Modular-invariant closed-string field theory

Michio Kaku
Physics Department, City College of the City University of New York, New York, New York 10031

Joseph Lykken
Santa Cruz Institute for Particle Physics, University of California at Santa Cruz, Santa Cruz, California 95064

(Received 21 April 1988)

Attempts so far at constructing a covariant closed-string field theory have been frustrated by the fact that modular invariance always appears to be violated. At both the tree and loop levels, moduli space is either overcounted an infinite number of times, or undercounted because of a missing region. We solve this problem by demonstrating that a new *closed four-string interaction* is necessary to reproduce the closed-string amplitude which precisely fills the missing region. This closed four-string interaction, which has the topology of a tetrahedron, is predicted by geometric string field theory. The tetrahedron graph is generated by gauge fixing the geometric theory's local gauge group, the unified string group, and is the exact counterpart of the instantaneous four-fermion Coulomb term found in QED. We prove the existence of this tetrahedron graph both analytically and by direct computer calculation and show that it is the key to reproducing the Shapiro-Virasoro amplitude.

I. INTRODUCTION

Modular invariance is a crucial component of string theory which is intimately linked to space-time supersymmetry and finiteness. The original closed-string field theory in the light-cone gauge[1] can be shown to be modular invariant.[2,3] At the tree level, this is just the statement that the closed-string field theory reproduces one conformally inequivalent Riemann surface, the complex plane. However, at the loop level, the proof of modular invariance is nontrivial, because of the existence of Dehn twists, global diffeomorphisms which cannot be connected to the identity map.

However, attempts at generalizing string field theory to the covariant case have so far failed to reproduce modular invariance either at the tree or loop level. Generalizations based on the "covariantized light-cone"[4] approach are plagued by an infinite overcounting of moduli space, even at the tree level. This is because the fictitious "parametrization length" is not treated as a gauge degree of freedom. Furthermore, proposals based on generalizing Witten's action [5] can be shown to undercount moduli space because of a missing region. For example, the s-channel four-point amplitude does not continuously deform into the t-channel or the u-channel graphs.[6] As a consequence, there is a large chunk of the complex plane that is missing. This is not a trivial question, because the heterotic superstring theory,[7] which is the leading candidate for a theory of all known forces, is a closed-string theory.

Recently, one of us proposed an entirely new approach to string field theory called *geometric string field theory*[8,9] which is free of the problem of overcounting. Because the σ parametrization (including the parametrization length) is treated as a local gauge degree of freedom, we are free to gauge fix the parametrization length, and hence overcounting is avoided. As in Yang-Mills theory or general relativity, geometric string field theory can be derived by gauging a certain local gauge group (the "unified string group," which contains the Virasoro group as a subgroup).

Geometric string field theory is based on the physical, invariant length of strings,

$$L=\int_0^{2\pi\alpha} d\sigma\sqrt{X_\mu'(\sigma)^2}\ , \tag{1.1}$$

rather than the fictitious parametrization length:

$$2\pi\alpha=\int_0^{2\pi\alpha} d\sigma \tag{1.2}$$

which can be changed at any time.

The essential feature of the geometric theory is that it is defined in loop space, i.e., the space of physical, space-time strings, rather than parameter space. In loop space, we can isolate the essential physical dynamics of the string theory without the gauge-dependent complications due to parametrization lengths, parametrization midpoints, ghost counting, ghost insertions, etc. Thus, the theory necessarily contains a "string vierbein"[8,9] $e_{\mu\sigma}^{\nu\rho}$ which allows us to change the parametrization of the string at will. The string vierbein separates the geometric theory from the usual Becchi-Rouet-Stora-Tyutin (BRST) theories, which are defined in parametrization space and hence are "rigid" string field theories.

For open strings, the fictitious parametrization lengths in geometric string field theory can be gauge fixed either to reproduce the "end-point gauge" (i.e., the covariantized light-cone gauge) or the "midpoint gauge" of Witten. In fact, geometric string field theory, when gauge fixed, yields a new vertex function with *arbitrary* parametrization lengths which allows us to smoothly interpolate between the end-point gauge and the midpoint gauge,[9] see Fig. 1.

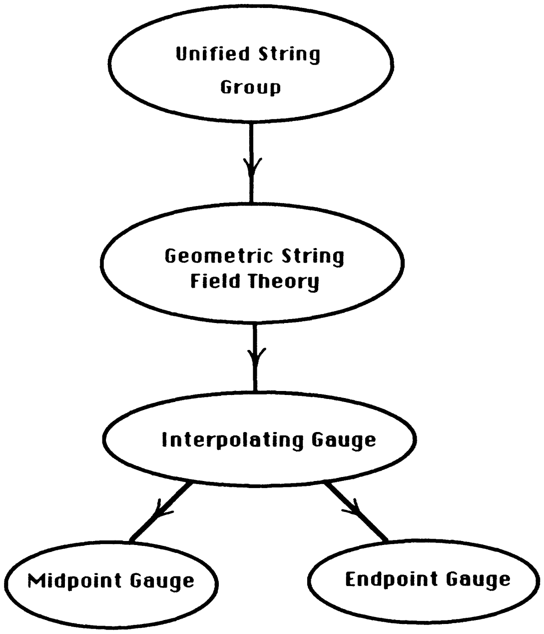

FIG. 1. The relationship between various closed-string field theories. The geometric string field theory is based entirely on a new "transfinite Lie group," which we call the unified string group. When gauged, we uniquely arrive at the geometric string field theory action. When we fix the parametrization of the various strings in an arbitrary way, we have the "interpolating gauge." When all strings are defined to have the same parametrization length $|\alpha|=\pi$, we have the "midpoint gauge," which necessarily has a four-string interaction. If the sum of the parametrization lengths equals zero $\sum \alpha=0$, then we have the "end-point gauge."

Geometric string field theory explains the origin of the open four-string interaction first introduced in Ref. 1. The meaning of this open four-string graph is simple: it is the counterpart of the instantaneous Coulomb term generated by gauge fixing the unified string group. In QED, for example, choosing the Coulomb gauge generates the instantaneous four-fermion term

$$\mathcal{L}_I=\bar{\psi}\gamma_0\psi\nabla^{-2}\bar{\psi}\gamma_0\psi \tag{1.3}$$

when we eliminate the gauge field A_0. Similarly, the open four-string interaction emerges when we eliminate the string vierbein $e_{\mu\sigma}^{\nu\rho}$. The open four-string interaction vanishes for the midpoint gauge but occupies a large portion of the integration region for the end-point gauge. In the interpolating gauge, the four-string interaction smoothly interpolates from the usual one found in the end-point gauge to the one for the midpoint gauge, where it vanishes.

We will see later that the origin of the open four-string interaction is due to the fact that the gauge group does not close properly in the end-point gauge unless the four-string interaction is added to the action.

For the closed-string case, however, the picture changes dramatically. Let us define the end-point gauge for closed-strings to describe light-cone-like configurations (a ring, of course, has no end, so we define the "end point" of three closed strings to be the mutual joining or splitting point). We then find the *opposite* situation: the unified string group closes properly for the end-point gauge, but not for midpoint-type configurations. This means that the midpoint-type configurations must necessarily have closed four-string interactions, rather than the end-point configurations, which is the opposite of the open-string case. *Thus, from group theory alone, we can show the necessity for a new closed four-string interaction which smoothly interpolates from midpoint-type configurations to end-point configurations, where it vanishes.*

The new closed four-string interaction, which we call the "tetrahedron graph," is shown in Fig. 2. Notice that the diagram is defined in physical loop space, so the interaction is spatially extended over several physical dimensions. Notice that the four strings collide at the center, forming the outline of a four-sided tetrahedron. If we parametrize this tetrahedron in σ space, we find the topology of Fig. 3. It is important to notice that Fig. 3 is defined totally in σ space, not σ,τ space. Geometric string field theory predicts that the new interaction in the midpoint gauge can be written in the form

$$\mathcal{L}_4=\int \prod_{i=1}^{4} \Phi_i[X_i]\mathcal{D}X_i d^2z\, \delta_{1234}\ , \tag{1.4}$$

where δ_{1234} is the overlap function for Fig. 3, where d^2z indicates that the tetrahedron graph occupies a nontrivial portion of the complex plane, and Φ_i is the usual closed-string field functional projected out by the $L_0-\tilde{L}_0$ operator. [We will also present the closed four-string interaction term in the interpolating gauge, which has a smaller region of integration than the midpoint contribution. We

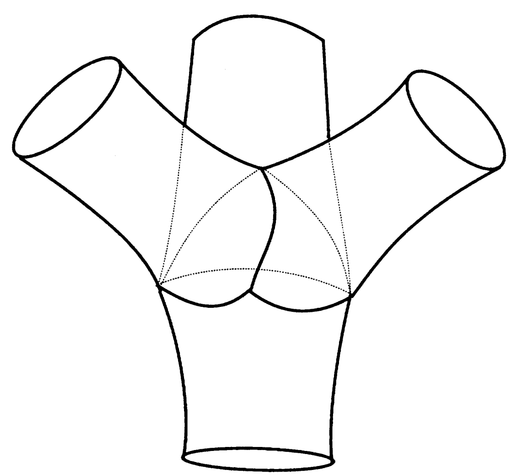

FIG. 2. The tetrahedron graph. In the midpoint gauge and the interpolating gauge, this four-string interaction must necessarily be added to action to restore modular invariance. This graph is a gauge artifact, a by-product of breaking the unified string group and choosing the midpoint gauge or interpolating gauge. It is the counterpart of the instantaneous four-fermion term found in QED, created by breaking U(1) invariance and choosing the Coulomb gauge.

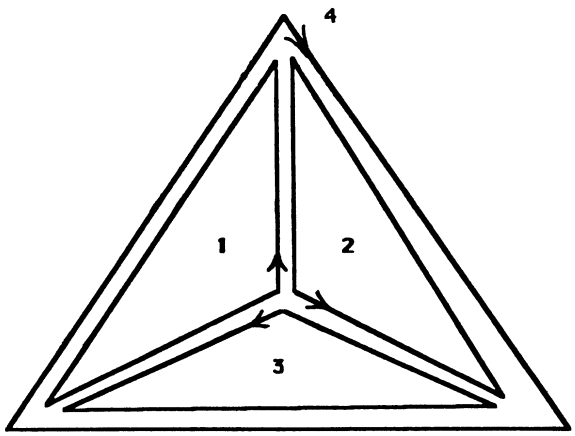

FIG. 3. In parameter space, the four-string graph has the topology of a tetrahedron. (The arrows point away from the origin.)

have omitted the contribution from the tangent space, i.e., the "ghost sector," which can be easily inserted into (1.4).]

The fact that this tetrahedron graph must be added to the action in the midpoint gauge can be seen intuitively. In Figs. 4(a) and 4(b), for example, we see two graphs which contribute to the s-channel pole for $1+2\rightarrow 3+4$ for strings of equal parametrization length 2π in the midpoint gauge. Figure 4(c) is contained in the t- and u-channel poles. Notice, however, that there is a missing piece of the integration region. Specifically, one may suspect that Fig. 5(a), which is an s-channel graph, might allow us to smoothly continue in the complex plane to the t- or u-channel graphs. But this is not so. If we move point b and point c past each other, keeping all string lengths constant, then we can topologically deform Fig. 5(a) into Fig. 5(b) such that the position of the second and fourth strings are reversed. Although Fig. 5(a) and 5(b) are topologically the same, the assignments of string lengths in Fig. 5(b) are not the usual ones. Thus, the string lengths are all wrong for it to be a t- or u-channel graph. It is easy to show, by this argument, that no diagrams of the type Fig. 5(a) can ever be smoothly continued into the t- or u-channel region. Thus, it is impossible to construct a modular-invariant theory based on the three-string vertex found in the midpoint gauge.

However, we have checked by explicit calculation, both analytically and on computer, that we reproduce the correct region of integration for the Shapiro-Virasoro amplitude once we include the four-string interaction for the midpoint gauge and the interpolating gauge. Thus, the missing piece of moduli space is given by the tetrahedron graph. The missing region, as expected, vanishes in the end-point gauge, smoothly increases in size in the interpolating gauge, and finally occupies most of the complex sphere in the midpoint gauge. By neglecting the tetrahedron graph, therefore, we accidentally throw away most of the integration region for the midpoint gauge.

For closed-strings, we can summarize the situation as follows.

End-point gauge:

$$\sum_{i=1}^{3} \alpha_i = 0, \text{ no four-string interaction;}$$

Midpoint gauge:

$$|\alpha_i| = \alpha, \text{ four-string interaction;}$$

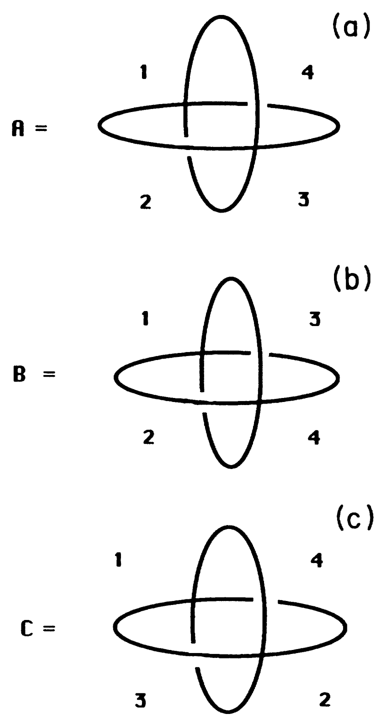

FIG. 4. The three "Rubik's cube" points. These three graphs correspond to the collision of four closed strings of equal parametrization length in the midpoint gauge. Notice that there are four ways in which the graph can be smoothly deformed by sliding strings of equal parametrization length past each other (vertically and horizontally, clockwise and counterclockwise), as in a Rubik's cube.

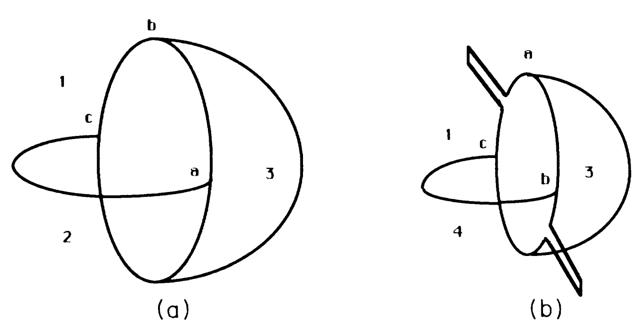

FIG. 5. In the midpoint gauge, it is impossible to continuously deform s-, t-, and u-channel graphs into each other, which means that the midpoint gauge with three-string interactions necessarily violates modular invariance. This missing region is filled by the tetrahedron graph. In (a), we see a typical s-channel graph. However, when we deform it into either a t- or u-channel graph (keeping all parametrizations length constant) notice that the parametrization lengths are all wrong.

Interpolating gauge:

α_i arbitrary, four-string interaction .

We should emphasize, however, that our end-point gauge is different from the formalism of Ref. 4 for several reasons, although they use essentially the same vertex function. First, there is no problem with overcounting of the parametrization length, because α is now treated as a genuine gauge degree of freedom in the geometric theory. Second, our higher loop graphs are totally different from those of Ref. 4. Our geometric theory has a proper time parameter, which yields the same number of diagrams as the old light-cone theory. By the results of Refs. 2 and 3, we know that the resulting theory is modular invariant. The theory of Ref. 4, however, because it does not have the usual light-cone counting for higher graphs, apparently has seven Teichmüller parameters (rather than the usual six) for each loop. Thus, even if the problem of overcounting of α could be solved, it is not clear whether the remaining theory is modular invariant because of this overcounting of Teichmüller parameters.

We stress that, by contrast, using the results of Refs. 2 and 3, our theory is *modular invariant to all orders in loops.* We will demonstrate this in a later paper.

II. TETRAHEDRON GRAPH

The advantage of geometric string field theory[8,9] is that it is based on an entirely new infinite-dimensional Lie algebra defined in physical loop space, rather than fictitious parameter space. We have the freedom of either choosing the end-point gauge[4] or the midpoint gauge.[5] Thus, the problem of overcounting moduli space, which is a persistent problem for the end-point gauge, is not a problem for geometric string field theory. The parametrization length of a string is a gauge artifact in the geometric theory. When we choose a different gauge, such as the midpoint gauge, geometric string field theory predicts that a new interaction must be added to the action, a tetrahedron graph as in Fig. 2 and Fig. 3. For details, see Refs. 8 and 9.

Earlier, we found that the midpoint gauge prevented us from going from one channel to the next. Let us now check that this tetrahedron graph allows us to go smoothly between the s-, t-, and u-channel graphs.

Let us define the symmetric a_{ij} as the parametrization length of the line common to both the ith and jth strings in Fig. 6. Let the parametrization length of each external closed-string be $2\pi\alpha_i$, where the α_i are arbitrary, so that we are in the interpolating gauge. (When $|\alpha_i|$ are all equal, we are in the midpoint gauge. When the sum of the α_i equals zero, we are in the end-point gauge. The interpolating gauge allows us to smoothly interpolate between these two gauge choices.) We have

$$\begin{aligned} &a_{12}+a_{13}+a_{14}=2\pi|\alpha_1| , \\ &a_{21}+a_{23}+a_{24}=2\pi|\alpha_2| , \\ &a_{31}+a_{32}+a_{34}=2\pi|\alpha_3| , \\ &a_{41}+a_{42}+a_{43}=2\pi|\alpha_4| . \end{aligned} \tag{2.1}$$

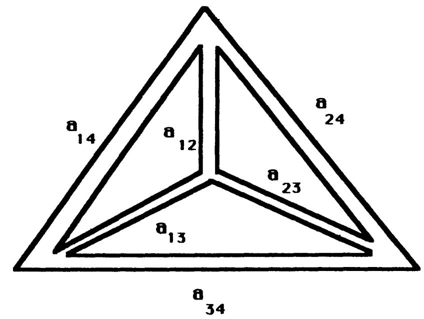

FIG. 6. The length a_{ij} is equal to the parametrization distance which is common to strings i and j. Notice that there are six such a_{ij}'s, but there are four constraints on them because the external string lengths are all fixed. Thus, the tetrahedron graph is parametrized by precisely two degrees of freedom, exactly the correct number of moduli to fill the missing region of the Shapiro-Virasoro amplitude.

Notice that there are six independent a_{ij} lengths. But there are only four constraint equations. This means that there are $6-4=2$ degrees of freedom in constructing the tetrahedron graph. This is precisely the number of degrees of freedom needed to account for the two-dimensional missing region of the complex plane. For example, let us parametrize the configurations shown in Figs. 4(a)–4(c) in the midpoint gauge. Let us label them by the points A, B, and C:

$$\begin{aligned} &\text{A:}\ a_{12}=\pi , a_{14}=\pi , \\ &\text{B:}\ a_{12}=\pi , a_{14}=0 , \\ &\text{C:}\ a_{12}=0 , a_{14}=\pi . \end{aligned} \tag{2.2}$$

If we plot this region in the a_{12}-a_{14} plane in Fig. 7, then we find that the region covered by the tetrahedron graph is a right triangle. The lines which connect these three points can be parametrized as

$$\begin{aligned} &\text{A}\to\text{B:}\ a_{12}=\pi , a_{14}=\pi\to 0 , \\ &\text{A}\to\text{C:}\ a_{12}=\pi\to 0 , a_{14}=\pi , \\ &\text{B}\to\text{C:}\ a_{12}+a_{14}=\pi . \end{aligned} \tag{2.3}$$

We will see that the region within this right triangle is the missing region of the complex plane. An integration

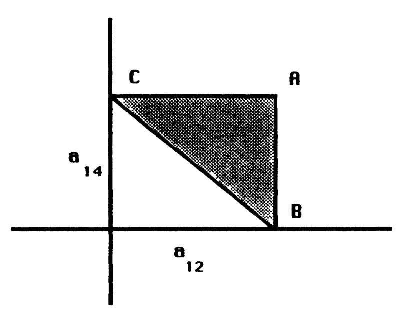

FIG. 7. In a_{ij} parameter space, the missing region can be shown to be a triangle. The corners of the triangle are the Rubik's cube points.

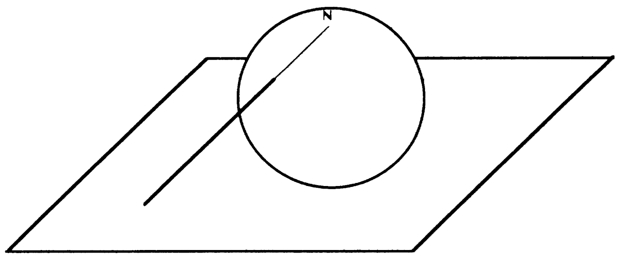

FIG. 8. A stereographic projection. The complex z plane of the Shapiro-Virasoro amplitude is mapped to the sphere.

over this two-dimensional region can be performed using the following:

$$da \equiv \prod_{i<j} da_{ij} \prod_{i=1}^{4} \delta \left[2\pi - \sum_{j\,(j\neq i)} a_{ij} \right] \theta(\pi - a_{ij}) \ . \quad (2.4)$$

To see how this parametrization can be viewed symmetrically, it is helpful to visualize the two-dimensional complex plane as a sphere. We map the two-dimensional complex plane to the sphere by a stereographic projection as in Fig. 8. The resulting figure is shown in Fig. 9, where the original triangle in Fig. 7 is now mapped to a triangular region in the northern hemisphere. Points inside the regions I, II, and III correspond to the s-, t-, and u-channel graphs. These two-dimensional regions represent the state of four strings before they collide. They are parametrized by two variables: the angle θ which the two three-vertices make with respect to each other before they collide and the separation τ between them. Notice that these regions do not make up the complete conformal sphere, and hence modular invariance is broken. It is impossible, as we saw, to smoothly go from one region to the next in the midpoint gauge with only three-point vertices.

Points A, B, and C, which correspond to Figs. 4(a), 4(b), and 4(c) and which used to form the vertices of a right triangle in Fig. 7, are now mapped to the equator of the sphere in Fig. 9. Most important is the "north pole,"

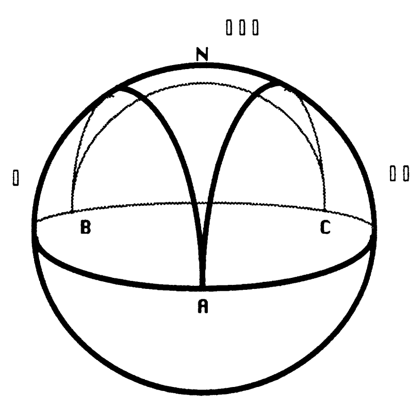

FIG. 9. Regions I, II, and III are filled by the midpoint gauge with only three-string interactions. Notice that there is missing region in the northern and southern hemispheres. This missing region is precisely filled by the tetrahedron graph. Points A, B, and C are the Rubik's cube points of Fig. 4.

the point in which all lines have the same parametrization length, i.e.,

$$\text{north pole:} \quad a_{ij} = \frac{2\pi}{3} \ . \quad (2.5)$$

(Notice that the north pole is clearly outside the region parametrized by the three-string vertex in the midpoint gauge without the tetrahedron graph.) Thus, a large portion of the conformal sphere is actually outside the region parametrized by the midpoint gauge vertex.

Similarly, there is also the "south pole" on the sphere, which corresponds to reversing the cyclic ordering of the external lines on the tetrahedron, which can be accomplished by reversing the order of two external lines. To see visually how we can enter the southern hemisphere, notice that there are two ways point A can be converted into point B in (2.3). We can either rotate strings 3 and 4 clockwise or counterclockwise with respect to 1 and 2, which remain stationary. Depending on our parametrization, one rotation puts us into the northern hemisphere, while the other puts us in the southern hemisphere.

Notice that from point A, there are four distinct rotations that we can make among the four strings in the midpoint gauge. We can rotate strings 3 and 4 clockwise (or counterclockwise), which will take us to point B on the sphere via the northern (southern) hemisphere. Or we can rotate strings 2 and 3 clockwise (counterclockwise) and reach point C via the northern (southern) hemisphere.

Points A, B, and C on the equator of Fig. 9 thus correspond to what we call "Rubik's cube" points. As in the Rubik's cube puzzle, we have the freedom of rotating the object clockwise or counterclockwise via an axis which is either horizontal or vertical. The four different rotations we can make at the Rubik's cube point correspond exactly to the four ways we can move from point A to either point B or C via the northern or southern hemisphere in Fig. 9.

Putting everything together, we can now write down an expression for the closed four-string interaction in the midpoint gauge (see Fig. 3):

$$\mathcal{L}_4 = \int \prod_{i=1}^{4} \Phi_i[X_i] \mathcal{D}X_i \delta_{1234} \mu \, da \ , \quad (2.6)$$

where

$$\begin{aligned} \delta_{1234} &= \prod_{i<j} \prod_{0<\sigma_i<2\pi} \delta(X_i(\sigma_i) - \theta_{ij} X_j(b_{ij} - \sigma_i)) \ , \\ \theta_{ij} &= \theta(c_{ij} - \sigma_i)\theta(\sigma_i - d_{ij}) \ , \end{aligned} \quad (2.7)$$

and where

$$\begin{aligned} & b_{12} = 2\pi, \ b_{13} = 2\pi, \ b_{14} = 2\pi + a_{12} \ , \\ & b_{23} = 2\pi, \ b_{24} = 2\pi - a_{12}, \ b_{34} = 2\pi - a_{14} + a_{13} \ , \\ & c_{12} = a_{12}, \ c_{13} = 2\pi, \ c_{14} = a_{12} + a_{14} \ , \\ & c_{23} = a_{23}, \ c_{24} = a_{23} + a_{24}, \ c_{34} = a_{13} + a_{34} \ , \\ & d_{12} = 0, \ d_{13} = a_{12} + a_{14}, \ d_{14} = a_{12} \ , \\ & d_{23} = 0, \ d_{24} = a_{23}, \ d_{34} = a_{13} \ . \end{aligned} \quad (2.8)$$

μ is the Jacobian which takes us from the sphere to the usual Koba-Nielsen plane. We will show how to write down an explicit representation for μ, as well as show how to derive (2.6) from geometric string field theory, in Section V. There, we will also show how to write down the closed four-string interaction in the interpolating gauge.

It is essential to point out that the closed four-string interaction is defined in loop space in Fig. 2 and in σ space in Fig. 3 with no extension in τ space. It is nontrivial, therefore, to show that a conformal map exists from the complex plane to this configuration such that the four-string interaction occupies no extension in the τ direction. We will now show this, both analytically and by computer calculation.

III. CONFORMAL MAPS

Now that we have exhibited the topology of the four-string interaction, let us verify, by explicit construction, that there is a missing region in the complex plane which is filled by the tetrahedron graph. Let us first begin by discussing the end-point configuration for the $4\to 1+2+3$ and the $1+2\to 3+4$ processes. There is no missing region for the end-point configuration.[1] When we smoothly make the transition from the end-point configuration to the interpolating gauge, the four-string interaction begins to occupy an increasingly larger portion of the complex plane. Finally, in the midpoint gauge, the four-string interaction actually occupies most of the complex plane.

The light-cone mapping[10] is given by

$$\rho(z)=\ln(z-1)-\delta\ln(z-x)+\ln z \ , \tag{3.1}$$

where the lengths of the four strings are given by

$$\begin{aligned}&\alpha_1=\alpha_3=\pi \ ,\\ &\alpha_2=-\pi\delta, \quad \alpha_4=\pi\delta-2\pi \ .\end{aligned} \tag{3.2}$$

By varying δ from -1 to 1, we make the transition from the $4\to 1+2+3$ process to the $1+2\to 3+4$ process in the end-point gauge.

Let us call w_+ and w_- the turning points in the complex z plane where $\rho'(z)=0$. Then

$$\begin{aligned}w_\pm=\{2x+1-\delta\pm[(2x+1-\delta)^2\\ -4(2-\delta)x]^{1/2}\}[2(2-\delta)]^{-1} \ .\end{aligned} \tag{3.3}$$

Setting $w_+=w_-$, and solving for x, we find

$$x_0=\tfrac{1}{2}\pm\tfrac{1}{2}(2\delta-\delta^2)^{1/2} \ . \tag{3.4}$$

Let us now choose $\delta=-1$ for the process $4\to 1+2+3$. In Fig. 10(a) we sketch the three regions of the complex x plane which describe the various physical processes. Each of the three regions described in the figure corresponds to the values of x which enter into the various channels of the Shapiro-Virasoro amplitude. The most important point is

$$x_0=e^{\pm\pi i/3} \tag{3.5}$$

which is the point common to all three regions. (Notice that the entire complex plane is filled, without the necessity of any closed four-string interaction. This was first pointed out in Ref. 1.) At this point, we have

$$z_0=w_+=w_-=\tfrac{1}{2}+\frac{i}{2\sqrt{3}} \ . \tag{3.6}$$

The boundary of all three regions can be found analytically. It corresponds to an equipotential line formed if we place equal electric charges at 0, x, and 1:

$$\begin{aligned}\ln|w_+|\,|w_+-1|\,|w_+-x|\\ =\ln|w_-|\,|w_--1|\,|w_--x| \ .\end{aligned} \tag{3.7}$$

We can rewrite this equation totally in terms of x by inserting (3.3) into (3.7):

$$|(x-\tfrac{1}{2})(x-2)(x+1)+(x^2-x+1)^{3/2}| = |(x-\tfrac{1}{2})(x-2)(x+1)-(x^2-x+1)^{3/2}| \ . \tag{3.8}$$

We can see that some of the solutions of this equation are $x=-1$, $\tfrac{1}{2}$, 2, $\pm i$, $1\pm i$ and $e^{\pm\pi i/3}$. By solving this equation, we find

$$(x-\tfrac{1}{2})^*(x-2)^*(x+1)^*(x^2-x+1)^{3/2}+(x-\tfrac{1}{2})(x-2)(x+1)(x^2-x+1)^{*3/2}=0 \ . \tag{3.9}$$

We now use the equations

$$\begin{aligned}&(x-\tfrac{1}{2})^*(x-2)^*(x+1)^*=x^{-3}(x-\tfrac{1}{2})(x-2)(x+1) \ ,\\ &(x^2-x+1)^{*3/2}=\pm(x^{-3})(x^2-x+1)^{3/2} \ ,\end{aligned} \tag{3.10}$$

which are satisfied if $xx^*=1$ or $(x-1)(x-1)^*=1$. (The particular branch of the cut we are using can be determined by placing simple values into the defining equation.) Thus, we find that the boundary separating the three regions is given by a circle of unit radius centered around the origin, as well as centered around the number 1, and the vertical boundary between them, as in Fig. 10(a).

The map (3.1) with $\delta=-1$ corresponds to having three external charges with equal magnitude, with the charge at infinity having charge -3π. We could also have chosen $\delta=3$, which would correspond to having three equal charges with the charge at infinity having charge π. Since conformal transformations treat the point at infinity as an ordinary point, we must also analyze this configuration as well. If we solve (3.3) for this choice and then solve (3.7) to find the lines which separate the three regions, we find that (3.10) is again the solution. Thus, the three regions still remain the same, even with this choice of δ.

Next, let us make the projective transformation which

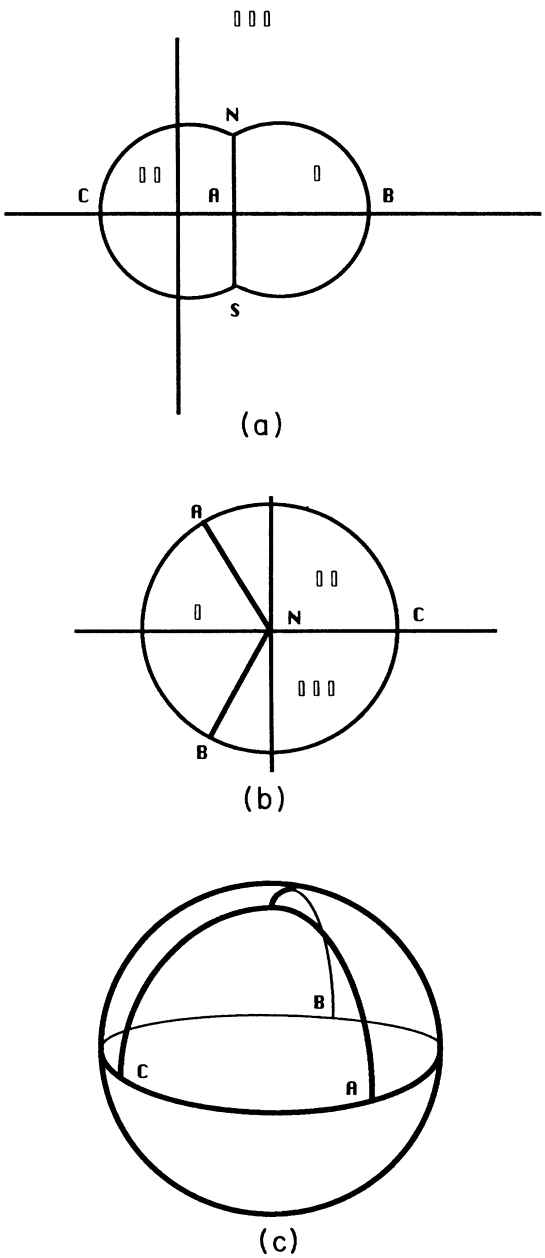

FIG. 10. In (a), the complex z plane in the end-point gauge has no missing region. Regions I, II, and III, which are generated by three-string interactions in the end-point gauge, fill up the complex plane entirely. The missing region has degenerated to the north and south poles. In (b), the complex z plane in the end-point gauge has been remapped, such that the north pole is now the origin. In (c), the complex z plane has been stereographically mapped to the sphere.

maps x_0 to the origin and A, B, and C to the circumference of a circle:

$$z' = -e^{2\pi i/3}\frac{z-e^{\pi i/3}}{z-e^{-\pi i/3}} . \tag{3.11}$$

Notice that the real axis of the complex plane is mapped to a circle, and that the circles of the complex plane are mapped to straight radial lines, as in Fig. 10(b). Then, we stereographically map the new complex plane to the sphere, as in Fig. 10(c). The origin of the new complex plane is mapped to the north pole, and the point at infinity in the new complex plane is mapped into the south pole.

In summary, we have now mapped the original complex plane to the sphere such that the sphere is divided symmetrically into three regions, with each region representing a different process:

I: $(1+2)+3\rightarrow 4$,

II: $1+(2+3)\rightarrow 4$,

III: $(1+3)+2\rightarrow 4$,

such that the configurations A, B, and C lie on the equator.

Now let δ gradually change to $+1$. The three regions then slowly collapse onto the real axis. At the point $\delta=+1$, the three circles have collapsed totally onto the real axis from the origin to the point $+1$. Thus, for $\delta=+1$, the three separate regions of the complex plane collapse into one region.

Now that we have mapped out the complex plane for the end-point gauge (where the four-string interaction vanishes), let us investigate the mapping for the interpolating gauge, when the external strings can have arbitrary parametrization length. By the Schwartz-Christoffel transformation, we have the following map:[11]

$$\frac{d\rho}{dz} = k\frac{\prod_{i=1}^{N-2}[(z-w_i)(z-\bar{w}_i)]^{1/2}}{\prod_{j=1}^{N-1}(z-x_j)} = \frac{f(z)}{g(z)} , \tag{3.12}$$

where $x_1=1$, $x_2=x$, $x_3=0$, $x_N=\infty$ and

$$k=-\alpha_N . \tag{3.13}$$

This transformation guarantees that the points x_i correspond to the points at $\pm\infty$ in the ρ plane, and that w_i corresponds to points which are the zeros of ρ', the "turning points." We connect the pairs of points w_i and $\bar{w}_i$ by Riemann cuts, which are in general not straight lines.

Furthermore, we must impose boundary conditions, such as those at infinity, which will fix the lengths of each string in terms of the w's. If each string has parametrization length α_i, then the conditions are

$$\alpha_i = \lim_{z\rightarrow x_i}\frac{f(z)(z-x_i)}{g(z)} . \tag{3.14}$$

Furthermore, we must impose

$$\text{Re}\rho(w_i)=\text{Re}\rho(\bar{w}_i) . \tag{3.15}$$

This guarantees that the interaction of three strings takes place instantaneously in τ space. Finally, we can also impose

$$\text{Im}\rho(w_i)-\text{Im}\rho(\bar{w}_i)=a_{ij} , \tag{3.16}$$

where the constant a_{ij} in (3.16) and Fig. 6 determines the parametrized distance over which the ith and jth strings merge during the interaction, for some j. ($a_{12}=0$ for the end-point vertex.)

We have defined our conformal map in the interpolating gauge with arbitrary length strings. To reach the end-point configuration, in (3.12) we set

$$\text{end-point gauge:}\quad w_i=\bar{w}_i,\ \sum_{i=1}^{N}\alpha_i=0\ . \tag{3.17}$$

When we smoothly change our parametrization to (3.17), the map (3.12) smoothly turns into the usual end-point map (3.1) and the Riemann cuts that extend from w_i to $\bar{w}_i$ smoothly collapse into a point. (We have analytically shown that the transition to the end-point gauge is a smooth one in Ref. 9. There are no singularities in the map as we approach the end-point gauge.)

Similarly, to reach the midpoint gauge, we impose the following on the map (3.12):

$$|\alpha_i|=\alpha\ . \tag{3.18}$$

It is obvious that this map (3.12) allows us to smoothly go from the midpoint gauge in (3.18) to the end-point gauge in (3.17).

For $N=4$, we can perform the integration exactly in terms of elliptic functions of the first and third kind. The calculation is similar to but different from the calculation for the open-string case[9] in several important ways. Notice that we are allowing the w_i to roam over the entire complex plane, while $\bar{w}_i=w_i^*$ for the open-string case. This small but important distinction will vastly increase the complexity of the closed-string theory.

To solve the problem, let us change parameters from w_i to the following:

$$f^2(z)=[B_1(z-\beta)^2+C_1(z-\gamma)^2]\times[B_2(z-\beta)^2+C_2(z-\gamma^2)]\ , \tag{3.19}$$

where $B_{1,2},C_{1,2},\gamma,\beta$ are constants. Let us now change variables from z to ξ:

$$\xi=\delta^{-1}\frac{z-\gamma}{z-\beta},\quad z=\frac{\beta\delta\xi-\gamma}{\delta\xi-1}\ , \tag{3.20}$$

where

$$\delta=(-B_1/C_1)^{1/2}\ . \tag{3.21}$$

The point of changing variables is that we can now write

$$f(z)=\eta(B_1B_2)^{1/2}(z-\beta)^2\ , \tag{3.22}$$

where

$$\eta^2=(1-\xi^2)(1-k^2\xi^2),\quad k^2=B_1C_2/B_2C_1\ .$$

The Jacobian of the change of integration is

$$dz=\delta\frac{(z-\beta)^2}{\gamma-\beta}d\xi\ . \tag{3.23}$$

For example, with this change of variables, we can write the following integral as an elliptic integral:

$$\int^z\frac{d\bar{z}}{f(\bar{z})}=(-B_2C_1)^{-1/2}(\gamma-\beta)^{-1}\times\int^{\xi(z)}[(1-\xi^2)(1-k^2\xi^2)]^{-1/2}d\xi\ . \tag{3.24}$$

To actually perform the integration, let us now decompose the original expression for ρ':

$$\frac{d\rho}{dz}=\frac{f}{g}=f^{-1}(z)\left[\sum_{i=1}^{3}\frac{a_i}{z-x_i}+a_4z+a_5\right] \tag{3.25}$$

where the a_i are functions of x. For later convenience, let us also define the following function:

$$\phi(A,B,C,D,k,z)=\int^{\xi(z)}\frac{d\xi}{\eta}\frac{A\xi+B}{C\xi+D}\ , \tag{3.26}$$

where A, B, C, and D are constants. All integrations can be performed, and we find

$$\phi(A,B,C,D,k,z)=\frac{A}{C}F(k,z)-\frac{\Delta}{CD}\Pi(C^2D^{-2},k,z)+\Delta D^{-2}\int\frac{\xi}{\eta}(1-C^2D^{-2}\xi^2)^{-1}d\xi\ , \tag{3.27}$$

where $\Delta=AD-BC$ and F and Π are the usual elliptic integrals of the first and third kind:

$$F(k,z)=\int^{\xi(z)}d\xi(1-\xi^2)^{-1/2}(1-k^2\xi^2)^{-1/2}\ , \tag{3.28}$$

$$\Pi(\nu,k,z)=\int^{\xi(z)}d\xi(1-\xi^2)^{-1/2}(1-k^2\xi^2)^{-1/2}\times(1-\nu\xi^2)^{-1}\ . \tag{3.29}$$

The last integral in ϕ is easily calculated totally in terms of rational functions:

$$\int(x-n)^{-1}(ax^2+bx+c)^{-1/2}dx=(-an^2+bn+c)^{-1/2}\ln\{[(ax^2+bx+c)^{1/2}+(an^2+bn+c)^{1/2}](x-n)^{-1}+\tfrac{1}{2}(b+2an)(an^2+bn+c)^{-1}\}\ . \tag{3.30}$$

Now define, for $i=1-3$,

$$\begin{aligned}&A_i=\delta,\quad B_i=-1,\quad C_i=\delta(\beta-z_i),\quad D_i=z_i-\gamma\ ,\\&A_4=\delta\beta,\quad B_4=-\gamma,\quad C_4=-\delta,\quad D_4=-1\ ,\\&A_5=C_5=0,\quad B_5=1,\quad D_5=1\ .\end{aligned} \tag{3.31}$$

Our final result, written in our new coordinates, is given by

$$\rho(z)=(-B_2C_1)^{-1/2}(\gamma-\beta)^{-1}\sum_{i=1}^{5}\phi(A_i,B_i,C_i,D_i,k_i,z)\ . \tag{3.32}$$

In summary, we have now found the conformal map which takes us from the complex z plane to the ρ plane, where the external closed strings have arbitrary parametrization lengths. All parameters, including the w_i, can now be uniquely determined once the value of x and the various external parametrization lengths are fixed.

Although the transformation itself is not very difficult, the main problem is enforcing the real and imaginary constraints on the Riemann cuts, i.e., (3.15) and (3.16). Both analytically and by computer, the Riemann cuts are the main difficulty. There is, however, a very simple technique which allows us to draw the qualitative features of the ρ plane almost by inspection. This is the method of "equipotential lines," first introduced in Ref. 1. This method makes possible the visual resolution of ρ configurations which are quite opaque or even intractable analytically. The "equipotential line" method is based on the simple observation that, in the end-point gauge,

$$\tau=\mathrm{Re}\rho(z)=\sum_i \alpha_i \ln|z-x_i| \ . \tag{3.33}$$

Physically, this is equivalent to the statement that τ is equal to the potential function generated by static charges α_i placed at positions x_i in the two-dimensional plane. Lines of equal τ are therefore equipotential lines. However, lines of equal τ correspond to vertical lines in the ρ plane. Since the evolution of vertical lines swept out in the ρ plane is precisely the evolution of the string itself, and since these lines are equipotential lines in the z plane, then it follows that the evolution of equipotential lines are topologically equivalent to the evolution of the string itself. In this fashion, the open four-string interaction was first isolated in the string model in Ref. 1.

Using the method of equipotential lines, we can now, almost by inspection, show the existence of the closed four-string interaction. Our first problem is to give an electrostatic analogue of the Riemann cuts. We begin with the Neumann function N, which satisfies

$$\nabla^2 N(\sigma,\tau;\sigma',\tau')=2\pi\delta(\sigma-\sigma')\delta(\tau-\tau') \ . \tag{3.34}$$

Now multiply both sides by τ', integrate over $d\sigma' d\tau'$. The right-hand side just becomes $2\pi\tau$. The left-hand side contains a surface term in τ' space. By carefully analyzing the surface terms, we find

$$\tau=\frac{1}{2\pi}\sum_i \int N(\sigma,\tau;\sigma_i',\tau_i')d\sigma_i' \Big|_{\tau=\tau_i} \ . \tag{3.35}$$

Normally, in the end-point gauge, the integral over σ_i' is taken at the points at infinity, so we conveniently rederive the usual equation for τ in (3.33). However, for the interpolating gauge, we also have the surface term defined at the points at the instant the strings merge. Thus, the sum over terms taken at the points τ_i contains both the points at infinity as well as the points in the ρ plane when closed strings collide. By examining the previous equation, we now see that it is obvious that τ is the potential function created by external charges placed at infinity with charge $\int d\sigma=2\pi\alpha_i$ as well as a string of vertical charges proportional to a_{ij} placed along the vertical line where the strings interact at some instant τ. But this vertical line is nothing but the Riemann cut, so we see that we must smear electric charges along the cut. Thus, not only is the cut the analogue of a vertical smear of charges along the lines where strings merge, they are also equipotential lines in themselves. (In a later paper, we will show how the equipotential method can be generalized to the case of multiloop diagrams in the interpolating gauge.)

Once we know that Riemann cuts are sources of line charges and are also equipotential lines, it is then a simple matter to verify the correctness of the topology of the sphere in Fig. 9 given in the previous section. For example, the equipotential lines formed by the charges create precisely the topology of Fig. 3. *In fact, we can even show that two of the lines which form Fig. 3 are actually Riemann cuts themselves.*

In (3.32) we have an explicit, although complicated, expression for the mapping of the complex plane into the configuration for four-string scattering. However, let us now describe how the end-point configuration can be smoothly deformed into the midpoint gauge. In Fig. 10(a), we saw that the complex plane for the light-cone choice of parametrization can be divided into three regions, depending on the pole structure of the amplitude. Now let $|\alpha_4|$ decrease by an infinitesimal amount, keeping the other lengths the same. (This, of course, cannot be achieved within the parametrization of the end-point gauge. This parametrization is only possible in the interpolating gauge.)

In Fig. 11, we see the missing four-string region begin to emerge when we choose the following parametrization:

$$\begin{aligned} &\alpha_1=\alpha_2=\alpha_3 \ , \\ &\alpha_4=-\sum_{i=1}^{3}\alpha_3+\delta \ , \\ &\sum_{i=1}^{4}\alpha_i=\delta \ , \end{aligned} \tag{3.36}$$

where δ measures the deviation from the end-point gauge. When $\delta=0$, we have the end-point gauge. When $|\alpha_i|=\alpha_1$, we have the midpoint gauge.

The north pole at x_0 now slowly opens up and becomes a small triangular region, which represents the missing

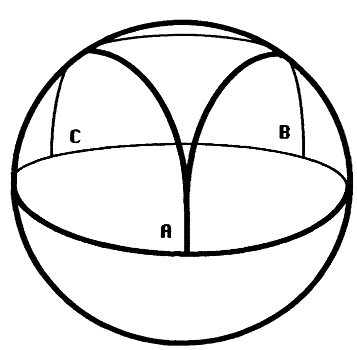

FIG. 11. In the interpolating gauge, when smoothly going from the end-point gauge to the midpoint gauge, the northern polar region begins as a point, then starts to open up, and then finally fills up most of the northern hemisphere.

region of the complex plane corresponding to a four-string interaction. Likewise, the point at z_0 splits into four distinct points w_1, $\tilde{w}_1$, w_2, and $\tilde{w}_2$. We now see two Riemann cuts emerging where previously there was only the point at z_0.

As δ gets larger, the triangular region surrounding x_0 slowly gets larger, until the configuration eventually approaches the midpoint gauge, when all strings have equal parametrization length. When this happens, the small triangular region around the north pole gradually gets larger, until it swallows up the entire northern (southern) part of the complex plane and becomes Fig. 9. Notice that the three points marking the vertices of the small triangular region all go to the equator. These three points, in fact, are precisely the Rubik's cube points A, B, and C found earlier.

Fortunately, this behavior can actually be shown analytically, without using complicated elliptic functions. We can calculate the precise location of the three vertices of the triangular missing region. Let us assume, for example, that x is sitting at one of the three vertices. Then the Riemann cuts degenerate, so that two of the w_i merge into a single point. For example,

$$\tilde{w}_1=\tilde{w}_2=w_3 \ . \qquad (3.37)$$

For this special case, the integration can be solved explicitly. The conformal mapping reduces to

$$\frac{d\rho}{dz}=k\frac{(z-w_1)^{1/2}(z-w_2)^{1/2}(z-w_3)}{z(z-1)(z-x)} \ . \qquad (3.38)$$

Fortunately, we know that

$$\begin{aligned} F(w_1,w_1,d)=& \int (ax^2+bx+c)^{1/2}(x-d)^1 dx \\ &=\sqrt{\phi}+(ad+b)a^{-1/2}\ln(2ad+b+2ax+2\sqrt{a\phi}) \\ &\quad -(ad^2+bd+c)^{1/2}\ln x^{-1}[2(ad^2+bd+c)+x(2ad+b)+2\sqrt{\phi}(ad^2+bd+c)^{1/2}] \end{aligned} \qquad (3.39)$$

where

$$\phi=ax^2+(2ad+b)x+ad^2+bd+c \ .$$

Then the integration yields

$$\begin{aligned} \rho(z)=&-w_3x^{-1}F(w_1,w_2,0) \\ &+(1+w_3)(x-1)^{-1}F(w_1,w_2,1) \\ &-(x_3+x)x^{-1}(x-1)^{-1}F(w_1,w_2,x) \ . \end{aligned} \qquad (3.40)$$

Now let δ be infinitesimally small, so that the three vertices of the triangular region around the north pole are located infinitesimally close to x_0, i.e.,

$$x_i=e^{\pi i/3}+\epsilon_i \ . \qquad (3.41)$$

The points w_i are also located infinitesimally close to z_0:

$$w_i=\tfrac{1}{2}+\frac{i}{2\sqrt{3}}+\eta_i \ . \qquad (3.42)$$

It is now straightforward to insert these equations into (3.40), and we can now solve explicitly for ϵ_i and η_i in terms of δ, showing the location of the vertices of the triangular region without elliptic functions.

So far, we have only discussed the missing region analytically. We have also shown, by direct computer calculation, that the closed four-string interaction corresponds to the missing region. We have performed this analysis for the process $1+2\rightarrow 3+4$ for the parametrization

$$\alpha_1=\alpha_2=-\alpha_3=-\alpha_4 \ .$$

As we noted earlier, in the end-point gauge, the process $1+2\rightarrow 3+4$ corresponds to having all three regions collapse into one. As we can see from (3.3) and (3.4), the regions I and II in Fig. 10(a) collapse into the real axis from the points 0 to $+1$ as we let δ approach $+1$.

Now let us slowly make the transition from the end-point gauge to the midpoint gauge. If we change the parametrization lengths a_{ij} infinitesimally to the interpolating gauge, then the points which lie between 0 and $+1$ slowly separate, creating a squashed ellipse. The two halves of the ellipse, above and below the real axis, are the precise analogues of the two triangular missing regions appearing around the north and south poles in the $4\rightarrow 1+2+3$ case.

We have performed a numerical analysis of the conformal map (3.12) by computer to trace the development of the missing region for the interpolating gauges which differ by finite amounts from the end-point gauge. As expected, the squashed ellipse expands, reaching maximum size for the midpoint gauge, as in Fig. 12. Points A and B, which are the points where the ellipse intersects the real axis, move smoothly from 0 and $+1$ near the end-point gauge to the points -1 and $+2$ for the midpoint gauge. -1 and $+2$ are the other Rubik's cube points. Point C, at $x=\frac{1}{2}$, does not move at all, since it already

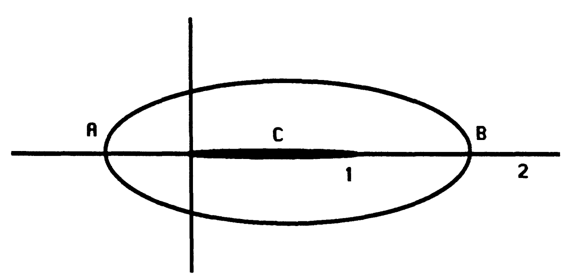

FIG. 12. By computer, we have independently checked the correctness of our results. The diagram shows the z plane structure for four-string scattering: $1+2\rightarrow 3+4$.

lies both at the boundary of regions I, II, and III, its proper end-point gauge location, and at the third Rubik's cube point, its proper midpoint gauge location. On the sphere, the missing region expands from a narrow slice around $\frac{1}{3}$ of the equator to a large region covering most of the sphere, both poles, and extending around $\frac{2}{3}$ of the equator.

The computer calculation is straightforward, with most of the computer time going toward maintaining the crucial constraints (3.15) and (3.16). This is important, because these constraints maintain the topology of Fig. 3. for the entire missing region. This is a nontrivial result. For example, if the missing region corresponded to a region of the complex plane which could only be described by a four-string interaction which had a finite extension over τ as well as σ, then our analysis would collapse totally. The would mean that string field theory can never produce a modular-invariant closed-string theory. The essential feature of this computer program is that it proves that the missing region can be totally parametrized by configurations which are instantaneous in τ, i.e., which can be generated by string field theory.

IV. UNIFIED STRING GROUP

So far, we have not yet demonstrated the deeper, group-theoretical meaning behind the closed four-string interaction. The essential point is that, in geometric string field theory, the group theory alone dictates that the open four-string interaction must be added to the end-point gauge but not the end-point gauge, and conversely, that the closed four-string interaction must be added to the midpoint gauge, but not the end-point gauge. This is because Jacobi identities for the unified string group do not close properly for open strings with end-point gauge vertices, or for closed strings with midpoint type vertices.

To show this, let us review the basic algebra which underlies geometric string field theory, which is called the "string algebra." The string group is defined totally in *loop space* without any parametrization or any background classical gravitational field. It is defined in the space of physical, space-time strings in 26 dimensions, not two-dimensional conformal space. Let us define a "triplet" as three oriented open strings which can be arranged in loop space as in Fig. 13. Let us define an antitriplet as a triplet with the opposite cyclic ordering. Let us define

$$f_{C_1C_2C_3}=\begin{cases}+1 \text{ for triplets },\\ -1 \text{ for antitriplets },\\ 0 \text{ otherwise }.\end{cases} \tag{4.1}$$

Let us define $\tilde{C}$ as the string C with reversed orientation. Then the structure constants for the string group are

$$f_{C_1C_2}^{C_3}=f_{C_1C_2\tilde{C}_3}\,,\qquad f_{C_1C_2}^{C_3}=-f_{C_2C_1}^{C_3}\,. \tag{4.2}$$

Let us define L_C to be an operator associated with a physical string C. Then the algebra for the string group is defined as

$$[L_{C_1},L_{C_2}]=f_{C_1C_2}^{C_3}L_{C_3}\,. \tag{4.3}$$

(We sum over repeated indices. Notice that the joining points of several strings do not necessarily have to meet.) The essential point is that the Jacobi identities are satisfied for this algebra. For example, the three strings in Fig. 14(a) can be inserted into the Jacobi identity,

$$[L_{C_{[1}},[L_{C_2},L_{C_{3]}}]]=0\,, \tag{4.4}$$

so that

$$f_{C_1C_2}^{C_4}f_{C_3C_4}^{C_5}+f_{C_2C_3}^{C_4}f_{C_1C_4}^{C_5}+f_{C_3C_1}^{C_4}f_{C_2C_4}^{C_5}=0\,. \tag{4.5}$$

When we analyze the configuration of Fig. 14(a), we find

$$1-1+0=0\,. \tag{4.6}$$

The string algebra (4.3) has cardinality

$$\aleph_1\sim 2^{\aleph_0}\,.$$

We call it a *transfinite Lie algebra*. We can now define how covariant and contravariant fields transform under

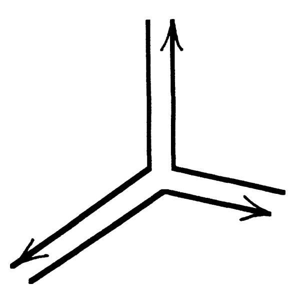

FIG. 13. A "triplet" of open strings. These strings are unparametrized and exist in physical space-time. There is no such thing as a parametrization midpoint in this space.

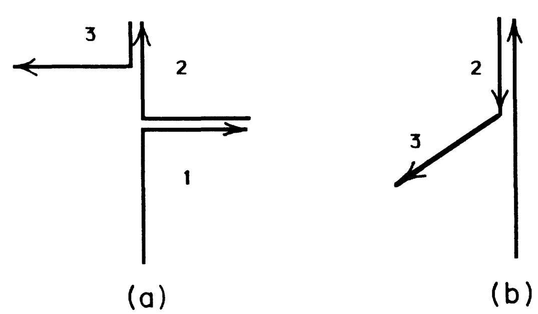

FIG. 14. In (a), we see three strings in physical space which are used to prove the Jacobi identity. However, we can show that a different definition of a triplet does not satisfy the Jacobi identity. Let us define a new triplet such that the physical string lengths sum to zero. Now calculate the Jacobi identity for this new triplet for three strings given in (b). A simple check shows that the Jacobi identity is not satisfied (unless higher string interactions are added).

the string group:

$$\delta\phi_{C_1} = -f_{C_1C_2}^{C_3}\phi_{C_3}\Lambda^{C_2} ,$$
$$\delta\phi^{C_1} = f_{C_2C_3}^{C_1}\phi^{C_2}\Lambda^{C_3} . \tag{4.7}$$

Notice that we now have two invariants under the string group:

$$\delta(\phi^C\phi_C) = 0 ,$$
$$\delta(\phi^{C_1}\phi^{C_2}\phi^{C_3}f_{C_1C_2C_3}) . \tag{4.8}$$

The crucial step is now to observe that a theory based only on "covariantized-light-cone"-type configurations (in loop space) does not obey the Jacobi identities. Let us now restrict the structure constants to those triplets where

$$|L_1| + |L_2| = |L_3| , \tag{4.9}$$

where L_i is the physical length (not parametrization length) of the ith string. In Fig. 14(b), we see a typical arrangement of three strings. By checking the Jacobi identity for these three strings, we see

$$1+0+0 \neq 0 . \tag{4.10}$$

Furthermore, we find that the ϕ^3 interaction is not gauge invariant:

$$\delta(\phi^3) \neq 0 . \tag{4.11}$$

There is a way out of this problem, which is to add a four-string interaction, such that

$$\delta(\phi^3 + g\phi^4) \sim 0 . \tag{4.12}$$

This general effect (the nonclosure of the algebra) was first noticed in Ref. 4. Even with the introduction of the four-string interaction, however, the resulting action only closes on-shell.

The important observation is that when we generalize these results to the closed-string case, we find precisely the *opposite* situation, that the end-point gauge closes without any four-string interaction.

There are several key differences when we discuss the closed-string group.

(a) Triplets are their own antitriplets. It is easy to show that, by rotating a triplet of closed strings in physical space, we can convert it into the anticyclic ordering. Thus, we cannot define antitriplets distinct from triplets. Thus, the algebra collapses.

(b) It is impossible to construct constant antisymmetric tensors like $f_{C_1C_2}^{C_3}$ in loop space for symmetric configurations as in Ref. 5. By rotating a triplet, we can always invert the first and second strings and thus show that the tensor is symmetric. Thus, antisymmetric structure constants are impossible to define.

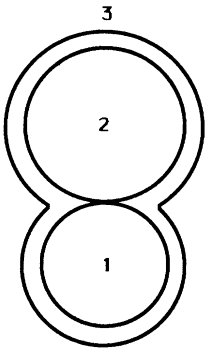

FIG. 15. A "triplet" of closed strings. This arrangement of three unparametrized closed strings in physical space-time (and not the symmetric arrangement of Ref. 5) allows us to construct a Lie algebra in which the Jacobi identities are satisfied. This Lie algebra, which we call a "transfinite Lie algebra," is the foundation upon which we build the geometric string field theory of closed strings.

(c) The Jacobi identities do not close for symmetric triplet configurations considered in Ref. 5. Because this configuration is fully symmetric, it produces four-string diagrams in the Jacobi identity with the topology of a tetrahedron, as in Fig. 3. Because these tetrahedron graphs are symmetrical, we find

$$1+1+1 \neq 0 . \tag{4.13}$$

No matter how we may modify the numerical assignments of the structure constants with a symmetric three-string vertex, we can never satisfy the Jacobi identity.

Thus, we are now forced to adopt a new approach to closed-strings.

In order to cure the problem with Jacobi identities, we are now forced to take diagrams such that $|L_1| + |L_2| = |L_3|$, as in Fig. 15. This new definition of a triplet will guarantee that diagrams like Fig. 3 never appear in the Jacobi identity. Let us define the constant tensor for triplets in Fig. 15. Notice that we can define an outer and inner string in Fig. 15 because the configuration is not totally symmetric in three strings. Then

$$f_{C_1C_2C_3} = \begin{cases} +1\ (-1) & \text{for triplets if } C_3 = \text{outer (inner) string} , \\ 0 & \text{otherwise} . \end{cases} \tag{4.14}$$

Notice, as we have stressed, that only symmetric matrices in C_1 and C_2 are possible for closed strings. Now define

$$f_{C_1C_2}^{C_3} = f_{C_1C_2\tilde{C}_3} , \tag{4.15}$$

where $\tilde{C}$ is a closed string with reversed orientation from C.

Our string algebra is now defined as

$$\{L_{C_1}, L_{C_2}\} = \hat{f}\,{}_{C_1C_2}^{C_3} L_{C_3} . \tag{4.16}$$

There are several surprising features to this equation. Notice that we are forced to introduce *anticommutators* rather than commutators. Thus, the generators are Grassmann type. In order for the right-hand side to have the same statistics as the left-hand side, we are forced to consider $\hat{f}$ as a Grassman operator. Thus, the algebra is not a Lie algebra in the usual sense. (It becomes a standard Lie algebra only when we introduce a parametrization, which is done in Ref. 9.)

We will take the following choice:

$$\hat{f}\,{}^{C_3}_{C_1C_2}=f^{C_3}_{C_1C_2}\hat{i}_{C_1\cap C_2\cap C_3}\ . \tag{4.17}$$

Notice that $\hat{i}$ is a Grassmann operator which is defined only at the interaction point of three strings (the end point) in a triplet as in Fig. 15. If we now calculate the Jacobi identities, we find[9]

$$1-1+0=0\ . \tag{4.18}$$

The crucial minus sign does not come from the antisymmetry of the structure constants, because the structure constants are symmetric. It comes from moving the Grassmann variable defined at the interaction points past each other. (This is the origin of the "ghost insertion operator" found in BRST theories.)

We can show that

$$\begin{aligned}&\delta(\phi^C\phi_C)=0\ ,\\&\delta(\phi_{C_1}\phi_{C_2}\phi^{C_3}\hat{f}\,{}^{\tilde{C}_3}_{\tilde{C}_1\tilde{C}_2})=0\ ,\end{aligned} \tag{4.19}$$

where ϕ^C and ϕ_C are independent fields. Thus, ϕ^2 and ϕ^3 terms are again invariant. As in the open-string case, we can retrieve the group theory for the midpoint gauge, but only at the price of adding the four-string interaction (and then only achieving local invariance on-shell).

The next step is to introduce parametrizations into the algebra and write down the full universal string group and unified string group. We wish to expand the original string group to include the semidirect product of the Virasoro group with the string group. Thus

$$\frac{\text{USG}}{\text{Diff}(S_1)_-}=\text{SG} \tag{4.20}$$

where $\text{Diff}(S_1)_-$ is the reparametrization group generated by $L_n-\tilde{L}_{-n}$ and is a subgroup of the full $\text{Diff}(S_1)$, which is generated by L_n and $\tilde{L}_m$.

It is straightforward to write down the universal string group (USG) and then gauge it, in the same way we construct the Yang-Mills theory and general relativity.

Let us briefly summarize the main points of geometric string field theory.[8,9]

(1) The entire theory, as in Yang-Mills theory or gravity, can be deduced by gauging a single local gauge group. For string theory, it is the unified string group. There is no ambiguity or guesswork in constructing the action.

(2) The tangent space of the theory, when gauge fixed, becomes the "ghost sector" of the BRST theory, which gives a simple, geometrical meaning to the origin of the ghost sector.

(3) The arcane rules of bosonization and ghost counting are a representation-dependent way of calculating the Clebsch-Gordan coefficients for various tensor products in the geometric theory. Thus, conformal field theory is incorporated into geometric string theory.

(4) We do not have overcounting of the fictitious string length α because α is now a local gauge parameter.

(5) When gauge fixed, the geometric theory produces a vertex function with arbitrary string lengths, which we call the interpolating gauge. We can show, at the level of Neumann functions, that

$$\begin{aligned}&\lim_{a_{12}\to 0}|V\rangle_{\text{interpolating}}=|V\rangle_{\text{end point}}\ ,\\&\lim_{|\alpha_i|\to 2\pi}|V\rangle_{\text{interpolating}}=|V\rangle_{\text{midpoint}}\ .\end{aligned} \tag{4.21}$$

(6) The integration rules for $\times$ and $\int$ do not have to be postulated. They are uniquely defined by the tensor calculus of the group itself.

(7) The number of connection fields for the theory is dictated by the structure of the group.

For the final action, see Refs. 8 and 9.

V. DERIVATION OF THE FOUR-STRING INTERACTION

So far, we have exhibited the structure of the four-string interaction, demonstrated that there is a missing region of the complex plane which is filled by the four-string vertex, and shown that the four-string vertex arises from group-theoretical considerations alone.

In this section, we will explicitly show how the four-string interaction can be derived in the same way that the instantaneous four-fermion Coulomb term can be derived in ordinary QED. The origin of the Coulomb term is that the propagator is not invariant under a gauge transformation. For example, in a covariant gauge the A_0 field may propagate as $1/\Box$, while in the Coulomb gauge the propagator for A_0 becomes ∇^{-2}, i.e.,

$$\text{QED:}\ \ \Box^{-1}\to\nabla^{-2}\ . \tag{5.1}$$

Thus, because the A_0 no longer has any time derivatives on it, we may functionally integrate out over A_0, which immediately leaves us with the four-fermion interaction.

An identical situation happens with the geometric string field theory. The propagator for closed-strings,

$$D=\int_{|z|<1}z^{L_0-2}\bar{z}^{\tilde{L}_0-2}d^2z\ , \tag{5.2}$$

is not gauge invariant under a transformation generated by the universal string group. We find

$$UDU^{-1}\neq D\ . \tag{5.3}$$

In fact, as in QED, we can actually gauge rotate the theory so that the propagator term $z^{L_0}\bar{z}^{\tilde{L}_0}$ becomes the number 1. Instead of eliminating the A_0 field, as in QED, we eliminate the string vierbein[8] $e^{\nu\rho}_{\mu\sigma}$ that exists in the theory. *The elimination of the vierbein creates the closed four-string interaction. Since the vierbein couples to the three-vertex, the four-string interaction is just the* *square of the three-vertex in the interpolating gauge.*

This conveniently explains the reason for the shape of the open four-string interaction. In the end-point gauge for the open-string theory, the three-string interactions take place at the end points of strings. Hence, we call it the end-point gauge. However, the four-string interaction takes place within the interior of the string, which seems highly unusual. The explanation of this is that the four-vertex is actually the square of the three-vertex in the interpolating gauge.

Let us begin by first writing down the interpolating gauge vertex:

$$\prod_{i<j}^{3} \prod_{0<\sigma_i<2\pi|\alpha_i|} [X_i(\sigma_i)-\theta_{ij}X_j(b_{ij}-\sigma_j)]\,|\,V_{\alpha_1\alpha_2\alpha_3}\rangle=0 , \tag{5.4}$$

where

$$\begin{aligned}
&\theta_{12}=\theta(a_{12}-\sigma_1) ,\\
&\theta_{13}=\theta(2\pi|\alpha_3|-\sigma_1)\theta(\sigma_1-a_{12}) ,\\
&\theta_{23}=\theta(a_{23}-\sigma_2) ,\\
&b_{12}=2\pi\alpha_2 ,\\
&b_{13}=2\pi|\alpha_3|+a_{12} ,\\
&b_{23}=2\pi|\alpha_3| .
\end{aligned} \tag{5.5}$$

We will use the key identity, which describes the action of L_n on a vertex function:

$$(L_n^r/\alpha_r)\,|\,V\rangle=\left[\sum_{s=1}^{3}\sum_{m=0}^{\infty} n\tilde{N}_{nm}^{rs}L_{-m}^{s}/\alpha_s+c_n^r/\alpha_r\right]|\,V\rangle , \tag{5.6}$$

where

$$c_n^r=\frac{D}{2}\sum_{p=1}^{n-1}p(n-p)N_{p,n-p}^{rr}+\sum_{p=1}^{n-1}(p^2-n^2)\tilde{N}_{p,n-p}^{rr}-n^2\tilde{N}_{n0}^{rr} . \tag{5.7}$$

(This identity can easily be proved. We know that L_n transforms as a tensor of weight 2, so its action on $|\,V\rangle$ can be determined classically. The only problem is the quantum anomaly, which can be calculated by hitting this identity with $\langle 0\,|$. The proof that this anomaly can be ignored for the interpolating gauge is quite involved, requiring a rather elaborate mathematical apparatus involving intricate identities on Neumann functions. For a discussion, see Ref. 9.)

This identity alone is sufficient to determine the complete vertex function, without the necessity of appealing to "ghost" fields. Let us define the vertex function as

$$f^{\alpha\beta\gamma}=\langle \mathbf{e}^{\alpha}\,|\,\langle \mathbf{e}^{\beta}\,|\,\langle \mathbf{e}^{\gamma}\,|\,V\rangle . \tag{5.8}$$

By moving the various L_n contained within each Verma module to the right, and then reflecting them on the vertex function, we obtain a series of L_{-m} terms which eventually annihilate on the vacua on the left. By successive reflections, we can eventually eliminate all of the L_n contained within the Verma module, leaving only the c-number Clebsch-Gordan coefficient left. The point is that the transformation properties of L_n on the vertex function alone are sufficient to determine the entire three-string interaction.

However, it is also possible to obtain an explicit representation of the vertex function in the more familiar harmonic-oscillator representation. Let us now write down an explicit representation of the vertex function:

$$\begin{aligned}
&|\,V\rangle=|\,V_0\rangle\otimes|\,\tilde{V}_0\rangle\\
&|\,V_0\rangle=[\gamma^{(+)}(\sigma_i)+\gamma^{(-)}(\sigma_i)]\\
&\qquad\times\exp(\tfrac{1}{2}\alpha_{-n}^{r}N_{nm}^{rs}\alpha_{-m}^{s}+\gamma^{-nr}n\tilde{N}_{nm}^{rs}\gamma_{-m}^{s})\,|\,0\rangle_{123} ,
\end{aligned} \tag{5.9}$$

which satisfies the usual continuity equations for three strings, where σ_I is the interaction point, and $\tilde{N}$ are the Neumann coefficients for the following:[12]

$$\begin{aligned}
\tilde{N}(z,z')&=\frac{-1}{z-z'}\left[\frac{d\rho}{dz}\right]\left[\frac{d\tilde{\rho}}{d\tilde{z}}\right]^{-2}\prod_{r=1}^{3}(z-x_r)(\tilde{z}-x_r)^{-1}\\
&=(1/\alpha_r)\delta_{rs}\left[\theta(\tilde{\xi}_s-\xi_r)\sum_{n=0}^{\infty}e^{-n(\zeta_s-\zeta_r)}-\theta(\xi_r-\tilde{\xi}_s)\sum_{n=1}^{\infty}e^{-n(\zeta_r-\zeta_s)}\right]-(\alpha_r/\alpha_s^2)\sum_{\substack{n=0\\m=1}}\tilde{N}_{nm}^{rs}e^{n\xi_r+m\xi_s} .
\end{aligned} \tag{5.10}$$

The structure constants of the geometric theory are based on the vertex (5.6) except that we take $a_{12}\to 0$ and the parameter σ is now replaced by $\sigma+\zeta^{\sigma}$, where summing over the field ζ^{σ} allows us to sum over all possible parametrizations of the vertex function.

As explained in Refs. 8 and 9, gauge fixing implies taking various values of the ζ^{σ} field. For example, $\zeta^{\sigma}=0$ corresponds to taking the end-point gauge. In general, for different gauges and different values of ζ^{σ}, we must insert the determinant of the string vierbein $e_{\mu\sigma}^{\nu\rho}$ into the measure. The end-point and interpolating gauge, therefore, have measure different from one.

This is the reason why we can use the end-point gauge freely in the construction of N-point functions, but we must be careful to insert $\det|e^{\nu\rho}_{\mu\sigma}|$ for all other measures.

To eliminate this troublesome measure term, we will extract out the ζ^σ by taking a Taylor expansion of the vertex function around the end-point gauge.

We wish, therefore, to write down the operator which changes the parametrization lengths of the three strings via a Taylor expansion. In general,

$$U_{\bar{\alpha}_1}^{\alpha_1}U_{\bar{\alpha}_2}^{\alpha_2}U_{\bar{\alpha}_3}^{\alpha_3}|V_{\alpha_1\alpha_2\alpha_3}\rangle=|V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_3}\rangle . \tag{5.11}$$

The operator U is easy to construct (and, in fact, is the remnant of the string vierbein after gauge fixing and after extracting out the Taylor series). In order to generate the reparametrization which changes the string length, we write

$$\sigma\to\sigma+\epsilon(\sigma),\quad \epsilon(\sigma)=\frac{\delta\alpha}{\alpha}\sigma . \tag{5.12}$$

We use the operator

$$U=\exp\left[\sum_{-\infty}^{\infty}\epsilon^n(L_n-\tilde{L}_{-n})\right] \tag{5.13}$$

where

$$\epsilon(\sigma)=-i\sum_{-\infty}^{\infty}\epsilon^n e^{in\sigma/\alpha} . \tag{5.14}$$

Now let us write down the four-string scattering amplitude in the end-point gauge, keeping strings 1, 2, 3, and 4 on shell with parametrization length α_i:

$$\begin{aligned}A_4&=\langle V_{\alpha_1\alpha_2\alpha_5}|D|V_{\alpha_6\alpha_3\alpha_4}\rangle\\&=\langle V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_5}|U_5^{-1}DU_6|V_{\bar{\alpha}_6\bar{\alpha}_3\bar{\alpha}_4}\rangle ,\end{aligned} \tag{5.15}$$

where

$$\begin{aligned}U_5&=\prod_{i=1,2,5}U_{\alpha_i}^{\bar{\alpha}_i} ,\\U_6&=\prod_{i=3,4,6}U_{\alpha_i}^{\bar{\alpha}_i} ,\end{aligned} \tag{5.16}$$

and $\bar{\alpha}_i$ are the parametrization lengths of the strings in the interpolating gauge. We can always write equations which are valid for arbitrary α_i, but we are specifically interested in the special case when the external strings have equal parametrization length:

$$\begin{aligned}&\alpha_1=\bar{\alpha}_1=\alpha_2=\bar{\alpha}_2=-\alpha_3=-\bar{\alpha}_3=-\alpha_4=-\bar{\alpha}_4 ,\\&\alpha_1+\alpha_2+\alpha_5=0 ,\\&\alpha_3+\alpha_4+\alpha_6=0 .\end{aligned} \tag{5.17}$$

The new string vertex is defined in the interpolating gauge:

$$\begin{aligned}&\bar{\alpha}_1+\bar{\alpha}_2+\bar{\alpha}_5=2\delta ,\\&\bar{\alpha}_3+\bar{\alpha}_4+\bar{\alpha}_6=-2\delta ,\end{aligned} \tag{5.18}$$

where δ is a small number.

Infinitesimally, we can write

$$U=\left[1+\sum_{r=1}^{3}\sum_{-\infty}^{\infty}\epsilon^{rn}(L_n-\tilde{L}_{-n})\right] . \tag{5.19}$$

Although the U operator contains factors of $L_n-\tilde{L}_{-n}$, we can always convert this to a factor which only contains L_{-n} and $\tilde{L}_{-n}$. We can always show, via (5.6), that

$$\begin{aligned}|V_{\alpha_1\alpha_2\alpha_3}\rangle&=\prod_{i=1}^{3}U_{\alpha_i}^{\bar{\alpha}_i}|V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_3}\rangle\\&=e^{L^r_{-n}N^{rs}_{nm}\delta\alpha^s_m}e^{\tilde{L}^r_{-n}N^{rs}_{nm}\delta\alpha^s_m}|V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_3}\rangle ,\end{aligned}$$

where $\delta\alpha^r_n$ is linearly related to ϵ^{rn} in a simple way (see Ref. 9). Physically, this means that the on-shell matrix elements of the vertex function multiplied by any combination of U matrices will always produce the usual matrix elements of the Shapiro-Virasoro model. Off-shell, however, remarkable identities are made possible by the presence of these L_{-n} factors.

In order to eliminate the U factors, we will find it convenient to split the propagator D into two pieces. Let the unit circle surrounding the origin be broken up into two regions: R_1 which contains the origin and R_2 being everything else. We will define these two regions more precisely in a moment. Then

$$\begin{aligned}D&=\int_{|z|<1}z^{L_0-2}\bar{z}^{\tilde{L}_0-2}d^2z\\&=\left[\int_{R_1}+\int_{R_2}\right]z^{L_0-2}\bar{z}^{\tilde{L}_0-2}d^2z .\end{aligned} \tag{5.20}$$

Then the amplitude A_4 can be broken up into two pieces as well, which we call I and II. We find for the first region

$$\mathrm{I}=\left\langle V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_5}\middle|\left[1+\sum_{r=1,2,5}\epsilon^{rn}(L^r_n-L^r_n)\right]\int_{R_1}z^{L_0-2}\bar{z}^{\tilde{L}_0-2}d^2z\left[1+\sum_{s=3,4,6}\epsilon^{sn}(L^s_n-\tilde{L}^s_{-n})\right]\middle|V_{\bar{\alpha}_3\bar{\alpha}_4\bar{\alpha}_6}\right\rangle \tag{5.21}$$

where $\bar{\alpha}_i$ is defined in the interpolating gauge and where ϵ^n is the remnant of the vierbein which changes the length of the strings.

Now comes the critical step.

In Ref. 9, we introduced the "method of reflections." We will move each L_n to the right, until it reflects off the vertex and turns into L_{-m}. This reflection process also generates other terms as well, such as anomalous terms and terms containing L_0. Then, we move L_{-m} to the left, until it reflects off the vertex function and turns into L_p, also leaving other terms as well. We continue this process an infinite number of times. Each reflection, we pick up a factor of z^n. But since $|z|<1$, after an infinite number of reflections, we find that this term disappears completely.

The only terms left are the anomalous terms and the L_0 factors. To see this reflection process more clearly, let us write

$$\begin{aligned}\langle V \mid L_n z^{L_0} \mid V\rangle &= \langle V \mid z^{L_0} L_n z^n \mid V\rangle \\ &= \langle V \mid z^{L_0} n \tilde{N}^{66}_{nm} L_{-m} \mid V\rangle + \cdots \\ &= \langle V \mid L_{-m} z^m z^{L_0} \mid V\rangle z^n n \tilde{N}^{66}_{nm} + \cdots \\ &= \langle V \mid L_p z^{L_0} \mid V\rangle z^n n \tilde{N}^{66}_{nm} z^m m \tilde{N}^{55}_{mp} + \cdots ,\end{aligned} \tag{5.22}$$

where the ellipsis represent anomalous terms and L_0 terms.

(We also have terms like (5.22) coming from the tangent space of the geometric theory. The four-string amplitude (5.15) contains a factor of $\gamma_0 \tilde{\gamma}_0$ in the propagator, a leftover from gauge fixing (see Ref. 9 for definitions and notation). When L_n moves past this factor, we use

$$[L_n, \gamma_0] = -n\gamma_n . \tag{5.23}$$

This factor of γ_n, in turn, moves to the right. Since it has weight 2, it also reflects off the vertex function as in (5.6), without the anomalous term. Thus, this term also generates its own infinite series of L_0 terms. The only difference between these terms and the terms found in (5.21) is that they contain one extra factor of n^2 rather than n appearing in the series.)

Now let us collect this infinite number of terms created by the method of reflection. We find

$$\begin{aligned} I &= \left\langle V_{\alpha_1\alpha_2\alpha_5} \left| \int_{R_1} z^{L_0-2} \bar{z}^{\bar{L}_0-2} d^2 z \right| V_{\alpha_3\alpha_4\alpha_6} \right\rangle \\ &= \left\langle V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_5} \left| \int_{R_1} (z')^{L_0} (\bar{z}')^{\bar{L}_0} f(z) f(\bar{z}) \frac{d^2 z}{|z\bar{z}|^2} \right| V_{\bar{\alpha}_3\bar{\alpha}_4\bar{\alpha}_6} \right\rangle ,\end{aligned} \tag{5.24}$$

where

$$z' = z[1+\xi(z,\delta)], \quad f(z) = 1+\eta(z) , \tag{5.25}$$

where

$$\begin{aligned}\xi(z) = &\sum_{r=1,2,5} \sum_{i=2}^{\infty} \sum_{n_i=1}^{\infty} \delta\alpha^{rn_1} \tilde{N}^{r5}_{n_1 n_2} \alpha_5 z^{n_2} n_2 \tilde{N}^{66}_{n_2 n_3} \cdots z^{n_i} n_i \tilde{N}^{r_i r_i}_{n_i o} \frac{1}{\alpha_{r_i}} \\ &+ \sum_{s=3,4,6} \sum_{j=2}^{\infty} \sum_{n_j=1}^{\infty} \delta\alpha^{sn_1} \tilde{N}^{s6}_{n_1 n_2} \alpha_6 z^{n_2} n_2 \tilde{N}^{55}_{n_2 n_3} \cdots z^{n_j} n_j \tilde{N}^{s_j s_j}_{n_j o} \frac{1}{\alpha_{s_j}} ,\end{aligned} \tag{5.26}$$

and

$$\begin{aligned}\eta(z) = &\sum_p \sum_{r=1,2,5} \sum_{i=2}^{\infty} \sum_{n_i=1}^{\infty} \delta\alpha^{rn_1} \tilde{N}^{r5}_{n_1 n_2} \alpha_5 z^{n_2} \cdots \tilde{N}^{r_i p}_{n_i o} \frac{1}{\alpha_p} \\ &+ \sum_q \sum_{s=2,3,6} \sum_{j=2}^{\infty} \sum_{n_j=1}^{\infty} \delta\alpha^{sn_1} \tilde{N}^{s6}_{n_1 n_2} \alpha_6 z^{n_2} \cdots \tilde{N}^{s_j q}_{n_j o} \frac{1}{\alpha_q} \\ &+ \sum_{r=1,2,5} \sum_{i=2}^{\infty} \sum_{l=2}^{i} \sum_{n_i=1}^{\infty} \delta\alpha^{rn_1} \tilde{N}^{r5}_{n_1 n_2} \alpha_5 z^{n_2} \cdots z^{n_l} n_l^2 \tilde{N}^{r_l r_{l+1}}_{n_l n_{l+1}} \cdots z^{n_i} n_i \tilde{N}^{r_i r_i}_{n_i o} \frac{1}{\alpha_{r_i}} \\ &+ \sum_{s=3,4,6} \sum_{j=2}^{\infty} \sum_{l=2}^{j} \sum_{n_j=1}^{\infty} \delta\alpha^{sn_1} \tilde{N}^{s6}_{n_1 n_2} \alpha_6 z^{n_2} \cdots z^{n_l} n_l^2 \tilde{N}^{r_l r_{l+1}}_{n_l n_{l+1}} \cdots z^{n_j} n_j \tilde{N}^{s_j s_j}_{n_j o} \frac{1}{\alpha_{s_j}} ,\end{aligned}$$

where

$$\begin{aligned} &i = \text{even} \rightarrow r_i = 6, \quad s_i = 5, \quad p = 3,4, \quad q = 1,2 , \\ &i = \text{odd} \rightarrow r_i = 5, \quad s_i = 6, \quad p = 1,2, \quad q = 3,4 .\end{aligned} \tag{5.27}$$

(Notice that the last two terms appearing in the expression for η differ from the previous two terms because they have factors of n^2; i.e., they come from the γ_0 contribution to the power series.)

Fortunately, it is possible to sum the entire series and obtain rather simple results. A careful examination of the power series shows that *we are simply reproducing the Neumann function for the four-string scattering amplitude.* For example, we have the well-known formula (in the coherent-state basis) which is often used to calculate the four-string interaction amplitude:

$$\langle 0|\exp[\tfrac{1}{2}(a|M_2|a)+(a|L_2)]\exp[\tfrac{1}{2}(a^\dagger|M_1|a^\dagger)+(a^\dagger|L_1)]|0\rangle$$
$$=\det{}^{-D/2}(1-M_2M_1)\exp\left[\left[L_1\left|\frac{1}{1-M_2M_1}\right|L_2\right]+\frac{1}{2}\left[L_2\left|M_1\frac{1}{1-M_2M_1}\right|L_2\right]+\frac{1}{2}\left[L_1\left|\frac{1}{1-M_2M_1}\right|M_2L_1\right]\right]. \quad (5.28)$$

We will find it convenient to define our Neumann functions in terms of

$$\hat{N}^{rs}_{mn}=\sqrt{m}\,N^{rs}_{mn}\sqrt{n}\,z_r^m z_s^n\ , \quad (5.29)$$

where

$$z_i=e^{-(\tau_i+i\theta_i)/\alpha_i}\ ,$$
$$z=z_6z_5\ . \quad (5.30)$$

In terms of $\hat{N}$, we can now explicitly contract the four-string scattering amplitude $\langle V|D|V\rangle$. We find

$$L_2=\hat{N}^{5r}|a^{\dagger r})\ ,$$
$$L_1=\hat{N}^{6s}|a^{\dagger s})\ ,$$
$$M_2=\hat{N}^{55}\ ,$$
$$M_1=\hat{N}^{66}\ , \quad (5.31)$$

where r = 1,2 and s = 3,4. Putting these factors together, we find that the four-string scattering amplitude can be written in terms of Neumann functions M^{ij} defined over the two-dimensional four-string scattering surface:

$$\prod_{i=1}^{4}e^{(a_i^\dagger|M^{ij}|a_j^\dagger)/2}|0\rangle\ , \quad (5.32)$$

where

$$M^{rr'}=\hat{N}^{rr'}+\hat{N}^{r5}\hat{N}^{66}\frac{1}{1-\hat{N}^{55}\hat{N}^{66}}\hat{N}^{5r'}\ ,$$
$$M^{ss'}=\hat{N}^{ss'}+\hat{N}^{s6}\frac{1}{1-\hat{N}^{55}\hat{N}^{66}}\hat{N}^{55}\hat{N}^{6s'}\ ,$$
$$M^{rs}=\hat{N}^{r5}\frac{1}{1-\hat{N}^{66}\hat{N}^{55}}\hat{N}^{6s}\ . \quad (5.33)$$

Now we are in a position to sum the series for ξ and η and compare them with the Neumann functions M^{ij} defined over the four-string scattering surface:

$$\xi(z,\delta)=\sum_{r=1,2}\sum_{n=0}^{\infty}\delta\tilde{\alpha}^{rn}(M^{r5}_{no}-M^{r6}_{no})$$
$$+\sum_{s=3,4}\sum_{n=0}^{\infty}\delta\tilde{\alpha}^{sn}(M^{s6}_{no}-M^{s5}_{no})\ , \quad (5.34)$$

where

$$\delta\tilde{\alpha}^{rn}=\frac{\delta\alpha^{rn}}{\sqrt{n}}z_5^{-1}\ ,$$
$$\delta\tilde{\alpha}^{sn}=\frac{\delta\alpha^{sn}}{\sqrt{n}}z_6^{-1}\ , \quad (5.35)$$

and

$$\eta(z,\delta)=-\xi(z)+z\frac{d\xi}{dz}\ . \quad (5.36)$$

[At first glance, it may appear that the factor of $d\xi/dz$ appearing in (5.36) does not appear in the power expansion for η in (5.25). Actually, this identity is satisfied because of the n^2 terms which originate from the γ_0 factors. The z derivative of ξ in (5.25) simply pulls down the exponent of z, which is n. This, in turn, produces factors of n^2 which precisely cancel those n^2 terms coming from η.]

With these identities, we can now show that

$$\mathrm{I}=\left\langle V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_5}\left|\int_{R_1}z'^{L_0}\bar{z}'^{\bar{L}_0}\frac{d^2z}{|z\bar{z}|^2}f(z)f(\bar{z})\right|V_{\bar{\alpha}_3\bar{\alpha}_4\bar{\alpha}_6}\right\rangle$$
$$=\langle V_{\bar{\alpha}_1\bar{\alpha}_2\bar{\alpha}_5}|D|V_{\bar{\alpha}_3\bar{\alpha}_4\bar{\alpha}_5}\rangle\ . \quad (5.37)$$

Notice that I is now defined in the interpolating gauge, not the end-point gauge.

At this point, we will now establish the link between I and II. Notice that the method of reflections altered the complex variable appearing in the propagator, from z to z'. This also affects the integration region as well. In general, the region of integration shrinks infinitesimally, from the unit circle to the region R_1. Let us call C_{12} the boundary between region R_1 and R_2. Let $z=xe^{i\theta}$. Then

$$R_1\!:x<1-\epsilon(\theta)\ ,$$
$$C_{12}\!:x=1-\epsilon(\theta)\ ,$$
$$R_2\!:x>1-\epsilon(\theta)\ , \quad (5.38)$$

where $\epsilon(\theta)$ is a calculable function. The essential point is that the boundary between R_1 and R_2 is the line formed by $|z'|=1$. This means that the line C_{12} is defined by

$$|z[1+\xi(z,\delta)]|=1\ , \quad (5.39)$$

which means that we can solve for $\epsilon(\theta)$ explicitly:

$$\epsilon(\theta)=\xi(x=1,\theta,\delta)\ . \quad (5.40)$$

The effect of this constraint is to make the absolute value of the propagator between the vertices vanish in (5.24):

$$|z'^{L_0}\bar{z}'^{\bar{L}_0}|=1\ . \quad (5.41)$$

This is the counterpart of (5.1).

Now comes the important step. We will find that II corresponds precisely to the tetrahedron graph.

Let us define $\bar{\delta}$ to be between zero and δ. Although we can write down the arbitrary case, we will find it convenient to take the case where the external strings have equal length. We will fix the value of $\bar{\delta}$ by the criteria that $|z'|=1$ throughout the entire R_2, not just the boundary between R_1 and R_2. This will fix the value of x as a function of $\bar{\delta}$. We now find

$$|z[1+\xi(z,\bar{\delta})]|=1, \quad x=x(\theta,\bar{\delta}) . \tag{5.42}$$

With this restriction, we now find that II can be written as

$$\begin{aligned} \mathrm{II} &= \int_{R_2} \frac{d^2z}{|z\bar{z}|^2} \langle V_{\alpha,\alpha,-2\alpha+2\bar{\delta}} | \tilde{U}_{\bar{\delta}}^{-1} z^{L_0} \bar{z}^{\bar{L}_0} U_{\bar{\delta}} | V_{-\alpha,-\alpha,2\alpha-2\bar{\delta}} \rangle \\ &= \int_{R_2} \frac{d^2z}{|z\bar{z}|^2} \langle V_{\alpha,\alpha,-2\alpha+2\bar{\delta}} | e^{i\theta'(L_0-\bar{L}_0)} | V_{-\alpha,-\alpha,2\alpha-2\bar{\delta}} \rangle , \end{aligned} \tag{5.43}$$

where

$$z'=e^{i\theta'} \tag{5.44}$$

and

$$U_{\bar{\delta}}=U_{2\alpha}^{2\alpha-2\bar{\delta}}, \quad \tilde{U}_{\bar{\delta}}=\tilde{U}_{-2\alpha}^{-2\alpha+2\bar{\delta}} . \tag{5.45}$$

Notice that the propagator between the two vertices has completely collapsed. Instead of an x integration between the vertices, we only have an angular integration which simply spins the relative orientation of these two vertices at the same τ. *Thus, II has now become an instantaneous four-string interaction with the topology of a tetrahedron.* This is the desired tetrahedron graph. To rewrite the tetrahedron graph in familiar coordinates, let us use a_{12} and a_{13} as the independent variables. Then

$$a_{12}=2\pi\bar{\delta}, \quad a_{13}=\alpha\theta' . \tag{5.46}$$

Putting everything together, our final result is therefore

$$\begin{aligned} A_4 &= \langle V_{\alpha,\alpha,-2\alpha} | D | V_{-\alpha,-\alpha,2\alpha} \rangle \\ &= \mathrm{I}+\mathrm{II} \\ &= \langle V_{\alpha,\alpha,-2\alpha+2\delta} | D | V_{-\alpha,-\alpha,2\alpha-2\delta} \rangle + \int_{R_2} \mu\, da_{12}\, da_{13} \langle V_{\alpha,\alpha,-2\alpha+a_{12}/\pi} | e^{i(a_{13}/\alpha)(L_0-\bar{L}_0)} | V_{-\alpha,-\alpha,2\alpha-a_{12}/\pi} \rangle , \end{aligned} \tag{5.47}$$

where

$$\mu=|z\bar{z}|^{-2}\frac{\partial(x,\theta)}{\partial(a_{12},a_{13})} . \tag{5.48}$$

This is our final result. Notice that the closed four-string interaction is simply the square of the interpolating vertex.

There is a simple way in which to check the validity of our results. Notice that (5.34) gives us an exact expression for the variation of our coordinates when we make this reparametrization. This change in coordinates can be checked directly against a conformal transformation which also changes the parametrization lengths of strings. For example, (3.35) gives us an expression for τ in terms of an integral over the boundary of the Riemann surface. Now let us calculate the variation of τ given an infinitesimal change in the parametrization lengths at infinity and at the interaction points. A careful analysis of the surface terms (which cancel against each other) shows that

$$\delta\tau=\frac{1}{2\pi}\sum_i \int N(\sigma,\tau;\sigma_i',\tau_i')\delta d\sigma_i' , \tag{5.49}$$

where the last factor symbolically means that we are changing the integration region over the parametrization lengths. (The variation of the N itself and the variation of τ within N all cancel.) Now perform the integration in (5.49), resulting in an expression for the change in the interaction point due to a change in external parametrization. If we compare this against (5.34), we find that we have an exact correspondence.

For example, in the end-point gauge, the variation of the τ^r created by a reparametrization of the end-point vertex $\delta\alpha_s$ which maps the vertex back to the end-point configuration is given by

$$\delta\tau^r=\sum_n \cos(n\sigma^r/\alpha_r)N_{n0}^{rs}\delta\alpha_s ,$$

where $\sum\delta\alpha_s=0$.

In fact, we can always check each of our equations by going back to the end-point configuration, where the technology of Neumann functions is quite developed and we can perform almost all the summations explicitly. This check on our results is nontrivial. Notice that the change in τ described by (5.49) is a two-dimensional conformal transformation. However, the change (5.34) in our coordinates was due to a one-dimensional reparametrization. Thus, it is not obvious that a one-dimensional reparametrization should reproduce a two-dimensional conformal transformation. We have found, however, that they are indeed the same.

Several things should be noted here.

(1) The method of reflections can be carried out with any choice of parametrization lengths for the vertex functions. We have chosen equal lengths for the external legs only for convenience.

(2) Although we have only considered infinitesimal changes in parametrization lengths, the method can be iterated to produce finite changes in the parametrization lengths.

(3) Only the end-point gauge is free of the closed four-string interaction. All other gauges must have the closed four-string interaction in order to maintain modular invariance.

(4) Our result is modular invariant because it reproduces the covariantized light-cone configurations without four-string interactions, which are known to be modular invariant at the tree level because it reproduces the Shapiro-Virasoro amplitude and modular invariant to all loop levels because of the arguments of Refs. 2 and 3.

(5) The method of reflections generalizes easily to the multiloop amplitude. Because the various factors of L_n now reflect an infinite number of times within a loop, the factors of η and ξ simply reproduce the Neumann function for the multiloop amplitude.

(6) Although we have performed our method of reflections in the S matrix, we can always perform this process in the action itself. Then the derivation of the four-string interaction proceeds almost exactly as in QED. This will help resolve the question of whether there are five- or six-point closed-string vertices at higher levels. This will be presented in a later paper.

VI. CONCLUSION

In summary, we have seen that the key to modular invariance in closed-string field theory is to formulate the theory via geometric string field theory, which is defined in loop space, rather than parametrization space. The advantage of this formulation is that we can change the parametrization lengths of our strings at will via string vierbeins $e^{\nu\rho}_{\mu\sigma}$. Because the parametrization lengths are fictitious, the advantage of formulating the theory in loop space is that the extraneous complications due to ghosts, midpoints, ghost counting, etc., are all stripped clean.

The essential point is that the geometric theory has more fields than the "rigid" BRST string field theories. The string vierbein $e^{\nu\rho}_{\mu\sigma}$ is specifically designed to allow us to change the parametrization at any time.

In the geometric theory, the ghost sector is the tangent space of the theory, and conformal field theory is a powerful method of calculating the Clebsch-Gordan coefficients of irreducible representations of the local gauge group, the unified string group. In geometric field theory, the action arises as the unique gauge-invariant action.

The closed four-string interaction arises when we choose a specific gauge. As in QED, it arises when we eliminate a gauge field. In QED, the instantaneous four-fermion term arises when we eliminate A_0. In string theory, it arises when we eliminate the string vierbein $e^{\nu\rho}_{\mu\sigma}$. Thus, the four-string interaction (which is essential for modular invariance) is a gauge-dependent artifact. In fact, in some gauge it disappears altogether. In other gauges (e.g., the midpoint gauge for closed strings) it actually occupies most of the integration region.

We have seen that the Shapiro-Virasoro amplitude arises naturally when we include the four-string interaction. In fact, our "method of reflection" is powerful enough to solve for the four-string interaction in *any* gauge.

In a later paper, we will prove modular invariance to all orders at higher loop levels. We will show, at the level of the action rather than S-matrix elements, that we can always extract out the four-string Coulomb term when we change the gauge. As stressed in Ref. 9, the geometric theory for closed strings includes a proper time parameter, so that the counting of diagrams is the same as in Ref. 1. However, it has been shown in Refs. 2 and 3 that the light-cone counting yields precisely one modular-invariant amplitude (due to the imaginary periods of the Abelian integrals on the light-cone surface).

In particular, we will demonstrate the nontrivial way in which moduli space is triangulated when transforming from the end-point gauge to the midpoint gauge. We have three distinct ways in which this triangulation takes place. In the midpoint gauge, moduli space is triangulated by cylinders of equal circumference but unequal and uncorrelated lengths via three and four cylinder interactions. In the end-point gauge, however, moduli space is triangulated by cylinders of unequal circumference and unequal (but correlated) lengths via three cylinder interactions. In the interpolating gauge, moduli space is triangulated by cylinders of unequal circumference and unequal (but correlated) lengths via three and four cylinder interactions.

Thus, our counting differs considerably from that of Ref. 4, where there are apparently seven Teichmüller parameters for each closed-string loop (rather than six), because they have no proper time parameter.

ACKNOWLEDGMENTS

We are happy to acknowledge fruitful conversations with S. Giddings, B. Sakita, and A. Jevicki. M.K. was supported in part by Grants Nos. NSF-PHY-86-15338 and CUNY-FRAP-13873. J.L. was supported in part by DOE Contract No. DE-AA03-76SF00010.

[1]M. Kaku and K. Kikkawa, Phys. Rev. D **10**, 1100 (1974); **10**, 1823 (1974).
[2]S. Giddings and S. Wolpert, Commun. Math. Phys. **109**, 177 (1987).
[3]S. Giddings and E. D'Hoker, Nucl. Phys. **B291**, 90 (1987).
[4]H. Hata, K. Itoh, T. Kugo, H. Kunitomo, and K. Ogawa, Phys. Rev. D **35**, 1318 (1987); A. Neveu and P. West, Nucl. Phys. **B278**, 601 (1986).

[5]E. Witten, Nucl. Phys. **B276**, 291 (1986); see also J. Lykken and S. Raby, *ibid.* **B278**, 256 (1986); S. Sen and R. Holman, Phys. Rev. Lett. **58**, 1304 (1987); A. Strominger, Nucl. Phys. **B294**, 93 (1987); J. Feng, Report No. CALT-68-1466, 1987 (unpublished); G. Siopsis, Report No. CTP-TAMU-63, 1987 (unpublished).

[6]S. Giddings and E. Martinec, Nucl. Phys. **B278**, 91 (1986).

[7]D.J. Gross, J.A. Harvey, E. Martinec, and R. Rohm, Phys. Rev. Lett. **54**, 502 (1985).

[8]M. Kaku, *Introduction to Superstrings* (Springer, New York 1988); Int. J. Mod. Phys. **A2**, 1 (1987); Phys. Lett. B **200**, 22 (1988).

[9]M. Kaku, Reports Nos. CCNY-HEP-14-1986, CCNY-HEP-3-1987, CCNY-HEP-7-1988 (unpublished).

[10]S. Mandelstam, Nucl. Phys. **B64**, 205 (1973).

[11]S. Giddings, Nucl. Phys. **B278**, 242 (1986).

[12]K. Suehiro, Nucl. Phys. **B296**, 333 (1988).

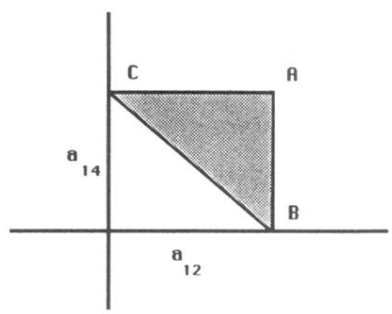

FIG. 7. In a_{ij} parameter space, the missing region can be shown to be a triangle. The corners of the triangle are the Rubik's cube points.