{k_i \cdot k_j + \Delta_{ij}}{z_i - z_j}, \quad (249)$$

where the extra parameters  $\Delta_{ij}$  are subjected to the following conditions:

$$\Delta_{ij} = \Delta_{ji}, \quad \sum_{j \neq i} \Delta_{ij} = k_i^2, \quad (250)$$

then the scattering equations remain invariant, since the  $-C \sum_{j \neq i} k_i \cdot k_j = C k_i^2$  term in equation (245) is cancelled by the new  $\Delta_{ij}$  term.

And the conditions (250) also ensure that, when performing the substitution (249), the following linear combinations of the rational functions  $S_i$  vanish, just as in the massless case:

$$\sum_{i=1}^n S_i, \quad \sum_{i=1}^n z_i S_i, \quad \sum_{i=1}^n z_i^2 S_i. \quad (251)$$

It is therefore possible to extend the CHY formalism to massive particles if one can find a set of allowed values for  $\Delta_{ij}$  such that the propagators become massive. But setting

$$\Delta_{ij} = \begin{cases} \frac{1}{2}m^2, & j = i \pm 1 \\ 0, & \text{otherwise} \end{cases}, \quad (252)$$

where  $m$  is the mass of each particle, will achieve exactly that. For with the following replacement, first observed by Dolan and Goddard to fix the massive scattering equations<sup>[8]</sup>,

$$k_i \cdot k_{i+1} \rightarrow k_i \cdot k_{i+1} + \frac{m^2}{2}, \quad (253)$$

each propagator carrying legs  $a, a+1, \dots, a+b$  is altered thus:

$$\begin{aligned} \frac{1}{2 \sum_{i=a}^{a+b-1} \sum_{j=a+1}^{a+b} k_i \cdot k_j} &\rightarrow \frac{1}{2 \sum_{i=a}^{a+b-1} \sum_{j=a+1}^{a+b} k_i \cdot k_j + b m^2} = \\ &\frac{1}{2 \sum_{i=a}^{a+b-1} \sum_{j=a+1}^{a+b} k_i \cdot k_j + (b+1) m^2 - m^2} = \\ &\frac{1}{\left(\sum_{i=a}^b k_i\right)^2 - m^2}. \end{aligned} \quad (254)$$

The string theory integrand must also be modified in order to retain Möbius invariance when the particles are massive. This is due to the Koba-Nielsen factor  $\Lambda_n(\alpha', k, z)$ , which transforms as follows under Möbius transformations:

$$\begin{aligned} \Lambda_n = \prod_{i=1}^{n-1} \prod_{j=2}^n (z_i - z_j)^{\alpha' k_i \cdot k_j} &\rightarrow \prod_{i=1}^{n-1} \prod_{j=2}^n [(z_i - z_j)(C z_i + D)(C z_j + D)]^{\alpha' k_i \cdot k_j} = \\ &\Lambda_n \prod_{i=1}^n (C z_i - D)^{\alpha' \sum_{j \neq i} k_i \cdot k_j}. \end{aligned} \quad (255)$$

We see that the Koba-Nielsen factor is only Möbius invariant when the particles are massless. However, we also see that the substitution

$$k_i \cdot k_j \rightarrow k_i \cdot k_j + \Delta_{ij} \quad (256)$$

with  $\Delta_{ij}$  subjected to exactly the same conditions (250) as in the CHY formalism, fixes the Möbius invariance of  $\Lambda_n$ . We conclude that in string theory as well as in the CHY formalism, the substitution (253) converts a massless amplitude into a massive one.



The substitution (256) can also be used to add masses to only some of the external particles. For example, if for some two numbers  $a, b \in \mathbb{Z}_n$  we set

$$\Delta_{ij} = \begin{cases} m^2, & \text{when } (i = a \text{ and } j = b) \text{ or } (i = b \text{ and } j = a) \\ 0, & \text{otherwise} \end{cases}, \quad (257)$$

then the external legs  $a$  and  $b$  acquire mass  $m$ , while the remaining particles remain massless. Denoting the massless fields  $\phi$  and the massive fields  $\phi_m$ , the amplitude obtained by the substitution (257) describes a theory where the Lagrangian has interaction terms  $\phi^3$  and  $\phi_m^2\phi$ .

## 8 Conclusion

In this thesis we have scrutinized two formalisms:

$$\text{string theory} \quad \leftrightarrow \quad \text{the CHY formalism.}$$

In the two formalisms, an amplitude can be represented as a Möbius invariant

$$\text{divergent integral} \quad \leftrightarrow \quad \text{complex contour integral.}$$

The computation of amplitudes is greatly facilitated by the fact that each integral can be re-written as a sum of

$$\text{divergent regions of the domain} \quad \leftrightarrow \quad \text{residues.}$$

A large class of these integrals are well-behaved and can be calculated with a simple set of rules. There are, however, integrals that are not so well-behaved, and which therefore cannot be computed directly by using the rules. These are the integrals that:

$$\text{require analytical extension} \quad \leftrightarrow \quad \text{have higher-order poles.}$$

In scalar theories, the integrals are always of the well-behaved kind, but in gravity and Yang-Mills theory it is necessary to compute integrals that cannot be directly evaluated with the rules. These more difficult integrals can be calculated by reducing them to well-behaved integrals via

$$\text{integration by parts} \quad \leftrightarrow \quad \text{Pfaffian identities.}$$

These statements evince the great similarity between amplitude computations in string theory and the CHY formalism – a similarity that in at least two instances becomes an exact duality. The resemblance extends so far that precisely the same prescription converts massless amplitudes into massive ones.

Among the avenues one could pursue to build on the results in this thesis, one of the most fruitful might be to employ complex analysis to find a rule for directly evaluating CHY integrals with higher-order poles. Such a rule would provide a powerful tool for calculating amplitudes by eliminating the need to iteratively reduce integrals into simpler ones and could also shed light on the intricate analytically-extended functions that complicate string theory calculations. Another field of enquiry that has been left open is a systematic

treatment of the identities relating CHY integrals. The few Pfaffian identities that are the easiest to visualize and apply have been described, but the full extent of the identities has not been ascertained, nor how they tie into the general KLT relations in [11] or into the integrations by parts by which tachyonic poles are removed in [15]. It would also be desirable to have a more succinct rule for the overall sign of CHY integrals.

While it is clear that there are many features of string theory and CHY tree-amplitudes that can be studied further, it is my hope that this thesis has in some measure served as a proof of principle to the effect that in calculating the tree-level scattering of scalars, gluons, and gravitons, there is no need to integrate or take limits or draw Feynman diagrams or solve rational equations – they can all be determined with graph theory alone. In the final analysis, the computation of tree-amplitudes reduces to discrete mathematics and becomes trivial. And so the stage is set for loops.

A CHY formula for supergravity at one loop was postulated in ref. [25] in 2014 by Adamo, Casali, and Skinner (ACS) after employing ambitwistor string theory to study scattering equations on a torus. The formula involves elliptic curves and the question of how to extract rational expressions from it remained a puzzle until Geyer, Mason, Monteiro, and Tourkine (GMMT) in ref. [26] managed to derive loop formulas on the Riemann sphere by applying a residue theorem to the results of ACS. If simple algorithmic rules could be found for loop calculations with the GMMT formulas, it could have a significant impact on the future of amplitude computations.

## 9 Acknowledgement

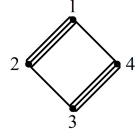
On a closing note, I should like to express my gratitude to Emil Bjerrum-Bohr and Poul Henrik Damgaard, who have proved to be the most inspiring mentors a student could hope for. I am also much obliged to Jacob Bourjaily for his bountiful guidance and encouragement.

# Appendices

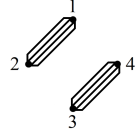
## A Table of CHY integrals with higher order poles

Below is an exhaustive list, for four to six particles, of the 4-regular CHY graphs that cannot immediately be evaluated with the integration rules of section 3. They have been calculated via the method explained in section 4.

### A.1 $n = 4$

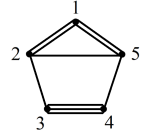


$$= \frac{s_{13}}{s_{34}^2}$$

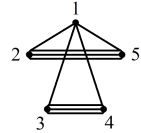


$$= \frac{s_{23}(s_{23}s_{13})}{s_{34}^3}$$

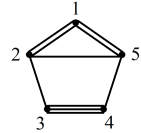
### A.2 $n = 5$



$$= \frac{1}{s_{34}^2} - \frac{s_{24}}{s_{15}s_{34}^2} - \frac{s_{35}}{s_{12}s_{34}^2}$$

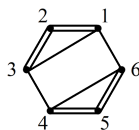


$$= \frac{s_{23}s_{45} - s_{24}s_{35}}{s_{25}^2 s_{34}^2}$$



$$= -\frac{1}{s_{34}^2} - \frac{s_{13}s_{14}}{s_{25}s_{34}^3} - \frac{s_{23}s_{24}}{s_{15}s_{34}^3} - \frac{s_{35}s_{45}}{s_{12}s_{34}^3}$$

### A.3 $n = 6$



$$= \frac{s_{14}}{s_{56}s_{23}s_{123}^2} + \frac{s_{14} + s_{24}}{s_{12}s_{56}s_{123}^2} + \frac{s_{14} + s_{15}}{s_{23}s_{45}s_{123}^2} + \frac{s_{36}}{s_{12}s_{45}s_{123}^2} - \frac{1}{s_{45}s_{123}^2} - \frac{1}{s_{12}s_{123}^2}$$

$$\begin{array}{c} 2 \\ \parallel \\ 3 \end{array} \begin{array}{c} 1 \\ \parallel \\ 6 \end{array} = \frac{1}{s_{12}^2 s_{45}} - \frac{s_{23}}{s_{123} s_{45} s_{12}^2} - \frac{s_{16}}{s_{126} s_{45} s_{12}^2} - \frac{1}{s_{123} s_{45} s_{12}} - \frac{1}{s_{126} s_{45} s_{12}}$$

$$\begin{array}{c} 2 \\ \parallel \\ 3 \end{array} \begin{array}{c} 1 \\ \parallel \\ 6 \end{array} = \frac{s_{14} - s_{24}}{s_{12}^2 s_{56} s_{123}} + \frac{s_{23} s_{24} - s_{13} s_{14}}{s_{12}^2 s_{56} s_{123}^2} + \frac{s_{26} - s_{16}}{s_{12}^2 s_{45} s_{123}} + \frac{s_{13} s_{16} - s_{23} s_{26}}{s_{12}^2 s_{45} s_{123}^2}$$

$$\begin{array}{c} 2 \\ \parallel \\ 3 \end{array} \begin{array}{c} 1 \\ \parallel \\ 6 \end{array} = \frac{s_{13}}{s_{56} s_{123} s_{12}^2} + \frac{s_{26}}{s_{34} s_{345} s_{12}^2} - \frac{1}{s_{34} s_{56} s_{12}} - \frac{s_{156}}{s_{34} s_{56} s_{12}^2}$$

$$\begin{array}{c} 2 \\ \parallel \\ 3 \end{array} \begin{array}{c} 1 \\ \parallel \\ 6 \end{array} = \frac{s_{13}}{s_{12}^2 s_{45} s_{123}} + \frac{s_{26}}{s_{12}^2 s_{45} s_{345}} + \frac{s_{13}}{s_{12}^2 s_{56} s_{123}} + \frac{s_{26}}{s_{12}^2 s_{34} s_{345}} \\ - \frac{1}{s_{12}^2 s_{34}} - \frac{1}{s_{12}^2 s_{45}} - \frac{1}{s_{12}^2 s_{56}} + \frac{s_{134}}{s_{12}^2 s_{34} s_{56}}$$

$$\begin{array}{c} 2 \\ \parallel \\ 3 \end{array} \begin{array}{c} 1 \\ \parallel \\ 6 \end{array} = \frac{s_{23} s_{46}}{s_{12}^2 s_{123} s_{45}^2} + \frac{s_{46}}{s_{12} s_{123} s_{45}^2} + \frac{s_{14} s_{25} - s_{15} s_{24}}{s_{12}^2 s_{36} s_{45}^2} + \frac{s_{24}}{s_{12}^2 s_{45}^2} + \frac{s_{26} s_{34}}{s_{12}^2 s_{345} s_{45}^2} + \frac{s_{26}}{s_{12}^2 s_{345} s_{45}}$$

$$\begin{array}{c} 2 \\ \parallel \\ 3 \end{array} \begin{array}{c} 1 \\ \parallel \\ 6 \end{array} = \frac{s_{15} s_{24} - s_{14} s_{25}}{s_{12}^2 s_{36} s_{45}^2} + \frac{s_{15} s_{24} - s_{14} s_{25}}{s_{12} s_{123}^2 s_{45}^2} \\ + \frac{s_{15} s_{24} - s_{14} s_{25} + s_{24} s_{35} - s_{14} s_{35} + s_{15} s_{34} - s_{25} s_{34}}{s_{12}^2 s_{123} s_{45}^2} \\ + \frac{s_{13} s_{14} s_{35} + s_{23} s_{25} s_{34} - s_{23} s_{24} s_{35} - s_{13} s_{15} s_{34}}{s_{12}^2 s_{123}^2 s_{45}^2}$$

$$\begin{array}{c} 2 \\ \parallel \\ 3 \end{array} \begin{array}{c} 1 \\ \parallel \\ 6 \end{array} = - \frac{s_{15} s_{24}}{s_{234} s_{34}^2 s_{56}^2} - \frac{s_{13} s_{46}}{s_{12}^2 s_{56}^2 s_{456}} - \frac{s_{26} s_{35}}{s_{12}^2 s_{34}^2 s_{345}} - \frac{1}{s_{12} s_{34} s_{56}} + \frac{s_{13} s_{345}}{s_{12}^2 s_{34} s_{56}^2} \\ + \frac{s_{35}(s_{16} + s_{56}) + s_{36} s_{16}}{s_{12}^2 s_{34}^2 s_{56}} + \frac{s_{15} s_{456} + s_{13} s_{45} - s_{14} s_{35}}{s_{12} s_{34}^2 s_{56}^2} + \frac{s_{13}}{s_{12}^2 s_{34} s_{56}} \\ + \frac{s_{15}}{s_{12} s_{34} s_{56}^2} - \frac{s_{23}}{s_{12} s_{34}^2 s_{56}} + \frac{(s_{13} + s_{14})(s_{35} s_{46} - s_{36} s_{45})}{s_{12}^2 s_{34}^2 s_{56}^2}$$

$$\begin{aligned}
\begin{array}{c} \text{2} \\ \text{1} \\ \text{3} \quad \text{6} \\ \text{4} \quad \text{5} \end{array} &= \frac{1}{s_{12}^2 s_{45}} + \frac{1}{s_{12}^2 s_{36}} + \frac{s_{13} s_{23}}{s_{12}^3 s_{123} s_{45}} + \frac{s_{14} s_{24}}{s_{12}^3 s_{124} s_{36}} + \frac{s_{15} s_{25}}{s_{12}^3 s_{125} s_{36}} + \frac{s_{16} s_{26}}{s_{12}^3 s_{126} s_{45}} \\
&\quad - \frac{s_{245}}{s_{12}^2 s_{36} s_{45}} - \frac{(s_{13} + s_{16})(s_{24} + s_{25})}{s_{12}^3 s_{36} s_{45}}
\end{aligned}$$

$$\begin{aligned}
\begin{array}{c} \text{2} \\ \text{1} \\ \text{3} \quad \text{6} \\ \text{4} \quad \text{5} \end{array} &= \frac{1}{s_{12}^2} \left( \frac{1}{s_{36}} + \frac{1}{s_{56}} \right) + \frac{(s_{13} + s_{16})(s_{23} + s_{26})}{s_{12}^3 s_{36} s_{45}} + \frac{(s_{15} + s_{16})(s_{25} + s_{26})}{s_{12}^3 s_{34} s_{56}} \\
&\quad + \frac{s_{13} s_{23}}{s_{12}^3 s_{123}} \left( \frac{1}{s_{45}} + \frac{1}{s_{56}} \right) + \frac{s_{14} s_{24}}{s_{12}^3 s_{124}} \left( \frac{1}{s_{36}} + \frac{1}{s_{56}} \right) + \frac{s_{15} s_{25}}{s_{12}^3 s_{125}} \left( \frac{1}{s_{34}} + \frac{1}{s_{36}} \right) \\
&\quad + \frac{s_{16} s_{26}}{s_{12}^3 s_{126}} \left( \frac{1}{s_{34}} + \frac{1}{s_{45}} \right)
\end{aligned}$$

$$\begin{aligned}
\begin{array}{c} \text{2} \\ \text{1} \\ \text{3} \quad \text{6} \\ \text{4} \quad \text{5} \end{array} &= \frac{(s_{14} s_{25} - s_{15} s_{24})(s_{13} s_{26} - s_{16} s_{23})}{s_{12}^3 s_{36}^2 s_{45}^2} - \frac{s_{13} s_{23} s_{46}}{s_{12}^3 s_{123} s_{45}^2} \\
&\quad - \frac{s_{14} s_{24} s_{35}}{s_{12}^3 s_{124} s_{36}^2} - \frac{s_{15} s_{25} s_{46}}{s_{12}^3 s_{125} s_{36}^2} - \frac{s_{16} s_{26} s_{35}}{s_{12}^3 s_{126} s_{45}^2} \\
&\quad + \frac{s_{15} s_{25}(s_{34} + s_{45}) + s_{14} s_{25}(s_{23} + s_{34}) + s_{15} s_{24}(s_{13} + s_{34}) - s_{14} s_{24} s_{35}}{s_{12}^3 s_{36}^2 s_{45}} \\
&\quad + \frac{s_{13} s_{26}(s_{24} + s_{34}) + s_{16} s_{23}(s_{14} + s_{34}) + s_{16} s_{26}(s_{34} + s_{36}) - s_{13} s_{23} s_{46}}{s_{12}^3 s_{36}^2 s_{45}^2} \\
&\quad + \frac{s_{14} s_{25} + s_{16} s_{23} - s_{13} s_{24} - s_{15} s_{26}}{s_{12}^3 s_{36} s_{45}} - \frac{s_{246}}{s_{12}^2 s_{36} s_{45}}
\end{aligned}$$

$$\begin{aligned}
\begin{array}{c} \text{2} \\ \text{1} \\ \text{3} \quad \text{6} \\ \text{4} \quad \text{5} \end{array} &= \frac{s_{35} s_{46}}{s_{34} s_{56}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{3} \quad \text{6} \\ \text{5} \quad \text{4} \end{array} + \frac{s_{12} s_{36} - s_{36} s_{45}}{s_{34} s_{56}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{3} \quad \text{6} \\ \text{6} \quad \text{5} \end{array} - \frac{s_{13} s_{26}}{s_{34} s_{56}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{6} \quad \text{3} \\ \text{5} \quad \text{4} \end{array} \\
&\quad + \frac{s_{16} s_{23}}{s_{34} s_{56}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{6} \quad \text{3} \\ \text{4} \quad \text{5} \end{array}
\end{aligned}$$

$$\begin{aligned}
\begin{array}{c} \text{2} \\ \text{1} \\ \text{3} \quad \text{6} \\ \text{4} \quad \text{5} \end{array} &= -\frac{s_{14}}{s_{456}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{6} \quad \text{3} \\ \text{6} \quad \text{5} \end{array} - \frac{s_{15}}{s_{456}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{6} \quad \text{3} \\ \text{6} \quad \text{4} \end{array} - \frac{s_{24}}{s_{456}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{6} \quad \text{3} \\ \text{4} \quad \text{5} \end{array} - \frac{s_{25}}{s_{456}} \begin{array}{c} \text{2} \\ \text{1} \\ \text{6} \quad \text{3} \\ \text{5} \quad \text{4} \end{array}
\end{aligned}$$

## B Expanding diverging integrals beyond leading order

If one wishes to determine the non-leading order terms in the  $\alpha'$  expansion of a divergent integral, one can do so by a procedure that is somewhat more involved than the integration rules of section 2. The procedure basically consists in rewriting the integral as a sum of terms that either do not diverge or that factor into beta functions, whose expansions are well known – a trick also used in [24].

As an example, consider the integral  $I_6[H]$  with  $H$  given as follows:

$$H(z) = \frac{1}{(z_1 - z_2)(z_1 - z_4)(z_2 - z_6)(z_3 - z_5)^2(z_4 - z_6)}.$$

In our standard gauge, the integral takes on the following form:

$$I_6[H] = \int_0^1 dz_3 \int_0^{z_3} dz_4 \int_0^{z_4} dz_5 \frac{\Lambda_6(\alpha', k, z)}{z_4(z_3 - z_5)^2}.$$

Now we introduce the following variables:

$$x = z_3, \quad y = \frac{z_4}{z_3}, \quad z = \frac{z_5}{z_4}.$$

In terms of the new variables we can rewrite the integral as follows:

$$I_6[H] = \int_0^1 dx \int_0^1 dy \int_0^1 dz x^{\alpha'A-1} (1-x)^{\alpha'B} y^{\alpha'C} (1-y)^{\alpha'D} z^{\alpha'F} (1-z)^{\alpha'G} \times \\ (1-xy)^{\alpha'J} (1-yz)^{\alpha'K-2} (1-xyz)^{\alpha'L},$$

where

$$\begin{aligned} A &= s_{3456} \\ B &= s_{23} \\ C &= s_{456} \\ D &= s_{34} \\ F &= s_{56} \\ G &= s_{45} \\ J &= s_{24} \\ K &= s_{35} \\ L &= s_{25}. \end{aligned}$$

We can write the integral as the sum of two terms as follows

$$I_6[H] = I_1 + I_2,$$

where

$$I_1 = \int_0^1 dx \int_0^1 dy \int_0^1 dz x^{\alpha'A-1} (1-x)^{\alpha'B} y^{\alpha'C} (1-y)^{\alpha'D} z^{\alpha'F} (1-z)^{\alpha'G} (1-yz)^{\alpha'K-2} \times \\ \left( (1-xy)^{\alpha'J} (1-xyz)^{\alpha'L} - (1-x)^{\alpha'(J+L)} \right),$$

$$I_2 = \int_0^1 dx x^{\alpha'A-1} (1-x)^{\alpha'(B+J+L)} \int_0^1 dy \int_0^1 dz y^{\alpha'C} (1-y)^{\alpha'D} z^{\alpha'F} (1-z)^{\alpha'G} (1-yz)^{\alpha'K-2}.$$

The point of this rewriting is that the integration over  $x$  has been factored out of  $I_2$ , while  $I_1$  is non-divergent. For if  $x$  tends to zero or  $y$  and  $z$  tend to one so that  $x^{\alpha'A-1}$  or  $(1-yz)^{\alpha'K-2}$  will diverge, then  $(1-xy)^{\alpha'J}(1-xyz)^{\alpha'L} - (1-x)^{\alpha'(J+L)}$  will tend to zero, so that the integral remains finite when integrating over this domain. We can therefore perform the expansion of  $I_1$  by expanding the integrand and then integrating term by term:

$$\begin{aligned} I_1 &= \int_0^1 dx \int_0^1 dy \int_0^1 dz \left( \frac{-(J+L)\log(1-x) + \log(1-xy) + L\log(1-xyz)}{x(1-yz)^2} \alpha' + \mathcal{O}(\alpha'^2) \right) \\ &= (2J+3L)\zeta(3)\alpha' + \mathcal{O}(\alpha'^2). \end{aligned}$$

In order to expand  $I_2$  we still need to separate the integrations over  $y$  and  $z$ , which I shall denote  $I'_2$ :

$$I'_2 = \int_0^1 dy \int_0^1 dz y^{\alpha'C} (1-y)^{\alpha'D} z^{\alpha'F} (1-z)^{\alpha'G} (1-yz)^{\alpha'K-2}.$$

In its current form,  $I'_2$  cannot straight-forwardly be written as a sum of a non-divergent term and a term that factors into beta functions. There is, however, a shift of variables that enables such a rewriting, namely the shift to the following variables:

$$\begin{aligned} s &\equiv 1 - yz, \\ t &\equiv \frac{1-y}{1-yz}. \end{aligned}$$

To more easily determine what the integration domain is in terms of the new variables, one can perform the variable shift in a sequence of steps:

$$\begin{aligned} I'_2 &= \int_0^1 dy \int_0^y dw y^{\alpha'(C-F-G)-1} (1-y)^{\alpha'D} w^{\alpha'F} (y-w)^{\alpha'G} (1-w)^{\alpha'K-2}, \quad w \equiv yz, \\ &= \int_0^1 ds \int_0^s dr r^{\alpha'D} (1-r)^{\alpha'(C-F-G)-1} s^{\alpha'K-2} (1-s)^{\alpha'F} (s-r)^{\alpha'G}, \quad s \equiv 1-w, \quad r \equiv 1-y, \\ &= \int_0^1 ds \int_0^1 dt s^{\alpha'(D+G+K)-1} (1-s)^{\alpha'F} t^{\alpha'D} (1-t)^{\alpha'G} (1-st)^{\alpha'(C-F-G)-1}, \quad t \equiv \frac{r}{s}. \end{aligned}$$

It is now clear that we can rewrite  $I'_2$  as the following sum:

$$I'_2 = I'_{21} + I'_{22}$$

with

$$I'_{21} = \int_0^1 ds \int_0^1 dt s^{\alpha'(D+G+K)-1} (1-s)^{\alpha'F} t^{\alpha'D} (1-t)^{\alpha'G} \left( (1-st)^{\alpha'(C-F-G)-1} - 1 \right),$$

$$I'_{22} = \int_0^1 ds s^{\alpha'(D+G+K)-1} (1-s)^{\alpha'F} \int_0^1 dt t^{\alpha'D} (1-t)^{\alpha'G}.$$

Because  $I'_{21}$  is non-divergent, we can expand the integrand in  $\alpha'$  and perform the integration term by term:

$$\begin{aligned}
I'_{21} &= \int_0^1 ds \int_0^1 dt \left( \frac{t}{1-st} + \right. \\
&\frac{Fst \log(1-s) + (D+G+K)st \log(s) + Gst \log(1-t) + Dst \log(t) + (C-F-G) \log(1-st)}{s(1-st)} \alpha' \\
&\quad \left. + \mathcal{O}(\alpha'^2) \right) \\
&= 1 + \left( K - D - G - (C+K) \frac{\pi^2}{6} \right) \alpha' + \mathcal{O}(\alpha'^2).
\end{aligned}$$

Because  $I'_{22}$  can be expressed as a product of beta functions, it can be expanded straightforwardly:

$$\begin{aligned}
I'_{22} &= B(\alpha'(D+G+K), \alpha'F+1) B(\alpha'D+1, \alpha'G+1) \\
&= \frac{1}{D+G+K} \frac{1}{\alpha'} - \frac{D+G}{D+G+K} - \left( \frac{F\pi^2}{6} - \frac{6D^2 + 12DG + 6G^2 - DG\pi^2}{6(D+G+K)} \right) \alpha' + \mathcal{O}(\alpha'^2).
\end{aligned}$$

We now know the expansions of  $I'_{21}$  and  $I'_{22}$ , and we can therefore expand the integral  $I$  since

$$I_6[H] = I_1 + B(\alpha'A, \alpha'(B+J+L)+1)(I'_{21} + I'_{22})$$

We conclude that to next-to-next-to leading order,  $I$  is given as follows:

$$\begin{aligned}
I_6[H] &= \\
&\frac{1}{A(D+G+K)} \frac{1}{\alpha'^2} + \frac{K}{A(D+G+K)} \frac{1}{\alpha'} + \\
&\frac{6K^2 - (AB + AJ + AL + CD + CG + CK + DF + DG + DK + FG + FK + GK + K^2)\pi^2}{6A(D+G+K)} \\
&+ \mathcal{O}(\alpha'^2).
\end{aligned}$$



## References

- [1] front page illustration: *Points* by Wassily Kandinsky (1920), oil on canvas, Ohara Museum of Art, Kurashiki.
- [2] Martin, William Alexander Parsons. *Hanlin Papers: Or, Essays on the Intellectual Life of the Chinese*. Trubner & Company (1880) p. 252
- [3] F. Cachazo, S. He, and E. Y. Yuan, “Scattering Equations and Kawai-Lewellen-Tye Orthogonality,” *Phys. Rev.* **D90** (2014) no. 6, 065001, arXiv:1306.6575 [hep-th].
- [4] F. Cachazo, S. He, and E. Y. Yuan, “Scattering of Massless Particles in Arbitrary Dimensions,” *Phys. Rev. Lett.* **113** (2014) no. 17, 171601, arXiv:1307.2199 [hep-th].
- [5] F. Cachazo, S. He, and E. Y. Yuan, “Scattering of Massless Particles: Scalars, Gluons and Gravitons,” *JHEP* **1407** (2014) 033, arXiv:1309.0885 [hep-th].
- [6] F. Cachazo, S. He, and E. Y. Yuan, “Einstein-Yang-Mills Scattering Amplitudes From Scattering Equations,” *JHEP* **1501** (2015) 121, arXiv:1409.8256 [hep-th].
- [7] F. Cachazo, S. He, and E. Y. Yuan, “Scattering Equations and Matrices: From Einstein To Yang-Mills, DBI and NLSM,” arXiv:1412.3479 [hep-th].
- [8] L. Dolan and P. Goddard, “Proof of the Formula of Cachazo, He and Yuan for Yang-Mills Tree Amplitudes in Arbitrary Dimension,” *JHEP* **1405** (2014) 010, arXiv:1311.5200 [hep-th].
- [9] S. G. Naculich, “Scattering equations and BCJ relations for gauge and gravitational amplitudes with massive scalar particles,” *JHEP* **1409** (2014) 029, arXiv:1407.7836 [hep-th]  
S. G. Naculich, “CHY representations for gauge theory and gravity amplitudes with up to three massive particles,” *JHEP* **1505** (2015) 050, arXiv:1501.03500 [hep-th]  
S. G. Naculich, “Amplitudes for massive vector and scalar bosons in spontaneously-broken gauge theory from the CHY representation,” arXiv:1506.06134 [hep-th]
- [10] C. Kalousios, “Scattering Equations, Generating Functions and All Massless Five Point Tree Amplitudes,” arXiv:1502.07711 [hep-th]
- [11] F. Cachazo and H. Gomez, “Computations of Contour Integrals on  $\mathcal{M}_{0,n}$ ,” arXiv:1505.03571 [hep-th]
- [12] J. Scherk, “Zero-slope limit of the dual resonance model”, *Nucl. Phys. B* **222** (1971) 31.2
- [13] L. Mason and D. Skinner, “Ambitwistor Strings and the Scattering Equations,” *JHEP* **1407** (2014) 048, arXiv:1311.2564 [hep-th].
- [14] N. Berkovits, “Infinite Tension Limit of the Pure Spinor Superstring,” *JHEP* **1403** (2014) 017, arXiv:1311.4156 [hep-th].
- [15] N. E. J. Bjerrum-Bohr, P. H. Damgaard, P. Tourkine, and P. Vanhove, “Scattering Equations and String Theory Amplitudes,” *Phys. Rev.* **D90** (2014) no. 10, 106002, arXiv:1403.4553 [hep-th].
- [16] C. Baadsgaard, N. Bjerrum-Bohr, J. L. Bourjaily, and P. H. Damgaard, “Integration Rules for Scattering Equations,” arXiv:1506.06137 [hep-th].
- [17] C. Baadsgaard, N. Bjerrum-Bohr, J. L. Bourjaily, and P. H. Damgaard, “Scattering Equations and Feynman Diagrams,” arXiv:1507.00997 [hep-th].

- [18] M. B. Green, J. H. Schwarz, E Witten, *Superstring Theory*, vol 1. Cambridge University Press (1987).
- [19] F. Cachazo, S. He, and E. Y. Yuan, “Scattering in Three Dimensions from Rational Maps,” *Phys. Rev.* **113** (2014) no. 17, 171601, arXiv:1306.2962 [hep-th].
- [20] N. Nakanishi, “Crossing-Symmetric Decomposition of the Five-Point and Six-Point Veneziano Formulas into Tree-Graph Integrals,” *Prog. Theor. Phys.* **45** (1971) 436–451.  
 N. Nakanishi, “Crossing-Symmetric Decomposition of the  $n$ -Point Veneziano Formula into Tree-graph Integrals. 1.,” *Prog. Theor. Phys.* **45** (1971) 451–470.  
 N. Nakanishi, “Crossing-Symmetric Decomposition of the  $n$ -Point Veneziano Formula into Tree-Graph Integrals. 2. Koba-Nielsen Representation,” *Prog. Theor. Phys.* **45** (1971) 919–926.
- [21] R. Marotta and F. Pezzella, ”String-generated quartic scalar interactions”, arXiv:0003044 [hep-th].
- [22] R. Marotta and F. Pezzella, ”Two-Loop  $\Phi^4$ -Diagrams from String Theory”, arXiv:9912158 [hep-th].
- [23] A. Frizzo, L. Magnea, and R. Russo, ”Scalar field theory limits of bosonic string amplitudes”, *Nucl. Phys. B* **379** (2000) 579.1, arXiv:9912183 [hep-th].
- [24] D. Oprisa and S. Stieberger, “Six Gluon Open Superstring Disk Amplitude, Multiple Hypergeometric Series and Euler-Zagier Sums,” arXiv:0509042 [hep-th].
- [25] T. Adamo, E. Casali and D. Skinner, “Ambitwistor strings and the scattering equations at one-loop,” *JHEP* **1404** (2014) 104 [arXiv:1312.3828 [hep-th]].
- [26] Y. Geyer, L. Mason, R. Monteiro and P. Tourkine, “Loop Integrands from the Riemann Sphere,” arXiv:1507.00321 [hep-th].