Wilson loops and minimal surfaces

Nadav Drukker*
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106
and Department of Physics, Princeton University, Princeton, New Jersey 08544

David J. Gross†
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106

Hirosi Ooguri‡
Department of Physics, University of California, Berkeley, California 94720-7300
and Lawrence Berkeley National Laboratory, Mail Stop 50A-5101, Berkeley, California 94720

(Received 13 May 1999; published 15 November 1999)

The AdS-CFT correspondence suggests that the Wilson loop of the large N gauge theory with $\mathcal{N}=4$ supersymmetry in four dimensions is described by a minimal surface in $AdS_5\times S^5$. We examine various aspects of this proposal, comparing gauge theory expectations with computations of minimal surfaces. There is a distinguished class of loops, which we call BPS loops, whose expectation values are free from ultraviolet divergence. We formulate the loop equation for such loops. To the extent that we have checked, the minimal surface in $AdS_5\times S^5$ gives a solution of the equation. We also discuss the zigzag symmetry of the loop operator. In the $\mathcal{N}=4$ gauge theory, we expect the zigzag symmetry to hold when the loop does not couple the scalar fields in the supermultiplet. We will show how this is realized for the minimal surface. [S0556-2821(99)08718-4]

PACS number(s): 11.15.Tk, 11.25.$-$w

I. INTRODUCTION

The remarkable duality between four dimensional supersymmetric gauge theories and type IIB string theory on an $AdS_5\times S^5$ background [1] has been studied extensively over the past year and a half. This conjecture is difficult to test. As with many dualities, it relates a weakly coupled string theory to a strongly coupled gauge theory. Weakly coupled string theory is well defined, even though there are technical problems in doing calculations with Ramond-Ramond backgrounds. But how can one compare the results to the gauge theory, which is strongly coupled? Even if there is no phase transition in going from weak to strong coupling in the gauge theory, there is little that can be said about the strongly coupled gauge theory. By virtue of nonrenormalization theorems, it is possible to calculate some quantities in perturbation theory and extrapolate to strong coupling. Such techniques, however, raise the question of whether these comparisons can be regarded as strong evidence for the conjecture or whether the result is dictated by symmetry alone.

Gauge theory without fermions has a nonperturbative formulation on the lattice. This allows one to define, if not compute, quantities at arbitrarily large bare couplings. The lattice formulation of gauge theory enables one to derive a rigorous form of the loop equation [2], for the large N limit of the theory. These equations are satisfied on the lattice and are solved by the master field of the theory. The only case where the loop equation has been explicitly solved is two dimensions, where the theory is soluble [3].

The loop equation can also be derived formally in the continuum field theory. It has been shown that the perturbative expansion of the theory yields a solution to the loop equation. This is also the case for supersymmetric theories. Thus, although there is no formulation of supersymmetric theories on the lattice, we assume that those theories still satisfy a large N loop equation. Since this equation holds for all couplings, we can use it for strong coupling as well. One of the goals of this paper is to check if the AdS_5 ansatz for the expectation value of the Wilson loop operator satisfies the loop equation. To the extent that we were able to reliably estimate properties of string in AdS_5, the loop equation is satisfied. However, we were unable to test them in all interesting cases. In the course of our investigation we will also learn new facts about Wilson loops and strings in anti–de Sitter space (Ads).

We discuss the best understood and most studied case of the AdS conformal field theory (CFT) correspondence between type IIB superstring on $AdS_5\times S^5$ and $\mathcal{N}=4$ super Yang-Mills theory with gauge group $SU(N)$ in four dimensions. We will concentrate on the case with Euclidean signature metric. Let us review some basic facts about this duality.[1]

The near horizon geometry of N D3-branes is given by the metric

*Email address: drukker@itp.ucsb.edu
†Email address: gross@itp.ucsb.edu
‡Email address: hooguri@lbl.gov

[1]For more complete reviews, see [4,5].

$$\frac{ds^2}{\alpha'}=\frac{U^2}{\sqrt{4\pi g_s N}}\sum_{\mu=0}^{3} dX^\mu dX^\mu + \sqrt{4\pi g_s N}\,\frac{dU^2}{U^2}+\sqrt{4\pi g_s N}d\Omega_5^2, \tag{1.1}$$

where g_s is the string coupling constant and the string tension is $(2\pi\alpha')^{-1}$. The background contains N units of Ramond-Ramond flux. The X and U are coordinates on AdS_5, and $d\Omega_5^2$ is the metric on S^5 with unit radius. The curvature radii of both AdS_5 and S^5 are given by $(4\pi g_s N)^{1/4} l_s$ where $\alpha' = l_s^2$. We will find it more convenient to rescale the coordinates X^μ by $1/\sqrt{4\pi g_s N}$ and introduce new coordinates $Y^i=\theta^i/U\,(i=1,\ldots,6)$, where θ^i are the coordinates on S^5 and $\theta^2=1$. The metric in this coordinate system is

$$\frac{ds^2}{\alpha'}=\sqrt{4\pi g_s N}Y^{-2}\left(\sum_{\mu=0}^{3} dX^\mu dX^\mu+\sum_{i=1}^{6} dY^i dY^i\right). \tag{1.2}$$

It is interesting to note that $AdS_5\times S^5$ is conformal to flat $\mathbb{R}^{10}$ if the radii of AdS_5 and S^5 are the same. In this coordinate system, the boundary of AdS_5 is mapped to the origin $Y^i=0$ of $\mathbb{R}^6$.

The gauge theory coupling $g_{\rm YM}$ and the string coupling g_s are related by $g_{\rm YM}^2=4\pi g_s$. We are interested in the limit of $N\to\infty$ while keeping the 't Hooft coupling $\lambda=g_{\rm YM}^2 N$ finite [6]. After taking the large N limit, we will consider the region $\lambda\gg 1$, where the curvature is small compared to the string scale and stringy excitations are negligible. In this case, the supergravity approximation is reliable. According to the AdS-CFT correspondence, every supergravity field has a corresponding local operator in the gauge theory. Correlators of local operators are given by the supergravity action for fields with point sources on the boundary of AdS_5 [7,8]. In the classical limit one just solves the equations of motion with such sources.

An interesting set of nonlocal operators in a gauge theory is composed of Wilson loops. It was proposed in [9,10] that the Wilson loop is defined by an open string ending on the loop at the boundary of AdS_5. In the classical limit, the string is described by a minimal surface. As a result of the curvature of AdS_5, the minimal surface does not stay near the boundary, but goes deep into the interior of space, where the area element can be made smaller. Because of this the behavior of the Wilson loop, for a large area, is that of a conformal theory, and the area law does not produce confinement.

The gauge theory under discussion does not contain quarks or other fields in the fundamental representation of the gauge group. To construct the Wilson loop describing the phase associated with moving a particle in the fundamental representation around a closed curve, we place one of the D-branes very far away from the others. The ground states of the string stretched from the distant D-brane to the others consist of the W bosons and their superpartners in the fundamental representation of the gauge group of the remaining branes. Thus, for large λ, the expectation value of the Wilson loop is related to the classical action of the string, with appropriate boundary conditions. To the leading order in λ, we can ignore the effect of the Ramond-Ramond flux and use the Nambu-Goto action, namely, the area of the minimal surface:

$$A=\int \frac{d\sigma_1 d\sigma_2}{2\pi\alpha'\sqrt{\lambda}}\sqrt{g} =\int \frac{d\sigma_1 d\sigma_2}{2\pi Y^2}\sqrt{\det(\partial_\alpha X^\mu\partial_\beta X^\mu+\partial_\alpha Y^i\partial_\beta Y^i)}. \tag{1.3}$$

Because of the Y^{-2} factor, this area is infinite. After regularizing the divergence, the infinite part was identified as due to the mass of the W boson and subtracted [9]. Taking two parallel lines (with opposite orientation) as a quark-antiquark pair, the remaining finite part defines the quark-antiquark potential. Such calculations were used to study the phases of the $\mathcal{N}=4$ super Yang-Mills theory and to demonstrate confinement in nonsupersymmetric generalizations [11,12].

We will argue below that the correct action of the Wilson loop is not the area of the minimal surface, but the Legendre transform of it with respect to some of the loop variables. The reason is that some of the string coordinates satisfy Neumann conditions rather than Dirichlet conditions. For a certain class of loops, this Legendre transform exactly removes the divergent piece from the area. As a result, the expectation values of such loops are finite.

The appropriate Wilson loop for $\mathcal{N}=4$ super Yang-Mills theory is an operator of the form (suppressing all fermion fields for the moment)

$$W/[C]=\frac{1}{N}\mathrm{Tr}\,\mathcal{P}\exp\left(\oint(iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\right), \tag{1.4}$$

where A_μ are the gauge fields and Φ_i are the six scalars in the adjoint representation, and C represents the loop variables $(x^\mu(s),y^i(s))$. Here $(x^\mu(s))$ determines the actual loop in four dimensions; $(y^i(s))$ can be thought of as the extra six coordinates of the ten-dimensional $\mathcal{N}=1$ super Yang-Mills theory, of which our theory is the dimensionally reduced version. It turns out that minimal surfaces terminating at the boundary of AdS_5 correspond only to loops that satisfy the constraint $\dot{x}^2=\dot{y}^2$. This constraint was derived before, and we study in greater depth its origin and meaning. In [9], the constraint was introduced as a consequence of the fact that the mass of the open string and the Higgs vacuum expectation value (VEV) are proportional to each other. We will show that the constraint also has a geometric interpretation in terms of a minimal surface in $AdS_5\times S^5$. Another interpretation of the constraint has to do with the $\mathcal{N}=4$ supersymmetry; the loops obeying the constraint are Bogomol'nyi-Prasad-Sommerfield-(BPS)-type objects in loop space. After discussing various aspects of loops obeying the constraint, we present some idea on how to extend the calculation to a more general class of loops.

The loop equation is a differential equation on the loop space. We evaluate, using string theory on AdS_5, the action of the loop differential operator $\hat{L}$ on a certain class of Wilson loops. On a smooth loop C, we find that the differential operator annihilates the vacuum expectation of the loop $\langle W\rangle$, in accordance with the loop equation as derived in the gauge theory. On the other hand, for a loop with a self-intersection point, the gauge theory predicts that $\hat{L}\langle W\rangle$ is nonzero and proportional to $g_{\mathrm{YM}}^2 N$. We point out the gauge theory also predicts that a cusp (a sharp turning point) in a loop gives a nonzero contribution to the loop equation, proportional to $g_{\mathrm{YM}}^2 N$. We will show that $\hat{L}\langle W\rangle$ for a loop with a cusp evaluated by the minimal surface in $AdS_5\times S^5$ is indeed nonvanishing and proportional to $g_{\mathrm{YM}}^2 N$. We have not been able to reproduce the precise dependence on the angle at the cusp due to our lack of detailed understanding of loops not obeying the constraint $\dot{x}^2=\dot{y}^2$. For the same reason we were unable to reproduce the expected result at an intersection.

The paper is organized as follows.

In Sec. II, we start with a brief review of the Wilson loop operator in the pure Yang-Mills theory. We then point out an important subtlety in performing the Wick rotation in the supersymmetric theory. We will present some results from the perturbation theory where the subtlety in the Wick rotation plays an interesting role.

In Sec. III, we turn to string theory in $AdS_5\times S^5$. We will give a precise specification of boundary conditions on the string world sheet and the geometric origin of the constraint $\dot{x}^2=\dot{y}^2$. For some cases, we can compute the area of minimal surfaces explicitly. These include loops with intersections or cusps. For such loops, the areas have logarithmic divergences. After calculating those areas, we explain the need for the Legendre transform and show that it removes the linear divergence. The absence of a linear divergence fits well with what we expect for the supersymmetric gauge theory. We will clarify the issue of zigzag symmetry and end the section with a discussion of loops that do not satisfy the constraint.

In Sec. IV, we give a review of the loop equation in the pure Yang-Mills theory and derive its generalization to the case of $\mathcal{N}=4$ super Yang-Mills theory in four dimensions.

In Sec. V, we will discuss to what extent the minimal surface calculation in AdS_5 is consistent with the loop equation.

To make the body of the paper more readable, some details are presented in appendixes. In Appendix A we derive the Wilson loop as the first quantized action of the W boson. In Appendix B we calculate the area of a minimal surface near a cusp. In Appendix C we present some more details on the loop equation of the $\mathcal{N}=4$ theory.

II. WILSON LOOPS IN $N=4$ GAUGE THEORY

We define the Wilson loop operator in the supersymmetric gauge theory and review some of its basic properties. We pay particular attention to its coupling to the scalar fields in the supermultiplet.

A. Definition

One of the most interesting observables in gauge theories is the Wilson loop, the path-ordered exponential of the gauge field,

$$W=\frac{1}{N}\operatorname{Tr}\mathcal{P}\exp\left(i\oint A_\mu dx^\mu\right), \tag{2.1}$$

with the trace in the fundamental representation. The Wilson loop can be defined for any closed path in space, providing a large class of gauge invariant observables. In fact, these operators, and their products, form a complete basis of gauge-invariant operators for pure Yang-Mills theory. An appropriate definition of the loop operator for the $\mathcal{N}=4$ super Yang-Mills theory in four dimensions will be given below.

One of physical applications of Wilson loops stems from the fact that an infinitely massive quark in the fundamental representation moving along the loop will be transformed by the phase factor in Eq. (2.1). Thus the dynamical effects of the gauge dynamics on external quark sources is measured by the Wilson loop. In particular, for a parallel quark-antiquark pair, the Wilson loop is the exponent of the effective potential between the quarks and serves as an order parameter for confinement [13].

The Maldacena conjecture states that type IIB string theory on $AdS_5\times S^5$ is dual to $\mathcal{N}=4$ super Yang-Mills theory in four dimensions. This gauge theory does not contain quarks in the fundamental representation. To construct the Wilson loop, we separate a single D-brane from the N D-branes and take it very far away. For large N, we can ignore the fields on the distant D-brane, except for open strings stretching between it and the other N. The ground states of the open string are the W bosons and their superpartners of the broken, $SU(N)$, gauge group. Their trajectories should give the same effect as that of an infinitely massive particle in the fundamental representation.

The correlation functions of the W boson can be written in the first quantized formalism as an integral over paths. This description is studied in detail in Appendix A. When the four-dimensional space has the Lorentzian signature metric, the phase factor associated with the loop is given by the vacuum expectation value of the operator

$$W=\frac{1}{N}\operatorname{Tr}\mathcal{P}\exp\left(i\oint(A_\mu\dot{x}^\mu+|\dot{x}|\Phi_i\theta^i)ds\right). \tag{2.2}$$

When the metric is Euclidean, there is an important modification to this formula as

$$W=\frac{1}{N}\operatorname{Tr}\mathcal{P}\exp\left(i\oint(A_\mu\dot{x}^\mu-i|\dot{x}|\Phi_i\theta^i)ds\right). \tag{2.3}$$

Notice the presence of i in the second term in the exponent. The "phase factor" in the Euclidean theory is not really a phase, but contains a real part.

In the above, θ^i are angular coordinates of magnitude 1 and can be regarded as coordinates on S^5. In the gauge theory, we may consider a more general class of Wilson loops of the form

$$W=\frac{1}{N}\mathrm{Tr}\,\mathcal{P}\,\exp\left(\oint(iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\right), \tag{2.4}$$

with an arbitrary function $y^i(s)$. This is the general loop we would get by dimensional reduction from the ten dimensional gauge theory, where Φ_i would be the extra six components of the gauge field. Equation (2.3) restricts us to the case of $\dot{x}^2-\dot{y}^2=0$. This suggests that the metric on the loop variables $(x^\mu(s),y^i(s))$ has the signature (4,6). It is important to stress that this is not the signature of $AdS_5\times S^5$, but of the space where the loops are defined.[2] As we will show later, the signature of the loop space metric is related to the fact that the six loop variables $y^i(s)$ correspond to T-dual coordinates on the string world sheet. The constraint $\dot{x}^2-\dot{y}^2=0$ is also related to supersymmetry.

Gauge invariance in four dimensions requires that the Wilson loop close in four dimensions; i.e. the loop variables $x^\mu(s)$ are continuous and periodic around the loop. This is not the case for the other six variables $y^i(s)$, and the loop may have a jump in these six directions.

B. Perturbation theory

As a warm-up, we study properties of the Wilson loops in perturbation theory. To first order in g_{YM}^2N, the expectation value of the loop $\langle W\rangle$ is given by

$$\langle W[C]\rangle=1-g_{\mathrm{YM}}^2\,N\oint ds\oint ds'\,P\{\dot{x}^\mu(s)\dot{x}^\nu(s')G_{\mu\nu}[x(s)-x(s')]-\dot{y}^i(s)\dot{y}^j(s')G_{ij}[x(s)-x(s')]\}, \tag{2.5}$$

where $G_{\mu\nu}$ and G_{ij} are the gauge field and scalar propagators. The relative minus sign comes from the extra i in front of the scalar piece in the exponent in Eq. (2.3). This integral is linearly divergent. With a regularization of the propagator with cutoff ϵ [i.e., replacing $1/x^2$ with $1/(x^2+\epsilon^2)$], the divergent piece coming from the exchange of the gauge field A_μ is evaluated as

$$-\frac{\lambda}{8\pi^2}\oint ds\int_{e/|\dot{x}|}^{e/|\dot{x}|}ds'\,\dot{x}^\mu(s)\dot{x}^\nu(s')\frac{\delta_{\mu\nu}}{\epsilon^2}=-\frac{\lambda}{(2\pi)^2\epsilon}\oint ds|\dot{x}|=-\lambda\frac{L}{(2\pi)^2\epsilon}, \tag{2.6}$$

where L is the circumference of the loop. The divergent contribution from the exchange of the scalars Φ_i is

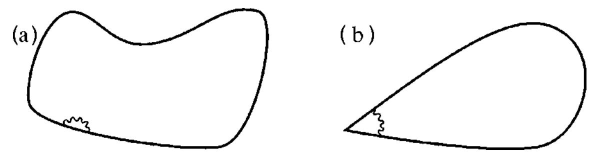

FIG. 1. (a) At one loop, there is a linear divergence from the propagator connecting coincident points. The divergence is proportional to the circumference of the loop. (b) At cusps and intersections, an additional logarithmic divergence appears when the two external legs approach the singular point.

$$\frac{\lambda}{8\pi^2}\oint ds\int_{e/|\dot{x}|}^{e/|\dot{x}|}ds'\,\dot{y}^i(s)\dot{y}^j(s')\frac{\delta_{ij}}{\epsilon^2}=\frac{\lambda}{(2\pi)^2\epsilon}\oint ds|\dot{x}|\frac{\dot{y}^2}{\dot{x}^2}. \tag{2.7}$$

Combining these terms together, we find

$$W=1+\frac{\lambda}{(2\pi)^2\epsilon}\oint ds|\dot{x}|\left(1-\frac{\dot{y}^2}{\dot{x}^2}\right)+\text{finite}. \tag{2.8}$$

We note that the linear divergence cancels when the constraint $\dot{x}^2=\dot{y}^2$ is satisfied.

At nth order in the $\lambda=g_{\mathrm{YM}}^2N$ expansion, one finds a linear divergence of the form

$$\frac{\lambda^n}{\epsilon}\oint ds|\dot{x}|G_n\left(\frac{\dot{y}^2}{\dot{x}^2}\right), \tag{2.9}$$

for some polynomial $G_n(z)$. We now argue that $G_n(1)=0$; namely the linear divergence cancels when $\dot{x}^2=\dot{y}^2$, to all order in the perturbative expansion. The nth order term is calculated by connected Feynman diagrams with external legs attached to the loop. The linear divergence appears when all the external legs come together in four dimensions. Since the Feynman rule of the $\mathcal{N}=4$ gauge theory is obtained by the dimensional reduction of the ten-dimensional theory, the ten-dimensional rotational invariance of the Feynman rule is recovered in the coincidence limit. Therefore the contractions of the external indices by the Feynman rule produce only rotational invariant combinations of $(\dot{x}^\mu,i\dot{y}^i)$, namely, a polynomial of $(\dot{x}^2-\dot{y}^2)$. The polynomial does not have a constant term since a connected Feynman diagram for $\langle W\rangle$ needs to have at least two external lines attached to the loop. Therefore the polynomial vanishes when $\dot{x}^2-\dot{y}^2=0$.

When the loop has a cusp, there is an extra logarithmic divergence from graphs as shown in Fig. 1. Let us denote the angle at the cusp by Ω. We choose the angle so that $\Omega=\pi$ at a regular point of the loop. A one-loop computation with the gauge field gives

$$\frac{\lambda}{(2\pi)^2}[(\pi-\Omega)\cot\Omega+1]\log\frac{L}{\epsilon}. \tag{2.10}$$

A cusp is a discontinuity of $\dot{x}^\mu$. There may also be a discontinuity in $\dot{y}^i$, which we measure by an angle Θ. We choose Θ so that $\Theta=0$ when $\dot{y}^i$ is continuous. A one-loop computation with the scalar fields gives

[2] One may regard the extra factor of i in the Euclidean case (2.3) as a Wick rotation of the six y coordinates so that we can express the constraint as $\dot{x}^\mu\dot{x}_\mu+\dot{y}^i\dot{y}_i=0$, both in the Lorentzian and Euclidean cases. To avoid confusion, we will not use this convention and write the i explicitly in all our expressions in the Euclidean case.

$$-\frac{\lambda}{(2\pi)^2}\left(-\frac{\pi-\Omega}{\sin\Omega}\cos\Theta+1\right)\log\frac{L}{\epsilon}. \tag{2.11}$$

Combining Eqs. (2.10) and (2.11) together, we obtain

$$\frac{\lambda}{(2\pi)^2}\frac{\pi-\Omega}{\sin\Omega}(\cos\Omega+\cos\Theta)\log\frac{L}{\epsilon}. \tag{2.12}$$

A similar computation at an intersection gives

$$\frac{\lambda}{2\pi}\frac{1}{\sin\Omega}(\cos\Omega+\cos\Theta)\log\frac{L}{\epsilon}. \tag{2.13}$$

III. MINIMAL SURFACES IN ANTI–de SITTER SPACE

According to the Maldacena conjecture, the expectation value of the Wilson loop is given by the action of a string bounded by the curve at the boundary of space:

$$\langle W[C]\rangle=\int_{\partial X=C}\mathcal{D}X\exp(-\sqrt{\lambda}S[X]), \tag{3.1}$$

for some string action $S[X]$. Here X represents both the bosonic and the fermionic coordinates of the string. For large λ, we can estimate the path integral by the steepest descent method. Consequently, the expectation value of the Wilson loop is related to the area A of the minimal surface bounded by C as

$$\langle W\rangle\simeq\exp(-\sqrt{\lambda}A). \tag{3.2}$$

The motivation for this ansatz is that the W boson considered in Sec. II A is described in the D-brane language by an open string going between the single separated D-brane and the other N D-branes. In the near-horizon limit, the N D-branes are replaced by the AdS_5 geometry and the open string is stretched from the boundary to the interior of AdS_5.

To be precise, this argument only tells us that the Wilson loop and the string in AdS_5 are related to each other. The expression (3.1) is schematic at best, and there may be an additional loop-dependent factor in Eq. (3.2). A similar problem exists in computation of correlation functions of local operators; there is no known way to fix the relative normalization of local operators in the gauge theory and supergravity fields in AdS_5. To determine the normalization factor, one has to compute the two-point functions [14,15]. In our case, the normalization factor in Eq. (3.2) may depend on the loop variables $C=(x^\mu(s),y^i(s))$. In fact, we will argue below that the correct action to be used in Eq. (3.2) is not the area A of the surface, but the Legendre transform of it. This modification does not change the equations of motion, and the solutions are still minimal surfaces. However, the values of the classical action for these surfaces are different than their areas.

We will assume that, to the leading order in λ, there is no further C-dependent factor. Otherwise, the conjecture would be meaningless as it would produce no falsifiable predictions. On the other hand, one expects a C-dependent factor in the subleading order, such as the fluctuation determinant of the surface in AdS_5. There can also be a factor in the relation between the W-boson propagation amplitude and the Wilson loop computed in Appendix A. Such a factor would be kinematic in nature and independent of λ, and therefore negligible in our analysis.

A. Boundary conditions and BPS loop

The Wilson loop discussed in [9] obeys the constraint

$$\dot{x}^2=\dot{y}^2. \tag{3.3}$$

This constraint was originally derived by using the coupling of the fundamental string to the gauge fields and to the scalars. In our derivation of the loop operator from the phase factor for the W-boson amplitude in Appendix A, the constraint arises from the saddle point in integrating over different reparametrizations of the same loop, essentially for the same reason as in [9].

In this section, we will give another interpretation of the constraint (3.3), in terms of the string theory in $AdS_5\times S^5$. For this interpretation we need to give a precise specification of the boundary condition on the string in $AdS_5\times S^5$.

We begin with super Yang-Mills theory in ten dimensions, which is realized on space-filling D9-branes. We ignore the fact that this theory is anomalous since we will reduce it to the anomaly-free theory in four dimensions. Moreover, we are only interested in the boundary conditions on bosonic variables.[3] The Wilson loop in ten dimensions corresponds to an open string world sheet bounded by the loop; i.e., we should impose full Dirichlet boundary conditions on the string world sheet. This is natural since, without the Wilson loop operator, the string end point obeys fully Neumann boundary conditions along the D9-brane. The conditions imposed by the Wilson loop are complementary to the boundary conditions on the D9-brane.

To reduce the theory to four dimensions, we perform T-duality along six directions. An open string ending on the D3-brane obeys four Neumann and six Dirichlet boundary conditions. Consequently, the Wilson loop operator in the four-dimensional gauge theory imposes complementary boundary conditions, namely, four Dirichlet and six Neumann boundary conditions. If the Wilson loop is parametrized by the loop variables $(x^\mu(s),y^i(s))$, where $\dot{y}^i(s)$ couples to the six scalar fields, then the six loop variables $\dot{y}^i(s)$ are to be identified with the six Neumann boundary conditions on the string world sheet.

We are ready to specify the boundary condition on the string world sheet residing in $AdS_5\times S^5$, with line element

$$\frac{ds^2}{\alpha'}=\sqrt{\lambda}Y^{-2}\left(\sum_{\mu=0}^{3}dX^\mu dX^\mu+\sum_{i=1}^{6}dY^i dY^i\right). \tag{3.4}$$

Choose the string world-sheet coordinates to be (σ^1,σ^2) such that the boundary is located at $\sigma^2=0$. Since X^μ is iden-

[3] Boundary conditions for fermionic variables are not relevant in our analysis of the loop for large λ.

tified with the four-dimensional coordinates where the gauge theory resides, it is natural to impose Dirichlet conditions on X^μ, so that

$$X^\mu(\sigma_1,0)=x^\mu(\sigma_1). \tag{3.5}$$

The remaining six string coordinates $Y^i(\sigma^1,\sigma^2)$ obey Neumann boundary conditions. We propose that these boundary conditions are

$$J_1^\alpha\partial_\alpha Y^i(\sigma^1,0)=\dot{y}^i(\sigma^1), \tag{3.6}$$

where $J_\alpha{}^\beta(\alpha,\beta=1,2)$ is the complex structure on the string world sheet given in terms of the induced metric $g_{\alpha\beta}$,

$$J_\alpha{}^\beta=\frac{1}{\sqrt{g}}\,g_{\alpha\gamma}\epsilon^{\gamma\beta}. \tag{3.7}$$

Although we do not have a derivation of the boundary condition (3.6) from first principles, it can be motivated as follows. Because of the identification of the SO(6) symmetries in the AdS-CFT correspondence, it is clear that Neumann boundary conditions must set $\dot{y}^i$ equal to $J_1{}^\alpha\partial_\alpha Y^i$ up to a relative normalization of the two. The use of the induced complex structure $J_\alpha{}^\beta$ in the Neumann boundary condition is required by the reparametrization invariance on the world sheet. The fact that the condition $\dot{x}^2=\dot{y}^2$ has a natural interpretation in terms of the minimal surface, as we will explain below, suggests that the normalization factor is 1, as in Eq. (3.6).

For a generic choice of the loop variables $(x^\mu(s),y^i(s))$, there is a unique minimal surface in Euclidean space obeying the ten boundary conditions, Eqs. (3.5) and (3.6). However, the resulting minimal surface does not necessarily terminate at the boundary $Y^i=0$ of AdS_5. The condition $Y^i=0$ would be additional Dirichlet conditions, which may or may not be compatible with Eq. (3.6). In fact, one can show that, for a smooth loop, the additional condition $Y^i(\sigma^1,0)=0$ is satisfied by the minimal surface if and only if the loop variables obey the constraint $\dot{x}^2=\dot{y}^2$. To see this consider the Hamilton-Jacobi equation[4] for the area A of a minimal surface bounded by a loop $(X^\mu(s),Y^i(s))$ in $AdS_5\times S^5$:

$$\left(\frac{\delta A}{\delta X^\mu}\right)^2+\left(\frac{\delta A}{\delta Y^i}\right)^2=\frac{1}{(2\pi)^2Y^4}[(\partial_1X^\mu)^2+(\partial_1Y^i)^2]. \tag{3.8}$$

Since the momenta conjugate to the X^μ's and the Y^i's are given by

$$\frac{\delta A}{\delta X^\mu}=\frac{1}{2\pi Y^2}J_1{}^\alpha\partial_\alpha X^\mu,\qquad \frac{\delta A}{\delta Y^i}=\frac{1}{2\pi Y^2}J_1{}^\alpha\partial_\alpha Y^i, \tag{3.9}$$

we obtain

$$(J_1{}^\alpha\partial_\alpha X^\mu)^2+(J_1{}^\alpha\partial_\alpha Y^i)^2=(\partial_1X^\mu)^2+(\partial_1Y^i)^2. \tag{3.10}$$

If the minimal surface obeys the boundary conditions (3.5) and (3.6), this becomes

$$\dot{x}^2-\dot{y}^2=(J_1{}^\alpha\partial_\alpha X^\mu)^2-(\partial_1Y^i)^2. \tag{3.11}$$

Now impose the additional condition that the string world sheet terminate at the boundary of AdS_5, i.e., $Y^i(\sigma^1,0)=0$. Obviously, $\partial_1Y^i(\sigma^1,0)=0$. This alone tells us that $\dot{x}^2-\dot{y}^2\geqslant 0$. Moreover, if the boundary is smooth, it costs a large area to keep $J_1{}^\alpha\partial_\alpha X^\mu$ nonzero near the boundary of AdS_5, so it has to vanish at the boundary $Y=0$ [9]. Therefore the condition that the minimal surface terminate at the boundary of AdS_5 requires $\dot{x}^2=\dot{y}^2$.

When the constraint $\dot{x}^2=\dot{y}^2$ is satisfied, one can reinterpret the six Neumann conditions (3.6) as Dirichlet conditions on S^5. To see this, it is useful to decompose the six coordinates Y^i

$$Y^i=Y\theta^i \tag{3.12}$$

where θ^i are coordinates on S^5 and $Y=U^{-1}$ is one of the coordinates on AdS_5. Since for a smooth loop the classical solution has $\partial_\alpha Y^i=(\partial_\alpha Y)\theta^i$ at the boundary $Y=0$ of AdS_5, the Neumann conditions (3.6) turn into the Dirichlet conditions on S^5 as

$$\theta^i(\sigma^1,0)=\frac{\dot{y}^i}{|\dot{y}|}. \tag{3.13}$$

This justifies the boundary conditions used in [9].

There is yet another interpretation of the constraint $\dot{x}^2=\dot{y}^2$, and it has to do with supersymmetry. The loops we have considered so far couple only to bosonic fields: the gauge field A_μ and scalars Φ^i. We also need to allow coupling to the fermionic fields in the exponent. Fermionic variables $\zeta(s)$ along the loop couple to the gauginos Ψ as

$$\overline{\zeta}(\dot{x}^\mu\Gamma_\mu-i\dot{y}^i\Gamma_i)\Psi, \tag{3.14}$$

where we are using ten-dimensional gamma matrices Γ_μ and Γ_i with signature (10,0). This is derived in Appendix C. Exactly when the constraint is satisfied this combination of gamma matrices becomes nilpotent. Consequently, only half the components of ζ couple to Ψ, putting the loop in a short representation of local supersymmetry in super loop space. The simplest example is when the Wilson loop is a straight line, when $\dot{x}$ and $\dot{y}$ are independent of s. If ζ is also constant, this loop is the phase factor associated with the a trajectory of a free BPS particle.

---

[4]In general, the Hamilton-Jacobi equation for the area of a minimal surface on a Riemannian manifold with a metric G_{IJ} takes the form

$$G^{IJ}(\delta A/\delta X^I)(\delta A/\delta X^J)=G_{IJ}\delta_1X^I\delta_1X^J.$$

B. Calculating the area

The computation of the Wilson loop in AdS_5 requires an infrared regularization, since the area of the minimal surface terminating at the boundary of AdS_5 is infinite due to the factor Y^{-2} in the metric. In order to make sense of the ansatz (3.2), we need to regularize the area. One natural way to do so is to impose the boundary conditions (3.5) and (3.6) at $Y=0$, but integrate the area element only over the part of the surface with $Y \geqslant \epsilon$. On the gauge theory side, the Wilson loop requires regularization in the ultraviolet. According to the UV-IR relation in the AdS-CFT correspondence [16], the IR cutoff ϵ in AdS_5 should be identified with the UV cutoff in the gauge theory.

There are a few cases when minimal surfaces can be studied analytically.

1. Parallel lines

The minimal surface for parallel lines, each of length L and separated by a distance R, was obtained in [9,10]. The area of the loop is

$$A=\frac{2L}{2\pi\epsilon}-\frac{4\pi\sqrt{2}}{\Gamma(1/4)^4}\frac{L}{R}. \tag{3.15}$$

2. Circular loop

The minimal surface in AdS_5 bounded by a circle of radius R is found in [17,18] as

$$Y(r,\varphi)=\sqrt{R^2-r^2}, \tag{3.16}$$

where r and φ are radial coordinates on a plane in the four dimensions, and we use them as coordinates on the string world sheet also. The area of the surface with the cutoff ϵ is

$$\begin{aligned} A&=\frac{1}{2\pi}\int dr\, r d\varphi Y^{-2}\sqrt{1+Y'^2}\\ &=R\int_0^{\sqrt{R^2-\epsilon^2}}\frac{r dr}{(R^2-r^2)^{3/2}}=\frac{2\pi R}{2\pi\epsilon}-1. \end{aligned} \tag{3.17}$$

3. Cusp

Another family of minimal surfaces we can solve analytically is a surface near a cusp on $\mathbb{R}^4$ and its generalization including a jump on S^5. We can find analytical solutions in this case since the boundary conditions are scale invariant. Using radial coordinates in the vicinity of the cusp, r and φ, as world-sheet coordinates, the scale-invariant ansatz

$$Y(r,\varphi)=\frac{r}{f(\varphi)} \tag{3.18}$$

reduces the determination of the minimal surface to a one-dimensional problem. The resulting surface is depicted in Fig. 2. When there is also a jump on S^5, one needs to introduce another variable. An analytical solution in this case is found in a similar way. These solutions are presented in Appendix B. The result is that the area of the surface has a logarithmic divergence as well as a linear divergence. It behaves as

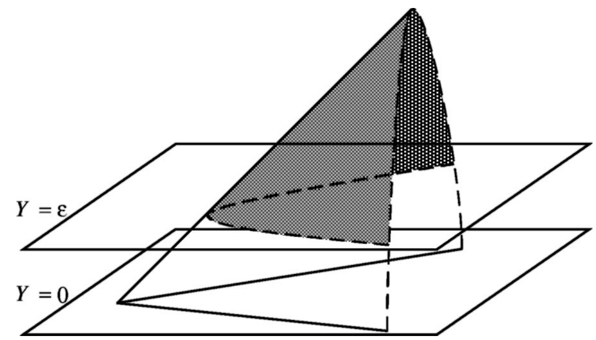

FIG. 2. A minimal surface for a Wilson loop with a cusp. The regularized area is evaluated over the shaded region.

$$A=\frac{L}{2\pi\epsilon}-\frac{1}{2\pi}F(\Omega,\Theta)\log\frac{L}{\epsilon}+\cdots, \tag{3.19}$$

where Ω and Θ are the cusp angles in R^4 and S^5, respectively.

When either Θ or Ω vanishes, we can express $F(\Omega,\Theta)/2\pi$ in terms of elliptic integrals. In Fig. 3 we show the numerical evaluation of the function $F(\Omega,0)$ as the solid curve. This is to be compared with the perturbative expression (2.12) shown as the dashed curve. The function $F(\Omega,0)$ is zero at $\Omega=\pi$ and has a pole at $\Omega=0$. As the angle $\Omega \to 0$ at the cusp, the loop goes back along its original path, or backtracks. Regularizing the extra divergence from the pole turns it into a linear divergence which cancels part of the linear divergence from the length of the loop. This is related to issues discussed in the section on the zigzag symmetry.

Away from the cusp, the surface approaches the boundary along the Y direction without a momentum in the X direction. Right at the cusp, however, the surface has momentum in

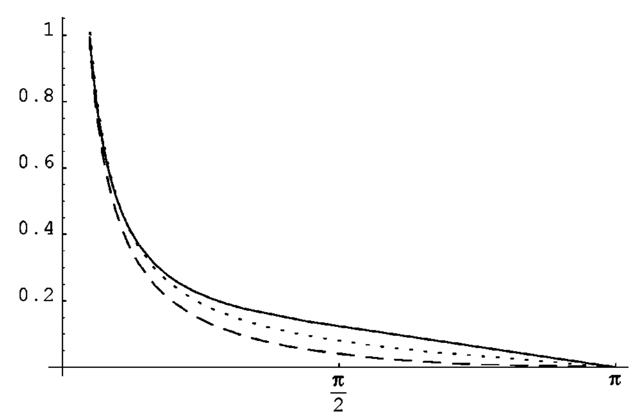

FIG. 3. The solid curve shows the function $F(\Omega,0)/2\pi$, which appears in the logarithmic divergence of the minimal surface with the cusp of angle Ω. This is compared with the perturbative result (2.12) at a cusp shown as the dashed curve. The dotted curve is half of the perturbative result (2.13) at an intersection.

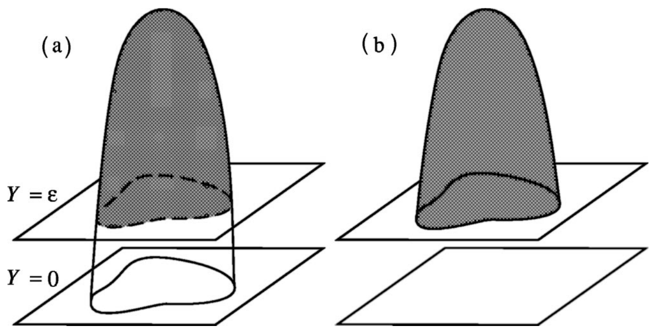

FIG. 4. The comparison of the two regularization prescriptions. The boundary conditions are imposed at $Y=0$ in (a) and at $Y=\epsilon$ in (b). The shaded regions represent the regularized areas.

both the Y and r directions. This means that, although the constraint $\dot{x}^2=\dot{y}^2$ is obeyed almost everywhere, it is modified at the cusp as

$$\dot{x}^2=(1+f_0^2)\dot{y}^2, \tag{3.20}$$

where $f_0=f(\varphi=\Omega/2)$ is the minimal value of $f(\varphi)$.

4. Intersection

The minimal surface for a self-intersecting loop is just the sum of two cusps. The only difference is that, by the exchange symmetry of the two components of the loop, the intersection forces

$$\frac{\dot{y}}{|\dot{x}|}=0 \tag{3.21}$$

instead of Eq. (3.20)

In all the examples above, there is a linear divergence $(2\pi\epsilon)^{-1}$ in the regularized area. This is true for any loop. As explained in [9], this leading divergence in the area of the minimal surface in AdS_5 is proportional to the circumference of the loop.[5] The linear divergence arises from the leading behavior of the surface at small Y, i.e., near the boundary of AdS_5.

In this section, we have computed the regularized area by imposing the boundary condition at the boundary $Y=0$ of AdS_5 and integrating the area element over the part of the surface $Y\geqslant\epsilon$. This is not a unique way to regularize the area. Another reasonable way to compute the minimal surface is to impose the boundary conditions, not at $Y=0$, but at $Y=\epsilon$. The area bounded by the loop on $Y=\epsilon$ is then by itself finite. A comparison of the two regularization prescriptions is illustrated in Fig. 4. These two regularizations give the same values for the area, up to terms which vanish as $\epsilon\rightarrow 0$. For example, consider the circular loop. The solution (3.16) can also be regarded as a minimal surface with the boundary condition on $Y=\epsilon$, except that the radius of the circle on $Y=\epsilon$ is now $R_0=\sqrt{R^2-\epsilon^2}$. The area computed in this new regularization is then

$$A=\frac{1}{\epsilon}\sqrt{R_0^2+\epsilon^2}-1=\frac{2\pi R_0}{2\pi\epsilon}-1+\frac{\epsilon}{2R_0}+\cdots. \tag{3.22}$$

Thus the results of the two regularizations are the same up to terms which vanish as $\epsilon\rightarrow 0$. It is straightforward to show that this is also the case for the parallel lines. We have also verified that when the loop has a cusp or an intersection, the two regularizations give the same area modulo terms which are finite as $\epsilon\rightarrow 0$, which are subleading compared to the logarithmic divergence.

When we impose the boundary condition at $Y=\epsilon$, the constraint on the loop variables is not exactly $\dot{x}^2=\dot{y}^2$, but it is modified. If the loop is smooth, the modification is only by $O(\epsilon)$ terms.[6] Therefore most of the results in this paper are independent of the choice between the two ways of imposing the boundary conditions. The only exception to this rule is the discussion of the zigzag symmetry. The zigzag symmetry of the string world sheet on AdS_5 seems to fit well with our expectations about the gauge theory when we use the boundary conditions at $Y=\epsilon$ rather than at $Y=0$.

C. Legendre transformation

The Maldacena conjecture implies that the Wilson loop is related to a string ending along the loop on the boundary of space. In the classical limit, we expect that the string world sheet is described by a minimal surface. This argument, however, does not completely determine the value of $\langle W\rangle$ for large λ since there are many actions whose equations of motion are solved by minimal surfaces. They differ by total derivatives, or boundary terms. Since the surface has boundaries, such terms can be important. In [9,10] it was assumed that one should use the Nambu-Goto action, so the Wilson loop was given in terms of the area A of the minimal surface. This is what we have studied so far. In this section, we argue that $\langle W\rangle$ is in fact given not by A but by an appropriate Legendre transform.

We have shown that the loop variables $\dot{y}^i$ impose Neumann boundary conditions (3.6) on the coordinates Y^i. Therefore $\langle W\rangle$ should be regarded as a functional of the coordinates X^μ and the momenta P_i conjugate to Y^i, defined by

$$P_i=\frac{\delta A}{\delta\partial_2 Y^i}=\frac{1}{2\pi\sqrt{\lambda}\,\alpha'}\sqrt{g}g^{2\alpha}\partial_\alpha Y^j G_{ij}. \tag{3.23}$$

[5] We are using the coordinates X^μ in Eq. (1.2) to describe the configurations of the Wilson loops. With these coordinates, there is no factor of λ in the relation between the IR cutoff ϵ in AdS_5 and the UV cutoff of the gauge theory [16]. These coordinates are different from the coordinates on the D3-brane probe, by a factor of $\sqrt{\lambda}$ [19].

[6] If the loop has a cusp or an intersection, as we saw earlier, the boundary conditions imposed at $Y=0$ imply that the constraint $\dot{x}^2=\dot{y}^2$ holds almost everywhere along the loop, except at a cusp or an intersection point. When we impose the boundary conditions at $Y=\epsilon$, the constraint is modified in regions of size ϵ near the cusp and the intersection point.

The Nambu-Goto action is a natural functional of $X^\mu(s)$ and $Y^i(s)$ and is more appropriate for the full Dirichlet boundary conditions. To replace it with a functional of $X^\mu(s)$ and $P^i(s)$, we need to perform the Legendre transform

$$\tilde{L}=L-\partial_2(P_iY^i) \tag{3.24}$$

or

$$\tilde{A}=A-\oint d\sigma_1 P_iY^i. \tag{3.25}$$

To show that $\tilde{A}$ is a natural functional of (X^μ,P^i), we use Hamilton-Jacobi theory. Under a general variation of the Y coordinates, the variation of the area A of the minimal surface is given by

$$\begin{aligned}\delta A=&\int d\sigma_1 d\sigma_2\left(\frac{\delta A}{\delta Y^i}-\partial_\alpha\frac{\delta A}{\delta\partial_\alpha Y^i}\right)\delta Y^i(\sigma_1,\sigma_2)\\&+\oint d\sigma_1\frac{\delta A}{\delta\partial_2 Y^i}\,\delta Y^i(\sigma_1,0)\\=&\oint d\sigma_1 P_i(\sigma_1,0)\,\delta Y^i(\sigma_1,0).\end{aligned} \tag{3.26}$$

Here we used the equations of motion. Therefore, after performing the Legendre transformation, we obtain

$$\delta\tilde{A}=-\oint d\sigma_1 Y^i(\sigma_1,0)\,\delta P_i(\sigma_1,0). \tag{3.27}$$

Thus $\tilde{A}$ is a functional of the momenta P^i at the boundary, not the coordinates Y^i.

The Neumann boundary conditions (3.6) are conditions on the momenta P^i:

$$\frac{\dot{y}^i}{2\pi}=P^i=Y^2P_i\,. \tag{3.28}$$

In fact, if the loop variables $\dot{y}^i(s)$ are continuous, the coordinates Y^i are parallel to the momenta P_i, as we saw in Eq. (3.13). In this case, the Legendre transform gives

$$\begin{aligned}\tilde{A}&=A-\frac{1}{2\pi}\oint d\sigma_1\frac{\dot{y}^i}{Y^2}Y^i\\&=A-\frac{1}{2\pi}\oint d\sigma_1\frac{|\dot{y}|}{Y}=A-\frac{1}{2\pi\epsilon}\oint ds|\dot{y}|,\end{aligned} \tag{3.29}$$

where ϵ is the regulator. In the last step, we have set $Y=\epsilon$ since the regularized action is evaluated for $Y\geqslant\epsilon$.

In the previous section, we saw that the area A of minimal surface has a linear divergence proportional to the circumference of the boundary. By combining it with Eq. (3.29), we find

$$\tilde{A}=\frac{1}{2\pi\epsilon}\oint ds(|\dot{x}|-|\dot{y}|)+\text{finite} \tag{3.30}$$

for a smooth loop. Therefore the linear divergence cancels when the constraint $\dot{x}^2=\dot{y}^2$ is satisfied. The minimal surface in AdS_5 is supposed to describe the Wilson loop for large coupling λ. We saw in Sec. II B that the cancellation of the divergence also takes place to all order in the perturbative expansion λ. This suggests that the cancellation of the linear divergence is exact, and a smooth loop obeying $\dot{x}^2=\dot{y}^2$ does not require regularization. We suspect that this is a consequence of the BPS property of the loop. When the loop is a straight line, it preserves a global supersymmetry, not only the local one. In that case the lowest order perturbation calculation is exact. The modified action is zero; the expectation value of the Wilson loop is 1.

We were not able to find an explicit expression for $\tilde{L}$ as a function of X^μ, P^i, and their derivatives. We only know how to evaluate it for classical solutions in terms of the old variables.

By definition, the area A of the minimal surface is positive. On the other hand, its Legendre transform $\tilde{A}$ may be negative and the expectation value of the loop, $\langle W\rangle=\exp(-\sqrt{\lambda}\tilde{A})$, may be larger than 1. In the pure Yang-Mills theory, the Wilson loop is a trace of a unitary operator (divided by the rank N of the gauge group), and its expectation value has to obey the inequality $\langle W\rangle\leqslant 1$. This is not the case in the supersymmetric theory in the Euclidean signature space since W in Eq. (2.3) is not a pure phase, and there is no unitarity bound on its expectation value.

We have shown that the expectation value of a smooth Wilson loop obeying $\dot{x}^2=\dot{y}^2$ is finite. If the loop has a cusp or an intersection, the cancellation is not exact and we are left with the logarithmic divergence[7]

$$\tilde{A}=-\frac{1}{2\pi}F(\Omega,\Theta)\log\frac{L}{\epsilon}+\text{finite}. \tag{3.31}$$

It is interesting to note that the constraint $\dot{x}^2=\dot{y}^2$ is not satisfied either at a cusp

$$\dot{x}^2=(1+f_0^2)\dot{y}^2 \tag{3.32}$$

or at an intersection point

$$\frac{\dot{y}^i}{|\dot{x}|}=0. \tag{3.33}$$

We suspect that the logarithmic divergences at the cusp and the intersection are caused by the failure of the loop to satisfy the BPS condition at these points.

D. Zigzag symmetry

A Wilson loop of the form

$$W=\frac{1}{N}\mathrm{Tr}\,\mathcal{P}\exp\left(i\oint ds A_\mu\dot{x}^\mu\right) \tag{3.34}$$

[7]If $\Theta\neq 0$, the function $F(\Omega,\Theta)$ gets a contribution from the Legendre transformation.

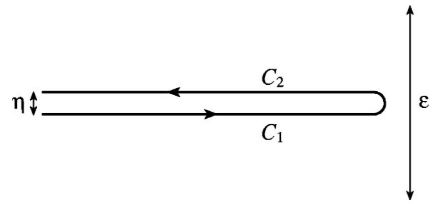

FIG. 5. The zigzag loop. The loop goes in one direction along C_1 and comes back along C_2. The two segments C_1 and C_2 are parallel and their distance η is less than the gauge theory UV cutoff ϵ.

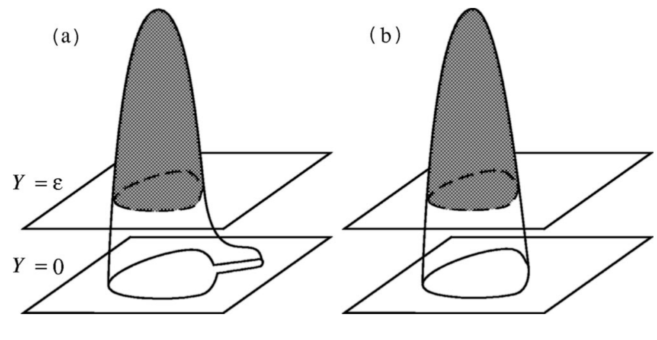

FIG. 6. The area of a loop with a zigzag (a) is roughly the same as the loop without it (b).

is reparametrization invariant, in s, namely, unchanged by $s \to f(s)$. Formally, it is even invariant under reparametrizations which backtrack [namely, when $\dot{f}(s)$ is not always positive] since the phase factor going forward and then backwards will cancel. Polyakov has argued in [20] that this ''zigzag symmetry'' is one of the basic properties of the QCD string. One must, however, be careful, even in pure Yang-Mills theory, since the loop requires regularization. Zigzag symmetry, in fact, is only true perturbatively for regularized loops, where the backtracking paths are closer than the ultraviolet cutoff. It was pointed out in [9] that the Wilson loop in the supersymmetric theory (2.3), with the constraint $\dot{x}^2 = \dot{y}^2$, does not have this symmetry. This is because the couplings of the Wilson loop to the scalar fields Φ^i are proportional to $|\dot{x}|$, which does not change the sign when the loop backtracks. Thus, if the loop stays at the same point θ^i on S^5, there is no cancellation of the coupling to the scalar fields.

In perturbation theory, one can easily prove that the zigzag symmetry holds for the Wilson loop (2.4) when $\dot{y}^i = 0$. Suppose we have a segment C_1 of a loop which goes in one direction and another segment C_2 which comes back parallel to C_1, but in the opposite direction, as shown in Fig. 5. If the distance η between C_1 and C_2 is much less than the UV regularization ϵ of the gauge theory, there is one-to-one cancellation between a Feynman diagram Γ which has one of its external leg ending on C_1 and another diagram Γ' which is identical to Γ except that the corresponding leg ends on C_2. Therefore, to all order in the perturbative expansion, the segments C_1 and C_2 do not contribute to the expectation value of the Wilson loop. On the other hand, if $\dot{y}^i = |\dot{x}|\theta^i$ and θ^i is fixed at a point on S^5, a diagram with a leg coupled to $\dot{y}^i$ on C_1 and one with the corresponding leg coupled to $\dot{y}^i$ on C_2 add up, rather than cancel each other. The perturbative computation therefore shows no zigzag symmetry in this case.

When the coupling λ is large, we expect that $\langle W \rangle$ is related to the minimal surface. The area functional, and as a matter of fact any other functional which is an integral over a minimal surface, has zigzag symmetry. The proof is simple. If we look at the region $Y \geq \epsilon$, the minimal surface bounded by a backtracking loop is almost identical to the surface bound by the curve without backtracking if the separation η between C_1 and C_2 is much less than the cutoff ϵ. This is illustrated in Fig. 6. Therefore an action on the surface given by an integral over the part of the surface in $Y \geq \epsilon$ is the same with or without the backtracking.

At first sight, the zigzag symmetry of the minimal surface appears in contradiction with the gauge theory expectation since we know the minimal surface ending along a smooth loop on the boundary of AdS_5 obeys the constraint $\dot{x}^2 = \dot{y}^2$ and therefore $\dot{y}^i \neq 0$. In the gauge theory, we do *not* expect zigzag symmetry when $\dot{y}^i$ is nonzero and constant. A close examination of the boundary condition, however, reveals that the situation is more subtle. It is true that, if we impose the boundary conditions at $Y=0$, the part of the surface connecting C_1 and C_2 does not reach $Y=\epsilon$ and does not contribute to the regularized area for $Y \geq \epsilon$. Therefore zigzag symmetry holds for $\langle W \rangle$. This is also the case when we impose the boundary condition at $Y=\epsilon$. In this case, if $\epsilon \gg \eta$, the minimal surface goes from C_1 to C_2 along the $Y=\epsilon$ surface. Therefore the contribution of the segments to the regularized area is proportional to η/ϵ^2 times the length of the segment and vanish in the limit $\eta \to 0$.

However, the physical interpretations of the two computations are quite different. If the boundary conditions are imposed at $Y=0$, the constraint $\dot{x}^2 = \dot{y}^2$ holds provided the segments C_1 and C_2 are smooth. On the other hand, if the conditions are imposed on the $Y=\epsilon$ hypersurface, the minimal surface bounded by C_1 and C_2 stays within η from $Y=\epsilon$, and $\dot{y}^2$ vanishes as $\eta/\epsilon \to 0$. If we take the latter point of view, the apparent contradiction with the gauge theory expectation disappears since the minimal surface in question is related to the Wilson loop which does not couple to the scalar fields in the segments C_1 and C_2. This is exactly the situation in which zigzag symmetry arises in the gauge theory.

One may argue that the boundary condition at $Y=\epsilon$ gives a more precise definition of the Wilson loop $\langle W \rangle$ as a functional of the loop variables $(x^\mu(s), y^i(s))$. The Legendre transformation of the area A in Sec. III C, for example, is a way to define a functional of the momenta P^i evaluated at $Y=\epsilon$ and not at $Y=0$. It does not make sense to perform this procedure at $Y=0$ since the factor $1/\epsilon$ on the right-hand side of Eq. (3.25) needs to be replaced by ∞. In most of the cases discussed in this paper, whether we impose the boundary conditions at $Y=0$ or $Y=\epsilon$ does not make much difference since the value of the momenta P^i stays almost the same in the region $0 \leq Y \leq \epsilon$. The analysis of zigzag symmetry, however, seems to be an exception to this rule. If we use the boundary condition at $Y=\epsilon$, the existence of the minimal surface requires the constraint $\dot{y}^i(s) = 0$ rather than $\dot{x}^2 = \dot{y}^2$ for the backtracking loop, and the result fits well with the

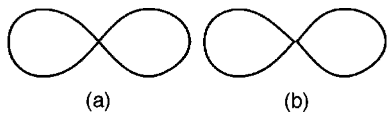

FIG. 7. (a) A self-intersecting loop which corresponds to a single trace operator and (b) a pair of loops obtained by reconnecting the loop at the intersection.

gauge theory expectation. Clearly, the regularization-dependent nature of zigzag symmetry needs to be clarified further.

An analysis similar to the one given above leads to the following observations about the Wilson loop, which we find interesting. Consider a self-intersecting loop as in Fig. 7. The area calculated on the minimal surface bound by the loop (a) is the same as the sum of the two areas bounded by the separated loops (b). In the gauge theory, these loops are very different objects. One is a single trace operator and the other a multitrace operator.

We can even connect two distant closed loops by a long neck without changing the value of the loop since the minimal surface spanning the neck region does not contribute to the area. Graphically, this can be written as

$$\left\langle {}^{i}_{j}\;\overline{\overline{\qquad\qquad}}\;{}^{k}_{l} \right\rangle = \frac{1}{N}\delta_{ij}\delta_{kl}. \tag{3.35}$$

This suggests that the parallel transport $U = \mathcal{P}\exp(i\int A_\mu dx^\mu)$ along an open curve behaves as a random matrix. As in the case of the zigzag symmetry, if we impose the boundary condition at $Y=\epsilon$, the minimal surface exists only when $\dot{y}^i(s)=0$, and we are considering a loop which does not couple to the scalar fields in the neck region.

E. Removing the constraint

So far we considered loops of the form (2.3) which satisfy the constraint $\dot{x}^2-\dot{y}^2=0$. When the loop has a cusp or an intersection, this constraint is modified as in Eqs. (3.20) and (3.21). In the gauge theory, we can define the loop operator for any $(x^\mu(s),y^i(s))$, not necessarily obeying the constraint. Consequently, we need to find a way to calculate an expectation value of such a loop in AdS_5 so that the relation between the gauge theory and string theory is complete.

The reason given by Maldacena for the constraint (and also in Appendix A) is that the W bosons are BPS particles and their charges and masses are related. To break the constraint, one needs a non-BPS object with an arbitrary mass. Fortunately, string theory contains many such objects. Instead of considering the ground state of the open string corresponding to the W boson, one may use excited string states, which have extra mass from the string oscillations. As shown in the Appendix A, an excited string indeed generates a loop obeying the modified constraint

$$\dot{y}^2=\dot{x}^2\frac{M^2}{M^2+m^2}, \tag{3.36}$$

where $M=\epsilon^{-1}$ is the original W-boson mass and m is the mass of the excitations. This makes it possible to relax the constraint, at least for $\dot{x}^2 \geqslant y^2$.

For the loop obeying the original constraint $\dot{x}^2=\dot{y}^2$, the regularized area has the linear divergence of the form

$$A=\frac{1}{2\pi\epsilon}\oint ds|\dot{x}|+\cdots=\frac{1}{2\pi}\oint M|\dot{x}|+\cdots \quad . \tag{3.37}$$

We expect that the corresponding computation using the string excitation replaces M by $\sqrt{M^2+m^2}$ as

$$A=\frac{1}{2\pi}\oint ds\sqrt{M^2+m^2}|\dot{x}|+\cdots=\frac{1}{2\pi\epsilon}\oint ds\frac{\dot{x}^2}{|\dot{y}|}+\cdots. \tag{3.38}$$

The Legendre transformation turns this into

$$\begin{aligned}\tilde{A}&=A-\frac{1}{2\pi\epsilon}\oint ds|\dot{y}|\\&=\frac{1}{2\pi\epsilon}\oint ds\left(\frac{\dot{x}^2}{|\dot{y}|}-|\dot{y}|\right)+\cdots.\end{aligned} \tag{3.39}$$

This shows that the linear divergence is not completely canceled for $|\dot{x}|\neq|\dot{y}|$. Since a highly excited string state may be sensitive to stringy corrections, we can trust this estimate of the linear divergence only for small deviation from the constraint. In the following, we will use an approximate expression for $|\dot{x}|\sim|\dot{y}|$ as

$$\tilde{A}=\frac{1}{\pi\epsilon}\oint ds(|\dot{x}|-|\dot{y}|)+\cdots. \tag{3.40}$$

IV. LOOP EQUATION

Since the expectation value of the Wilson loop is a measure of confinement, much attention has been given to calculating them. In particular, in the large N limit of gauge theory, they satisfy a closed set of equations [2]. In this section, we first give a review of the loop equation for pure Yang-Mills theory (for more details see [21,22]). The equation is easy to write down and is formally satisfied, order by order, in the perturbative expansion of the gauge theory. The lattice version of the loop equations is also satisfied in the nonperturbative lattice formulation of the theory. However, the only case where one can solve explicitly for Wilson loops is in two dimensions. There indeed they do satisfy the loop equation. We will then formulate the loop equation for the $\mathcal{N}=4$ super Yang-Mills theory in four dimensions. As far as we know, the loop equation in this case has not been derived before. We will find that the BPS condition (3.3) will play a crucial role. We will discuss details of the construction in Appendix C and present only the general ideas here.

A. Bosonic theories

The action of pure gauge theory in any number of dimensions is[8]

$$S=\frac{1}{4g_{\rm YM}^2}\int dx\,{\rm Tr}\,F_{\mu\nu}F^{\mu\nu}, \tag{4.1}$$

and the Wilson loop is given by

$$W=\frac{1}{N}{\rm Tr}\,\mathcal{P}\exp\left(i\oint A_\mu dx^\mu\right), \tag{4.2}$$

where the integral is over a path parametrized by x^μ. The main observation is that there is a differential operator on loop space which brings down the variation of the action $D^\nu F_{\mu\nu}$ as

$$\hat{L}\langle W\rangle=-i\oint ds\,\dot{x}^\mu\left\langle (D^\nu F_{\mu\nu})^a(s)\frac{1}{N}{\rm Tr}\,\mathcal{P}T^a(s)\times\exp\left(i\oint A_\mu dx^\mu\right)\right\rangle, \tag{4.3}$$

where $T^a(s)$ is the generator of the gauge group inserted at the point s along the loop.

There are a few equivalent definitions of $\hat{L}$. We will use

$$\hat{L}=\lim_{\eta\to 0}\oint ds\int_{s-\eta}^{s+\eta}ds'\frac{\delta^2}{\delta x^\mu(s')\,\delta x_\mu(s)}. \tag{4.4}$$

As we will explain below, η has to be taken much shorter than the UV cutoff scale ϵ in order to extract the term $D^\nu F_{\mu\nu}$. The insertion of $D^\nu F_{\mu\nu}$ into the loop would be zero if we use the classical equation of motion, but quantum corrections produce contact terms. To see that, one can write the equations of motion as the functional derivative of the action $\mathcal{S}$ and use the Schwinger-Dyson equations, i.e., integration by parts in the functional integral,

$$\begin{aligned}\hat{L}\langle W\rangle&=ig_{\rm YM}^2\int\mathcal{D}A\oint ds\frac{1}{N}{\rm Tr}\,\mathcal{P}T^a(s)\\&\quad\times\exp\left(i\oint A_\mu dx^\mu\right)\dot{x}^\mu(s)\frac{\delta e^{-s}}{\delta A^{\mu a}(x(s))}\\&=-ig_{\rm YM}^2\left\langle\oint ds\dot{x}^\mu(s)\frac{\delta}{\delta A^{\mu a}(x(s))}\frac{1}{N}{\rm Tr}\,\mathcal{P}T^a(s)\right.\\&\quad\left.\times\exp\left(i\oint A_\mu dx^\mu\right)\right\rangle.\end{aligned} \tag{4.5}$$

The functional derivative $\delta/\delta A_\mu(x(s))$ in this equation is formally evaluated as

$$\begin{aligned}\hat{L}\langle W\rangle&=\frac{\lambda}{N^2}\oint ds\oint ds'\,\delta(x^\mu(s')-x^\mu(s))\dot{x}_\mu(s)\dot{x}^\mu(s')\\&\quad\times\left\langle{\rm Tr}\,\mathcal{P}T^a(s)T^a(s')\exp\left(i\oint A_\mu dx^\mu\right)\right\rangle.\end{aligned} \tag{4.6}$$

We then use the relation between the generators of SU(N),

$$T^a_{nm}T^a_{kl}=\delta_{nk}\delta_{ml}-\frac{\delta_{nm}\delta_{kl}}{N}. \tag{4.7}$$

Ignoring the $1/N$ term, the trace is broken into two. This gives the correlation function of two loops. In the large N limit, the correlator factorizes and we obtain

$$\begin{aligned}\hat{L}\langle W\rangle&=\lambda\oint ds\oint ds'\,\delta(x^\mu(s')-x^\mu(s))\\&\quad\times\dot{x}_\mu(s)\dot{x}^\mu(s')\langle W_{ss'}\rangle\langle W_{s',s}\rangle.\end{aligned} \tag{4.8}$$

Here $W_{ss'}$ is a Wilson loop that start at s and goes to s' and $W_{s's}$ goes from s' to s. They are closed due to the delta function.[9]

Equation (4.8) shows that $\hat{L}\langle W\rangle$ receives contributions from self-intersections of the loop. Since the derivation of the equation is rather formal, it is not clear whether we need to count the trivial case of $s=s'$, in which case $W_{ss'}=1$ and $W_{s's}=W$. In most of the literature on the loop equation, this trivial self-intersection is ignored. In any case, it can be taken care of by multiplicative renormalization of the loop operator. In the supersymmetric gauge theory, the leading contribution from the trivial self-intersection cancels when $\dot{x}^2=\dot{y}^2$.

In the definition of the loop derivative $\hat{L}$, it is important to take the limit $\eta\to 0$. This procedure isolates the term $D^\nu F_{\nu\mu}$, which is a contact term of the double functional derivative. If η is of the order of the UV cutoff ϵ, there will be other contributions to the loop equation such as $F_{\mu\nu}F^{\nu\rho}\dot{x}_\rho$. When calculating the loop equation in perturbation theory, we can take η to be arbitrarily small, and in particular $\eta\ll\epsilon$. This is how we view the loop equation in the continuum theory. In fact, it was shown that the perturbative expansion of the Wilson loop solves the loop equation [23]. When we study the loop equation the string in AdS_5, we will consider the same limit $\eta\to 0$.

In the lattice regularization, it is not possible to calibrate the variation of the loop in distance shorter than the lattice spacing ϵ. In this case, a different definition of $\hat{L}$ is used which does not require taking such a limit.

[8]The complete action contains a gauge-fixing term and ghosts. Those appear also in the equations of motion, but can be dropped by a Ward identity [23].

[9]The delta function is not sharp, but is regularized by the cutoff ϵ. That means that the loops $W_{ss'}$ and $W_{s's}$ are not exactly closed loops, and the two ends may be separated by a distance ϵ. This does not contradict gauge invariance since one may consider only gauge transformations which do not vary much over that scale, so the ''almost'' closed loops are ''almost'' gauge invariant. We expect those loops to be equal to the closed loops up to $O(\epsilon)$ corrections.

It is possible to define a loop derivative localized at a point on the loop, instead of the integrated version considered above. The entire derivation goes through by simply dropping one $\oint ds$.

B. Supersymmetric case

We briefly summarize how to derive the loop equation in the supersymmetric theory, leaving the details to Appendix C. We derive them only for variations from constrained loops $\dot{x}^2=\dot{y}^2$. One important modification is due to the extra factor of i in front of the scalars in the Wilson loop operator in the Euclidean theory:

$$W=\frac{1}{N}\operatorname{Tr}\mathcal{P}\exp\left(\oint(iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\right). \tag{4.9}$$

Another novelty is the need to include the fermions. The fermions are important even when the loop equation is evaluated at the body part $\zeta(s)=0$ of super loop space since the fermions appear as source terms in the equations of motion for the gauge fields and the scalars. Here we will explain the effect of the extra i. In Appendix C, we will discuss how to deal with the fermions.

If we define loop derivative

$$\hat{L}=\lim_{\eta\to 0}\oint ds\int_{s-\eta}^{s+\eta}ds'\left(\frac{\delta^2}{\delta x^\mu(s')\,\delta x_\mu(s)}-\frac{\delta^2}{\delta y^i(s')\,\delta y_i(s)}\right), \tag{4.10}$$

then the relative minus sign combines with the extra i to give

$$\hat{L}\langle W\rangle=-i\frac{g_{\rm YM}^2}{N}\oint ds\left\langle\left(\dot{x}^\mu\frac{\delta}{\delta A^{\mu a}}-i\dot{y}^i\frac{\delta}{\delta\Phi^{ia}}\right)\operatorname{Tr}\mathcal{P}T^a\times\exp\left(\oint(iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\right)\right\rangle$$
$$=\lambda\oint ds\oint ds'[\dot{x}^\mu(s)\dot{x}_\mu(s')-\dot{y}^i(s)\dot{y}_i(s')]\times\delta^4(x(s)-x(s'))\langle W_1\rangle\langle W_2\rangle. \tag{4.11}$$

A simple way to obtain this is by considering the extra i as the Wick rotation of the y^i coordinates and repeat the derivation from Eq. (4.4) to (4.8). The right-hand side of the bosonic loop equation contains a cubic divergence proportional to the circumference of the loop. In the supersymmetric case this ''zero-point energy'' cancels for a smooth loop by the constraint $\dot{x}^2=\dot{y}^2$.

C. Predictions

In this subsection, we evaluate the right-hand side of the loop equation (4.11) for various types of loops. In the next section, we will compare it with computations of the loop using the minimal surface spanned by the loop in AdS_5.

In the supersymmetric theory, the trivial self-intersection at $s=s'$ does not contribute to the right-hand side if the loop is smooth and obeys the constraint $\dot{x}^2=\dot{y}^2$. This is related to the fact that such a loop does not require regularization. To be precise, the constraint only cancels the leading divergence proportional to ϵ^{-3}. Since the delta function in Eq. (4.11) has a width ϵ, the Taylor expansion of $x(s')$ at $s'=s$ gives subleading terms in ϵ such as

$$-\frac{\lambda}{3\epsilon}\oint ds(\ddot{x}^2-\ddot{y}^2). \tag{4.12}$$

However, this expression is highly regularization dependent. Moreover, there are other contributions of the same order due to the fact that the loops $W_{ss'}$ and $W_{s's}$ are not precisely closed, as explained in footnote 9. At any rate, these terms are negligible (by a factor ϵ) compared to the terms we will find at cusps and intersections, and we will ignore them for the rest of the paper.

For a loop with an intersection, the integral over the regularized delta function on the right-hand side of the loop equation gives

$$\lambda(\cos\Omega+\cos\Theta)\oint ds\oint ds'|\dot{x}(s)||\dot{x}(s')|\times\delta_\epsilon^4(x^\mu(s)-s^\mu(s'))$$
$$=\lambda(\cos\Omega+\cos\Theta)\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dx'\,\delta_\epsilon^4(\sin\Omega(x-x'))$$
$$=\lambda\frac{\cos\Omega+\cos\Theta}{2\pi\epsilon^2\sin\Omega}. \tag{4.13}$$

It is important to note that the result depends explicitly on the UV cutoff ϵ^{-2}. Here we have evaluated the leading term in the ϵ^{-1} expansion only. There are subleading terms in the expansion which are comparable to Eq. (4.12) at the trivial self-intersection.

A cusp also gives an interesting contribution to the loop equation. This may be regarded as a special case of the trivial self-intersection. In fact, in the literature, this effect is ignored together with that of the trivial self-intersection.[10] In the supersymmetric theory, the contribution from the trivial self-intersection at a smooth point on the loop is canceled by the constraint $\dot{x}^2=\dot{y}^2$. The situation is more interesting at the cusp since the tangent vector $\dot{x}^\mu(s)$ is discontinuous there. If there is a jump on S^5, $\dot{y}^i(s)$ is also discontinuous. A simple calculation [identical to Eq. (2.12), where we found the logarithmic divergence in perturbation theory] shows that the cusp contribute to the right-hand side of the loop equation as

$$2\lambda(\cos\Omega+\cos\Theta)\int_{-\infty}^{0}dx\int_{0}^{\infty}dx'\,\delta_\epsilon^4(\sin\Omega(x-x'))$$
$$=\lambda\frac{(\pi-\Omega)(\cos\Omega+\cos\Theta)}{(2\pi\epsilon)^2\sin\Omega}. \tag{4.14}$$

To summarize, we can express the loop equation as

[10]In the lattice formulation, the effect of the cusp to the loop equation is not seen since there is no local definition of a cusp.

$$\hat{L}\langle W\rangle=\frac{\lambda}{2\pi\epsilon^2}\left(\sum_{n:\quad\text{cusps}}\frac{(\pi-\Omega_n)(\cos\Omega_n+\cos\Theta_n)}{2\pi\sin\Omega_n}\langle W\rangle\right.$$
$$\left.+\sum_{m:\quad\text{intersections}}\frac{\cos\Omega_m+\cos\Theta_m}{\sin\Omega_m}\langle W_m\rangle\langle\tilde{W}_m\rangle\right)$$
$$+O\left(\frac{\lambda}{\epsilon}\right), \tag{4.15}$$

where W_m and $\tilde{W}_m$ are Wilson loops one obtains by detaching the original loop into two at the intersection point m.

V. LOOP EQUATION IN $AdS_5\times S^5$

A. General case

In this section, we will examine whether the computation of the loop using string theory in AdS_5 agrees with the predictions of the loop equation. A general form of the loop expectation value is

$$\langle W\rangle=\Delta\exp(-\sqrt{\lambda}\tilde{A}). \tag{5.1}$$

We assume that the dependence of the prefactor Δ on the loop variables is subleading for large λ. Since the loop derivative $\hat{L}$ does not commute with the constraint $\dot{x}^2=\dot{y}^2$, we need an expression for $\tilde{A}$ when the constraint is not satisfied. As we saw in Sec. III E the exponent $\tilde{A}$ has a linear divergence of the form

$$\tilde{A}(x,y)=\frac{1}{\pi\epsilon}\oint ds(|\dot{x}|-|\dot{y}|)+\cdots \tag{5.2}$$

to the leading order in $(|\dot{x}|-|\dot{y}|)$. The loop derivative is a second order differential operator. When the derivatives act on the exponent and bring it down twice, the result is proportional to λ. On the other hand, when they act on Δ or on the same $\tilde{A}$ twice, we get things only of order λ or less. In the following, we will pay attention to the leading term in λ only. The exact expression we have to evaluate is, therefore,

$$\lambda\lim_{\eta\to 0}\oint ds\int_{s-\eta}^{s+\eta}ds'\left(\frac{\delta\tilde{A}}{\delta x^\mu(s')}\frac{\delta\dot{A}}{\delta x_\mu(s)}-\frac{\delta\tilde{A}}{\delta y^i(s')}\frac{\delta\tilde{A}}{\delta y_i(s)}\right). \tag{5.3}$$

We do not have to include the fermionic derivative. When it acts once on a bosonic loop, it gives a fermion whose expectation value is zero. There are also nonzero contributions when it acts twice on $\tilde{A}$, but they are subleading in λ.

Let us evaluate Eq. (5.3). Although the linear divergence $(1/2\pi\epsilon)\oint ds(|\dot{x}|-|\dot{y}|)$ in $\tilde{A}(x,y)$ vanishes for the loop obeying the constraint, the variation $\hat{L}$ does not commute with the constraint. Thus the linear divergence term gives an important contribution to Eq. (5.3). Since the variation of the length functional

$$L=\oint ds\sqrt{\dot{x}^2} \tag{5.4}$$

gives the acceleration $\ddot{x}^\mu$ (in the parametrization where $|\dot{x}|=1$) and the same for y, we obtain

$$\lambda\left(\frac{\delta\tilde{A}}{\delta x^\mu(s')}\frac{\delta\dot{A}}{\delta x_\mu(s)}-\frac{\delta\tilde{A}}{\delta y^i(s')}\frac{\delta\tilde{A}}{\delta y_i(s)}\right)$$
$$=\frac{\lambda}{\pi^2\epsilon^2}[\ddot{x}_\mu(s)\ddot{x}^\mu(s')-\ddot{y}_i(s)\ddot{y}^i(s')]+\cdots. \tag{5.5}$$

Note that it has the same divergence, ϵ^{-2}, as the right-hand side of the loop equation. Moreover, the powers of λ match up in the loop equation and in Eq. (5.5). The ellipsis on the right-hand side represents variations of the remaining terms in $\tilde{A}$, which are finite for a smooth loop. To compute $\hat{L}\langle W\rangle$, we integrate Eq. (5.5) over $s-\eta\leqslant s'\leqslant s+\eta$. When the loop is smooth, the acceleration $(\ddot{x}^\mu,\ddot{y}^i)$ itself is finite. Therefore, by taking $\eta\to 0$, one finds that $\hat{L}\langle W\rangle=0$ in this case. This is consistent with the loop equation. Therefore we reach the first conclusion, that a minimal surface in AdS_5 bounded by a smooth loop solves the loop equation.

B. Loops with cusps

If the loop has a cusp of angle Ω, the tangent vector is discontinuous and $\ddot{x}$ has a delta function pointing along the unit vector bisector $\hat{e}$:

$$\ddot{x}^\mu=2\cos\frac{\Omega}{2}\,\delta(s)\hat{e}^\mu. \tag{5.6}$$

A similar thing happens when $\dot{y}$ is discontinuous, with the angle Θ replacing Ω in the above. This delta function is regularized by η, not ϵ, since it is related to the shortest length scale on which the loop is defined. Thus the integral of Eq. (5.5) over s and s' gives a nonzero result as

$$\frac{\lambda}{\pi^2\epsilon^2}\oint ds\int_{s-\eta}^{s+\eta}ds'[\ddot{x}_\mu(s)\ddot{x}^\mu(s')-\ddot{y}_i(s)\ddot{y}^i(s')]$$
$$=\frac{4\lambda}{\pi^2\epsilon^2}\left(\cos^2\frac{\Omega}{2}-\sin^2\frac{\Theta}{2}\right)$$
$$=\frac{2\lambda}{\pi^2\epsilon^2}(\cos\Omega+\cos\Theta). \tag{5.7}$$

In comparison with the prediction of Eq. (4.14) of the loop equation, we are missing the factor of $(\pi-\Omega)/\sin\Omega$. This, however, is not a contradiction. The expression for the linear divergence term in Eq. (5.2) is an approximation for small $(|\dot{x}|-|\dot{y}|)$. Since $\dot{x}^2=(1+f_0)\dot{y}^2$ with $f_0=f(\Omega/2)$ at the cusp, this approximation is valid only when f_0 is small. Apart from this factor, Eq. (5.7) agrees with the prediction of the loop equation, that the cusp gives a nonzero contribution to the loop equation proportional to $\lambda=g_{YM}^2N$ times ϵ^{-2}.

When $(|\dot{x}|-|\dot{y}|)$ is not small, the expression (5.2) needs to be modified as

$$\tilde{A}(x,y)=\frac{1}{\pi\epsilon}\oint ds|\dot{x}|G\left(\frac{\dot{y}^2}{\dot{x}^2}\right)+\cdots \tag{5.8}$$

for some function $G(z)$. By repeating the computations that lead to Eq. (5.7), we find that the contribution of the cusp takes the form

$$\hat{L}\exp(-\sqrt{\lambda}\tilde{A})=\lambda\mathcal{G}(f_0)(\cos\Omega+\cos\Theta)\exp(-\sqrt{\lambda}\tilde{a})+\cdots, \tag{5.9}$$

where $\mathcal{G}(f_0)$ is a function related to $G(z)$. The agreement with Eq. (4.14) requires

$$\mathcal{G}(f(\Omega/2))=\frac{\pi-\Omega}{8\sin\Omega}. \tag{5.10}$$

Proving this would be a very strong evidence for the conjecture.

Loops with cusps have also logarithmic divergences, which could contribute to the loop equations. To see that, one may write the logarithmically divergent term as

$$\frac{1}{2\pi}F(\Omega)\log\frac{L}{\epsilon}=\frac{1}{2\pi}\int ds\int ds'|\dot{x}(s)||\dot{x}(s')| \times\frac{\sin\varphi}{\pi-\varphi}F(\varphi)\frac{1}{(x-x')^2+\epsilon^2} \tag{5.11}$$

where $\pi-\varphi$ is the angle between $\dot{x}(s)$ and $\dot{x}(s')$. To check this equation one should integrate over two straight lines meeting at a point. Differentiating Eq. (5.11) gives a few terms, among them

$$\ddot{x}(s)\frac{1}{\epsilon}\frac{\sin\Omega}{\pi-\Omega}F(\Omega), \tag{5.12}$$

which has the same divergence as the piece that gave Eq. (5.7).

C. Self-intersecting loops

The situation at a self-intersection is more mysterious since $\dot{x}$ and $\dot{y}$ are both continuous at the intersection point. However, we have problems in our ability to test the loop equation in this case. First of all, $\dot{y}^i=0$ at the intersection, and the function $G(z)$ which appears in the linear divergence term in Eq. (5.8) may be singular at $z=|\dot{y}|/|\dot{x}|=0$. Since we do not know about the function $G(z)$ except for its behavior near $z=1$, it is difficult to tell whether there is a contribution from the intersection.

The presence of the unknown factor Δ in Eq. (5.1) makes the situation worse. As we explained before, the Wilson loop is

$$\langle W\rangle=\Delta\exp(-\sqrt{\lambda}\tilde{A}). \tag{5.13}$$

For a self-intersecting loop we expect

$$\hat{L}\langle W_{1+2}\rangle=\lambda\frac{\cos\Omega+\cos\Theta}{\sin\Omega}\langle W_1\rangle\langle W_2\rangle, \tag{5.14}$$

where W_{1+2} is the self-intersecting loop and W_1 and W_2 its two pieces. In order for this to be consistent with the AdS_5 computation, we need to find

$$\hat{L}\exp(-\sqrt{\lambda}\tilde{A}_{1+2})=\lambda\frac{\cos\Omega+\cos\Theta}{\sin\Omega}\frac{\Delta_1\Delta_2}{\Delta_{1+2}} \times\exp[-\sqrt{\lambda}(\tilde{A}_1+\tilde{A}_2)]. \tag{5.15}$$

Since we do not know the relation between the factors Δ_1, Δ_2, and Δ_{1+2}, a quantitative test is difficult in this case although it seems unlikely that the ration would be zero.

It would be very interesting to determine the function $G(z)$ which appears in the linear divergence as it would settle the question as to whether the intersection gives the contribution to $\hat{L}\exp(-\sqrt{\lambda}\tilde{A})$ predicted by the loop equation.

VI. DISCUSSION

The AdS-CFT correspondence allows us to calculate certain Wilson loops in terms of minimal surfaces in anti–de Sitter space. We presented a few reasons why only loops satisfying the constraint $\dot{x}^2=\dot{y}^2$ (generically) are given in terms of minimal surfaces. For more general loops we run into the problem of inconsistent boundary conditions.

The constrained loops are invariant under half of the local supersymmetry in super loopspace. As such, they are BPS objects and are free from divergences. The area of the minimal surface is divergent, so it is not the correct functional that yields the Wilson loop. Since the minimal surface satisfies Neumann boundary conditions, its natural to take for the action the Legendre transform of the area. We showed this yields a finite result.

In other examples of the AdS-CFT correspondence the action has to be modified as well. In nonsupersymmetric cases, such as the near extremal D3-brane, the effect of adding the boundary term is to subtract $L/(2\pi\epsilon)$. The result is finite, but contains a piece proportional to the circumference times the radius of the horizon. This may be considered a mass renormalization of the W boson. The scale of the renormalization is not the UV cutoff, but rather the scale of supersymmetry breaking. In addition, if $\dot{x}^2\neq\dot{y}^2$, the Wilson loop will contain a linear divergence proportional to the UV cutoff.

The surface observables on the M5-brane theory, as calculated in $AdS_7\times S^4$ have quadratic and logarithmic divergences [9,18,28]. Taking the Legendre transformation will eliminate the quadratic divergence, but we are not sure whether it will also remove the logarithmic divergence.

Recently, there were some attempts to go beyond the classical calculation and include fluctuations of the minimal surfaces [24,25,26]. One of the goals was to find the ''Lüscher term,'' the Coulomb-like correction to the linear potential in a confining phase [27]. Any attempt to perform such a calculation will require using the correct Neumann boundary conditions on the spherical coordinates and including the appropriate boundary terms.

Finally, we formulated the loop equations for those loops and checked if the AdS ansatz satisfies them. For smooth loops, as a result of the supersymmetry, the loop equations should give zero. This is indeed the result we find also from variation of the minimal surface.

This calculation actually requires extending the prescription to loops that do not satisfy the constraint. We propose that the natural extension for small deviation from the constraint gives a linear divergence proportional to $\sqrt{\lambda}L$. This term is particularly important when we consider the loop equations for loops with cusps. The expected result is finite and proportional to λ. This is in fact what we find, but we do not have enough control over the calculation to compare the coefficients.

The situation with self-intersecting loops is more mysterious: we expect a nonzero answer, but cannot reproduce that. There are, however, some reasons why this test is more difficult than the other cases. In particular, the constraint is broken by a large amount at the intersection.

Classical string theory tells us only how to calculate loops satisfying the constraint. These are BPS objects in loop space and, therefore, easier to control. As we argued, non-BPS Wilson loops are related to excited open strings, but we are unable to evaluate them reliably. A similar statement is true for local opeartors: one has control only over the chiral operators. Nonchiral operators should be given by excited closed string states. Despite the large effort devoted to testing the Maldacena conjecture, there is still no good understanding of non-BPS objects.

ACKNOWLEDGMENTS

We thank Korkut Bardakci, Sunny Itzhaki, Juan Maldacena, Joe Polchinski, and Bruno Zumino. N. D. thanks Theory Group of UC Berkeley and LBNL for their hospitality. H. O. thanks the Institute for Theoretical Physics at Santa Barbara, Department of Mathematics and Physics at University of Amsterdam, and Theory Group at CERN, for their hospitality and for providing excellent working environments. The work of H.O. was supported in part by NSF grant PHY-95-14797 and DOE grant DE-AC03-76SF00098. This work was supported in part by the NSF under grant PHY94-07194.

APPENDIX A: DERIVATION OF THE WILSON LOOP

In this appendix, we will define the coupling of the Wilson loop to the bosonic fields, A_μ and Φ^i, in the $\mathcal{N}=4$ super Yang-Mills theory. We will pay special attention to the effect of Wick rotation on the Euclidean signature space. In a gauge theory containing a matter field in the fundamental representation of the gauge group, the Wilson loop is derived by writing a correlation function of the matter fields in terms of the first quantized path integral over trajectories of the corresponding particle. The resulting phase factor dictates the proper coupling of the Wilson loop to the gauge field. The $\mathcal{N}=4$ super Yang-Mills theory in four dimensions does not contain such fields. Instead, we use W bosons that appear when we break $\mathrm{SU}(N+1)\to\mathrm{SU}(N)\times\mathrm{U}(1)$.

The bosonic action for the $\mathrm{SU}(N+1)$ theory is

$$\hat{S}=\frac{1}{4}\hat{F}_{\mu\nu}^2+\frac{1}{2}(\hat{D}_\mu\hat{\Phi}_i)^2-\frac{1}{4}[\hat{\Phi}_i,\hat{\Phi}_j]^2. \tag{A1}$$

By decomposing the gauge group to $\mathrm{SU}(N)\times\mathrm{U}(1)$ as

$$\hat{A}_\mu=\begin{pmatrix} A_\mu & w_\mu \\ w_\mu^\dagger & a_\mu \end{pmatrix},\quad \hat{\Phi}_i=\begin{pmatrix} \hat{\Phi}_i & w_i \\ w_i^\dagger & M\theta_i \end{pmatrix}, \tag{A2}$$

with $\theta^2=1$, the action can be written as

$$\begin{aligned}\hat{S}=&+\frac{1}{4}F_{\mu\nu}^2+\frac{1}{2}(D_\mu\Phi_i)^2-\frac{1}{4}[\Phi_i,\Phi_j]^2+\frac{1}{2}(\partial_\mu M\theta_i)^2\\ &+(\partial_{[\mu}a_{\nu]})^2+\frac{1}{2}w_i^\dagger[(\Phi_k-M\theta_k)^2\delta_{ij}\\ &-(\Phi_i-M\theta_i)(\Phi_j-M\theta_j)]w_j+\frac{1}{2}[(D_\mu-ia_\mu)w_i]^2+\cdots\\ =&S_{\mathrm{SU}(N)}+\frac{1}{2}(\partial_\mu M\theta_i)^2+\frac{1}{4}f_{\mu\nu}^2+\frac{1}{4}[(D_\mu-ia_\mu)w_i]^2\\ &+\frac{1}{2}w_i^\dagger[(\Phi_k-M\theta_k)^2\delta_{ij}-(\Phi_i-M\theta_i)(\Phi_j-M\theta_j)]w_j\\ &+\cdots,\end{aligned} \tag{A3}$$

where $F_{\mu\nu}$ and $f_{\mu\nu}$ are the field strengths of the $\mathrm{SU}(N)$ and $\mathrm{U}(1)$ factors, respectively. The ellipsis in the action represents terms in higher powers of w_i, etc. If θ_i is in the one-direction, the mass term for w_i with $i\neq 1$ becomes

$$w_i^\dagger(\Phi_1-M\theta_1)^2w_i-w_i^\dagger\Phi_i(\Phi_1-M\theta_1)w_1 \tag{A4}$$

with approximate mass eigenvalues $\Phi_1-M\theta_1$. To simplify the following analysis, we replace these terms with

$$w^\dagger(\Phi_1-M\theta_1)^2w. \tag{A5}$$

Let us consider the correlation function

$$\langle w(x)^\dagger w(x)w(y)^\dagger w(y)\rangle. \tag{A6}$$

We can integrate over the w field and find

$$\begin{aligned}&\int \mathcal{D}A_\mu\mathcal{D}\Phi_i\mathcal{D}w\mathcal{D}a_\mu\mathcal{D}M\theta_i e^{-\hat{S}}w(x)^\dagger w(x)w(y)^\dagger w(y)\\ &=\int \mathcal{D}M\theta_i\mathcal{D}a_\mu \exp\left(\int \frac{1}{2}(\partial_\mu M\theta_i)^2+\frac{1}{4}(f_{\mu\nu})^2\right)\\ &\quad\times\int \mathcal{D}A_\mu\mathcal{D}\Phi_i e^{-S_{\mathrm{SU}(N)}}\\ &\quad\times\left\langle x\left|\frac{1}{-\frac{1}{2}(D_\mu-ia_\mu)^2+\frac{1}{2}(\Phi_i-M\theta_i)^2}\right|y\right\rangle\\ &\quad\times\left\langle y\left|\frac{1}{-\frac{1}{2}(D_\mu-ia_\mu)^2+\frac{1}{2}(\Phi_i-M\theta_i)^2}\right|x\right\rangle.\end{aligned} \tag{A7}$$

The correlation functions in this expression can be written as

$$\left\langle x\left|\frac{1}{-\frac{1}{2}(D_\mu-ia_\mu)^2+\frac{1}{2}(\Phi_i-M\theta_i)^2}\right|y\right\rangle$$
$$=\int dT\left\langle x\left|\exp\left[T\left(\frac{1}{2}(D_\mu-ia_\mu)^2-\frac{1}{2}(\Phi_i-M\theta_i)^2\right)\right]\right|y\right\rangle$$
$$=\int dT\int_{x(0)=x}^{x(T)=y}\mathcal{D}x(s)\mathcal{D}p(s)\exp\left[\int_0^T ds\left(-i\dot{x}_\mu p^\mu-\frac{1}{2}(p_\mu+A_\mu+a_\mu)^2-\frac{1}{2}(\Phi_i-M\theta_i)^2\right)\right]$$
$$\times\int dT\int_{x(0)=x}^{x(T)=y}\mathcal{D}x(s)\exp\left[\int_0^T ds\left(-\frac{1}{2}\dot{x}_\mu^2+iA_\mu\dot{x}^\mu+ia_\mu\dot{x}^\mu-\frac{1}{2}(\Phi_i-M\theta_i)^2\right)\right].\tag{A8}$$

Combining everything together and integrating over y, we obtain

$$\int dy\langle w(x)^\dagger w(x)w(y)^\dagger w(y)\rangle=\int\mathcal{D}M\theta_i\exp\left(-\int\frac{1}{2}(\partial_\mu M\theta_i)^2\right)\int dT\int_{x(0)=x}^{x(T)=x}\mathcal{D}x(s)$$
$$\times\exp\left(-\frac{1}{2}\int_0^T ds(x_\mu^2+M^2)\right)\int\mathcal{D}a_\mu\exp\left(\int\frac{1}{4}(f_{\mu\nu})^2 e\oint ds\, ia_\mu\dot{x}^\mu\right)\int\mathcal{D}A_\mu\mathcal{D}\Phi_i e^{-S_{\mathrm{SU}(N)}}$$
$$\times\exp\left[\int ds\left(iA_\mu\dot{x}^\mu-\frac{1}{2}\Phi_i^2+M\Phi_i\theta^i\right)\right].\tag{A9}$$

Let us examine Eq. (A9) carefully. The first term $M^2\int\frac{1}{2}(\partial_\mu\theta_i)^2$ is the action of the θ_i field, which for large M becomes classical. The second term includes an integral over all the closed paths through x. To define the Wilson loop we just look at one such path, leaving the integration over paths for latter. The next term in the exponent breaks reparametrization invariance and will set $\dot{x}_\mu^2=\theta_i^2$, as shown below. The next term is the action for the Abelian gauge field on the single brane and the effect of the Wilson loop on it. Since $N\gg 1$ and we are taking the probe approximation, we should ignore this term. As we will see, for large M the Φ^2 term will be subleading, so the last term is simply the Wilson loop

$$\langle W(x^\mu,\theta_i)\rangle=\int\mathcal{D}A_\mu\mathcal{D}\Phi_i e^{-S_{\mathrm{SU}(N)}}\times\exp\left(\int ds(iA_\mu\dot{x}^\mu-\Phi_i\theta^i)\right).\tag{A10}$$

The term with $\dot{x}^2+M^2$ is not reparametrization invariant. When we perform the integral over different parametrizations of the same path (including the integral over T), we find a saddle point. A general parametrization is $s\to\tilde{s}(s)$ such that $\tilde{s}(0)=0$ and $\tilde{s}(\tilde{T})=T$. To integrate over different parametrizations, we can perform the path integral over $c(s)=d\tilde{s}/ds$ with the action

$$-\int_0^T ds\frac{1}{2}\left(\frac{1}{c}\dot{\tilde{x}}_\mu^2+cM^2\right)+\int_0^T ds\left(iA_\mu\dot{\tilde{x}}^\mu-c\frac{1}{2}\Phi_i^2+cM\Phi_i\theta^i\right).\tag{A11}$$

For large M the first term dominates, so it will pick the saddle point

$$c(s)^2=\frac{\dot{x}_\mu^2}{M^2},\tag{A12}$$

and indeed the Φ^2 piece in the loop drops out.

Combining them together, we obtain

$$\int dy\langle w(x)^\dagger w(x)w(y)^\dagger w(y)\rangle$$
$$=\int\tilde{\mathcal{D}}x(s)\exp\left(-\int dsM|\dot{x}|\right)\int\mathcal{D}A_\mu\mathcal{D}\Phi_i e^{-S_{\mathrm{SU}(N)}}\times\exp\left(\int_0^1 ds(iA_\mu\dot{x}^\mu+|\dot{x}|\Phi_i\theta^i)\right).\tag{A13}$$

The integral $\int ds|\dot{x}|M$ is the length of the loop times the mass LM. Since it is a c number independent of λ, we can ignore it as subleading in the large λ analysis in this paper. For the same reason, possible determinant factors are also neglected in the above.

The calculation above can also be done in Lorentzian signature. The difference is an extra i in Eq. (A8):

$$\left\langle x\left|\frac{1}{-\frac{1}{2}(D_\mu-ia_\mu)^2+\frac{1}{2}(\Phi_i-M\theta_i)^2}\right|y\right\rangle$$
$$=\int dT\left\langle x\left|\exp\left[iT\left(+\frac{1}{2}(D_\mu-ia_\mu)^2-\frac{1}{2}(\Phi_i-M\theta_i)^2\right)\right]\right|y\right\rangle.\tag{A14}$$

The rest of the calculation carries through with this i showing up in different places. The final result is

$$\begin{aligned}\int \mathcal{D}\theta_i \exp\left(i\int \frac{1}{2}(\partial_\mu\theta_i)^2\right)\int \tilde{\mathcal{D}}x(s) \\ \times\exp\left(i\int_0^1 ds M|\dot{x}|\right)\int \mathcal{D}a_\mu \exp\left(i\int \frac{1}{4}(f_{\mu\nu})^2\right) \\ \times\exp\left(i\int ds\, a_\mu \dot{x}^\mu\right)\int \mathcal{D}A_\mu \mathcal{D}\Phi_i e^{iS_{\mathrm{SU}(N)}} \\ \times\exp\left(i\int_0^1 ds(A_\mu\dot{x}^\mu+|\dot{x}|\Phi_i\theta^i)\right)\end{aligned} \tag{A15}$$

though it is less clear now why the term $i(\dot{x}^2+M^2)$ should dominate the path integral to set the saddle point.

Instead of the W boson, we may consider a more general particle with an arbitrary mass with a propagator

$$\frac{1}{\frac{1}{2}(D_\mu-ia_\mu)^2+\frac{1}{2}(\Phi_i-M\theta_i)^2+\frac{1}{2}m^2}. \tag{A16}$$

By the same calculation as above, we obtain the exponent

$$-\int_0^1 ds\sqrt{\dot{x}_\mu^2(M^2+m^2)}+\int_0^1 ds\left(iA_\mu\dot{x}^\mu+\frac{M|\dot{x}|}{\sqrt{M^2+m^2}}\Phi_i\theta^i\right). \tag{A17}$$

Excited states of the open strings have this propagator and can be used to construct loops with $\dot{x}^2\neq\dot{y}^2$.

So far θ is a constant. To construct loops which move in the θ directions, we have to use many probe D-branes, one for each value of θ the loop goes through. We start with $\mathrm{SU}(N+M)$ and break to $\mathrm{SU}(N)\times\mathrm{SU}(M)$, which will then be broken to $\mathrm{SU}(N)\times\mathrm{U}(1)^M$. Likewise, one should be able to couple the loop to the fermions to get the supersymmetric loops used in Appendix C.

APPENDIX B: AREA OF A CUSP

1. At one point on S^5

Here we study the minimal surface near a cusp. We consider a loop on a two-dimensional plane in four dimensions, staying at the same point on S^5. We take the opening angle of the cusp to be Ω. We choose radial coordinates r and φ on the plane and use them to parametrize the world sheet also. The boundary conditions are (using the first regularization discussed in Sec. III B)

$$Y(r,0)=Y(r,\Omega)=0. \tag{B1}$$

To study the behavior of the surface near the cusp, we can use scale invariance to set

$$Y(r,\varphi)=\frac{r}{f(\varphi)}. \tag{B2}$$

Using this ansatz, the area is

$$A=\frac{1}{2\pi}\int dr d\varphi \frac{1}{r}\sqrt{f^4+f^2+f'^2}. \tag{B3}$$

This reduces the minimal surface to a one-dimensional problem with the effective Lagrangian

$$L=\int d\varphi\sqrt{f^4+f^2+f'^2}. \tag{B4}$$

Since L does not depend explicitly on φ, the energy E given by

$$E=\frac{f^2+f^4}{\sqrt{f^4+f^2+f'^2}} \tag{B5}$$

is conserved. At the minimum of f, the energy is given by

$$E=f_0\sqrt{1+f_0^2}\quad [f_0=f(\Omega/2)]. \tag{B6}$$

Substituting this back in Eq. (B5),

$$\begin{aligned}\frac{\Omega}{2}&=\int_0^{\Omega/2}d\varphi \\ &=f_0\sqrt{1+f_0^2}\int_{f_0}^\infty \frac{df}{f\sqrt{(1+f^2)(f^2-f_0^2)(f^2+f_0^2+1)}} \\ &=f_0\sqrt{1+f_0^2}\int_0^\infty \frac{dz}{(z^2+f_0^2)\sqrt{(z^2+f_0^2+1)(z^2+2f_0^2+1)}} \\ &=\frac{i}{f_0}\Pi\left(\arcsin i\infty, \frac{\sqrt{1+2f_0^2}}{f_0}, \sqrt{\frac{1+2f_0^2}{1+f_0^2}}\right),\end{aligned} \tag{B7}$$

where Π is an elliptic integral of the third kind. The regularized action is then

$$\begin{aligned}L&=\int_{r\geq\epsilon f(\varphi)} d\varphi\sqrt{f^4+f'^2} \\ &=\int dz\sqrt{\frac{z^2+f_0^2+1}{z^2+2f_0^2+1}} \\ &=i\sqrt{1+f_0^2}E\left(\arcsin i\sqrt{\frac{\frac{r^2}{\epsilon^2}-f_0^2}{1+2f_0^2}}, \sqrt{\frac{1+2f_0^2}{1+f_0^2}}\right),\end{aligned} \tag{B8}$$

where E is an elliptic integral of the second kind. For small ϵ, it diverges linearly as $2r/\epsilon-F(\Omega)$. The function F is obtained by solving Eq. (B7) for f_0 as a function of Ω and substituting it into L in the above. The total area is

$$A=\frac{1}{2\pi}\int^L dr\frac{1}{r}\left(\frac{2r}{\epsilon}-F(\Omega)\right)=\frac{2L}{2\pi\epsilon}-\frac{1}{2\pi}F(\Omega)\log\frac{L}{\epsilon}. \tag{B9}$$

This is the regular linear divergence plus a logarithmic divergence. After the Legendre transformation, we obtain

$$\tilde{A}=-\frac{1}{2>p}F(\Omega)\log\frac{L}{\epsilon}. \tag{B10}$$

2. With a jump on S^5

The same analysis can be done for a loop which jumps, at the cusp, to a different point on S^5 with a relative angle Θ. We parametrize the string world sheet by r and θ, where θ is a coordinate along the large circle connecting the two different points on S^5. Because of scale invariance, we can set

$$Y(r,\theta)=\frac{r}{\tilde{f}(\theta)}, \tag{B11}$$

for some function $\tilde{f}(\theta)$. The other angular parameter φ is a function of θ only. The area is therefore

$$A=\int dr\, d\theta \frac{1}{r}\sqrt{\tilde{f}'^2+(1+\tilde{f}^2)(1+\tilde{f}^2\varphi'^2)}. \tag{B12}$$

The problem is integrable since there are two conserved quantities

$$E=\frac{1+\tilde{f}^2}{\sqrt{\tilde{f}'^2+(1+\tilde{f}^2)(1+\tilde{f}^2\varphi'^2)}},$$

$$J=\frac{(1+\tilde{f}^2)\tilde{f}^2\varphi'}{\sqrt{\tilde{f}'^2+(1+\tilde{f}^2)(1+\tilde{f}^2\varphi'^2)}}. \tag{B13}$$

In general, the result cannot be written in terms of elliptic integrals, and we will leave it to the overmotivated reader to find simple expressions for those integrals. If we set $\Omega=\pi$, there is no cusp in the x plane. In this case, the integrals are simplified, and the results are expressed in terms of the elliptic integrals.

APPENDIX C: DETAILS OF THE LOOP EQUATION IN $\mathcal{N}=4$ SUPER YANG-MILLS THEORY

The bosonic part of the Euclidean Wilson loop is

$$W=\frac{1}{N}\mathrm{Tr}\mathcal{P}\exp\left(\int (iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\right). \tag{C1}$$

We can define the bosonic part of the loop derivative to be

$$\hat{L}=\lim_{\eta\to 0}\int ds\int_{s-\eta}^{s+\eta}ds'\left(\frac{\delta^2}{\delta x^\mu(s')\,\delta x_\mu(s)}-\frac{\delta^2}{\delta y^i(s')\,\delta y_i(s)}\right). \tag{C2}$$

The extra i in front of $\Phi_i\dot{y}^i$ in the exponent conspires with the relative minus sign in the loop derivative to give the bosonic part of the equations of motion

$$\begin{aligned}\hat{L}\langle W\rangle=-i\int ds\Big\langle\Big(&\dot{x}^\mu(D^\nu F_{\mu\nu})^a\\&+i\dot{x}^\mu[\Phi^i,D_\mu\Phi_i]^a+i\dot{y}^i(D^\nu D_\nu\Phi_i)^a\\&-i\dot{y}^i[\Phi^j,[\Phi_i,\Phi_j]]^a\Big)\mathrm{Tr}\,\mathcal{P}T^a(s)\\&\times\exp\int (iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\Big\rangle.\end{aligned} \tag{C3}$$

This is a linear combination of the bosonic equations of motion for A_μ and Φ^i, but we are missing source terms due to the fermions. What we would like to do here is to modify the functional differential operator $\hat{L}$, including derivatives of fermionic variables, so that the full equations of motion are reproduced. With such $\hat{L}$, the loop equation can be written as

$$\begin{aligned}\hat{L}\langle W\rangle&=-i\frac{g_{\mathrm{YM}}^2}{N}\int ds\Big\langle\Big(\dot{x}^\mu\frac{\delta}{\delta A^{\mu a}}-i\dot{y}^i\frac{\delta}{\delta\Phi^{ia}}\Big)\mathrm{Tr}\,\mathcal{P}T^a(s)\\&\quad\times\exp\Big(\int (iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\Big)\Big\rangle\\&=\lambda\int ds\int ds'[\dot{x}^\mu(s)\dot{x}_\mu(s')-\dot{y}^i(s)\dot{y}_i(s')]\\&\quad\times\delta^4(x(s)-x(s'))W_1W_2.\end{aligned} \tag{C4}$$

The Euclidean super Yang-Mills theory has fermionic fields Ψ which are Euclidean Majorana fermions [29] with 16 complex components. The gamma matrices Γ_M satisfy the Dirac algebra in ten dimensions with signature (10, 0), with the index $M=(\mu,i)$. The loop is parametrized by $(x^\mu(s),y^i(s))$ and their superpartner $\zeta(s)$ coupling to the gauginos Ψ.

A natural choice for the supersymmetrized loop is

$$\begin{aligned}W=\frac{1}{N}\mathrm{Tr}\,\mathcal{P}\Big[&\exp\Big(\int\bar{\zeta}(s)Q\,ds\Big)\\&\times\exp\Big(\int (iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)ds\Big)\exp(-\textstyle\int\bar{\zeta}(s)Q\,ds)\Big].\end{aligned} \tag{C5}$$

Here Q is the generator of supersymmetry of the gauge theory, which acts as

$$[Q,A_M]=\frac{i}{2}\Gamma_M\Psi,$$

$$\{Q,\Psi\}=-\frac{1}{4}\Gamma_{MN}F^{MN}, \tag{C6}$$

where we have combined the gauge field A_μ and the scalars Φ^i into the ten-dimensional gauge field A_M and computed the field strength F_{NM}. One may also include

$$[Q,\dot{x}_M]=\frac{i}{4}\Gamma_M\dot{\zeta} \tag{C7}$$

in the exponent, but it does not affect our analysis since we will only be interested at the top component of the Grassmann algebra and at the end of the calculation we set $\zeta=0$. The exponent of the Wilson loop is therefore given by

$$\begin{aligned}&e^{\bar{\zeta}Q}(iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)e^{-\bar{\zeta}Q}\\&\quad=(iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)-\frac{1}{2}\bar{\zeta}(\dot{x}^\mu\Gamma_\mu-i\dot{y}^i\Gamma_i)\Psi\\&\quad\quad-\frac{1}{16}\dot{x}^\mu F^{\nu\rho}\bar{\zeta}\Gamma_\mu\Gamma_{\nu\rho}\zeta+\cdots.\end{aligned} \tag{C8}$$

We will write the loop equation only for loops satisfying the constraint $\dot{x}^2=\dot{y}^2$. Therefore $\dot{x}^M\Gamma_M=\dot{x}^\mu\Gamma_\mu-i\dot{y}^i\Gamma_i$ is nilpotent. In this case, it is useful to work in the basis where

$$\dot{x}^\mu\Gamma_\mu-i\dot{y}^i\Gamma_i=\begin{pmatrix}0 & 0\\ 2|\dot{x}| & 0\end{pmatrix},\quad \zeta=\frac{1}{\sqrt{|\dot{x}|}}\begin{pmatrix}\zeta_1\\ \zeta_2\end{pmatrix},\quad \Psi=\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, \tag{C9}$$

and

$$\bar{\zeta}=\zeta^T\mathcal{C}, \tag{C10}$$

where $\mathcal{C}$ is the charge conjugation matrix. The Majorana spinor in Lorentzian signature space satisfies the reality condition $\bar{\zeta}=\zeta^\dagger\Gamma^0$. In the Euclidean case, we do not impose any reality condition [29]. The exponent of the loop Eq. (C8) in this basis becomes

$$(iA_\mu\dot{x}^\mu+\Phi_i\dot{y}^i)-\sqrt{|\dot{x}|}\bar{\zeta}_1\psi_1+\frac{1}{8}\sqrt{|\dot{x}|}F^{NM}\bar{\zeta}_1\Gamma_{NM}+\zeta+\cdots. \tag{C11}$$

By applying the fermionic derivative operator

$$\frac{1}{|\dot{x}|}\frac{\delta}{\delta\zeta(s')}\frac{\delta}{\delta\bar{\zeta}(s)}W\sim\frac{\delta}{\delta\zeta_1(s')}\frac{\delta}{\delta\bar{\zeta}_1(s)}W, \tag{C12}$$

we obtain the desired combination for the source terms in the equation of motion,

$$|\dot{x}|\bar{\psi}_1\psi_1=\bar{\Psi}(\dot{x}^\mu\Gamma_\mu+i\dot{y}^i\Gamma_i)\Psi. \tag{C13}$$

All other terms contain at least one $\zeta(s)$ and are not relevant for our analysis of the loop at $\zeta=0$. Thus we found that the supersymmetric loop derivative defined by

$$\hat{L}=\lim_{\eta\to 0}\int ds\int_{s-\eta}^{s+\eta}ds' \times\left(\frac{\delta^2}{\delta x^\mu(s')\,\delta x_\mu(s)}-\frac{\delta^2}{\delta y^i(s')\,\delta y_i(s)}+\frac{\delta}{\delta\zeta(s')}\frac{\delta}{\delta\bar{\zeta}(s)}\right) \tag{C14}$$

produces the variation of the action. For the loop at $\zeta=0$, this completes the loop equation for the $\mathcal{N}=4$ super Yang-Mills theory.

[1] J. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1997).
[2] Y. M. Makeenko and A. A. Migdal, Phys. Lett. **88B**, 135 (1979); Y. Makeenko and A. A. Migdal, Nucl. Phys. **B188**, 269 (1981).
[3] V. A. Kazakov and I. K. Kostov, Nucl. Phys. **B176**, 321 (1980).
[4] J. H. Schwarz, ''Introduction to *M* Theory and AdS/CFT Duality,'' hep-th/9812037.
[5] M. R. Douglas and S. Randjbar-Daemi, ''Two Lectures on the AdS/CFT Correspondence,'' hep-th/9902022.
[6] G. 't Hooft, Nucl. Phys. **B72**, 461 (1974).
[7] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, Phys. Lett. B **428**, 105 (1998).
[8] E. Witten, Adv. Theor. Math. Phys. **2**, 253 (1998).
[9] J. Maldacena, Phys. Rev. Lett. **80**, 4859 (1998).
[10] S.-J. Rey and J. Yee, ''Macroscopic Strings as Heavy Quarks in Large *N* Gauge Theory and Anti-de Sitter Supergravity,'' hep-th/9803001.
[11] E. Witten, Adv. Theor. Math. Phys. **2**, 505 (1998).
[12] A. Brandhuber, N. Itzhaki, J. Sonnenschein, and S. Yankielowicz, Phys. Lett. B **434**, 36 (1998).
[13] K. G. Wilson, Phys. Rev. D **10**, 2445 (1974).
[14] D. Z. Freedman, S. D. Mathur, A. Matusis, and L. Rastelli, Nucl. Phys. **B546**, 96 (1999).
[15] S. Lee, S. Minwalla, M. Rangamani, and N. Seiberg, Adv. Theor. Math. Phys. **B2**, 697 (1998).
[16] L. Susskind and E. Witten, ''The Holographic Bound in Anti-de Sitter Space,'' hep-th/9805114.
[17] N. Drukker, D. J. Gross, and H. Ooguri, announced by Gross at the Conference ''Strings '98,'' http://www.itp.ucsb.edu/online/strings98/gross/oh/10.html
[18] D. Berenstein, R. Corrado, W. Fischler, and J. Maldacena, Phys. Rev. D **59**, 105023 (1999).
[19] A. W. Peet and J. Polchinski, Phys. Rev. D **59**, 065006 (1999).
[20] A. M. Polyakov, Nucl. Phys. B (Proc. Suppl.) **68**, 1 (1997).
[21] A. A. Migdal, Phys. Rep. **102**, 199 (1984).
[22] A. M. Polyakov, *Gauge Fields and Strings, Contemporary Concepts in Physics* Vol. 3 (Harwood Academic, Chur, Switzerland, 1987), p. 301.
[23] R. A. Brandt, A. Gocksch, M. A. Sato, and F. Neri, Phys. Rev. D **26**, 3611 (1982).
[24] S. Forste, D. Ghoshal, and S. Theisen, ''Stringy Corrections to the Wilson Loop in $N=4$ Super Yang-Mills Theory,'' hep-th/9903042.
[25] J. Greensite and P. Olesen, J. High Energy Phys. **04**, 001 (1999).
[26] S. Naik, ''Improved Heavy Quark Potential at Finite Temperature from Anti-de Sitter Supergravity,'' hep-th/9904147.
[27] M. Lüscher, K. Symanzik, and P. Weisz, Nucl. Phys. **B173**, 365 (1980).
[28] C. R. Graham and E. Witten, Nucl. Phys. **B546**, 52 (1999).
[29] H. Nicolai, Nucl. Phys. **B140**, 294 (1978).