THE INTEGRABLE ANALYTIC GEOMETRY OF QUANTUM STRING ☆

Daniel FRIEDAN and Stephen SHENKER

Enrico Fermi and James Franck Institutes, and Department of Physics, University of Chicago, Chicago, IL 60637, USA

Received 13 May 1986

The quantum theory of closed bosonic strings is formulated as integrable analytic geometry on the universal moduli space of Riemann surfaces. Solutions of the equation of motion of quantum strings are flat hermitian metrics in holomorphic vector bundles over universal moduli space.

In this letter we formulate quantum string theory as integrable analytic geometry on the universal moduli space of Riemann surfaces, based on the analytic formulation of two-dimensional conformal field theory [1].

String theory currently provides the only viable candidate for a fundamental theory of nature [‡1]. Any such potentially fundamental theory should have a completely natural mathematical formulation. The gauge symmetries of string theory should be the transformations between mathematically equivalent descriptions of one underlying mathematical object. In particular, it is unnatural to assume a particular spacetime as a setting for string theory, since string theory is a quantization of general relativity. The structure of spacetime should be a property of the ground state.

One of our goals is a formulation of string theory which is sufficiently abstract to allow discussion of the nature of spacetime. A second goal is a formulation in which effective calculation might be possible. Even a potentially complete theory is unlikely to be useful, or even testable, unless effective nonperturbative calculations are possible. This argues for an analytic or even algebraic setting in which to do string theory.

The equation of motion of strings is essentially the two-dimensional conformal bootstrap equation [‡2]. The perturbation expansion of the string S-matrix is written as a sum over world surfaces in spacetime, which is the funcional integral version of a two-dimensional quantum field theory. The couplings of the two-dimensional quantum field theory express the background in which strings propagate. The perturbation expansion describes strings as the fluctuations around the ground state. The ground state itself is encoded as the two-dimensional field theory expressing the background.

A ground state of a string is a background in which the S-matrix of propagating strings is unitary. In bosonic string theory, unitarity requires conformal invariance of the world surface. Conformal invariance is expressed by the action of the Virasoro algebra in the Hilbert space of the two-dimensional field theory of the cylindrical world surface swept out by a closed string. The central charge must have the particular value $c = 26$. To solve the bosonic string equation of motion is to construct two-dimensional conformal field theories with $c = 26$. Understanding conformal field theory on all Riemann surfaces is the key to understanding quantum string theory.

Two-dimensional conformal field theory can be formulated as analytic geometry on the universal moduli space of Riemann surfaces [1]. In this letter we adapt the general analytic geometry of conformal field

☆ This work was supported in part by DOE grant DE-FG02-84ER-45144, NSF grant PHY-8451285 and the Sloan Foundation.

‡1 For references see ref. [2].

‡2 For background references in string theory and two-dimensional conformal field theory see refs. [1,2].

theory to the special case of string theory, where its application is especially natural. As the setting for the analytic formulation of both two-dimensional conformal field theory and string theory, we define the universal analytic moduli space of Riemann surfaces, which apparently has not been described previously. The fundamental object of quantum string theory is a hermitian metric h in a holomorphic, infinite-dimensional vector bundle W over universal moduli space. The quantum equation of motion of strings is the flatness equation on the metric h. Flatness equations are integrable. Here, the solutions are exactly the unitary representations of the universal modular group. String theory is thus reduced, in principle, to the representation theory of the universal modular group.

In fermionic string theory, unitarity is equivalent to *superconformal* invariance of the world surface. The fermionic theories are the only known string theories with consistent perturbative expansions. The equation of motion of the fermionic string is essentially the two-dimensional superconformal bootstrap equation. We concentrate in this letter on the closed bosonic string. The superanalytic formulation of the fermionic string parallels the bosonic theory. The fermionic string and two-dimensional superconformal field theory are set in the super analytic geometry of the universal super moduli space, based on the superconformal tensor calculus on super Riemann surfaces.

This work was motivated by the use of modular invariance as a crucial constraint in string theory [3,4] and in two-dimensional conformal field theory [5]. and by some of the recent discussions of the foundations of string theory and string field theory [6–9]. While engaged in the present work, we were encouraged by the paper of Belavin and Knizhnik [10] on the partition function of strings in flat spacetime, which also focuses on the complex analytic structure of moduli space, and by other recent studies of determinants of elliptic operators on surfaces and their factorization properties, on the moduli spaces of Riemann surfaces [11].

We begin by establishing some notation and defining the universal moduli space [1]. Let $\mathcal{M}_g$ be the moduli space of compact, connected, smooth, genus g Riemann surfaces [‡3]. Let $\overline{\mathcal{M}}_g$ be the moduli space of compact, connected, *stable* Riemann surfaces of genus g. The stable surfaces include the surfaces with nodes. $\overline{\mathcal{M}}_g$ is a compact, complex analytic orbifold or V-manifold, of complex dimension $n = 3g - 3$ for $g > 1$, $n = 1$ for $g = 1$ and $n = 0$ for $g = 0$. The space of surfaces with nodes is the compactification divisor $\mathcal{D}_g = \overline{\mathcal{M}}_g - \mathcal{M}_g$. The generic surface in $\mathcal{D}_g$ has exactly one node. The multiple self-intersections of $\mathcal{D}_g$, the subcomponents of $\mathcal{D}_g$, are the spaces of surfaces with multiple nodes. For every subcomponent of $\mathcal{D}_g$, for every configuration of k nodes, there is a collection of k independent analytic coordinates q^i, near $q^i = 0$, defining the subcomponent by the k independent equations $q^i = 0$. In more abstract language, all the self-intersections of $\mathcal{D}_g$ are transversal.

‡3 For explanations of moduli space see ref. [12]; for analytic geometry see ref. [13].

The generating functional of the connected part of the string S-matrix is

$$\mathcal{S}_{\mathrm{conn}} = \sum_{g=0}^{\infty} \int_{\mathcal{M}_g} Z(\overline{m}, m). \tag{1}$$

where $Z(\overline{m}, m)$ is the string partition function on the moduli spaces $\mathcal{M}_g$. The string coupling constant λ appears in a factor λ^{2g-2} which is absorbed into Z. $\mathcal{S}_{\mathrm{conn}}$ is a functional of the string background. The scattering amplitudes of strings are the coefficients of the infinitesimal variations of the background in the expansion of $\mathcal{S}_{\mathrm{conn}}$ around the string ground state.

The generating functional for the full S-matrix is the integral of the partition function over the moduli space of all compact, smooth Riemann surfaces, not necessarily connected:

$$\mathcal{S} = \exp(\mathcal{S}_{\mathrm{conn}}) = \int_{\mathcal{R}} Z(\overline{m}, m). \tag{2}$$

$\mathcal{R}$ is the universal moduli space

$$\mathcal{R} = \prod_{g=0}^{\infty} \left(\bigcup_{k=0}^{\infty} \mathrm{Sym}^k(\mathcal{M}_g) \right), \tag{3}$$

where Sym^k is the k-fold symmetric product. The symmetrization is due to the indistinguishability of conformally equivalent Riemann surfaces. Eq. (2) is a direct rewriting of eq. (1), given that the partition function of a disconnected surface is the product of the partition functions of the connected components. The factors $1/k!$ are provided by the symmetric products.

The space $\mathcal{R}$ is highly disconnected in its naive topology as a union of products of the individual moduli spaces $\mathcal{M}_g$. But the infinitely many disconnected components of $\mathcal{R}$ are connected by the formation of nodes. Any two surfaces in $\mathcal{R}$ can be obtained from a single connected surface of higher genus by analytic deformations which include formation and removal of nodes. The partition function extends to surfaces with nodes, so it lives on the universal moduli space of compact, *stable* Riemann surfaces

$$\overline{\mathcal{R}} = \prod_{g=0}^{\infty} \left(\bigcup_{k=0}^{\infty} \mathrm{Sym}^k(\overline{\mathcal{M}}_g) \right). \tag{4}$$

The compactification divisor $\mathcal{D} = \overline{\mathcal{R}} - \mathcal{R}$ is the space of compact Riemann surfaces with nodes. But nodes are invisible to a two-dimensional conformal field theory. That is, the partition function of a surface with nodes is exactly equal to the partition function of the smooth, possibly disconnected, surface which is made by removing the nodes and erasing the punctures which are left behind. This motivates defining an analytic structure on $\overline{\mathcal{R}}$ for which the removal of nodes is analytic. We write $m_{\mathcal{D}} \mapsto \pi(m_{\mathcal{D}})$ for the operation of removing the nodes from a surface $m_{\mathcal{D}} \in \mathcal{D}$ and erasing the punctures, leaving the smooth surface $\pi(m_{\mathcal{D}})$. We will see that $\overline{\mathcal{R}}$ is an effectively connected, compact analytic space on which the partition function is real analytic.

Let us now describe explicitly the partition function near $\mathcal{D}$. Take a particular coordinate neighborhood of a node, consisting of two disks $\{z_i : |z_i| < 1\}$ attached together at $z_1 = z_2 = 0$ to form the node. A surface with node is parametrized by $m_{\mathcal{D}} = (m, x_1, x_2)$, where $m = \pi(m_{\mathcal{D}})$ is the surface which remains when the node is removed and the punctures erased, and (x_1, x_2) is the unordered pair of punctures on the surface m which result from removal of the node. The opening of the node is parametrized by a single complex variable q. First remove a small neighborhood of the node, $|z_i| < |q|^{1/2}$ in both disks. Then attach the annular coordinate neighborhood $|q|^{1/2} < |z| < |q|^{-1/2}$ to the remaining surface by the identifications

$$\begin{aligned} z &= q^{1/2}/z_2 \quad \text{if } |q|^{1/2} < |z| \leqslant 1, \\ &= q^{-1/2} z_1 \quad \text{if } 1 \leqslant |z| < |q|^{-1/2}. \end{aligned} \tag{5}$$

The z annulus serves as a conformal plumbing joint. The moduli space near $\mathcal{D}$ is parametrized by $(m_{\mathcal{D}}, q) = (m, x_1, x_2, q)$, with $q = 0$ being the surface with node $m_{\mathcal{D}}$. In the limit $q \to 0$, z becomes a coordinate on the punctured complex plane, which is conformally equivalent to the infinite cylinder. To study the behavior of the partition function near $\mathcal{D}$, we only need to know the properties of conformal field theory on the annular plumbing joint in the limit $q \to 0$, which is just conformal field theory on the plane.

The conformal field theory on the plane is radially quantized, with Virasoro algebra $L_m, \overline{L}_m$. Near $\mathcal{D}$, the partition function of the surface $(m_{\mathcal{D}}, q)$ can be pictured as an expectation value in the radial quantization:

$$Z = \mathrm{tr}(q^{L_0} \overline{q}^{\overline{L}_0} \rho(m_{\mathcal{D}})) \tag{6}$$

where $\rho(m_{\mathcal{D}})$ is the density matrix of states on the boundary circles $|z_i| = 1$ which is prepared by the conformal field theory on the surface outside the region of the node $|z_i| < 1$. The density matrix ρ is independent of q. The trace in eq. (6) can be performed by summing over a complete set of states in the radial quantization. Use the one-to-one correspondence between the scaling fields $\varphi(x)$ of the conformal field theory and the eigenstates $|\varphi\rangle = \varphi(0)|0\rangle$, with L_0 eigenvalue h_φ and $\overline{L}_0$ eigenvalue $\overline{h}_\varphi$, to write the expansion in terms of two-point functions of the fields $\varphi(x)$ on the smooth surface $m = \pi(m_{\mathcal{D}})$:

$$\begin{aligned} Z &= \sum_{\varphi} q^{h_\varphi} \overline{q}^{\overline{h}_\varphi} \langle \varphi | \rho(m_{\mathcal{D}}) | \varphi \rangle \\ &= Z(\overline{m}, m) \sum_{\varphi} q^{h_\varphi} \overline{q}^{\overline{h}_\varphi} \langle \varphi(x_1) \varphi(x_2) \rangle_m . \end{aligned} \tag{7}$$

The correlation functions of the fields can be reconstructed from this expansion [1]. The leading contribution in eq. (7) comes from the ground state $|0\rangle$, which has $h_0 = \overline{h}_0 = 0$. The unique $\mathrm{SL}_2(\mathbb{C})$-invariant ground state $|0\rangle$ always corresponds to the identity operator, which is indifferent to its location, so the partition function on $\mathcal{D}$ itself is

$$Z(\overline{m}_{\mathcal{D}}, m_{\mathcal{D}}) = Z(\overline{m}, m), \quad Z = Z \circ \pi. \tag{8}$$

From the point of view of conformal field theory, as soon as a node forms all trace of it disappears.

We now put an analytic structure on $\overline{\mathcal{R}}$ which expresses precisely this indifference to nodes. The analytic structure on $\overline{\mathcal{R}}$ is the strongest analytic structure consistent with the naive analytic structure on $\mathcal{R}$, and for which the removal of nodes is an analytic operation. The strongest analytic structure is the one with fewest local holomorphic functions. Write the diagram

$$\begin{array}{ccc} \mathcal{D} & \to & \overline{\mathcal{R}} \\ \pi\downarrow & \nearrow & \\ \mathcal{R} & & \end{array} \tag{9}$$

in which $\mathcal{D} \to \overline{\mathcal{R}}$ and $\mathcal{R} \to \overline{\mathcal{R}}$ are the inclusion maps, and $\pi: \mathcal{D} \to \mathcal{R}$ is the map which removes the nodes and erases the punctures. The analytic structure on $\overline{\mathcal{R}}$ is the strongest for which (9) is a commuting diagram of analytic maps. A function f on $\overline{\mathcal{R}}$ is holomorphic if and only if it is holomorphic on the smooth surfaces $\mathcal{R}$, and satisfies $f_{/\mathcal{D}} = f \circ \pi$ on the compactification divisor. The topology on $\overline{\mathcal{R}}$ associated with this analytic structure is stronger than the naive topology. In particular, it is simple to show by induction in the genus that a globally holomorphic function on $\overline{\mathcal{R}}$ is constant. In this sense, $\overline{\mathcal{R}}$ is connected and compact. In such a rigid analytic setting, the global holomorphic objects are highly constrained and can be manipulated by analytic techniques with complete control. As far as we know, the universal moduli space has not previously been described. A universal Teichmüller space is known, but the corresponding moduli space is a single point [12].

The partition funtion for strings in spacetime has the form $Z = Z_{gh}Z_{st}$, where Z_{st} is the contribution from the two-dimensional conformal field theory of surfaces in d-dimensional spacetime, and Z_{gh} is the partition function of the chiral conformal ghost system of the world surface functional integral fixed in the natural conformal gauge [2].

The spacetime sector is an ordinary two-dimensional conformal field theory with central charge $c = d$. Formulated in the language of ref. [1], the partition function Z_{st} is a section of $\overline{E}_d \otimes E_d$, where, in general, $E_c = (\lambda_H)^{c/2}$, (λ_H) being the Hodge line bundle. E_c is the projective line bundle on $\overline{\mathcal{R}}$ which encodes the relationship between the two-dimensional conformal field theory and the surface geometry. The spacetime partition function is written in terms of a projectively flat hermitian metric h^{st} in a projective vector bundle W_{st} over $\overline{\mathcal{R}}$, and a holomorphic section $\psi_{st}(m)$ of the vector bundle $V_{st} = E_d \otimes W_{st}$ over $\overline{\mathcal{R}}$:

$$Z_{st}(\overline{m}, m) = h^{st}(\overline{\psi}_{st}, \psi_{st}) = \overline{\psi^a_{st}(m)} h^{st}_{\bar{a}b} \psi^b_{st}(m). \tag{10}$$

The ghost sector is not exactly a conformal field theory. The ghost partition function is $Z_{gh} = h^{gh}(\overline{\psi}_{gh}, \psi_{gh})$ where

$$\psi_{gh}(m) = \int db\, dc\, e^{S(b,c)}, \quad S(b,c) = \int_m b\bar{\partial}c \tag{11}$$

is the chiral fermionic ghost functional integral. Interpreted as a function on moduli space, $\psi_{gh}(m)$ would be identically zero, because of the $n = 3g - 3$ zero modes of $b(z)(dz)^2$ in genus $g > 1$, or the two constant zero modes b_0 and c_0 in genus $g = 1$. We whould instead interpret $\psi_{gh}(m)$ as a holomorphic half-density on $\overline{\mathcal{R}}$, a holomorphic section of the holomorphic line bundle $K = \wedge^{max}(T^*\overline{\mathcal{R}}^{1,0})$. Eq. (11) should be written

$$\psi_{gh}(m) = \int db\, dc\, e^{S(b,c)}$$

$$\times \int_m d^2z_1 \mu_1(\overline{z}_1, z_1) b(z_1) dm^1 \ldots$$

$$\times \int_m d^2z_n \mu_n(\overline{z}_n, z_n) b(z_n) dm^n, \tag{12}$$

where $\{dm^k\}$ is a basis of one-forms on $\overline{\mathcal{R}}$ at m, conjugate to a basis $\{\mu_k\}$ of Beltrami differentials on the surface m. To be more precise, again in the language of ref. [1], $\psi_{gh}(m)$ is a holomorphic section of $K \otimes E_{-26} \otimes W_{gh}$, where W_{gh} is a line bundle on $\overline{\mathcal{R}}$ with projectively flat hermitian metric h^{gh}.

In the critical dimension $d = 26$ the dependence on surface geometry disappears in the combined spacetime and ghost system. The curvature forms of the projectively flat metrics h^{st} and h^{gh} cancel. The product $h = h^{st}h^{gh}$ is a flat hermitian metric in $W = W_{st} \otimes W_{gh}$. The fibers of W are in general infinite dimensional. In the critical dimension, $E_d \otimes E_{-26} = E_0$ is the trivial line bundle on $\overline{\mathcal{R}}$, so $\psi(m) = \psi_{gh}(m)\psi_{st}(m)$ is a holomorphic section of $V = K \otimes W$. The string partition function is

$$Z(\overline{m}, m) = h(\overline{\psi}, \psi) = \overline{\psi^a(m)} h_{\bar{a}b} \psi^b(m), \tag{13}$$

which, as a section of $\overline{K} \otimes K$ is a density on $\overline{\mathcal{R}}$ which can be integrated, formally, over $\overline{\mathcal{R}}$ to give the S-matrix. Thus eq. (2) for the S-matrix is mathematically natural. This is the formal origin of gauge symmetry in the analytic quantum string theory. To go beyond *formal* gauge invariance we will need the super-analytic formulation of fermionic strings, since we do not know of a bosonic string ground state for which the integral in eq. (2) is finite.

The hermitian connection D in W is written in components

$$\bar{D} = \bar{\partial}, \quad D = \partial + A,$$

$$D h_{\bar{a}b} = \partial h_{\bar{a}b} - h_{\bar{a}c} A^c_b = 0. \tag{14}$$

A is the connection form of D. The quantum equation of motion of bosonic strings is the flatness equation

$$F = \bar{\partial} A = 0. \tag{15}$$

The connection in W can be thought of as the propagator of information across universal moduli space. Flatness, the local equation of motion of quantum strings, is the absence of ambiguity in the propagation of information.

Near a smooth, non-orbifold point in $\mathcal{R}$, a basis $\{w_a\}$ of covariant constant local holomorphic sections can be chosen for W. In this local frame, $A = 0$ and h is the constant hermitian matrix $h_{\bar{a}b} = h(\bar{w}_{\bar{a}}, w_b)$. In components, $Z = \overline{\psi^a(m)} h_{\bar{a}b} \psi^b(m)$, which is manifestly real analytic in m. Real analyticity of the partition function on $\bar{\mathcal{R}}$ is equivalent to locality on the world surface, and therefore to locality in spacetime. The real analyticity of the partition function follows from the local flatness of h and the local analyticity of ψ.

Because h and ψ are globally defined on $\bar{\mathcal{R}}$, Z is single valued under analytic continuation around the singular points in $\bar{\mathcal{R}}$, the orbifold points and the compactification divisor $\mathcal{D}$. To be single valued on $\bar{\mathcal{R}}$ can be interpreted as modular invariance. Let $\aleph$ be the universal analytic covering space of $\bar{\mathcal{R}}$. Define the universal modular group Γ to be the group of Deck or covering transformations of $\aleph$, $\bar{\mathcal{R}} = \aleph/\Gamma$. Γ is essentially the fundamental group of $\bar{\mathcal{R}}$ minus the singular points. Parallel transport of the covariant constant sections w_a around nontrivial closed curves in $\bar{\mathcal{R}}$, avoiding the singular points, gives a representation $\mathcal{H}$ of the universal modular group Γ: $\gamma w_b = \gamma^a_b w_a$. $\mathcal{H}$ is unitary because the metric $h_{\bar{a}b}$ is invariant under parallel transport. Therefore the string partition function is single valued:

$$Z = \overline{\psi^a(m)} h_{\bar{a}b} \psi^b(m) = \overline{\gamma^a_c \psi^c(m)} h_{\bar{a}b} \gamma^b_d \psi^d(m). \tag{16}$$

Modular invariance of the partition function becomes crossing symmetry of correlation functions on the world surface, when the correlation functions are reconstructed from the behavior of the partition function near $\mathcal{D}$. Crossing symmetry on the world surface is duality in string theory. The analytic formulation of quantum strings thus guarantees the two crucial properties of locality and duality.

The flat hermitian vector bundle W and the unitary representation $\mathcal{H}$ of Γ are equivalent. W is reconstructed from $\mathcal{H}$ as the quotient of $\aleph \times \mathcal{H}$ by Γ acting on both spaces simultaneously, $W = \aleph \times_\Gamma \mathcal{H}$. We say that the analytic geometry of quantum strings is integrable in the sense that this equivalence should eventually allow a purely group theoretic treatment of string theory.

The fiber of the universal analytic covering $\aleph \to \bar{\mathcal{R}}$ at a generic smooth point in $\bar{\mathcal{R}}$ is isomorphic to Γ itself. At a singular point in $\bar{\mathcal{R}}$, the fiber degenerates to a homogeneous space Γ/Γ_0, where Γ_0 is the little group at the singular point. At a generic smooth point, the fiber of W is modelled on $\mathcal{H}$. At a singular point, the fiber is W(0), modelled on the subspace $\mathcal{H}(0)$ consisting of the vectors in $\mathcal{H}$ which are invariant under the little group Γ_0.

The little group for a surface with nodes is the group generated by the Dehn twists $\gamma_k : q^k \to e^{2\pi i} q^k$ around the nodes. It is enough to describe the generic case of one node. In a basis of covariant constant sections near $\mathcal{D}$, the twist γ can be diagonalized, with eigenvalues $e^{2\pi i h}$. W(h) is the locally defined subbundle on which γ has eigenvalue $e^{2\pi i h}$. W degenerates to W(0) at $q = 0$.

The definition W as a holomorphic vector bundle over $\bar{\mathcal{R}}$ includes a transition map π^* corresponding to the removal of nodes map π in diagram (9). In terms of the local holomorphic sections of W over $\bar{\mathcal{R}}$, $\mathcal{R}$, and $\mathcal{D}$, written $\mathcal{W}(\bar{\mathcal{R}})$, $\mathcal{W}(\mathcal{R})$, $\mathcal{W}(\mathcal{D})$, the transition map π^* gives the commuting diagram

$$\begin{tikzcd} \mathcal{W}(\mathcal{D}) & \mathcal{W}(\bar{\mathcal{R}}) \arrow[l] \arrow[dl] \\ \mathcal{W}(\mathcal{R}) \arrow[u, "\pi^*"] & \end{tikzcd} \tag{17}$$

A local section of W must satisfy $\pi^* w(m_{\mathcal{D}}) = w(\pi(m_{\mathcal{D}}))$ on $\mathcal{D}$.

More concretely, let $m_{\mathcal{D}}$ be a surface with nodes, and write $m = \pi(m_{\mathcal{D}}) = m_1 \cup m_2 \cup \ldots \cup m_k$, the union of connected components m_i. Given sections w_{a_i} of

W at m_i, the symmetric product $w_{a_1} w_{a_2} \dots w_{a_k}$ is a section of W at m. The transition map π^* is given by the factorization structure tensors

$$\pi^*(w_{b_1} w_{b_2} \dots w_{b_k}) = F^a_{b_1 b_2 \dots b_k} w_a . \tag{18}$$

Note that a is an index for $W_{/\mathcal{D}} = W(0)$. Note also that $F^a_{b_1 \dots b_k}$ is symmetric (or graded symmetric because of the ghosts) in the indices b_i. The factorization tensors provide the data which is used to build the representation $\mathcal{H}$ of Γ from a collection $\mathcal{H}_g$ of representations of the finite genus modular groups Γ_g. Conversely, a representation of Γ determines a collection of factorization tensors.

The factorization tensors (18) can all be expressed in terms of the elementary structure tensor $F^a_{b_1 b_2}$, symmetric in b_1, b_2:

$$F^a_{b_1 b_2 \dots b_k} = F^a_{b_1 c_1} F^{c_1}_{b_2 c_2} \dots F^{c_{k-2}}_{b_{k-1} b_k} , \tag{19}$$

subject only to the associativity condition

$$F^a_{b_1 c} F^c_{b_2 b_3} = F^c_{b_1 b_2} F^a_{c b_3} . \tag{20}$$

The associativity condition is necessary and sufficient for π^* to be well defined, because all of the self-intersections of $\mathcal{D}$ are transversal.

The hermitian metric h, to be smooth on W at $\mathcal{D}$, must be consistent with the factorization structure:

$$h(\overline{\pi^* \psi}, \pi^* \psi) \circ \pi = h(\bar{\psi}, \psi),$$

$$h_{\bar{a}b} \overline{F^a_{c_1 c_2}} F^b_{d_1 d_2} = h_{\bar{c}_1 d_1} h_{\bar{c}_2 d_2} . \tag{21}$$

Eq. (21) makes h bear a strong formal resemblance to the metric of a Fock space, where F^a_{bc} is the symmetric tensor product.

To understand the behavior of ψ near $\mathcal{D}$, we only need some basic facts about the behavior of the conformal ghost fields on the complex plane [2]. The Fourier components of the ghost fields satisfy $[b_m, c_n]_+ = \delta_{m+n,0}$ and $c_n^\dagger = c_{-n}$, $b_n^\dagger = b_{-n}$. The SL_2 invariant state is written $|0\rangle$. Because of the background charge of the ghost system, $\langle 0|0\rangle = 0$, and the state conjugate to $|0\rangle$, satisfying $\langle 0'|0\rangle = 1$, is $|0'\rangle = c_{-1} c_0 c_1 |0\rangle$. Note that $|0'\rangle$ is not SL_2 invariant.

Insert a sum over states on each side of the node, at the circle $|z_1| = 1$ and also at the circle $|z_2| = 1$. The leading singularity comes from the insertion $|0\rangle\langle 0'| \dots |0'\rangle\langle 0|$, which has the SL_2 invariant states $|0\rangle$ facing the rest of the surface. From eq. (12), using explicit Beltrami differentials dual to dx_1, dx_2 and dq, we get the leading behavior

$$\psi_{gh}(m\mathcal{D})$$

$$\underset{q\to 0}{\sim} \psi_{gh}(m) \frac{1}{2\pi i} \oint dz_1 \frac{1}{2\pi i} \oint dz_0 \frac{1}{2\pi i} \oint dz_1$$

$$\times \langle 0'| q^{-1/2} z_1^2 b(z_1)\, dx_2 q^{-1} z_0 b(z_0)\, dq$$

$$\times (-q^{-1/2}) b(z_{-1})\, dx_1 |0'\rangle$$

$$\underset{q\to 0}{\sim} q^{-2} dx_1\, dq\, dx_2 \psi_{gh}(m). \tag{22}$$

The calculation in eq. (22) shows how K is constructed at $\mathcal{D}$. A holomorphic section χ of K near D must factorize on the double pole:

$$\chi(m, x_1, x_2, q) \underset{q\to 0}{\sim} q^{-2} dx_1\, dq\, dx_2 \chi(m). \tag{23}$$

The existence of the singular differential form $q^{-2} \times dx_1\, dq\, dx_2$ at $\mathcal{D}$ is the key to the definition of the canonical bundle K, allowing K at $\mathcal{D}$ to be expressed in terms of K on the smooth surfaces $\pi(\mathcal{D})$. The singular form $q^{-2} dx_1\, dq\, dx_2$ is natural at $\mathcal{D}$ exactly because each node turns into the punctured complex plane in the limit $q \to 0$.

As a secion of $V = K \otimes W$, ψ must factorize on the double pole:

$$\psi(m, x_1, x_2, q) \underset{q\to 0}{\sim} q^{-2} dx_1\, dq\, dx_2 \pi^* \psi(m),$$

$$\psi^a(m_1, m_2, x_1, x_2, q)$$

$$\underset{q\to 0}{\sim} q^{-2} dx_1\, dx_2\, dq F^a_{bc} \psi^b(m_1) \psi^c(m_2). \tag{24}$$

The factorization conditions (21) on $h_{\bar{a}b}$ and (24) on ψ give the factorization of Z:

$$Z(\bar{m}, \bar{x}_1, \bar{x}_2, \bar{q}, m, x_1, x_2, q)$$

$$\underset{q\to 0}{\sim} \bar{q}^{-2} q^{-2} |dq|^2 |dx_1\, dx_2|^2 Z(\bar{m}, m). \tag{25}$$

Belavin and Knizhnik [10] originally found the $\bar{q}^{-2} q^{-2}$ singularity of the genus g partition function in flat spacetime.

We have now described a quantum ground state of the string as a flat hermitian metric in a holomorphic vector bundle W over $\bar{\mathcal{R}}$, or equivalently a unitary representation $\mathcal{H}$ of Γ, along with a holomorphic section ψ of $V = K \otimes W$. The fundamental principles of string theory have been translated into flatness and analyticity on $\overline{\mathcal{R}}$. But this formulation of string theory is not concrete without a definite prescription for constructing $\mathcal{H}$. We will discuss possible criteria for choosing $\mathcal{H}$, but our understanding of this problem is still very incomplete. In particular, we lack a general abstract formulation of Wick rotation as analytic continuation of the representation $\mathcal{H}$. Until now we have implicitly been working in the abstract analog of "euclidean" spacetime, where the spacetime system is a unitary conformal field theory with straightforward factorization properties. In the Wick rotated system we expect there to be a manifold of solutions of the infinitesimal flatness equation (15), corresponding to physical fluctuations of the string. The crucial condition on $\mathcal{H}$ will be the positivity of the metric on physical states, after Wick rotation.

To deal with these issues it will surely be necessary to understand the gauge symmetries of the quantum string. The gauge symmetries of the classical string have been described in the context of classical string field theory [6,8]. Gauge symmetry in the quantum bosonic string is difficult to discuss because the partition function $Z = h(\overline{\psi}, \psi)$ cannot be integrated at $q = 0$. The integral could only be finite at $q = 0$ if the residue $\pi^*\psi$ at the double pole, and the single pole residue as well, were in the null space of h. But Z factorizes on the double pole, so the partition function itself would then be identically zero. It is intriguing to speculate that an argument of this type, for $Z = 0$, could be used to show the vanishing of the cosmological constant. In the super analytic formulation of fermionic strings, there must be an analogous factorization pole, but the integral over universal super moduli space cleverly suppresses the factorization singularity by the GOS projection, and suppresses the subleading singularity by a remarkable cancellation mechanism, which should in principle follow simply from smoothness of h and analyticity of ψ, on the super compactification divisor of super moduli space. We should wait to formulate the analytic geometry of the fermionic string before discussing gauge symmetries of the string S-matrix. But we proceed anyway, as practice for the fermionic string.

The obvious gauge symmetries are the unitary gauge transformations in the hermitian bundle W. But there is a larger class of gauge transformations. The vector bundle V is contained as a sub-bundle in the vector bundle Ω of differential forms on $\mathcal{R}$ with coefficients in W:

$$\Omega = \sum_{p,q=0}^{\max} (\wedge^{p,q}\overline{\mathcal{R}}) \otimes \mathrm{W},$$

$$\wedge^{p,q}\overline{\mathcal{R}} = \wedge^{p}(T^*\overline{\mathcal{R}}^{1,0}) \otimes \wedge^{q}(T^*\overline{\mathcal{R}}^{0,1}),$$

$$\Omega = \sum_{p,q=0}^{\max} \Omega^p_q, \quad \Omega^p_q = \wedge^p_q\overline{\mathcal{R}} \otimes \mathrm{V}$$

$$\mathrm{V} = \Omega^0_0, \quad \wedge^p_q\overline{\mathcal{R}} = \wedge^p(T\overline{\mathcal{R}}^{1,0})\wedge^q(T^*\overline{\mathcal{R}}^{0,1}). \quad (26)$$

The metric h on W, combined with the wedge product on forms, induces an inner product on sections ψ of Ω, $h\overline{\psi}\psi = h_{\overline{a}b}\overline{\psi}^{a}\psi^{b}$, whose value is a differential form on $\overline{\mathcal{R}}$. Integrating formally the volume element component in $h\overline{\psi}\psi$ gives an indefinite metric on the space of sections of Ω:

$$\langle\overline{\psi}\,|\,\psi\rangle = \int_{\overline{\mathcal{R}}} h\overline{\psi}\psi. \quad (27)$$

The connection D in W combines with the exterior derivatives $\overline{\partial}$, ∂ on forms to give exterior derivatives $\overline{\partial}$, D on sections of Ω. $\overline{\partial}$ is the ordinary antiholomorphic exterior derivative, $\overline{\partial}\Omega^p_q \subset \Omega^p_{q+1}$, and D is the covariant divergence operator $\mathrm{D}\Omega^p_q \subset \Omega^{p-1}_q$. The real exterior derivative with coefficients in W is Q $= \overline{\partial} + \mathrm{D}$. D being a metric connection is equivalent to $\mathrm{D}^2 = 0$. D being flat is equivalent to $[\overline{\partial}, \mathrm{D}]_+ = 0$. The compatibility of the metric h with the connection D is the formal self-adjointness condition $\mathrm{D}^\dagger = \overline{\partial}$. Thus the quantum equation of motion of the string can be written in the gauge invariant form

$$\mathrm{Q}^2 = 0, \quad \mathrm{Q}^\dagger = \mathrm{Q}. \quad (28)$$

If differential forms on $\overline{\mathcal{R}}$ are written in terms of ghost zero modes, the operator Q can be recognized as a quantum BRS operator. Thus eq. (28) is formally analogous to the classical field equation for open strings [8]. The gauge symmetries of the two equations have essentially the same structure. We stress, however, that eq. (28) is the *quantum* equation of motion of the *closed* string.

The flat hermitian geometry, or equivalently the unitary representation $\mathcal{H}$ of Γ, should be regarded as the primary object in the string theory. For each such representation $\mathcal{H}$ there is an associated complex $\Omega(\mathcal{H})$. The cohomology spaces of this complex are $H(\mathcal{H}) = H(\overline{\mathcal{R}}, W) = \Sigma_{p,q} H^p_q(\mathcal{H})$. The cohomology spaces $H(\mathcal{H})$ form a vector bundle over the space of representations of Γ. For the trivial representation 0, $H(0)$ is the cohomology ring $H(\overline{\mathcal{R}})$, which acts by multiplication of forms as an algebra on $H(\mathcal{H})$. As a section of Ω^0_0, ψ satisfies $\bar\partial\psi = 0$ because ψ is holomorphic, and $D\psi = 0$ because Ω^p_q is empty for $p < 0$. Therefore ψ determines a cohomology class, $[\psi](\mathcal{H})$ in $H^0_0(\mathcal{H})$.

We should be more precise about the analyticity condition $\bar\partial\psi = 0$. On the space of smooth surfaces, $\bar\partial\psi = 0$ is naively true. At $\mathcal{D}$, $\bar\partial\psi = 0$ is true despite, or rather because of, the double pole in ψ at $\mathcal{D}$. The $\bar\partial$ operator on K, and on V, should be written, in terms of the naive $\bar\partial_0$,

$$\bar\partial = \bar\partial_0 + (1/\pi)\partial\delta^2(q)\pi^* - (1/\pi)\delta^2(q)\mathrm{Res}_{\mathcal{D}},$$

$$\mathrm{Res}_{\mathcal{D}}\,\psi = \frac{1}{2\pi i}\oint \mathrm{d}q\,\psi(q). \tag{29}$$

The equation $\bar\partial\psi = 0$ thus enforces factorization on the double pole, which is necessary for analyticity at $\mathcal{D}$, while allowing ψ a subleading pole at $q = 0$. The single pole in ψ is associated with the fields of conformal weights $h = \bar h = 1$, corresponding to the massless excitations of the string at zero spacetime momentum.

We can interpret $\psi \to \psi + Q\eta$ as a formal gauge transformation, because the partition function is only changed by a total derivative:

$$h\bar\psi\psi \to h\bar\psi\psi + \mathrm{d}(h\bar\psi\eta + h\bar\eta\psi + h\bar\eta Q\eta). \tag{30}$$

The generating function

$$\mathcal{S} = \langle\bar\psi|\psi\rangle = \int_{\overline{\mathcal{R}}} h\bar\psi\psi \tag{31}$$

is formally gauge invariant, depending only on the cohomology class $[\psi](\mathcal{H}) \in H^0_0(\mathcal{H})$. The transformations $\psi \to \psi + Q\eta$ are thus formal gauge symmetries of the S-matrix. In particular, the single pole of ψ can be moved, or its residue modified, by a gauge transformation. Ground states with different massless excitations, and thus different spacetime geometries, might thus be gauge equivalent. The gauge *invariant* information is the cohomology class $[\psi](\mathcal{H})$. The choice of a representative ψ in the cohomology class is a choice of gauge.

We need to understand which unitary representations $\mathcal{H}$ are allowed ground states of the string. It is tempting to posit the uniqueness of $[\psi](\mathcal{H})$ as a basic condition on $\mathcal{H}$, since otherwise the ground state is not uniquely determined by the hermitian geometry alone. This condition would be $\dim H^p_q(\mathcal{H}) = 0$ unless $p = q = 0$, and $\dim H^0_0 = 1$. It would be interesting to try to interpret additional cohomology classes as some kind of BRS anomalies, thought of as obstructions to Wick rotation. The simplest resolution of the uniqueness problem would be that only one representation $\mathcal{H}$ exists which allows a Wick rotation giving a positive metric physical Hilbert space.

Unfortunately, there is no hope of uniqueness in finite genus, because many perturbative solutions of fermionic string theory are known, each with a finite number of free parameters. We might conjecture, however, that uniqueness of the string ground state can be attributed to nonperturbative effects. The nonperturbative effects should be accounted for by extending the sum over surfaces (2) from $\mathcal{R}$ to $\mathcal{R}_\infty$, the universal moduli space of surfaces, which includes an appropriate class of infinite genus surfaces. Here we only make some heuristic comments on this appropriate class. It should be based on some notion of effective compactness, which generalizes the compactness of the surfaces in $\overline{\mathcal{R}}$. We would then write $\overline{\mathcal{R}}_\infty$ as the space of effectively compact, stable Riemann surfaces. The compactification divisor $\mathcal{D}_\infty = \overline{\mathcal{R}}_\infty - \mathcal{R}_\infty$ should be dense in $\overline{\mathcal{R}}_\infty$. Adding or subtracting a handle, or splitting off a component surface, should be a small disturbance. The analytic structure on $\overline{\mathcal{R}}_\infty$ should be very smooth.

Every perturbative string ground state $\mathcal{H}$, ψ defined on the universal moduli space $\overline{\mathcal{R}}$ of finite genus Riemann surfaces, should have a well-defined extension to $\overline{\mathcal{R}}_\infty$. The uniqueness of the string ground state would require that these extensions all be gauge equivalent on $\overline{\mathcal{R}}_\infty$. We expect then that a uniqueness result is obtainable, if at all, only in the smooth geometry of $\overline{\mathcal{R}}_\infty$. We have considered the possible uniqueness of $\mathcal{H}$ in a setting analogous to euclidean signature spacetime. A puzzle common to all potentially complete theories is: should we expect a unique solution *after* Wick rotation? If there is a unique real time solution, up to gauge equivalence, then how is the S-matrix to be interpreted?

Spacetime geometry is associated with perturbative string theory, and presumably with the choice of gauge in $\overline{\mathcal{R}}_\infty$. The choice of gauge determines the residues of ψ at the single poles, and thus determines the massless physical particle spectrum. We do not know the details of the association between geometry and gauge choice because we do not have a theory of measurement. Measurement is the basic conceptual problem to be resolved in turning the formal structure of analytic quantum string theory into a theory of physics.

There should be no essential problem extending the analytic construction to fermionic string. The key ingredients are the local superconformal tensor calculus on super Riemann surfaces, and, because a node looks like the complex plane, the structure of the superconformal ghost system on the complex plane [2]. In the heterotic string, the partition function $Z = h\bar{\psi}_{\mathrm{ord}}\psi_{\mathrm{sup}}$ is an ordinary half-density in the antiholomorphic variables and a super half-density in the holomorphic variables. The spaces $\overline{\mathrm{W}}$ and W are not necessarily complex conjugate to each other. In fermionic string theory, the finiteness of the generating functional comes from cancellations of poles due to the sum over spin structures and the GOS projection, in the node viewed as a superconformal punctured plane.

In this work, we have defined the universal moduli space of Riemann surfaces, and we have written the quantum equation of motion of the bosonic string as the flatness equation (15). This equation can be integrated, in principle. The solutions are the unitary representations $\mathscr{H}$ of the universal modular group. We have given the gauge invariant form of the quantum equation of motion and have speculated on the relevance of infinite genus Riemann surfaces to the uniqueness of the ground state. Among the promising aspects of this analytic formalism is the possibility of precise control over string theory, so that certain key numbers like the cosmological constant might be determined exactly, even though it will probably remain impractical to describe the ground state explicitly and concretely in every detail. The remaining problems of immediate concern are: extending the analytic formalism to fermionic strings; giving a concrete description of the universal covering space $\aleph$ and the universal modular group; finding an abstract version of Wick rotation and specifying the physical positivity constraints on $\mathscr{H}$; and constructing a theory of measurement, so that the string states in the analytic formalism can be interpreted physically.

We should note that the strategy of the present work can be traced in several ways to the original S-matrix bootstrap program [14]. It is applied here in the highly constrained context of string theory. Analyticity, crossing symmetry, and unitarity of the S-matrix are expressed as flatness of h and analyticity of ψ on universal moduli space. Analyticity, crossing symmetry, and unitarity are abstracted away from spacetime, and adopted as dynamical principles of string theory, in the hope that the ground state of the string will resemble spacetime.

We thank Dan Freed, Emil Martinec, Edward Witten and Scott Wolpert for helpful conversations. We especially acknowledge discussions with Tom Banks on string field theory and on the foundations of string theory. A preliminary description of this work was given at the Workshop on Strings and Riemann Surfaces, MSRI, Berkeley, March 10–12, 1986.

References

[1] D. Friedan and S.H. Shenker, The analytic geometry of conformal field theory, EFI preprint 86-18A.
[2] E.g. M. Jacob, ed., Dual theory (North-Holland, Amsterdam, 1974);
J. Schwarz, ed., Superstrings, the first fifteen years (World Scientific, Singapore, 1985);
M. Green and D. Gross, eds., Proc. Workshop on Unified string theories (Institute for Theoretical Physics, Santa Barbara, CA, July 29–August 16, 1985) (World Scientific, Singapore, 1986), to be published;
D. Friedan, E. Martinec and S.H. Shenker, Conformal invariance, supersymmetry, and string theory, Chicago/Princeton preprint (1985), Nucl. Phys. B, to be published.
[3] J. Shapiro, Phys. Rev. D5 (1972) 1945.
[4] D. Gross, J. Harvey, E. Martinec and R. Rohm, Phys. Rev. Lett. 54 (1985) 502; Nucl. Phys. B256 (1985) 253; B267 (1986) 75.
[5] J. Cardy, UCSB preprint (1985), Nucl. Phys. B, to be published.
[6] W. Siegel, Phys. Lett. B 149 (1984) 157, 162; B 151 (1985) 391, 396;
W. Siegel and B. Zwiebach, Nucl. Phys. B263 (1986) 105;
T. Banks and M.E. Peskin, Nucl. Phys. B264 (1986) 573.
[7] D. Friedan, EFI preprint 85-27, Nucl. Phys. B, to be published; Phys. Lett. B 162 (1985) 102.
[8] E. Witten, Nucl. Phys. B268 (1986) 253.
[9] T. Banks, private communication.

[10] A.A. Belavin and V.G. Knizhnik, Landau Institute preprint (1986).
[11] J-M. Bismut and D. Freed, manuscript (1985);
R. Catenacci, M. Conalba, M. Martellini and C. Reina, Pavia preprint (1986);
J.B. Bost and T. Jolicoeur, Saclay preprint (1986);
N. Seiberg and E. Witten, Princeton preprint (1986);
L. Dixon and J. Harvey, Princeton preprint (1986);
L. Alvarez-Gaumé, P. Ginsparg, G. Moore and C. Vafa, Harvard preprint (1986);
L. Alvarez-Gaumé, G. Moore and C. Vafa, Harvard preprint (1986);
A. Cohen, G. Moore, P. Nelson and J. Polchinski, Harvard preprint (1986);
C. Gomez, IAS preprint (1986).
[12] E.g. L. Bers, Bull. London Math. Soc. 4 (1972) 257;
S. Wolpert, Ann. Math. 118 (1983) 491; On obtaining a positive line bundle from the Weil–Petersson class, University of Maryland Mathematics Department preprint (1983).
[13] E.g. P. Griffiths and J. Harris, Principles of algebraic geometry (Wiley, New York, 1978).
[14] J.A. Wheeler, Phys. Rev. 52 (1937) 1107;
W. Heisenberg, Z. Phys. 120 (1943) 513, 673;
G. Chew, *S*-matrix theory of the strong interactions (Benjamin, Reading, MA, 1961);
A.B. Zamolodchikov and Al.B. Zamolodchikov, Ann. Phys. 120 (1979) 253.