Review Article

Scattering Amplitudes, the AdS/CFT Correspondence, Minimal Surfaces, and Integrability

Luis F. Alday

School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA

Correspondence should be addressed to Luis F. Alday, alday@ias.edu

Received 22 March 2010; Accepted 8 April 2010

Academic Editor: Carlos Nunez

Copyright © 2010 Luis F. Alday. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We focus on the computation of scattering amplitudes of planar maximally supersymmetric Yang-Mills in four dimensions at strong coupling by means of the AdS/CFT correspondence and explain how the problem boils down to the computation of minimal surfaces in AdS in the first part of this paper. In the second part of this review we explain how integrability allows to give a solution to the problem in terms of a set of integral equations. The intention of the review is to give a pedagogical, rather than very detailed, exposition.

1. Introduction

The aim of this review is to study gluon scattering amplitudes of four-dimensional planar maximally supersymmetric Yang-Mills (MSYMs). We hope that the study of such amplitudes would teach us something about scattering amplitudes of QCD, but at the same time they are much more tractable. The reason for such tractability is twofold. On one hand, perturbative computations are much simpler than in QCD, due to the high degree of symmetry. In fact enormous progress has been made in the last few years. On the other hand, the strong coupling regime of the theory can be studied by means of the AdS/CFT duality, by studying a weakly coupled string sigma-model.

The aim of these lectures is to explain how to use the AdS/CFT duality in order to compute gluon scattering amplitudes of planar MSYM at strong coupling.

In the next section we briefly describe some weak coupling perturbative results for scattering amplitudes in planar MSYM. Scattering amplitudes of massless gauge theories in four dimensions suffer from IR divergences; hence, a regulator needs to be introduced in order to define them properly. A convenient choice is dimensional regularization, in which
the theory is studied in $D = 4 - 2\epsilon$ dimensions. As we will see, perturbative computations suggest an ansatz, the so-called BDS ansatz, according to which the amplitude exponentiates. By now, we know that the BDS ansatz is not correct at higher loops and for a large number of particles. Still, it was a useful guidance for recent developments, and perturbative results are often expressed in terms of deviations from this ansatz.

In Section 3 we proceed and explain how the AdS/CFT duality can be used to compute scattering amplitudes at strong coupling. As in the weak coupling computation, amplitudes are IR divergent as a regulator has to be introduced. In order to set up the computation we introduce a $D$-brane as regulator. Actual computations, however, are done by using the analogous of dimensional regularization, since we are interested in comparing our results to the expectations from the perturbative regime. The computation of scattering amplitudes at strong coupling reduces to a minimal area problem in AdS. We show in detail how the prescription works for the scattering of four gluons.

In Section 4 we focus on the mathematical problem of computing the area minimal surfaces in AdS relevant to the study of scattering amplitudes. This lead us to study classical strings on AdS. By means of a reduction, sometimes called Polchmeyer reduction, one can focus on the physical degrees of freedom of the problem. The problem possesses integrability, which in practical terms means that we can introduce a spectral parameter $\xi$ and study how physical quantities depend on such parameter. In particular, this allows to give a solution of the problem in terms of a set of integral equations that have the form of thermodynamic Bethe ansatz (TBA) equations. Furthermore, we argue that the area of the minimal surfaces, directly related to the scattering amplitude at strong coupling, is nothing but the free energy of the above-mentioned TBA system.

In Section 5 we end up with some conclusions and a list of open problems. Finally, we include an appendix with a Mathematica code solving numerically the TBA integral equations for a particular example.

**Disclaimer**

The present paper is not intended to be a complete/fully precise exposition of all the recent results regarding scattering amplitudes in planar maximally supersymmetric Yang-Mills. Rather, we attempt to explain in a pedagogical manner the computation at strong coupling and derive the final solution in the most straightforward possible manner.

### 2. Perturbative Scattering Amplitudes of Planar MSYM

In this section we briefly discuss the basics regarding the perturbative computation of planar gluon scattering amplitudes of MSYM. For more complete reviews see [1, 2] and references therein.

#### 2.1. Basics

Being massless particles of spin one in four dimensions, gluon states $|G\rangle = |h, p^\mu, a\rangle$ are characterized by their helicity, $h = \pm 1$, four momenta, $p^\mu$, with $p^2 = 0$, and color indices $a$ in the adjoint representation of the gauge group $SU(N_c)$. Generic scattering amplitudes depend in a complicated manner on these.
In the planar limit it is useful to write the amplitudes in the so-called color decomposed form

$$
\mathcal{A}^{(L)}_N \sim g^{2N-2} \left(g^2 N_c\right)^L \sum_{\rho} Tr(T^a_i \ldots T^a_i(N)) A^{(L)}_N (\rho(1), \ldots, \rho(N)),
$$

(2.1)

where we denoted by $\mathcal{A}^{(L)}_N$ the $L$-loop, $N$-point amplitude. The sum runs over noncyclic permutations, $N_c$ denotes the number of colors and $g$ the gauge theory coupling constant. Furthermore, we denoted by $T^a$ the gauge group generators in the adjoint representation. In the nonplanar case, there are also contributions with multi-traces, suppressed in the large $N_c$ limit. This decomposition separates the color structure from the kinematics. The term proportional to a given color structure is the so-called leading color ordered amplitude $A^{(L)}_N$, which will only depend on the momenta and the helicities of the particles undergoing the scattering.

As already mentioned, scattering amplitudes of massless particles in four dimensions suffer from infrared divergences. In order to see this more clearly, consider the scattering of four gluons at one loop. The amplitude contains a factor of the form

$$
A_{4, D=4}^{(1)} \sim \int \frac{d^4p}{p^2(p-k_1)^2(p-k_1-k_2)^2(p+k_4)^2},
$$

(2.2)

where $p$ is the momentum running along the loop and $k_i^2 = 0$ are the momenta of the external particles. We recognize two kinds of divergences. First, from the region $p^\mu \sim 0$, these are called soft divergences, since they are due to the interchange of gluons with very low (soft) momenta among external gluons. The second class of divergences comes from the region $p^\mu \sim a k_i^\mu$ and are called collinear divergences, since in this case the momentum of the gluon interchanged is parallel to the momentum of one of the external gluons. Notice that in addition we can have $p^\mu \sim a k_i^\mu$ with $a \sim 0$, so soft and collinear effects can combine, giving an enhanced divergence.

Such amplitude can be regularized in several ways. A popular choice is to use a version of dimensional regularization and consider the theory in $D = 4 - 2\epsilon$ dimensions. (Another possibility is to go to the Coulomb phase of the theory, staying in four dimensions [34]. This is equivalent to giving a small mass to the propagators in the outer loop when computing scattering amplitudes.) More precisely, we still consider four-dimensional external states but consider the momentum running inside the loop in a dimension slightly higher than four ($\epsilon$ is negative)

$$
A_{4, D=4}^{(1)} \rightarrow A_{4, D=4-2\epsilon}^{(1)},
$$

(2.3)

(In practice, we do the computation in four dimensions and once we arrive to the final result (2.2), we replace $d^4p \rightarrow d^{4-2\epsilon}p$.) Once a regulator is introduced, the amplitudes are finite, but the price we have to pay is an explicit dependence on such regulator. (and the consequent breaking of some of the symmetries is present in the original problem.) In the case of dimensional regularization, IR divergences manifest themselves as poles in $\epsilon$. From the explicit one loop expression (2.2) and the discussion above, we see that soft and collinear divergences combine to give a divergent factor $\approx 1/\epsilon^2$. In the general case of $L$ loops, we
expect the same phenomenon to occur for each subloop, hence we expect a divergence of the form

$$A_N^{(L)} \sim \frac{1}{\epsilon^{2L}} + \cdots$$  \hspace{1cm} (2.4)

As already mentioned, color-ordered amplitudes depend also on the helicities of the external gluons. As a consequence of supersymmetry, scattering amplitudes of MSYM satisfy supersymmetric Ward identities. In particular, these identities imply that certain amplitudes vanish at all loops, namely, an amplitude vanish if all helicities, or all but one, are plus

$$\mathcal{A}(++\cdots+) = \mathcal{A}(−+\cdots+) = 0.$$  \hspace{1cm} (2.5)

The first nontrivial amplitude is the so-called maximally helicity violating (MHV) amplitude, and it has the form $\mathcal{A}(−+\cdots−+\cdots+)$, namely, two external states have minus helicity and the rest plus. A great simplicity of studying MHV amplitudes is given by the fact that they contain a single Lorentz structure, which is already captured by the tree level amplitude, to all orders in perturbation theory. It is then convenient to factor out the tree level amplitude and study the so-called reduced amplitude

$$M_N^{(L)}(\epsilon) = \frac{A_N^{(L)}(\epsilon)}{A_N^{(0)}},$$  \hspace{1cm} (2.6)

which depends only on the kinematical invariants and the regulator. So, in summary, we consider the planar limit and furthermore we focus on MHV amplitudes only. The reduced amplitude will be simply a scalar function of the kinematical invariants (and of course $\epsilon$).

2.2. The BDS Ansatz

In the last years, a bast technology has been developed for the computation of perturbative scattering amplitudes; see, for instance, [1]. Such technology made it possible for the computation of scattering amplitudes for four (and to some extent for five) gluons to rather high order in the loop expansion. Bern, Dixon, and Smirnov (BDS) noticed in [3], after previous hints in [4], that all known MHV amplitudes satisfied a recursion relation, allowing to write amplitudes at a given order in terms of lower-order amplitudes. These recursions relations could be encoded by the following ansatz, which was then proposed by BDS for MHV amplitudes of any number of gluons at any loop order:

$$\mathcal{M} \equiv 1 + \sum_{\ell=1}^{N} \alpha^{\ell} M_n^{(\ell)} = \exp \left[ \sum_{\ell=1}^{N} \left( f^{(\ell)}(\epsilon) M_n^{(\ell)}(\ell\epsilon) + C^\ell + \mathcal{O}(\epsilon) \right) \right],$$  \hspace{1cm} (2.7)

where

$$f^{(\ell)}(\epsilon) = f_0^{(\ell)} + \epsilon f_1^{(\ell)} + \epsilon^2 f_2^{(\ell)}, \quad \alpha \approx \lambda \mu^{2\epsilon}$$  \hspace{1cm} (2.8)
\( \alpha \) is proportional to the 't Hooft coupling constant and keeps track of the perturbation order. We have also introduced an IR scale \( \mu \), in order to account for the dimensionfullness of the coupling constant in dimensions different from four.

The above relation, from now on called BDS ansatz, expresses the (log of the) amplitude in terms of the one loop amplitude \( M_1^\alpha(\ell \epsilon) \). Note that this relation is highly nontrivial, since the constants \( f_i^{(0)} \) and \( C_\ell \) do not depend on the number of gluons or the kinematics. Notice that in addition to the constant piece, there can be noniterating terms of order \( O(\epsilon) \). Due to the presence of these noniterating terms, it is more proper to say that the BDS ansatz is an ansatz for the log of the amplitude rather than for the amplitude itself, since these terms of order \( \epsilon \) could hit the divergent term in order to give finite contributions when expanding the exponential.

The structure of the IR divergent terms in (2.7) was determined in [5], where it was shown that indeed IR divergences factorize and exponentiate as in (2.7). The highly nontrivial content of the ansatz is that the finite pieces also exponentiate in a similar manner and are given, at all loops, by the same functions that characterize the IR divergent terms (times the finite pieces at one loop).

We will be particularly interested in the scattering of four gluons; in that case the BDS ansatz reduces to

\[
A_4 = A_{\text{tree}}(A_{\text{div,}s})^2(A_{\text{div,}t})^2 \exp \left( \frac{f(\lambda)}{8} \left( \log \frac{s}{t} \right)^2 + \text{const} \right), \tag{2.9}
\]

\[
A_{\text{div}} = \exp \left( -\frac{1}{8\epsilon^2} f^{(-2)} \left( \frac{\lambda \mu^{2\epsilon}}{s^{\epsilon}} \right) - \frac{1}{4\epsilon^\prime s^{(-1)}} \left( \frac{\lambda \mu^{2\epsilon}}{s^{\epsilon}} \right) \right), \tag{2.10}
\]

where \( s \) and \( t \) are the usual Mandelstan variables for the scattering of four particles. The amplitude has two pieces, a divergent one and a finite one, and is characterized by two functions. The so-called cusp anomalous dimension \( f(\lambda) = (\lambda \partial_\lambda)^2 f^{(-2)}(\lambda) \) controls both the leading divergent piece and the finite piece while the so-called collinear anomalous dimension, \( g(\lambda) \), controls the subleading divergent pole. Note that the BDS ansatz is giving the answer for the scattering amplitudes to all loops simply in terms of two functions of the coupling constant. Of course, one still would need to compute these functions, but that in principle is a much simpler task. In the next chapter we will show how to compute the four-point scattering amplitude at strong coupling and compare the result to the BDS ansatz.

The BDS ansatz satisfies a number of nontrivial tests:

(i) as already mentioned, it is consistent with the well-known exponential structure of IR divergences;

(ii) it agrees with explicit computations for four gluons at three loops and five gluons at two loops;

(iii) it has the correct collinear factorization properties, that is, the correct limit when two adjacent external momenta become collinear.

The last point is nontrivial and very restrictive. Color-ordered amplitudes \( A_N(k_1, k_2, \ldots, k_N) \) have a simple universal behavior as the momenta of two adjacent gluons, let us say that \( k_1 \) and \( k_2 \) become collinear, (very) schematically:

\[
A_N^{(L)}(k_1, k_2, k_3, \ldots, k_N) \rightarrow R(z)A_{N-1}^{(L)}(k_1, k_3, \ldots, k_N) \tag{2.11}
\]
The correct expression contains also terms involving the $N-1$ point function at lower order. with $k_i = k_1 + k_2$, $k_1 \to z k_i$ in the collinear limit and $R(z)$ is a universal function. The BDS ansatz turns out to be consistent with this behavior.

At the moment, there are strong reasons to believe that the BDS ansatz is correct for four and five gluons, while direct evidence against it was found by an explicit computation for the scattering of six gluons at two loops [6]. Still, even if not totally correct, such ansatz is useful as a guidance in perturbative and strong coupling computations. For instance, it is by now customary to express perturbative answers in terms of deviations from the BDS ansatz. We will further discuss this issue at the end of the next section.

Let us mention that the cusp anomalous dimension $f(\lambda)$ plays a very important role in recent developments regarding integrability in the AdS/CFT correspondence. It turns out that this function arises in many computations. In particular it controls the large spin limit of the anomalous dimension of the so-called twist two operators, of the form $\text{Tr} \phi D^s \phi$, namely,

$$\Delta - S = f(\lambda) \log S + \cdots$$

(2.12)

for large $S$, to all orders in perturbation theory. Using the integrability of the gauge theory dilatation operator [7] constructed an integral equation whose solution gives the cusp anomalous dimension to all values of the coupling constant. For instance, the equation can be solved at weak and strong values of the coupling [8–10], giving

$$f(\lambda) = \frac{\lambda}{2\pi^2} + \cdots, \quad \lambda \ll 1, \quad f(\lambda) = \frac{\sqrt{\lambda}}{\pi} + \cdots, \quad \lambda \gg 1.$$  

(2.13)

Such values agree with previous computations, done either at weak coupling by using perturbation theory [11] or at strong coupling by using the AdS/CFT duality [12].

As for the collinear anomalous dimension $g(\lambda)$, note from the expression (2.9) that it is scheme dependent. For instance, if we change $\mu \to a \mu$, then $g(\lambda)$ changes by a term proportional to the cusp anomalous dimension $f(\lambda)$. Hence the cusp anomalous dimension is a more “robust” quantity. This is also related to the fact that $f(\lambda)$ is related to the anomalous dimension of an operator, as described above, while $g(\lambda)$ is not.

In the next section we will explain how to compute gluon scattering amplitudes of planar MSYM at strong coupling by using the AdS/CFT duality. In particular, for the case of four gluons, the strong coupling results are in perfect agreement with (2.7).

### 3. Gluon Scattering Amplitudes at Strong Coupling

In order to attack the problem of computing scattering amplitudes at strong coupling (In this section, we follow closely [35], to which we refer the reader for the details,) we will make use of the AdS/CFT duality [13, 14]. This duality expresses the equivalence between four dimensional MSYM and type IIB string theory on $\text{AdS}_5 \times S^5$. The duality provides us with a dictionary between the parameters on both sides of the correspondence, more precisely

$$\frac{1}{N_c} \approx g_s, \quad \sqrt{\lambda} \equiv \sqrt{g^2 N_c} = \frac{R^2}{\alpha'}.$$  

(3.1)
where $g_s$ is the string coupling and $R$ is the radius of the AdS$_5$ that coincides with that of the S$^5$. We see that in the limit of a large number of colors and large $\lambda$, the strings are well described by a weakly coupled sigma model. A powerful feature of the AdS/CFT duality is that it allows to compute certain observables of MSYM at strong coupling by doing geometrical computations on AdS. A well-known example is the computation of the expectation value of supersymmetric Wilson loops, which reduces to a minimal area problem [15, 16]. In this section we will show that this is also the case for the computation of scattering amplitudes at strong coupling.

As in the gauge theory, we will need to introduce a regulator in order to define properly scattering amplitudes. In order to set up our computation we introduce a $D$-brane as IR regulator. Actual computations, however, will be done by using the supergravity analogous of dimensional regularization, since it is simpler to proceed in this scheme, and besides we are interested in comparing our results to gauge theory expectations.

### 3.1. Set-Up of the Computation

As a first IR regulator we consider a $D$-brane localized in the radial direction. In other words, we start with the AdS$_5$ metric written in Poincare coordinates

$$ds^2 = R^2 \frac{dx_{3+1}^2 + dz^2}{z^2},$$

and we place a $D$-brane at some fixed large value of $z = z_{IR}$ and extending along the $x_{3+1}$ coordinates. The asymptotic states are open strings that end on that $D$-brane. We then consider the scattering of these open strings, that will have the interpretation of the gluons.

The proper momentum of the strings is $k_{pr} = k z_{IR}/R$, where $k$ is the momentum conjugate to $x_{3+1}$, and it plays the role of gauge theory momentum and will be kept fixed as we take away the IR cut-off, $z_{IR} \to \infty$. Therefore, due to the warping of the metric, the proper momentum is very large, so we are considering the scattering of strings at fixed angle and with very large momentum.

Amplitudes in such regime were studied in flat space by Gross and Mende [17]. The key feature of their computation is that the amplitude is dominated by a saddle point of the classical action. In our case we need to consider classical strings on AdS. Hence, we need to consider a world-sheet with the topology of a disk with vertex operator insertions on its boundary, which correspond to the external states (see Figure 1). A disk amplitude with a fixed ordering of the open string vertex operators corresponds to a given color-ordered amplitude.

What are the boundary conditions for such world-sheet? First of all, in the vicinity of a vertex operator, the momentum of the external state fixes the form of the solution, since far away the solution looks like a free string propagating with a given momentum. Secondly, since the open strings are attached to the $D$-brane, $z = z_{IR}$ at the boundary.

It turns out to be hard to find classical solutions with the above properties. In order to state more simply the boundary conditions for the world-sheet, it is convenient to describe the solution in terms of $T$-dual coordinates $y^\mu$, defined as follows:

$$ds^2 = \omega^2(z) dx_\mu dx^\mu \to \partial_\mu y^\mu = i \omega^2(z) \epsilon_{\alpha \beta} \partial_\alpha x^\beta.$$
Figure 1: World-sheet corresponding to the scattering of four open strings. In the figure on the left we see four open strings ending on the IR $D$-brane; the world-sheet has then the topology of a disk, shown on the right, with four vertex operator insertions.

Note that we do not $T$-dualize along the radial direction. After defining $r = R^2/z$ we end up again with formally the same $AdS_5$ metric

$$ds^2 = R^2 \frac{dy_μ dy^μ + dr^2}{r^2}. \quad (3.4)$$

In terms of the $(3.4)$ coordinates, the boundary of the world-sheet is located at $r = R^2/z_{IR}$. Besides, the boundary conditions for the original coordinates $x^μ$, which are that they carry momentum $k^μ$, translate into the condition that $y^μ$ has “winding” $Δy^μ = 2πk^μ$. Summarizing, the boundary of the world-sheet is located at $r = R^2/z_{IR}$ and is a particular line constructed as follows (see Figure 3):

(i) for each particle of momentum $k^μ$, draw a segment joining two points separated by $Δy^μ = 2πk^μ$;

(ii) concatenate the segments according to the insertions on the disk (corresponding to a particular color ordering);

(iii) as gluons are massless, the segments are light-like. Due to momentum conservation, the segments form a closed polygon.

The world-sheet, when expressed in $T$-dual coordinates, will then end in such sequence of light-like segments (see Figure 2) located at $r = R^2/z_{IR}$.

Before $T$-dualizing, we can interpret the regulator $D$-brane as several $D3$-branes, one per each external particle, at a fixed radial distance and extended in the four directions $x^{1,3}$. The external states are then open strings ending in these branes, in such a way that two consecutive states end on each brane.

As we perform the four $T$-dualities, these $D3$-branes become $D(-1)$ branes or $D(-1)$ instantons, and the external states become open strings stretching a light-like distance $Δy^μ = 2πk^μ$ between two such $D(-1)$ instantons. These instantons are located at the vertices, or cusps, of the above mentioned polygons.

As we take away the IR cut-off, $z_{IR} → ∞$, the boundary of the world-sheet moves towards the boundary of the $T$-dual metric, at $r = 0$. At leading order in the strong coupling expansion, the computation that we are doing is formally equivalent to the one we would do
if we were to compute the expectation value of a Wilson loop given by a sequence of light-like segments.

Our prescription is then that the leading exponential behavior of the $N$-point scattering amplitude is given by the area $A$ of the minimal surface that ends on a sequence of light-like segments on the boundary

$$A_N \sim e^{-\left(\sqrt{\lambda}/2\pi\right)A(k_{1,\ldots,k_N})},$$

Note that at strong coupling it is more meaningful to talk about the geometric area of the world-sheet. This area is directly related to the log of the amplitude rather than to the amplitude itself. On the other hand, the BDS ansatz (2.7) is also an ansatz for the log of the amplitude. (As already mentioned in the previous section, the presence of noniterating terms $O(\epsilon)$ in the exponent of (2.7) allows only to write the log of the amplitude, and not the amplitude itself, up to terms of order $O(\epsilon)$.) Hence, the strong coupling limit of the BDS ansatz can be directly compared to the geometrical area, times a proportionality factor of the form $\sqrt{\lambda}$ in front of the whole answer.
We should stress that the strong coupling computation is blind to the type or polarization of the external particles and hence valid for a general scattering, MHV or non-MHV, among gluons or other particles. Such information will contribute to prefactors in (3.5) and will be subleading in a $1/\sqrt{\lambda}$. In particular, such prefactor could even vanish, as it should be the case if we were considering amplitudes of the form $\mathcal{A}(\pm, +, +, \ldots +)$. The difference between different scatterings of the same number of particles should be visible if we consider quantum corrections to the classical area. This is a very important, still open, problem.

We have then reduced the problem of computing scattering amplitudes at strong coupling to the problem of finding minimal surfaces in AdS. In the next section we will use the integrability of the problem in order to give a general solution in the form of a set of integral equations. However, in the following we will show that such surface can be found for the particular case of the scattering of four gluons without using any fancy technology. To find and understand this solution in detail will also be instructive for what follows.

### 3.2. Scattering of Four Gluons

Consider the scattering of two particles into two particles, $k_1 + k_3 \rightarrow k_2 + k_4$, and define the usual Mandelstam variables

$$s = -(k_1 + k_2)^2, \quad t = -(k_2 + k_3)^2.$$  \hspace{1cm} (3.6)

According to our prescription we need to find the minimal surface ending in the following light-like polygon, see Figure 4.

As a warm-up exercise it is instructive to consider the solution close to one of the cusps.

#### 3.2.1. The Single Cusp Solution

We start by considering the solution near the cusp where two of the light-like lines meet. So we consider two semi-infinite light-like lines meeting at a point. This case was already considered in [18]. The surface can be embedded into an AdS$_3$ subspace of the full AdS$_5$

$$ds^2 = -dy_0^2 + dy_1^2 + dr^2/r^2. \hspace{1cm} (3.7)$$

We are interested in computing the surface ending on a light-like Wilson loop along $y^1 = \pm y^0$, $y^0 > 0$; see Figure 5.

The problem has a boost and scaling symmetry that becomes explicit if we choose the following parameterization:

$$y_0 = e^\tau \cosh \sigma, \quad y_1 = e^\tau \sinh \sigma, \quad r = e^\tau w. \hspace{1cm} (3.8)$$
Figure 4: Polygon corresponding to the scattering of four gluons.

Figure 5: Single cusp solution.

Boosts and scaling transformations are simply shifts of \( \sigma \) and \( \tau \). The Nambu-Goto action becomes

\[
S_{NG} = \frac{R^2}{2\pi} \int d\sigma d\tau \sqrt{1 - \frac{(w(\tau) + w'(\tau))^2}{w(\tau)^2}}. \tag{3.9}
\]

One can explicitly check that \( w(\tau) = \sqrt{2} \) solves the equations of motion and has the correct boundary conditions. Hence the surface is given by

\[
r = \sqrt{2} \sqrt{y_0^2 - y_1^2}. \tag{3.10}
\]
3.2.2. Four Cusps Solution

Now, let’s come back to the case of the solution with four cusps. In order to write the Nambu-Goto action it is convenient to use Poincare coordinates \((r, y_0, y_1, y_2)\), setting \(y_3 = 0\), and parameterize the surface by its projection to the \((y_1, y_2)\) plane. In this case we obtain an action for two fields, \(r\) and \(t\), living in the space parameterized by \(y_1\) and \(y_2\)

\[
S = \frac{R^2}{2\pi} \int dy_1 dy_2 \sqrt{1 + (\partial_1 r)^2 - (\partial_1 y_0)^2 - (\partial_1 r \partial_2 y_0 - \partial_2 r \partial_1 y_0)^2},
\]

(3.11)

where \(r = r(y_1, y_2)\), \(\partial_1 r = \partial_{y_1} r\), and so forth. We then need to find a solution to the classical equations of motion with the appropriate boundary conditions. Let us consider first the case with \(s = t\), where the projection of the Wilson lines is a square. By scale invariance we can choose the edges of the square to be at \(y_1, y_2 = \pm 1\). The boundary conditions are then given by

\[
r(\pm 1, y_2) = r(y_1, \pm 1) = 0, \quad y_0(\pm 1, y_2) = \pm y_2, \quad y_0(y_1, \pm 1) = \pm y_1.
\]

(3.12)

Near each of the cusps the form of the solution should reduce to the single cusp solution (3.10). Making educated guesses satisfying the boundary conditions and with the correct properties near the cusps we propose

\[
y_0(y_1, y_2) = y_1 y_2, \quad r(y_1, y_2) = \sqrt{(1 - y_1^2)(1 - y_2^2)}.
\]

(3.13)

Remarkably it turns out to be a solution of the equations of motion.

On dimensional grounds, the finite part of the amplitude is some function of the form \(f(s/t)\). Hence, in order to capture the nontrivial dependence of the amplitude on the kinematics, we need to consider more general solutions with \(s \neq t\). In this case the projection of the surface to the \((y_1, y_2)\) plane will not be a square but a rhombus, with \(s\) and \(t\) given by the square of the distances between opposite vertices.

In order to find the solution for this more general case, it is instructive to study the surface (3.13) in terms of embedding coordinates. These are coordinates where we view \(\text{AdS}_5\) as the following surface embedded in \(\mathbb{R}^{2,4}\):

\[
-Y_{-1}^2 - Y_0^2 + Y_1^2 + Y_2^2 + Y_3^2 + Y_4^2 = -1.
\]

(3.14)
The relation between embedding and Poincare coordinates is

\[ Y^\mu = \frac{y^\mu}{r}, \quad \mu = 0, \ldots, 3, \]

\[ Y_1 + Y_4 = \frac{1}{r}, \quad Y_1 + Y_4 = \frac{r^2 + y_\mu y^\mu}{r}. \]  

(3.15)

The surface (3.13) is then given by

\[ Y_0 Y_{-1} = Y_1 Y_2, \quad Y_3 = Y_4 = 0. \]  

(3.16)

Once we have written our solution in embedding coordinates, we notice that we can apply \( SO(2,4) \) transformations, that are linearly realized in this coordinates, in order to obtain new solutions. This \( SO(2,4) \) symmetry is sometimes referred to as “dual conformal symmetry” and should not be confused with the original \( SO(2,4) \) symmetry associated to the original AdS space. Later we will comment on the implications of this symmetry.

Solutions with \( s \neq t \) can be obtained by starting from (3.13) and performing a boost in the \( 0-4 \) direction

\[ Y_0 Y_{-1} = Y_1 Y_2, \quad Y_4 = 0 \rightarrow Y_4 - v Y_0 = 0, \quad \sqrt{1 - v^2} Y_0 Y_{-1} = Y_1, Y_2. \]  

(3.17)

Hence, we end up with a two-parameter solution, the parameter \( a \), related to the size of the initial square and the parameter \( b \), related to the boost parameter. The solution can be conveniently written as

\[ r = \frac{a}{\cosh u_1 \cosh u_2 + b \sinh u_1 \sinh u_2}, \quad y_0 = \frac{a \sqrt{1 + b^2} \sinh u_1 \sinh u_2}{\cosh u_1 \cosh u_2 + b \sinh u_1 \sinh u_2}, \]  

(3.18)

\[ y_1 = \frac{a \sinh u_1 \cosh u_2}{\cosh u_1 \cosh u_2 + b \sinh u_1 \sinh u_2}, \quad y_2 = \frac{a \cosh u_1 \sinh u_2}{\cosh u_1 \cosh u_2 + b \sinh u_1 \sinh u_2}, \]  

(3.19)

where we have written the surface as a solution to the equations of motion in conformal gauge

\[ iS = -\frac{R^2}{2\pi} \int \mathcal{L} = -\frac{R^2}{2\pi} \int du_1 du_2 \frac{1}{2} \frac{(\partial r \partial y + \partial y_\mu \partial y^\mu)}{r^2}; \]  

(3.20)

\( a \) and \( b \) encode the kinematical information of the scattering as follows:

\[ s(2\pi)^2 = \frac{8a^2}{(1-b)^2}, \quad t(2\pi)^2 = \frac{8a^2}{(1+b)^2}, \quad \frac{s}{t} = \frac{(1+b)^2}{(1-b)^2}. \]  

(3.21)

According to the prescription, we should now plug the classical solution into the classical action in order to obtain the four-point scattering amplitude at strong coupling.
However, in doing so, we obtain a divergent answer. That is of course the case, since we have taken the IR regulator away. In order to obtain a finite answer we need to reintroduce a regulator. There are two basic possibilities.

(i) The first one, suggested by the strong coupling prescription, would be to set the boundary conditions at some fixed value of the radial direction $r = r_0$, namely, in (3.12) $r(\pm 1, y_2) = r(y_1, \pm 1) = r_0$, and then take $r_0 \to 0$; however, it seems pretty hard to find classical solutions with such boundary conditions. (Something that one could do is to plug the classical solution into the action and cut the integral giving the area at a finite radial distance, as done in [21], this will give the correct finite piece for the amplitude up to terms independent of the kinematics. However, It will not give the correct collinear anomalous dimension at strong coupling [34].)

(ii) The second one is to try to mimic what was done in perturbative computations and consider the amplitudes in $D = 4 - 2\epsilon$ dimensions. This route is quite convenient and besides it will allow a more direct comparison with the perturbative results.

3.2.3. Dimensional Regularization

Gauge theory amplitudes are regularized by considering the theory in $D = 4 - 2\epsilon$ dimensions. More precisely, one starts with $\mathcal{N} = 1$ in ten dimensions and then dimensionally reduce to $4 - 2\epsilon$ dimensions. For integer $2\epsilon$ this is precisely the low energy theory living on a $Dp$-brane, where $p = 3 - 2\epsilon$. We regularize the amplitudes at strong coupling by considering the gravity dual of these theories and then analytically continuing in $\epsilon$. The string frame metric is

$$ds^2 = f^{-1/2} dx_4^2 + f^{1/2} \left[ dr^2 + r^2 d\Omega_{5+2\epsilon}^2 \right], \quad f = \left( 4\pi^2 e^\epsilon \right)^\epsilon \Gamma(2 + \epsilon) \mu^{2\epsilon} \frac{\lambda}{r^{4+2\epsilon}}. \quad (3.22)$$

Following the steps described above, we are led to the following action:

$$S = \frac{\sqrt{c} \lambda \mu^{\epsilon}}{2\pi} \int \frac{L_{\epsilon=0}}{r^\epsilon}, \quad (3.23)$$

where $L_{\epsilon=0}$ is the Lagrangian density for AdS$_5$. The presence of the factor $r^\epsilon$ will have two important effects. On one hand, previously divergent integrals will now converge (if $\epsilon < 0$). On the other hand, the equations of motion will now depend on $\epsilon$, and we were not able to compute the full solution for arbitrary $\epsilon$. However, we are interested in computing the amplitude only up to finite terms as we take $\epsilon \to 0$. In that case, it turns out to be sufficient to plug the original solution into the $\epsilon$-deformed action (up to a contribution from the regions close to the cusps that add an unimportant additional constant term). After performing the integrals we obtain

$$S \approx \sqrt{\frac{\Gamma^{\epsilon}}{a^\epsilon}} \frac{\mu^{\epsilon}}{2\pi} \frac{1}{2} \left[ \frac{1}{2} - \frac{\epsilon}{2}, \frac{1}{2} - \frac{\epsilon}{2}, \frac{1}{2} ; b^2 \right]. \quad (3.24)$$
Expanding in powers of $\epsilon$ we get the final answer

$$A = e^{-\sqrt{\lambda}/2\pi} A, \quad -\frac{\sqrt{\lambda}}{2\pi} A = iS_{\text{div}} + \frac{\sqrt{\lambda}}{8\pi} \left( \log \frac{s}{t} \right)^2 + \tilde{C},$$

$$S_{\text{div}} = 2S_{\text{div},s} + 2S_{\text{div,t}},$$

$$iS_{\text{div},s} = -\frac{1}{\epsilon^2} \frac{1}{2\pi} \sqrt{\frac{\lambda \mu^2}{(-s)^2}} - \frac{1}{\epsilon} \frac{1}{4\pi} (1 - \log 2) \sqrt{\frac{\lambda \mu^2}{(-s)^2}}.$$

This should be compared with the field theory expectations (2.7). We notice that the general structure is in perfect agreement with the BDS ansatz. Once we use the strong coupling behavior for the cusp anomalous dimension $f(\lambda) = \sqrt{\lambda}/\pi + \cdots$, we see that the leading divergence, as well as the finite piece, has the correct kinematical dependence with the correct overall coefficient.

### 3.3. A Conformal Ward Identity

An important ingredient of the previous computation was the existence of a dual $SO(2,4)$ symmetry (Actually, this symmetry was first noticed in perturbative computations [36] and then independently in the strong coupling computation described here.), associated to the isometry group of the dual AdS$_5$ space. This symmetry allowed the construction of new solutions and fixed somehow the finite piece of the scattering amplitude. Naively, this conformal symmetry would imply that the amplitude is independent of $s$ and $t$, since they can be sent to arbitrary values by a dual conformal transformation. The whole dependence on $s$ and $t$ arises due to the necessity of introducing an IR regulator. However after keeping track of the dependence on the IR regulator, the amplitude is still determined by the dual conformal symmetry. Hence, this regulator breaks the dual conformal symmetry, but in a controlled way:

We can write a generic regularized area as follows

$$A_{\text{reg}}^n = \text{Div} + \text{Fin}(x_i).$$

As already mentioned, the structure of IR divergences of scattering amplitudes is known to all loops. As a result, the divergent piece of the area is well known. Here the $x_i$ denotes the position of the cusps. To a symmetry we associate a Ward identity, that will impose certain constraints on scattering amplitudes. The dual conformal symmetry under consideration will impose the following equation for the finite part of the area [19–21]

$$\sum_i \left( 2x_i^\mu (x_i \cdot \partial_{x_i}) - x_i^2 \partial_{x_i}^\mu \right) \text{Fin}(x_i) = f(\lambda) \sum_i \left( x_i^\mu + x_{i+1}^\mu - 2x_i^\mu \right) \log(x_{i+1,i+1}^2),$$

where $f(\lambda)$ is the value of the cusp anomalous dimension at strong coupling (properly normalized). This conformal Ward identity (3.27) was first written down in [19]. For the particular case of $n = 4$ and $n = 5$, the solution to this equation is actually unique; hence, the finite piece of the scattering of four and five gluons is fixed in case dual conformal symmetry persists to all values of the coupling constant.
What are the consequences of this symmetry at strong coupling? The Ward identity (3.27) implies that the strong coupling answer can be written as a solution to the above equation plus a solution of the homogeneous equation, namely, an arbitrary function of the cross-ratios. On the other hand, the one loop answer multiplied by the appropriate factor is a solution of the conformal Ward identities as it should be, since the symmetry is present at one loop. In the literature, this term is referred to as the BDS factor, hence

$$A_{\text{strong}} = A_{\text{BDS}} + F\left(\frac{x_{ij}^2 x_{kl}^2}{x_{ik}^2 x_{jl}^2}\right),$$

where $F$ is the so-called remainder function. The BDS ansatz, if true, would imply that $F$ is a constant. On the other hand, note that this is indeed the case for $n = 4, 5$, since in this case we cannot construct cross-ratios of the form $x_{ij}^2 x_{kl}^2/x_{ik}^2 x_{jl}^2$. On the other hand, we can construct such cross-ratios for $n = 6$, so in this case we can have a nontrivial remainder function. By explicit computations it was actually checked that such remainder function is indeed nontrivial at two loops [6, 22–24] and also at strong coupling [25, 26].

4. Minimal Surfaces on AdS

In the previous section we have shown how the problem of computing gluon scattering amplitudes at strong coupling reduces to the computation of the area of certain minimal surfaces in AdS. In this section we show how the integrability of the system can be used in order to give a solution to the problem, in the form of a set of integral equations. We will follow closely [26–29], to which we refer the readers for the details.

4.1. Mathematical Problem

The mathematical problem is to find the area of the minimal surface ending on the boundary of AdS at a polygon of light-like edges. Such polygon is parameterized by the position of its cusps $x_i$, which satisfy $x_{i,i+1}^2 = 0$. See Figure 6

As will be clear below, we will focus on certain regularized area that is invariant under conformal transformations. As such, it will depend only on cross-ratios, of the form $\chi_{ijkl} = x_{ij}^2 x_{kl}^2/x_{ik}^2 x_{jl}^2$. The mathematical problem is then to compute the minimal area as a function of such cross-ratios. As already mentioned, the full problem involves minimal surfaces on $\text{AdS}_5$; in which case there are $3N - 15$ cross-ratios if we are dealing with the scattering of $N$ gluons. However, for most of the exposition presented in this paper, we will restrict to special kinematical configurations in which the minimal surfaces involved are embedded in $\text{AdS}_3$; in this case, we have $N - 6$ cross-ratios (For the general scattering in four dimensions we have $4N$ coordinates, minus $N$, since the distance between consecutive points has to be light-like, minus 15, that is the dimension of the conformal group $SO(2,4)$. In the case of $\text{AdS}_3$, we have $2N - N$ minus 6, which is the dimension of $SO(2,2)$.), and the polygon is a zig-zaged polygon living in one-plus-one dimensions, see Figure 7

If we wanted a closed polygon we could consider the above contour on a cylinder. Note that we can consider only polygons with an even number of sides, hence $N = 2m$. As one can see in Figure 7, the contour is parameterized by $n x_j^+$ coordinates and $n x_i^-$ coordinates.
Figure 6: The light-like polygon in which the minimal surface ends is parameterized by the position of the cusps, $x_i$, with $x_{i+1}^+ = 0$.

Figure 7: A zig-zaged Wilson loop in $1 + 1$ dimensions is parameterized by $nx_i^+$ coordinates and $nx_i^-$ coordinates. If you want a closed polygon, you can fold the figure in a cylinder.

With each set we can form $n - 3$ invariant cross-ratios, of the form $\chi_{ijkl}^\pm = x_{ij}^\pm x_{kl}^\pm / x_{ik}^\pm x_{jl}^\pm$ and all together we have the expected $2n - 6$ cross-ratios.

In order to consider minimal surfaces in AdS$_3$ we need to consider the world-sheet of classical strings on AdS$_3$. This is the subject of the following subsection.

### 4.2. Strings on AdS$_3$

Classical strings on AdS$_3$ can be described in terms of embedding coordinates, where we see AdS$_3$ as the following surface embedded in $R^{2,2}$:

$$Y \cdot Y \equiv -Y_0^2 + Y_1^2 + Y_2^2 = -1.$$ (4.1)

We take the world-sheet to be the whole complex plane and $Y = Y(z, \bar{z})$. Since we are interested in classical solutions, the fields have to satisfy the conformal gauge equations of motion

$$\bar{\partial} \partial Y - (\partial Y \cdot \bar{\partial} Y) Y = 0.$$ (4.2)
Furthermore, the equations of motion are supplemented by the Virasoro constraints

$$\partial Y \cdot \partial Y = \bar{\partial} Y \cdot \bar{\partial} Y = 0. \quad (4.3)$$

An efficient way to focus only on the physical degrees of freedom is by performing a so-called Pohlmeyer kind of reduction; see, for instance, [30–32] and consider the “reduced” fields

$$\alpha = \log \partial Y \cdot \bar{\partial} Y, \quad p^2 = \partial^2 Y \cdot \bar{\partial}^2 Y. \quad (4.4)$$

As a consequence of the equations of motion and Virasoro constraints, $p$ can be seen to be a holomorphic function, and $p = p(z)$ and $\alpha(z, \bar{z})$ can be seen to satisfy a generalized version of the Sinh-Gordon equation

$$\partial \bar{\partial} \alpha - e^\alpha + p(z)\bar{p}(\bar{z})e^{-\alpha} = 0. \quad (4.5)$$

Hence, we have reduced the initial set of equations to a holomorphic function $p(z)$ and to a field $\alpha(z, \bar{z})$ which satisfies a generalized Sinh-Gordon equation. Eventually, we will be interested in the area of the classical world-sheet. This can be simply written as the conformal gauge action in terms of the reduced fields

$$\mathcal{A} = \int e^\alpha d^2z. \quad (4.6)$$

An important property of the reduced fields is that they are invariant under space-time conformal transformation. In other words, they contain only the “nontrivial” information about the space-time solution.

Summarizing, a solution of classical strings on AdS3 is mapped to a holomorphic function $p(z)$ and to a field $\alpha(z, \bar{z})$ which satisfies (4.5). We can also wonder whether given a holomorphic function $p(z)$ and a field $\alpha(z, \bar{z})$ satisfying (4.5) it is possible to construct a solution of classical strings on AdS3. The answer is affirmative, and there is a well-know procedure to construct such solution. What we need to do is to solve two auxiliary linear problems, which we denote as left and right

$$\left( d + B^L \right) q^L_{\dot{a}} = 0, \quad \left( d + B^R \right) q^R_{\dot{a}} = 0, \quad (4.7)$$

where the flat connections $B^{L,R}$ are simple two-by-two matrices constructed from $p(z)$ and $\alpha(z, \bar{z})$, for instance

$$B^L_z = \begin{pmatrix} \frac{\partial \alpha}{4} & \frac{1}{\sqrt{2}} e^{\alpha/2} \\ \frac{1}{\sqrt{2}} e^{-\alpha/2} & -\frac{\partial \alpha}{4} \end{pmatrix}; \quad (4.8)$$

we will denote the different components of the connections by $B^L_{\dot{a} \dot{b}}$ and $B^R_{\dot{a} \dot{b}}$. On the other hand, the indices $a$ and $\dot{a}$ in (4.7) denote independent solutions of the auxiliary linear
problems. Each \(q_{a}^{L}\) or \(q_{a}^{R}\) is then a doublet. We denote the components of this doublet by \(q_{a}^{L,\alpha}\) and so forth.

Given then the solutions of these two auxiliary linear problems, one can show that the space-time solution is simply given by

\[
Y_{a,\dot{a}}(z) = \frac{Y_{1} - Y_{0}}{Y_{0} + Y_{1}} = q_{a}^{L,\alpha}M_{\alpha,\beta}q_{\beta,\dot{a}}^{R}
\]

for some constant matrix \(M\) (which can be taken, for instance, to be the identity matrix). One can show that \(Y\) constructed this way satisfies all the required properties. In order to write down the answer, we have written the space-time coordinates in an \(SL(2) \times SL(2) \approx SO(2,2)\) notation.

An important observation will be very relevant for what we will do next: it is possible to introduce a spectral parameter \(\zeta\) in such a way that the flat connection is still flat, more precisely

\[
B_{z} \rightarrow B_{z}(\zeta) = \frac{1}{4} \left( \begin{array}{cc}
\partial_{\alpha} & 0 \\
0 & -\partial_{\alpha}
\end{array} \right) + \frac{1}{\zeta} \left( \begin{array}{cc}
1 & 1 \\
\sqrt{2} & \sqrt{2}
\end{array} \right) \left( \begin{array}{cc}
0 & e^{a/2} \\
e^{-a/2} & 0
\end{array} \right),
\]

\[
B_{\bar{z}} \rightarrow B_{\bar{z}}(\zeta) = \frac{1}{4} \left( \begin{array}{cc}
\bar{\partial}_{\alpha} & 0 \\
0 & \bar{\partial}_{\alpha}
\end{array} \right) + \frac{1}{\bar{\zeta}} \left( \begin{array}{cc}
1 & 1 \\
\sqrt{2} & \sqrt{2}
\end{array} \right) \left( \begin{array}{cc}
0 & \bar{e}^{a/2} \\
\bar{e}^{-a/2} & 0
\end{array} \right).
\]

One can check that the connections are still flat, namely, they satisfy

\[
dB + B \wedge B = \partial B_{z} - \bar{\partial} B_{z} + [B_{z}, B_{\bar{z}}] = 0,
\]

for any value of \(\zeta\). Another important fact is that the left and right connections are just particular cases of the above flat connection, more precisely

\[
B(\zeta = 1) = B^{L}, \quad B(\zeta = i) = B^{R}.
\]

Before proceeding, let us make the following remark. Since \(p(z)\) is a holomorphic function, it is possible to make a change of coordinates from the \(z\)-plane to the \(w\)-plane, where \(dw = \sqrt{p(z)}dz\). In the \(w\)-plane, the generalized Sinh-Gordon equation takes the usual form (after a simple field redefinition of \(a\))

\[
a = \tilde{a} + \frac{1}{4} \log p\bar{p} \rightarrow \partial_{w} \bar{\partial}_{w} \tilde{a} = \sinh 2\tilde{a}.
\]
questions we want to answer, one description may be more convenient than another. In the \(w\)-plane, the area is simply given by

\[
A = \int e^\alpha d^2w.
\]  

(4.14)

### 4.3. Classical Solutions Corresponding to Scattering Amplitudes

A natural question to ask is what are the properties of the holomorphic function \(p(z)\) and \(a(z, \bar{z})\) for solutions corresponding to scattering amplitudes. In order to gain some intuition we can consider the four cusps solution found in Section 3 and perform the Pohlmeyer reduction. We find

\[
p(z) = 1, \quad \alpha = \tilde{\alpha} = 0.
\]  

(4.15)

Hence, the four cusps solution simply corresponds to the vacuum solution of the Sinh-Gordon equation. What about solutions with a higher number of cusps? First of all we propose that the field \(\alpha\) is regular everywhere, since we are looking for smooth space-like solutions. Second, we expect a general solution to be similar to the four cusps solution when approaching the boundary, which corresponds to very large \(|z|\). So we expect that \(\tilde{\alpha} \to 0\) as \(|z|\) becomes large. Finally, if we are interested in a minimal surface ending on a polygon with \(2n\) cusps, we propose \(p(z)\) to be a polynomial of degree \(n - 2\):

\[
p(z) = z^{n-2} + c_{n-4}z^{n-4} + \cdots + c_0;
\]  

(4.16)

we have used rescalings and translations in order to set the coefficients of \(z^{n-2}\) and \(z^{n-3}\) to one and zero respectively. We see that such polynomial contains \(n-3\) complex coefficients, or \(2n-6\) real coefficients, which exactly agrees with the amount of expected independent cross-ratios. In conclusion, minimal surfaces ending on a light-like polygon with \(2n\) cusps correspond to a holomorphic polynomial of degree \(n - 2\) and a field \(\tilde{\alpha}\) satisfying the Sinh-Gordon equations, with boundary conditions such that it decays at infinity and diverges logarithmically at the zeroes of \(p(z)\).

Note that since \(\tilde{\alpha}\) decays at infinity, the integral defining the area (4.14) diverges. We define a regularized/finite area by subtracting the behavior at infinity

\[
A_{\text{reg}} = \int (e^{\tilde{\alpha}} - 1) d^2w.
\]  

(4.17)

The computation of this regularized area will be the main focus of the remaining of this paper. As already mentioned, all the fields entering into the expression for the regularized area are invariant under space-time conformal transformations. Hence, the regularized area will be a function of the cross-ratios only. (The full answer would include also the integral of the one we have subtracted. In order to compute it one would need to introduce a physical regulator and this part of the answer will not be conformal invariant. Anyway, its explicitly form can be worked out and turns out to be quite universal. In this paper we will focus on the “interesting” part of the answer \(A_{\text{reg}}\).)
4.3.1. Behavior of the Solution at Infinity

In the following, we want to understand the behavior of the solutions of the linear auxiliary problems for very large values of $|z|$, or $|w|$, since this will tell us the behavior of the worldsheet near its boundary. Suppose, by simplicity, that $p(z) = z^{n-2}$, hence $w \approx z^{n/2}$. As a result, as we go once around the $z$-plane, we go around the $w$-plane $n/2$ times.

Since for large values of $|w|$ $\tilde{\alpha}$ vanishes, the flat connections $B^{L-R}$ drastically are simplified and we can solve for the auxiliary linear problems. A general solution will be of the form

$$\begin{align*}
\psi^L_a &\approx c^+_a \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{w+\bar{w}} + c^-_a \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-(w+\bar{w})}, \\
\psi^R_a &\approx d^+_a \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{(w-\bar{w})/i} + d^-_a \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-(w-\bar{w})/i}.
\end{align*}
\tag{4.18}$$

We see that the $w$-plane divides naturally into quadrants; see Figure 8. In each quadrant one of the two solutions of each problem (left and right) dominates. For instance, if we focus on the upper right quadrant, the solution proportional to $c^+_a$ dominates in the left problem while the solution proportional to $d^+_a$ dominates in the right problem. For instance, for the first quadrant we have $Y_{aa} \approx (\text{Large}) \times c^+_a d^+_a$ at large values of $|w|$. This means that for large values of $|w|$, the whole quadrant corresponds to a single point in the boundary, which corresponds to a cusp.

Figure 8: When looking at the left problem, (each sheet of) the $w$-plane is naturally divided into two parts, according to the sign of $\text{Re}(w)$. In the same way, when looking at the right problem, the $w$-plane is naturally divided into two parts, according to the sign of $\text{Im}(w)$. Hence, the $w$-plane is naturally divided into four quadrants. Large values of $w$ in each of these angular sectors correspond to a cusp.
As we change quadrant, one (and only one) of the two dominant solution changes and we jump a light-like distance to the next cusp. Note that in each cusp we can write

$$Y_{a,a} \approx \lambda_a \tilde{\lambda}_a,$$  \hspace{1cm} (4.19)

where $\lambda$ is given by the leading contribution to the left problem and $\tilde{\lambda}$ by the leading contribution to the right problem. As we change quadrant, one of the two solutions, $\lambda$ or $\tilde{\lambda}$, changes. As we go around the $w$-plane $n/2$ times, we get the expected $2n$ cusps.

The above solution, in which $p(z)$ is a homogeneous polynomial, turns out to correspond to a regular polygon with an additional $Z_{2n}$ symmetry. In the general case in which the polynomial is not homogeneous, we get a polygon of different shapes but with the same number of cusps. Hence the degree of the polynomial determines the number of cusps while the coefficients on the polynomial determine the shape of the polygon.

### 4.4. $Y$-System for Minimal Surfaces

Let us focus on the left problem. We see that each sheet on the $w$-plane is naturally divided into two sectors, one with $\text{Re}(w) > 0$ and the other with $\text{Re}(w) < 0$. Note the following simple but important fact: in each sector the small solution is well defined (up to a normalization constant), on the other hand, the large solution is not, as we can add to it a part of the small solution. Let us then introduce the following terminology.

(i) The $w$-plane is divided into $n$ sectors, since each sheet contains two sectors. We label this sectors by $i = 0, \ldots, n-1$.

(ii) We call $s^L_i$ the small solution at the $i$th sector. By definition, this is the solution with the fastest decay along the line in the center of the $i$th sector, for large $|w|$. For instance, given solution (4.18), we see that the term proportional to $e^{-(\omega \pm i\pi)}$ has the fastest decay as we move along the positive real axis in the $w$-plane. Such small solution is well defined, up to a normalization constant.

In order to understand why these small solutions are important, we need to introduce a new element. Given that our connections are $SL(2)$ matrices, we can introduce an $SL(2)$ invariant product

$$q^L_a \wedge q^L_b \equiv \epsilon^{a \beta} q^L_{\alpha, a} q^L_{\beta, b}.$$  \hspace{1cm} (4.20)

Given a solution to the linear problem, also called a flat section, one can explicitly check that the above product is independent of the world-sheet coordinate $z$. Hence, we can normalize our solutions such that

$$q^L_a \wedge q^L_b = \epsilon_{ab}.$$  \hspace{1cm} (4.21)

As already seen, the location of the cusps is determined by the large solutions. The large component of a solution, on a given sector, can be extracted by using the small solution on such sector and the $SL(2)$ invariant product just introduced, more precisely

$$q^L_a \wedge s^L_i \approx \lambda^i_a.$$  \hspace{1cm} (4.22)
How do we construct space-time cross-ratios? We have seen that the location of the cusps is given by \( Y^a_{\bar{a}} = \lambda_{\bar{a}} \bar{\lambda}^i \). The space-time cross-ratios involve distances like

\[
Y^i \cdot Y^j = e^{ab} e^{\bar{a}b} Y^i_{\bar{a}} Y^j_{\bar{b}} = \left( \lambda^i \bar{\lambda}^j \right) \left( \bar{\lambda}^i \lambda^j \right), \quad \left( \lambda^i \bar{\lambda}^j \right) = e^{ab} \lambda_{\bar{a}} \lambda^i_{\bar{b}}.
\]

When seeing \( y^I_a \) as a two-by-two matrix, the normalization condition (4.21) simply means that the determinant of such matrix is one. Given this, one can easily show that

\[
\left\langle \lambda^i \bar{\lambda}^j \right\rangle \approx \left( \left( q^{\bar{a}}_I \wedge s_{\bar{i}}^I \right) \left( q^{\bar{a}}_I \wedge s_{\bar{i}}^I \right) \right) = s_I^i \wedge s_I^j.
\]

Which means that space-time cross-ratios can be constructed from inner products of the small solutions in the corresponding sectors, more precisely

\[
\frac{x^{+}_{ij} x^{+}_{kl}}{x^{+}_{ik} x^{+}_{jl}} = \left( \frac{s_i^k \wedge s_i^j}{s_i^k \wedge s_j^i} \right) \left( \frac{s_k^j \wedge s_i^j}{s_k^j \wedge s_j^i} \right)^{-1}.
\]

We mentioned that small solutions are defined up to a normalization constant. Note that such normalization constant would cancel out when computing cross-ratios.

The strategy we will follow is to introduce the spectral parameter \( \zeta \) as shown in the previous section and to study the small flat sections of the corresponding connection

\[
(d + B(\zeta)) s_i(\zeta) = 0; \quad (d + B(\zeta)) s_i(\zeta) = \sigma_3 B(\zeta) \sigma_3.
\]

then, we can consider the cross-ratios as a function of such spectral parameter

\[
\chi_{ijkl}(\zeta) = \frac{(s_i \wedge s_j)(s_k \wedge s_l)}{(s_i \wedge s_k)(s_j \wedge s_l)}.
\]

Then, if we set \( \zeta = 1 \) we obtain the physical plus/left cross-ratios while setting \( \zeta = i \) we obtain the physical minus/right cross-ratios:

\[
\chi_{ijkl}(\zeta = 1) = \chi_{ijkl}^+, \quad \chi_{ijkl}(\zeta = i) = \chi_{ijkl}^-.
\]

A property of the flat connection \( B(\zeta) \) that will be used in what follows is that it possesses a \( \mathbb{Z}_2 \) symmetry

\[
B(e^{i\pi} \zeta) = \sigma_3 B(\zeta) \sigma_3.
\]
where $\sigma_3$ is the usual Pauli matrix. This symmetry allows to relate small sections at different values of the spectral parameter, for instance, $s_{i+1}(\zeta) = \sigma_3 s_i(e^{i\pi} \zeta)$, and in particular, it implies

$$s_i \wedge s_j \left(e^{i\pi} \zeta\right) = s_{i+1} \wedge s_{j+1}(\zeta).$$

(4.30)

Besides, in order to simplify the subsequent derivation, we will assume that $s_i \wedge s_{i+1} = 1$.

Now we have all the elements to derive the so-called Hirota equations and subsequently the $Y$-system equations. The trick is to choose $s_0$ and $s_1$ as a complete basis of small flat sections and express two arbitrary consecutive flat sections $s_k$ and $s_{k+1}$ in terms of them:

$$s_k = (s_k \wedge s_1)s_0 - (s_k \wedge s_0)s_1,$$

$$s_{k+1} = (s_{k+1} \wedge s_1)s_0 - (s_{k+1} \wedge s_0)s_1.$$  

(4.31)

Next, we use (4.30) in order to express every wedge as a wedge involving $s_0$ and consider $1 = s_k \wedge s_{k+1}$; then we obtain

$$-(s_{k-1} \wedge s_0)^{++}(s_{k+1} \wedge s_0) + (s_k \wedge s_0)^{++}(s_k \wedge s_0) = 1,$$

(4.32)

where we have introduced the notations $f^+ = f(e^{\pm i\pi/2} \zeta)$, $f^{++} = f(e^{i\pi} \zeta)$, and so forth. This equation can be written in a more suggestive way by introducing $T_k = s_0 \wedge s_{k+1}(e^{-i(k+1)\pi/2} \zeta)$. We obtain

$$T_s^+ T_s^- = T_{s+1} T_{s-1} + 1$$

(4.33)

which has the form of the so-called Hirota equations. From the definition of $T_s$, we see that it is nontrivial for $s = 0, \ldots, n - 2$. The $Y$-system equations can be simply obtained by introducing $Y_s \equiv T_{s-1} T_{s+1}$; we then obtain

$$Y_s^+ Y_s^- = (1 + Y_{s+1})(1 + Y_{s-1}).$$

(4.34)

$Y_s$ is nontrivial for $s = 1, \ldots, n - 3$. Notice that this agrees with the amount of (complex) cross-ratios of our scattering problem. These are the $Y$-system equations. Note that they have simply followed from a chain of rather trivial facts. These are functional equations for $Y_s(\zeta)$ and are valid for any value of $\zeta$. One could reintroduce the normalized factors $s_i \wedge s_{i+1}$ and check that the $Y$-functions are given by the usual cross-ratios introduced above. The physical cross-ratios are then obtained by evaluating $Y_s(\zeta)$ at $\zeta = 1$ and $\zeta = i$.

Such equations were obtained by rather trivial considerations, and we do not expect them to be the whole story. In particular, note that the information of the holomorphic function $p(z)$ does not enter in such equations. There are many solutions to such equations. The correct solution is then picked by specifying the analytic properties and boundary conditions of $Y_s(\zeta)$ as we move on the $\zeta$-plane. This will be the subject of the following section.
4.5. Integral Equations

In order to pick the appropriate solution to the $Y$-system equations (4.34) we need to specify the analytic properties of $Y_s(\zeta)$. By analyzing the auxiliary linear problem and the definition of $Y_s(\zeta)$ one can show that

(i) $Y_s(\zeta)$ are analytically away from $\zeta = 0, \infty$,

(ii) as $\zeta \to 0, \infty$, the flat connection is simplified and the inverse problem can be solved by using a WKB approximation, where the role of $\hbar$ is played by $\zeta$ or $1/\zeta$.

More precisely, by calling $\zeta = e^\theta$, one can show that for large $\theta$

$$\log Y_s \approx -m_a \cosh \theta + \cdots,$$

(4.35)

where $m_a$ is given by the periods of $p(z)^{1/2}$ along the cycles $\gamma_a$, more precisely $m_a \approx -\int_{\gamma_a} \sqrt{p(z)} dz$. This is how the information of the polynomial $p(z)$ enters into the problem. These periods are usually complex, and there are $n - 3$ of them, which exactly agrees with the quantity of expected cross-ratios.

The strategy now is well known from the study of integrable systems. We can combine the $Y$-system equations with the analytic properties and boundary conditions for the $Y$-functions, in order to write a system of integral equations for them, whose solutions will automatically satisfy the $Y$-system equations and have the required boundary conditions.

The system of integral equations is given by

$$\log Y_s = -m_s \cosh \theta + K \star \log(1 + Y_{s+1})(1 + Y_{s-1}),$$

(4.36)

where the convolution operation is defined by

$$K \star f(\theta) = \int_{-\infty}^{\infty} d\theta' K(\theta - \theta') f(\theta'), \quad K(\theta) = \frac{1}{2\pi} \frac{1}{\cosh \theta}.$$

(4.37)

The system of equations (4.36) has the form of TBA equations, that arise when studying integrable models in finite volume; see, for example, [33]. Even though, for the sake of clarity, some coefficients have been suppressed in the derivation of these equations, the equations given have all the correct coefficients. Note that from the TBA point of view, the parameters $m_a$ enter as masses. Once the masses are given, the solution of the above system is unique. The physical cross-ratios can be read off by looking at $Y_s(\theta)$ for appropriate values of $\theta$.

For simplicity, in (4.36) we have assumed that all the masses are real and positive, and the same will be assumed for the rest of the paper. This can be achieved if all the zeroes of $p(z)$ are along the real line, such that $p(z)$ is positive for $z$ real, positive, and sufficiently large.

The last question is then how we compute the regularized area once we solve the above system of integral equations. It turns out that the area can be written in terms of the $Y$-functions and has the simple expression

$$A_{\text{reg}} = \sum_s \int d\theta \frac{m_s}{2\pi} \cosh \theta \log(1 + Y_s(\theta)).$$

(4.38)

This expression has exactly the form of the free energy of the TBA system.
The strategy to solve the full problem is then clear. The initial data (rather than the coefficients of the polynomial) are the masses $m_a$. For a given choice of these masses, we find a unique solution of the TBA integral equations. From this solution we can read off the space-time cross ratios and we can also compute the regularized area as the free energy of the TBA system. Hence, we obtain the regularized area for given space-time cross-ratios.

The above equations can be easily solved numerically, by an iterative method. See the appendix for more details.

### 4.6. Generalization to AdS$_5$

In the first part of this section we have seen how a solution to the problem of computing certain minimal surfaces in AdS$_3$ can be given by using the integrability of the system. In the following, we highlight the results for the general case of AdS$_5$ without deriving them. It turns out that, up to some subtleties, much of the previous analysis goes through, and the final results are a simple generalization of the results obtained for the AdS$_3$ case. We get a slightly more complicated $Y$-system for functions $Y_{a,s}$, with $a = 1, 2, 3$ and $s = 1, \ldots, N - 5$. The system reads

\[
\frac{Y_{-2,m}^+ Y_{2,m}^-}{Y_{1,m}^- Y_{-3,m}^+} = \frac{(1 + Y_{-2,m+1})(1 + Y_{2,m-1})}{(1 + Y_{1,m})(1 + Y_{-3,m})}, \quad m = 1, 2, \ldots, \quad (4.39)
\]

As before, in order to get a system of integral equations, we need to supplement these equations by the analytic properties of $Y_{a,s}(\zeta)$. As for the AdS$_3$ case, the large $\theta = \log \zeta$ behavior can be understood by a WKB analysis. The boundary conditions for the $Y$-functions are

\[
\begin{align*}
\log Y_{1,s} &\longrightarrow -m_s \cosh \theta - C_s \pm D_s + \cdots, \quad \theta \longrightarrow \pm \infty, \\
\log Y_{3,s} &\longrightarrow -m_s \cosh \theta - C_s \mp D_s + \cdots, \quad \theta \longrightarrow \pm \infty, \\
\log Y_{2,s} &\longrightarrow -\sqrt{2}m_s \cosh \theta + \cdots, \quad \theta \longrightarrow \pm \infty;
\end{align*}
\]

depending on the signature of space-time, the reality conditions of $C_s$ and $D_s$ are fixed. For instance, for (3,1) or (1,3) signature, the $D_s$ are real while the $C_s$ are purely imaginary. As for the AdS$_3$ case, these boundary conditions together with regularity imply the following set of integral equations:

\[
\begin{align*}
\log Y_{2,s} &= -m_s \sqrt{2} \cosh \theta - K_2 \ast \alpha_s - K_1 \ast \beta_s, \\
\log Y_{1,s} &= -m_s \cosh \theta - C_s - \frac{1}{2} K_2 \ast \beta_2 - K_1 \ast \alpha_s - \frac{1}{2} K_2 \ast \gamma_s, \\
\log Y_{3,s} &= -m_s \cosh \theta + C_s - \frac{1}{2} K_2 \ast \beta_2 - K_1 \ast \alpha_s + \frac{1}{2} K_2 \ast \gamma_s.
\end{align*}
\]
where

\[
\alpha_s = \log \left( \frac{1 + Y_{1,s}}{1 + Y_{2,s-1}} \right), \quad \gamma_s = \log \left( \frac{1 + Y_{1,s-1}}{1 + Y_{3,s-1}} \right),
\]

\[
\beta_s = \log \left( \frac{1 + Y_{2,s}}{1 + Y_{1,s-1} + Y_{3,s}} \right),
\]

and the kernels are given by

\[
K_1 = \frac{1}{2\pi \cosh \theta}, \quad K_2 = \frac{\sqrt{2} \cosh \theta}{\pi \cosh 2\theta}, \quad K_3 = \frac{i}{\pi} \tanh 2\theta.
\]

The data of the system of integral equations are the (in general complex) masses \( m_s \) and the purely imaginary constants \( C_s \). Since for the scattering of \( N \) gluons \( s = 1, \ldots, N - 5 \), we get \( 2(N - 5) + N - 5 = 3(N - 5) \) degrees of freedom, which exactly agrees with the expected amount of cross ratios.

Again, as for the case of the \( \text{AdS}_3 \) equations, these TBA equations can be solved iteratively. Once the solution is found, the space-time cross-ratios can be read off from the solution at particular values of \( \theta \). Furthermore, quite remarkably, the expression for the regularized area exactly agrees with the expression for the free energy of such TBA system

\[
A_{\text{reg}} = \sum_s \int d\theta \frac{m_s}{2\pi} \cosh \theta \log \left( \frac{(1 + Y_{1,s})(1 + Y_{3,s})(1 + Y_{2,s})}{\sqrt{2}} \right).
\]

This is the final solution for the general problem in \( \text{AdS}_5 \) (where for simplicity we have assumed real masses).

5. Conclusions

In the first part of this paper we have described how the computation of scattering amplitudes of planar MSYM at strong coupling reduces to a minimal area problem in \( \text{AdS} \). More precisely, one is to look for the area of minimal surfaces that is at the boundary on a light-like polygon whose edges are given by the momenta of the particles undergoing the scattering.

In the second part of this paper, we have focused on the above minimal area problem. The integrability of the system plays a crucial role in finding such solution. We have seen that it is convenient to introduce a one-parameter deformation (the spectral parameter \( \zeta \)) and study such deformed problem. One can then write down a system of functional equations, or \( Y \)-system, valid for any value of \( \zeta \). One can combine these functional equations with the knowledge of the analytic behavior of the \( Y \)-functions in the \( \zeta \)-plane, in order to write a set...
of integral equations which can be solved iteratively and give the desired answer. There are many directions one could try to follow.

(i) One possible direction would be to take the continuous limit of our construction, in order to study smooth Wilson loops.

(ii) It would be nice to find a physical connection between the integrable system that the TBA equations describe and the original integrable system.

(iii) Of course, the most interesting problem is the extension of this construction to the full quantum theory.

(iv) An interesting direction would also be to look for similar structures in perturbative computations.

Appendix

Numeric Implementation of TBA Equations for AdS$_3$

In this appendix we describe how to solve the TBA equations (4.36) numerically in Mathematica in a very simple way. We encourage the reader to try the code below by him/herself and to extend it to other cases, in order to understand how much better integral equations are compared to differential equations.

The algorithm is almost trivial: we simply iterate the integral equation plugging the $Y$-functions at iteration $k - 1$ in the right-hand side of (4.36) and reading from the left-hand side their values at the $k$th iteration, starting with the inhomogeneous terms in the TBA equations as the initial value (or zero in the previous step). We start by defining the kernel appearing in the integral equations,

$$K[x] = 1/(2 \pi \cosh[x]),$$

and specify how many gluons we want to consider. We do that by introducing a list of masses which appear in the integral equations. For example, let us focus on a polygon with 10 sides, so $n = 5$, and set the following numerical values of the masses:

$$m = \{1., 2.\}; M = \text{Length}[m].$$

We also introduce a cut-off for the several integrals and set the number of iterations. (The cut-off is chosen so that for the smallest mass we have $e^{-m \cosh(Cut)} = 10^{-8}$.)

$$\text{cut} = \text{ArcCosh}[8 \text{Log}[10] / \text{Min}[m]]; \text{ni} = 8.$$

At each iteration step we compute the new values of the $Y$-functions at a discrete set of points and construct the new $Y$-function as the function which interpolates through these points. For that we need an interpolating function and a function to perform the several integrals (The last command in the integration function ensures that we are really integrating something which is numerical once the integration variable $y$ takes some random numerical value. We chose 0.1 but anything would work.)

$$F[S_] := \text{FunctionInterpolation}[S, \{x, -\text{cut}, \text{cut}\}, \text{InterpolationPoints} -> 100]$$

$$\text{int}[S_] := \text{NIntegrate}[S, \{y, -\text{cut}, \text{cut}\}] /; \text{NumberQ}[S / . y -> .1].$$
This is all the machinery we need. The iterations are simply implemented by

\[
Y[a_,k_] := \begin{cases} (0^+) /; a == 0, & a == M+1, \\ \exp[-m[[i]] \cosh[x]] & a \neq 0 \end{cases}, \\
Y[i_,1] := \exp[-m[[i]] \cosh[x]] \\
Y[i_,k_] := Y[i,k] \\
F[\exp[-m[[i]] \cosh[x]] + \text{int}[K[x-y] \log((1+Y[i-1,k-1],[y])(1+Y[i+1,k-1],[y]))]].
\]

In the first line we set \( Y_0 = Y_{M+1} = 0 \); in the second line we set the seed values for the iteration procedure to be simply the asymptotic expressions for the \( Y \)-functions and in the last command we compute the \( Y \)-functions at the iteration step \( k \) using the right-hand side of the integral equations with the \( Y \)-functions at step \( k-1 \). We can now compute the regularized area from the free energy (4.38) of the TBA system and see how it converges to a particular value as we iterate (which gives a good idea of how good the numerics are)

\[
\text{energy} = \text{Table}[\int [\sum [\cosh[y] m[[i]] \log[1+Y[i,s][y]]/2, [i,M]]], [s,ni]] \text{ListPlot}[\text{energy}].
\]

For the masses chosen above we see that the free energy converges to 0.27. Finally, if we want to read the cross-ratios of the polygon whose area we have just computed, we simply need to evaluate the \( Y \)-functions at \( \theta = 0 \),

\[
\text{Table}[Y[i,ni][0], [i, M]].
\]

To compute the the free energy again for a different choice of masses we should first run \text{Clear}[Y].

**Acknowledgments**

The authors would like to thank Davide Gaiotto, Amit Sever, Pedro Vieira, and specially Juan Maldacena for collaboration on the material exposed in these lectures notes. This work was supported by U.S. Department of Energy Grant no. DE-FG02-90ER40542.

**References**


Submit your manuscripts at
http://www.hindawi.com