$$\mathcal{B}\begin{bmatrix}\delta'\\ \delta''\end{bmatrix}(z_i,\Omega;P_I^\mu)=Z_\Delta(\Omega)^{-16}\prod_{i<j}E(z_i,z_j)^{K_i\cdot K_j}\exp\left[i\pi P_I^a\Omega_{IJ}P_J^a+2\pi i P_I\cdot K_i\int_P^{z_i}\omega_I\right]\vartheta_\Lambda\begin{bmatrix}\delta'\\ \delta''\end{bmatrix}\left[K_i^a\int_P^{z_i}\omega_I,\Omega\right] . \quad (3.321)$$

Here the theta function for the lattice Λ with characteristics $\delta_I'^{\alpha}$ and $\delta_I''^{\alpha}$ is defined by

$$\vartheta_\Lambda\begin{bmatrix}\delta'\\ \delta''\end{bmatrix}(z_I^\alpha,\Omega)=\sum_{m_I^\alpha}\exp[i\pi(m_I^\alpha+\delta_I'^\alpha)\Omega_{IJ}g_{\alpha\beta}(m_J^\beta+\delta_J'^\beta)+2\pi i g_{\alpha\beta}(m_I^\beta+\delta_I'^\beta)(\delta_I''^\alpha+z_I^\alpha)] . \quad (3.322)$$

Under a modular transformation, we have

$$\vartheta_\Lambda\begin{bmatrix}D\delta'-C\delta''\\ -B\delta'+A\delta''\end{bmatrix}((C\Omega+D)^{-1}z,\Omega')=\zeta[\det(C\Omega+D)]^8 e^{i\pi z(C\Omega+D)^{-1}Cz}\vartheta_{\Lambda'}\begin{bmatrix}\delta'\\ \delta''\end{bmatrix}(z,\Omega) ,$$

where the image of Ω under the modular transformation Ω' is given by

$$\Omega'=(A\Omega+B)(C\Omega+D)^{-1} .$$

Moreover, Λ' is the lattice dual to Λ and

$$\zeta=\exp(-i\pi\delta'^{\mathrm{T}}B^{\mathrm{T}}D\delta'-i\pi\delta''^{\mathrm{T}}A^{\mathrm{T}}C\delta''+2\pi i\delta'^{\mathrm{T}}B^{\mathrm{T}}C\delta'') .$$

To have modular invariance, one clearly requires that the lattice be self-dual, $\Lambda'=\Lambda$, and that $\zeta=1$. The latter must hold for all modular transformations. The lattice must also be even, and this restricts the choice to $E_8\times E_8$ and Spin(32)/$Z(2)$ and causes the characteristics to vanish: $\delta'=\delta''=0$.

O. Inverse heterosis and general structure of amplitudes

Although the heterotic string was originally defined as the hybrid between a chiral half of the type-II string and another chiral half of the bosonic string, we have repeatedly witnessed the emergence of simplicity when working with the heterotic string directly. When calculations are initiated with the heterotic string action I_H of Eq. (3.301), using vertex operator insertions of the heterotic string for the relevant chirality, the final amplitude could be directly recast as an integral over internal momenta of the known ten-dimensional bosonic right-chirality part, times the left chirality of the heterotic string—alias the type-II string. Thus it appears that in practice the simplest way to compute in the heterotic or in the type-II superstring is to begin with Eq. (3.301) and, for fixed internal momenta, to decompose the amplitude into left- and right-chirality components.

Scattering amplitudes at fixed internal momenta, fixed supermoduli, and fixed spin structure are easily defined by inserting the constraints fixing the various internal momenta as in Eq. (2.97) for the bosonic string:

$$\langle V_1\cdots V_n\rangle(m_K,\overline{m}_K;p_I^\mu)=\int D(x\psi bc\beta\gamma)\prod_k|\langle\mu_k|B\rangle|^2\prod_b|\delta(\langle\mu_b|B\rangle)|^2V_1\cdots V_n e^{-I}\prod_{\mu,I}\delta\left[\oint_{A_I}\frac{dz}{2\pi}\partial_z x^\mu-p_I^\mu\right] . \quad (3.323)$$

This amplitude is Weyl invariant, local U(1) invariant, local reparametrization invariant, and local supersymmetry invariant. However, it is not modular invariant because we have picked out a preferred homology basis. Modular invariance is, however, recovered for the full amplitude after integrating over p_I^μ,

$$\langle V_1\cdots V_n\rangle=\int_{s\mathcal{M}_h}d^2m_K\int_{\mathfrak{F}}dp_I^\mu\langle V_1\cdots V_n\rangle(m_K,\overline{m}_K;p_I^\mu) . \quad (3.324)$$

The key feature of amplitudes at fixed momenta is that they will factorize as

$$\langle V_1\cdots V_n\rangle(m_K,\overline{m}_K;p_I^\mu)=(2\pi)^{10}\delta(k)|\mathcal{F}_\nu(m_K,z_i,\zeta;p_I^\mu)|^2 , \quad (3.325)$$

where $\mathcal{F}_\nu$ is holomorphic in moduli Ω_{IJ}, odd moduli $\chi_{\bar z}{}^+$, positions of vertex operators z_i, and parameters of the vertex operators ζ_i, and depends on left-chirality spin structure ν. When dealing with the heterotic string, there is an analogous statement,

$$\langle V_1\cdots V_n\rangle(m_K,\overline{m}_k;p_I^\mu)=\int D(x\psi bc\beta\gamma)\prod_k|\langle\mu_k|B\rangle|^2\prod_b\delta(\langle\mu_b|B\rangle)V_1\cdots V_n e^{-I}\prod_{\mu,I}\delta\left[\oint_{A_I}\frac{dz}{2\pi}\partial_z x^\mu-p_I^\mu\right] . \quad (3.326)$$

Here

$$\langle V_1\cdots V_n\rangle(m_K,\overline{m}_k;p_I^\mu)=(2\pi)^{10}\delta(k)\mathcal{F}_\nu(m_K,z_i,\zeta;p_I^\mu)\mathcal{B}(\overline{m}_k,\bar z_i,\bar\zeta;p_I^\mu) , \quad (3.327)$$

where $\mathcal{F}_\nu$ is the same as in Eq. (3.325). Since $\mathcal{B}$ is a known quantity, independent of supermoduli, we see that all the information of the type-II or heterotic strings is contained in the heterotic amplitudes.

P. Picture-changing formalism

Since all the information for the type-II and heterotic strings can be extracted from the study of the heterotic string (at fixed internal momenta), we shall restrict our attention to the heterotic string, keeping in mind that we may always fix the internal momenta preserving all symmetries but modular invariance. The formula for the amplitudes is then given in Eq. (3.178), where the heterotic string action is

$$I = I_H + I_{\rm sgh} \ , \tag{3.328}$$

with I_H given in Eq. (3.301) and $I_{\rm sgh}$ that of Eq. (3.175) but with $\bar\beta = \bar\gamma = 0$. It will be convenient to express it as

$$I = I_0 - \frac{1}{2\pi}\int d^2\xi \sqrt{g}\, \chi_{\bar z}{}^+ S \tag{3.329}$$

with

$$I_0 = \frac{1}{2\pi}\int d^2\xi\sqrt{g}\,(\tfrac{1}{2}D_z x^\mu D_{\bar z} x^\mu - \tfrac{1}{2}\psi_+^\mu D_{\bar z}\psi_+^\mu + bD_{\bar z}c + \bar b D_z \bar c + \beta D_{\bar z}\gamma) \tag{3.330}$$

and the full supercurrent

$$S = -\tfrac{1}{2}\psi_+^\mu D_z x^\mu - (D_z\beta)c - \tfrac{3}{2}\beta D_z c + \tfrac{1}{2} b\gamma \ . \tag{3.331}$$

It is always understood that anomalies are appropriately canceled by the presence of the internal degrees of freedom that we suppress here.

There is a BRST invariance inherited from the type-II string, and obtained by restricting Eq. (3.149) to left chirality only, using the bosonic BRST for the right components. In particular,

$$\delta_{\rm BRST}\beta = -\lambda S \ , \tag{3.332}$$

so that the full supercurrent is BRST invariant.

We shall now follow the treatment of Verlinde and Verlinde (1987a, 1987b) in order to make contact with the formulation of conformal field theory, usually expressed in terms of picture-changing operators. For the time being, our considerations will be local on moduli space; we shall need a better understanding of moduli space, and of its connection to supermoduli space, before being able to attack the global issues in Sec. VII. Locally, we can choose a slice for supermoduli space in which the super Beltrami differentials are of a special type. We assume that

$$\chi_{\bar z}{}^+ = \sum_{b=1}^{2h-2} m_b \mu_b^0 \ , \tag{3.333}$$

where m_b are the odd moduli and μ_b^0 are super Beltrami differentials independent of m_b. We can take the metric independent of m_b, so that $\mu_b^1 = 0$. This choice is justified by considerable simplifications in the superghost insertions in the ghost functional integral. We now have

$$\langle \mu_b \mid B\rangle = \langle \mu_b^0 \mid \beta\rangle \ . \tag{3.334}$$

By construction, the only thing that depends on odd moduli is $\chi_{\bar z}{}^+$ in the action, so that the odd moduli are easily integrated out. One finds

$$\langle V_1 \cdots V_n\rangle(m_K) = \int D(x\psi bc\beta\gamma) \times \prod_k |\langle \mu_k \mid B\rangle|^2 \times \prod_b \delta(\langle \mu_b^0 \mid \beta\rangle)\langle \mu_b^0 \mid S\rangle \times V_1 \cdots V_n e^{-I_0} \ . \tag{3.335}$$

The product

$$Y_{\mu_b} = \delta(\langle \mu_b^0 \mid B\rangle)\langle \mu_b^0 \mid S\rangle \tag{3.336}$$

is a BRST invariant, of ghost number 1, and can formally be thought of as the BRST transform of the step function

$$Y_{\mu_b} = \delta_{\rm BRST} H(\langle \mu_b^0 \mid \beta\rangle) \ , \tag{3.337}$$

though of course the step function is not well defined. From these quantum-number considerations, and the BRST invariance of Y, it is natural to guess that it is a generalization of the picture-changing operator. Indeed, for general μ_b, the operator Y is nonlocal, but if we choose

$$\mu_b(z) = \delta(z - z_b) \tag{3.338}$$

then it becomes local. Since z_b are points that could depend on moduli, they should properly be viewed as sections of the universal Teichmüller curve of Sec. IV.H. This bundle has no global smooth sections, which is another reason why these considerations should be considered as local on small patches of moduli space. Bosonization of superghost arguments, which we shall not present here (see, however, a brief discussion in Sec. VIII.E), give further evidence that this local version precisely coincides with the picture-changing operator of conformal field theory,

$$Y(z_b) = \delta(\beta(z_b))S(z_b) = e^{-i\sigma(z_b)}S(z_b) \ , \tag{3.339}$$

where $\sigma(z)$ is the bosonic field of Eq. (8.42). The expression for the general amplitude becomes

$$\langle V_1 \cdots V_n\rangle(m_K) = \int D(x\psi bc\beta\gamma) \times \prod_k |\langle \mu_k \mid B\rangle|^2 \times \prod_b Y(z_b) V_1 \cdots V_n e^{-I_0} \ . \tag{3.340}$$

This amplitude is formally BRST invariant. There are two issues that should, however, be investigated. The first is the usual possible contributions from the boundary of moduli space. The second is what happens when we make a different choice of insertions $\tilde{z}_b$, as will eventually be required by the topology of the universal Teichmüller curve. Using BRST arguments, Verlinde and Verlinde (1987) have argued that the difference will then be a total derivative on the patch where both z_b and $\tilde{z}_b$ are well defined. The effects of such total derivatives will be discussed in Sec. VII.G.

The prescription for the superstring multiloop measure based on BRST invariance and picture-changing operators is presented by Friedan, Martinec, and Shenker (1986) and Martinec (1987). In the path-integral formulation, that the ghost insertions can be recast as picture-changing operators [cf. Eq. (3.336)] was recognized by Witten (1986) in a superstring field theory context and later by Verlinde and Verlinde (1987). The last authors also provide key formulas for the conformal field theory of the superghosts and their bosonization.

IV. PARAMETRIZATIONS OF MODULI SPACE

In the previous sections we have considered the string partition function and scattering amplitudes as integrals over the moduli space $\mathcal{M}_h$ of compact Riemann surfaces or over the moduli space $\mathcal{M}_{h,n}$ of Riemann surfaces with n punctures as in Sec. II.J. These finite-dimensional spaces so far have been given abstract definitions as coset spaces, and it is imperative to describe them in a concrete manner, i.e., to provide some insight into their coordinates, curvatures, and function theory. A diverse choice of such parametrizations is available, and we shall here only describe some of those that have been used for the description of closed-string theory.

First, Riemann surfaces of constant curvature may be uniformized by the round sphere, the Euclidean plane, or the upper half plane. The natural geometry induced on $\mathcal{M}_h$ by such representation is by the Weil-Petersson metric. One of the remarkable aspects of the Weil-Petersson geometry is the abundance of completely explicit formulas, especially concerning the available coordinates and curvature formulas. In particular, we shall see how the Fenchel-Nielsen coordinates provide an elegant parametrization of Teichmüller space and yield explicit formulas for the Weil-Petersson geometry, though it is hard to identify moduli space, i.e., a fundamental domain for the mapping class group. If one is willing to formulate string theory on surfaces with at least one puncture, then the recently developed Penner decomposition (Penner, 1987a, 1987b) provides interesting formulas for the Weil-Petersson geometry on moduli space directly, identifying the boundaries of $\mathcal{M}_{h,n}$ as well.

The parametrization of moduli space with at least two punctures by Mandelstam diagrams has been discovered through string theory. In many ways, this is the parametrization diametrically opposed to constant curvature, since the curvature of Mandelstam diagrams vanishes everywhere except at some isolated interaction points where it is a Dirac δ function.

Another parametrization of the moduli space of surfaces with at least one boundary component, or in general for open strings, is provided by the open-string field theory of Witten (1986a, 1986b); in the mathematics literature it has been discussed independently in the work of Thurston (1980) and Bowditch and Epstein (1988). In a sense, it is analogous to the Penner decomposition, though it does not require constant curvature. We shall not discuss it further here, and instead refer the reader to the work of Witten (1986) and Giddings, Martinec, and Witten (1986), where these constructions were discussed.

Finally, Riemann surfaces may be parametrized by their $h\times h$ complex symmetric period matrix with the advantage of making the dependence on the complex structure of moduli space manifest. Every $h\times h$ complex symmetric matrix is not, however, the period matrix of a Riemann surface. The issue of which matrices do arise as period matrices of a Riemann surface (the so-called Schottky problem) still raises difficult questions. Actually, we shall deal with the complex structure of moduli space in its entirety in a completely separate section, VI.

A. Uniformization for constant-curvature geometry

Given a compact Riemann surface M of genus h and a metric of constant curvature R (recall that any metric is Weyl equivalent to a constant-curvature metric), it follows from the Gauss-Bonnet formula for the Euler number of the Riemann surface that R must be positive for $h=0$, zero for $h=1$, and negative for $h\geq 2$. For definiteness, we shall normalize the metric so that R is 1, 0, or -1, respectively, for the three cases.

The uniformization theorem states that M is isometric to a coset $\tilde{M}/\Gamma$, where $\tilde{M}$ is the simply connected covering of M, and Γ is a discrete subgroup of the isometry group of $\tilde{M}$, isomorphic to the first homotopy group of M:

$$\Gamma\sim\pi_1(M)\ .$$

Furthermore, the corresponding simply connected surfaces are unique for $h=0$, $h=1$ and $h\geq 2$, respectively, and are given by

the sphere $\mathbf{C}\cup\{\infty\}$, $ds^2=4(1+|z|^2)^{-2}dz\,d\bar{z}$, with $R=1$, isometry group: SU(2);

the plane $\mathbf{C}$, $ds^2=2\,dz\,d\bar{z}$, with $R=0$, isometry group: $z\mapsto az+b$, $|a|=1$;

the upper half plane, $\mathbf{H}=\{z=x+iy, y>0\}$, $ds^2=2y^{-2}dz\,d\bar{z}$, with $R=-1$, isometry group: SL(2, $\mathbf{R}$).

Actually, for compact surfaces, Γ should have no fixed points inside $\tilde{M}$ (note that Γ may have fixed points on the real line if $\tilde{M}=\mathbf{H}$). Thus, for $h=0$, Γ must be trivial, so for genus 0 there is only one sphere. We shall now study the cases for $h=1$ and $h\geq 2$ separately in the following sections.

B. The genus-1 and genus-2 cases and hyperelliptic surfaces

In the case of genus 1, moduli space and the Weil-Petersson metric can be identified easily. In fact it is readily seen that the only discrete fixed-point free subgroups of the isometry group of the plane are those generated by two translations, in two different directions if the quotient space is to be compact. Choosing two generators, we may characterize the complex structure by their ratio τ, which may be assumed to satisfy $\mathrm{Im}\tau > 0$. Different choices of generators lead to changes in τ which are generated by the transformations $\tau \mapsto -1/\tau$ and $\tau \mapsto \tau+1$. Thus the Riemann surface M can be identified with a parallelogram of sides 1 and τ, with opposite sides identified, Teichmüller space is just the upper half-space $\mathbf{H}=\{\tau_1+i\tau_2;\ \tau_2>0\}$, and the mapping class group is SL(2,Z). Up to a factor of $\frac{1}{2}$[cf. Eq. (2.123)], moduli space corresponds to a fundamental domain for SL(2,Z) within $\mathbf{H}$, which can be taken as (see Fig. 7)

$$\mathcal{M}_1=\{\,|\,\mathrm{Re}\tau\,|\leq\tfrac{1}{2},\ |\,\tau\,|\geq 1\}\ . \tag{4.1}$$

Note that SL(2,Z) admits fixed points, which will be crucial to the determination of phases of chiral determinants and space-time supersymmetry.

A quadratic differential on M must then be of the form $2\,\mathrm{Re}(\delta\kappa\, d\bar{z}^{\,2})$, with $\delta\kappa$ a complex constant, and the induced deformation of complex structures is

$$|\,dz\,|^2 \mapsto |\,dz\,|^2+2\,\mathrm{Re}(\delta\kappa\, d\bar{z}^{\,2}) = |\,dz+\delta\kappa\, d\bar{z}\,|^2 + O(\delta\kappa^2)\ . \tag{4.2}$$

This means that the complex coordinate z has undergone a "quasiconformal" deformation,

$$z \mapsto z+\delta\kappa\bar{z}\ , \tag{4.3}$$

and hence the new Riemann surface should be represented by a parallelogram of sides $1+\delta\kappa$, $\tau+\delta\kappa\bar{\tau}$, and the new ratio is $(\tau+\delta\kappa\bar{\tau})/(1+\delta\kappa)$. In particular $|\,\delta\tau\,|^2=4\,|\,\delta\kappa\,|^2\tau_2^2$, and since $\|2\,\mathrm{Re}(\delta\kappa\, d\bar{z}^{\,2})\|_{\mathrm{WP}}^2 = 8\,|\,\delta\kappa\,|^2$, it follows that

$$ds^2=2\,|\,\delta\tau\,|^2/\tau_2^2\ , \tag{4.4}$$

which is invariant under the mapping class group SL(2,Z).

At higher genus, such a simple parametrization is generally not available. However, when a surface can be represented as a double covering of the sphere—and is so-called *hyperelliptic*—then of course we have a polynomial equation for it, of the form

$$y^2=P(z)\ , \tag{4.5}$$

where $P(z)$ is a polynomial of degree $2h+2$ (or $2h+1$ if one point is sent to ∞). For the torus, this provides a representation familiar from the theory of elliptic functions.

Spin structures have a simple classification for hyperelliptic surfaces. Such surfaces have $2h+2$ branch points. Each partition of the branch points into two sets of $h+1+2k$ and $h+1-2k$ points correspond to a spin structure whose parity is that of k. The spin structures corresponding to $k=0$ and 1 are the nonsingular ones, where the Dirac operator has exactly no zero modes and one zero mode. More generally, k is the order of vanishing of the ϑ function with characteristics at $z=0$.

The set of hyperelliptic surfaces is a subvariety of moduli space of dimension $2h-1$ and hence is of measure zero for genus $h\geq 3$. At genus 2, however, every surface is hyperelliptic, and Eq. (4.5) is a good representation of a generic $h=2$ surface. Thus we have for genus 2

$$y^2=(z-a_1)(z-a_2)(z-a_3)(z-a_4)(z-a_5)(z-a_6)\ , \tag{4.6}$$

but of course, it should be realized that three distinct points can be fixed at will, so that we may set $a_4=0$, $a_5=1$, $a_6=\infty$. There then remain three complex coordinates: the three moduli of $\mathcal{M}_2$. The cut sphere is represented in Fig. 13, and the ramification points are exactly labeled by the a_i's. Actually, each distinct geometry of the cut plane provides a different Riemann surface, i.e., a different point in $\mathcal{M}_2$, and if all a_1, a_2, and a_3 run throughout $\mathbf{C}^3$, $\mathcal{M}_2$ is covered $6!=720$ times.

Holomorphic and meromorphic differentials for $h=2$ are completely explicit in this representation, and we have

- two holomorphic Abelian differentials

$$\omega_1=\frac{dz}{y},\quad \omega_2=\frac{z\,dz}{y}\ , \tag{4.7}$$

- three holomorphic quadratic differentials

$$\phi_j=\frac{z^{j-1}(dz)^2}{y^2},\quad j=1,2,3\ , \tag{4.8}$$

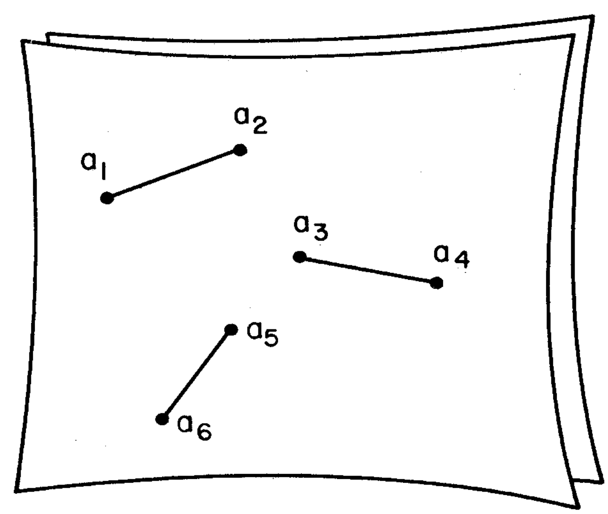


FIG. 13. Representation of a genus-2 surface as a square-root branched covering of the sphere (stereographically projected onto the complex plane).

- six holomorphic $\frac{1}{2}$ differentials, each corresponding to an odd-spin structure

$$v_i=\sqrt{(z-a_i)/y}\,(dz)^{1/2},\quad i=1,\ldots,6 \tag{4.9}$$

with single zeros at $z=a_i$,
- no holomorphic $\frac{1}{2}$ differentials with even-spin structure,
- two holomorphic $\frac{3}{2}$ differentials for each spin structure.

An odd-spin structure is just a selection of a branch point a_i, and the two holomorphic $\frac{3}{2}$ differentials are

$$(z-a_i)^{1/2}\left[\frac{dz}{y}\right]^{3/2},\quad z(z-a_i)^{1/2}\left[\frac{dz}{y}\right]^{3/2}.$$

Even-spin structures are partitions of the six branch points into two sets A and B of three elements each, and the holomorphic $\frac{3}{2}$ differentials are then

$$\prod_{a_i\in A}(z-a_i)^{1/2}\left[\frac{dz}{y}\right]^{3/2},$$

$$\prod_{a_i\in B}(z-a_i)^{1/2}\left[\frac{dz}{y}\right]^{3/2}.$$

A natural metric on the surface is obtained by using the holomorphic $\frac{1}{2}$ differentials

$$ds^2=|v_i v_{i'}|^2, \tag{4.10}$$

which has a double zero at a_i and $a_{i'}$ or a fourth-order zero at a_i if $i=i'$.

The above techniques extend to the case of hyperelliptic surfaces at higher genus as well, in a straightforward fashion, but we shall not discuss these here.

Work on explicit formulas for two-loop amplitudes includes that of Belavin *et al.* (1986), Kato, Matsuo, and Odake (1986), Moore (1986), and Lebedev and Morozov (1987). Conformal field theory on hyperelliptic surfaces has been dealt with by Zamolodchikov (1987). Two-loop studies were also carried out for the fermionic strings by Atick, Rabin, and Sen (1987), Atick and Sen (1987a), Moore and Morozov (1987), Morozov (1987), Parkes (1987), Lechtenfeld and Parkes (1988).

C. The higher-genus case

For $h\geq 2$, Γ is a subgroup of PSL(2,**R**) and its elements γ act on $z\in H$ by

$$z\to\gamma z=\frac{az+b}{cz+d},\quad \gamma=\begin{bmatrix} a & b\\ c & d\end{bmatrix},\quad ad-bc=1\ . \tag{4.11}$$

All γ's (except for the identity) act without fixed points, which requires that γ be *hyperbolic*, i.e.,

$$|\mathrm{tr}\gamma|>2\quad\text{for all }\gamma\in\Gamma-\{1\}\ . \tag{4.12}$$

Such groups are called *Fuchsian groups of the first kind.* We always choose representatives γ with $\mathrm{tr}\gamma>0$, without loss of generality. The Poincaré metric on **H** may be integrated, and we obtain the hyperbolic distance, given by

$$d(z,z')\geq 0,\quad \cosh d(z,z')=1-2\frac{(z-z')(\bar z-\bar z')}{(z-\bar z)(z'-\bar z')}\ . \tag{4.13}$$

Now since $\mathrm{tr}\gamma>2$, γ is conjugate within SL(2,**R**) to a pure *dilation*:

$$z\to e^{l_\gamma}z$$

whose fixed points are the origin and infinity, neither of which belongs to **H**. The scaling parameter l_γ is actually the shortest distance between any $z\in\mathbf{H}$ and its image γz:

$$l_\gamma=\min_{z\in H}d(z,\gamma z)\ . \tag{4.14}$$

On the surface M, the points z and γz are identified under the action of Γ, so the geodesic from z to γz is closed on M, and it belongs to the homotopy class of $\pi_1(M)$ corresponding to γ. Thus l_γ is the length of the shortest closed geodesic within the homotopy class of γ.

The group Γ is generated by its Fuchsian transformations around a canonical homology basis, say of A cycles and B cycles, so let us denote the corresponding generators by γ_{A_I} and γ_{B_I} for $I=1,2,\ldots,h$. Actually, these generators are only determined up to an overall isometry $\mathfrak{J}$ of **H**,

$$\gamma_{A_I}\to\mathfrak{J}^{-1}\gamma_{A_I}\mathfrak{J}\ ,\quad \gamma_{B_I}\to\mathfrak{J}^{-1}\gamma_{B_I}\mathfrak{J}\ ,$$

and the product of their commutators over all A and B cycles is the identity [the corresponding homotopy class of $\pi_1(M)$ is trivial]

$$\prod_{I=1}^{h}\gamma_{A_I}\gamma_{B_I}\gamma_{A_I}^{-1}\gamma_{B_I}^{-1}=1\ . \tag{4.15}$$

Since each generator depends on three real parameters (say a, b, and c), there are in total $6h$ real parameters, minus 3 for the global isometry and 3 more for the constraint, so altogether we are left with $6h-6$ parameters. Not surprisingly, this is exactly the dimension of Teichmüller and moduli space.

Does this means that we have an explicit parametrization of Teichmüller or moduli space? Suppose we pick γ_{A_I}'s and γ_{B_I}'s hyperbolic and satisfying Eq. (4.15); this can rather easily be done and implies some restrictions on the parameters of these matrices. Still, products of the γ_A and γ_B's need no longer be hyperbolic, leading to further restrictions on the parameter space for these matrices. The result for the $6h-6$ dimensional parameter space is a region of $\mathbf{R}^{6h-6}$ with a highly dented boundary. The corresponding coordinates are the so-called Fricke-Klein (1926) coordinates.

More on Fricke-Klein coordinates can be found in McKean (1972), Harvey (1978), Bers (1981), and Bers and Gardiner (1986).

D. Normal coordinates in the higher-genus case

For Riemannian manifolds there is a natural way of parametrizing a local neighborhood of a given point by tangent vectors at that point. In fact, to a tangent vector corresponds simply the point a unit amount of time away on the geodesic tangent to that vector. These are usually called normal coordinates and can be constructed as follows in the case of moduli space. Consider a fixed complex structure which will be identified with a Fuchsian group of the first kind. A tangent vector to moduli is a quadratic differential ϕ (equivalently a harmonic Beltrami differential $\bar{\mu}=\phi y^2$ with $y=\text{Im}z$), the geodesic with initial velocity ϕ a one-parameter family of Fuchsian groups Γ_ε. The Γ_ε's can be obtained by solving the Beltrami equation

$$\partial_{\bar{z}} w = \varepsilon\mu\partial_z w \tag{4.16}$$

for a mapping w sending $\mathbf{H}$ into itself, and setting $\Gamma_\varepsilon = w^{-1}\Gamma w$. Here we have extended μ by $\mu(z)=\mu(\bar{z})$ for z in the lower half-space. A related construction putting the complex structure of moduli better in evidence is based on extending μ to be 0 instead. The resulting w will then no longer preserve the real axis, so that $\mathbf{H}$ will be deformed into a quasi-half-space and Γ into "quasi-Fuchsian groups." Despite this difference in emphasis, the two ways lead to the same deformation, since the w's obtained either way differ only by a holomorphic mapping. Note that w is not conformal, and that this construction is the natural generalization of that described in Eq. (4.3) for the torus. Choosing now an orthonormal system of quadratic differentials ϕ_α, $\alpha=1,\ldots,3h-3$, and repeating the construction for $\phi=\sum_{\alpha=1}^{3h-3} t_\alpha\phi_\alpha$ we can parametrize a neighborhood of Γ by $(t_\alpha)\in\mathbf{C}^{3h-3}$. In this coordinate system the Weil-Petersson metric will satisfy $\partial_\gamma g_{\alpha\bar{\beta}}\,|_\Gamma=\partial_{\bar{\gamma}} g_{\alpha\bar{\beta}}\,|_\Gamma=0$, which implies in particular that it is Kählerian. We shall be especially interested in the Kähler form

$$\omega_{\text{WP}} = g_{\alpha\bar{\beta}}\, dt^\alpha \wedge dt^{\bar{\beta}} , \tag{4.17}$$

since its $(3h-3)$ power is the desired volume form.

Normal coordinates for the Weil-Petersson metric were introduced by Ahlfors (1966). A modern account including a detailed analysis of second variations of the area element of the surface under quasiconformal deformations (4.16) is that of Wolpert (1986). Normal coordinates for general Riemannian manifolds are also useful in background field calculations. See, for example, Alvarez-Gaumé, Freedman, and Mukhi (1981).

E. Fenchel-Nielsen coordinates

We now describe briefly Fenchel-Nielsen coordinates which are real coordinates for Teichmüller space. Although the complex structure is not evident in this system, they have the advantage of presenting Teichmüller space as $(\mathbf{R}\times\mathbf{R}^+)^{3h-3}$ and of providing a particularly simple formula for the Weil-Petersson Kähler form. To define these coordinates, one makes use of the following construction. Consider the maximal set of closed, nonintersecting geodesics on a given surface M. It is clear that $3h-3$ nonintersecting closed geodesics may always be drawn on a surface of genus h (see Fig. 14 for an example). It is not hard to see that any additional closed geodesic has to intersect at least one of the $3h-3$ initial ones. Thus the maximal number is $3h-3$; let us call their lengths l_i, $i=1,2,\ldots,3h-3$. Along each of these geodesics, we may now cut the surface apart and reglue it after a relative twist by an angle θ_i. For the range

$$\begin{aligned} &0<l_i<\infty, \quad i=1,2,\ldots,3h-3 , \\ &-\infty<\theta_i<\infty , \end{aligned} \tag{4.18}$$

these parametrize precisely one copy of Teichmüller space. Surfaces with nodes arise at the boundary of this domain when one of the l_i's goes to zero. Notice that the $3h-3$ closed geodesics divide M into $2h-2$ surfaces of genus 0 with three discs removed, which are called *pants*.

Now it is not difficult to see that the hyperbolic structure on a pant can be characterized by the lengths of the three boundaries. In fact, each pant can be viewed as built out of two copies of right hexagons (whose corners have 90° angles), and hexagons are characterized by three alternate sides. The gluing of these pants involves the relative twist of an angle θ_i along each geodesic.

The Weil-Petersson Kähler form has been shown to take the simple form

$$\omega_{\text{WP}} = \sum_{j=1}^{3h-3} l_j\, dl_j \wedge d\theta_j , \tag{4.19}$$

so that the Weil-Petersson measure is completely explicit and given by

$$d(\text{WP}) = \prod_{j=1}^{3h-3} l_j\, dl_j\, d\theta_j . \tag{4.20}$$

This very geometric viewpoint is also natural to path-integral quantization. In particular, D'Hoker and Phong (1986c) have exhibited string determinants in terms of these coordinates and Green's functions on pants. These in principle can be built of prime forms on hyperelliptic surfaces and may ultimately lead to rules for string amplitudes more closely analogous to the usual Feynman rules of field theory. The main difficulty of this approach is that the action of the mapping class group on the Fenchel-Nielsen coordinates and on the pant decomposition is extremely complicated, and unless some more direct way of representing this action can be found, their usefulness as a characterization of moduli space is unclear.

This decomposition is reminiscent of the division into

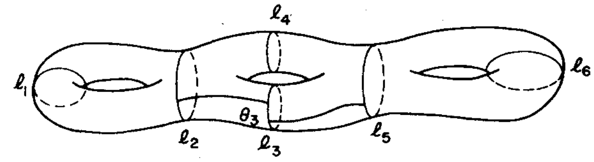


FIG. 14. Decomposition of a surface of genus 3 into four "pants" and corresponding Fenchel-Nielsen coordinates.

primitive graphs briefly discussed in the days of dual models by Gross and Schwarz (1970). In fact, it is quite tempting to construct a *fieldlike theory* this way, without propagators, however. Indeed, take the pant and now sew enough pants together so as to reconstruct the desired string diagram. Unfortunately, this approach does not seem to work, for essentially the same reasons as are given above: though one formally obtains the measure and the integrand, naive sewing will also yield an infinite factor in front, which is basically the cardinality of the mapping class group. One would have to factor out the proper (infinite) combinations right away, and this seems rather hopeless. The latter idea was explored by D'Hoker and Gross; similar attempts at string field theory based on pants may be found in Tseytlin (1986).

Equation (4.19), which shows that Fenchel-Nielsen coordinates are canonical coordinates with respect to the symplectic structure defined by Weil-Petersson Kähler form, is due to Wolpert (1982, 1983). These references also contain a great deal more on the interplay between the symplectic geometry of Teichmüller space and the hyperbolic geometry of the surface.

F. Penner decomposition

Recently, another description of the Teichmüller space $\mathcal{T}_{h,n}$ of Riemann surfaces endowed with constant-curvature metrics has emerged. However, this construction works only when there is at least one puncture, so that $n \geq 1$. The Penner decomposition exhibits a simple behavior under the action of the mapping class group, so that it may be used to describe the corresponding moduli space $\mathcal{M}_{h,n}$ of punctured surfaces. In addition, the Weil-Petersson Kähler form admits an explicit representation, and there is a reasonable description of the boundary of the fundamental domain of moduli space. In practice, we shall here restrict ourselves to Riemann surfaces with only one puncture, the generalization to the case with more punctures being straightforward.

One starts by representing a two-dimensional hyperbolic geometry by one copy $\mathcal{H}$ of the two-sheeted hyperboloid in $\mathbf{R}^3$, endowed with the Minkowski inner product $\xi\cdot\xi'=\xi^1\xi'^1+\xi^2\xi'^2-\xi^3\xi'^3$ for $\xi=(\xi^1,\xi^2,\xi^3)$:

$$\mathcal{H}=\{\xi\in\mathbf{R}^3;\xi\cdot\xi=-1,\ \xi^3>0\}\ . \tag{4.21}$$

The component connected to the identity $SO^+(2,1)$ of SO(2,1) leaves $\mathcal{H}$ invariant and acts isometrically. The metric induced on $\mathcal{H}$ by the flat Minkowski metric has curvature -1. Actually, there is a simple correspondence between the complex upper half-plane $\mathbf{H}=\{z=x+iy\in\mathbf{C},\, y>0\}$ and $\mathcal{H}$, given by

$$\xi^1=\frac{x}{y},\quad \xi^2=\frac{\Lambda^2-x^2-y^2}{2\Lambda y},\quad \xi^3=\frac{\Lambda^2+x^2+y^2}{2\Lambda y}\ , \tag{4.22}$$

where Λ is an arbitrary constant >0 and the hyperbolic metric on $\mathbf{H}$ is linked to the inner product on $\mathcal{H}$ by $\cosh d=-\xi\cdot\xi'$. The geodesics of $\mathbf{H}$, half-circles centered on the real line and arbitrary radius, are mapped onto the geodesics of $\mathcal{H}$, hyperbolas lying in planes that pass through $\xi=0$; see Fig. 15.

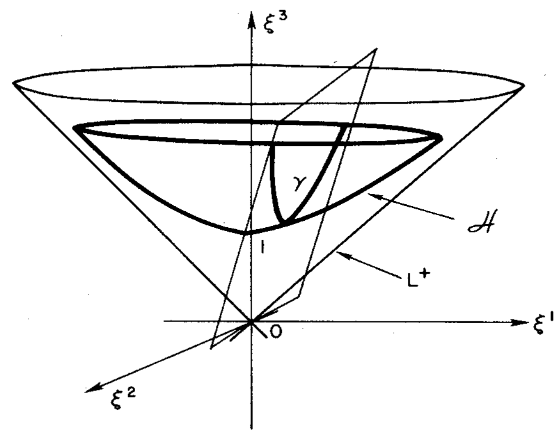


FIG. 15. Hyperbolic geometry as constructed from the three-dimensional hyperboloid $\mathcal{H}$. A geodesic γ is the intersection of $\mathcal{H}$ with a plane through the origin.

For every element of $\mathcal{M}_{h,1}$, we have a representation of $\pi_1(M)$ into $SO^+(2,1)$ by a Fuchsian group Γ, and the Riemann surface is represented by $M=\mathcal{H}/\Gamma$. The positive light cone is given by

$$L^+=\{\xi\in\mathbf{R}^3,\ \xi\cdot\xi=0,\ \xi^3>0\}\ . \tag{4.23}$$

Now consider a Riemann surface with one puncture P, so that Γ has a parabolic[27] generator $\gamma_P\in SO^+(2,1)$, corresponding to that puncture. A parabolic element $\neq 1$ is characterized by the light ray in L^+ it leaves invariant, and we may pick a particular point z in L^+ on this ray to represent γ_P.

A geodesic of $\mathcal{H}$ that starts at P and returns to P is called an *ideal arc*. The total length of an ideal arc is clearly infinite, but we shall now give a natural regularization. We draw a small circle around the puncture, orthogonal to all geodesics emanating from P; this circle, denoted by h, is called a *horocycle* and is characterized by its length (see Fig. 16 for the torus). If we had several punctures, each of them would inherit a horocycle. The horocycle defines a small disc with puncture P, and this disc may be viewed as the coset of $\mathcal{H}$ by the cyclic group generated by γ_P. The hyperbolic length of a geodesic c starting at P and returning to it, as measured from its intersections with the horocycle h, is now finite and denoted by $d_h(c)$. As the radius of the horocycle tends to zero $(h\to P)$, this length again diverges, but the difference between the lengths of two geodesics c_1 and c_2 converges,

$$\lim_{h\to P}\exp[d_h(c_1)-d_h(c_2)]=\left[\frac{\lambda(c_1)}{\lambda(c_2)}\right]^2\ . \tag{4.24}$$

[27]An element $\gamma\neq 1$ of $SO^+(2,1)$ is hyperbolic, parabolic, or elliptic provided the eigenvector with eigenvalue 1 lies outside L^+, on L^+, or inside L^+.

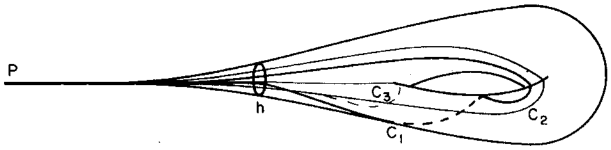


FIG. 16. An ideal triangulation of the once-punctured torus. The ideal arcs are C_1, C_2, and C_3. The horocycle h is also indicated.

If $\gamma(c)$ denotes the element of $SO^+(2,1)$ describing the ideal arc c starting at P and returning to it, and ξ is the element of L^+ fixed by γ_P, then one can show that the λ lengths are given by

$$\lambda^2(c) = -\xi \cdot [\gamma(c)\xi] \ . \tag{4.25}$$

Each length $\lambda(c) = \lambda_\Gamma(c)$ depends on the Fuchsian group, and the lengths are natural coordinates in the sense that elements φ of the mapping class group act simply on their ratios,

$$\frac{\lambda_{\varphi^*\Gamma}(c_1)}{\lambda_{\varphi^*\Gamma}(c_2)} = \frac{\lambda_\Gamma(\varphi^{-1}c_1)}{\lambda_\Gamma(\varphi^{-1}c_2)} \ . \tag{4.26}$$

It is the purpose of this construction to use the λ lengths as coordinates for moduli space.

1. Ideal triangulations

One obtains an *ideal triangulation* Δ of the Riemann surface $M_{h,1}$ by considering a maximal family of disjointly embedded simple ideal arcs Δ, so that no component of $M_{h,1} - \Delta$ is a mono-gon or bi-gon.

Thus an ideal triangulation is a decomposition of $M_{h,1}$ by ideal arcs into regions whose double is a sphere with three punctures. It is easy to see that $6h-3$ ideal arcs is the maximal number, and they divide the surface into $4h-2$ triangles, whose corners are all identified with the puncture P (see Fig. 16). Now given $M_{h,1}$ with its horocycle around the puncture P, the λ lengths of an ideal triangulation provide a one-to-one and onto map between $\mathcal{T}_{h,1}$ (together with its horocycle length) and $\mathbf{R}^q_+$, under which the ideal arcs $c_1, c_2, \ldots, c_q$ of Δ are sent into $\lambda(c_1), \lambda(c_2), \ldots \lambda(c_q)$. Here $q = 6h-3$, but recall that only ratios of λ lengths had a geometrical meaning, so that actually only $6h-4$ independent values survive, exactly the number of moduli of $\mathcal{M}_{h,1}$. (If we had n punctures, there would be $6h-6+3n$ ideal arcs, dividing the surface into $4h-4+2n$ triangles.)

2. Ideal cell decompositions

We now define a slightly refined decomposition of $M_{h,1}$ into triangles, which will naturally induce a mapping-class-group-invariant cell decomposition on $\mathcal{T}_{h,1}$.

Consider the orbit Γz for $z \in L^+$ and z invariant under the parabolic element γ_P, and consider its convex hull K, which is Γ invariant. The boundary ∂K consists of a countable set of faces, each of which is the convex hull of a finite number of points. Hyperbolicity of Γ guarantees that any face of ∂K intersects L^+ in an ellipse. Each edge of a face determines a geodesic in $\mathcal{H}$ and hence an ideal arc in $M_{h,1}$. The collection of ideal arcs in $M_{h,1}$ arising from all edges of faces of ∂K is a disjointly embedded collection Δ of arcs in $M_{h,1}$ connecting punctures, so that $M_{h,1} - \Delta$ is simply connected. The homotopy class of such a family of arcs is called an *ideal cell decomposition.* Thus with each $\Gamma \in \mathcal{T}_{h,1}$ we have associated an ideal cell decomposition $\Delta(\Gamma)$ of $M_{h,1}$.

A cell decomposition $\mathcal{C}$ of $\mathcal{T}_{h,1}$ is obtained by considering the collection of

$$\mathcal{C}(\Delta) = \{\Gamma \in \mathcal{T}_{h,1} \text{ such that } \Delta(\Gamma) \text{ belongs to the class } \Delta\} \tag{4.27}$$

and Δ ranges over all ideal cell decompositions of $M_{h,1}$. Each $\mathcal{C}(\Delta)$ is contractible, and $\mathcal{C}$ is a mapping-class-invariant decomposition.

The action of the mapping class group on Δ is generated by the following operation on two consecutive triangles with edges $\{a,b,e\}$ and $\{c,d,e\}$ of Δ, where the side e is common. Consider the quadrangle $\{a,b,c,d,e\}$ with one diagonal e. Remove e from the quadrangle and replace it by the other diagonal f; this yields a new triangulation Δ'. The corresponding λ lengths are polynomially related:

$$\lambda(e)\lambda(f) = \lambda(a)\lambda(c) + \lambda(b)\lambda(d) \ . \tag{4.28}$$

When Teichmüller space is described in terms of the λ coordinates, a cell $\mathcal{C}(\Delta)$ is described by the following inequalities. Let $\{a,b,e\}$ and $\{c,d,e\}$ be two consecutive triangles as before and let $\{\alpha,\beta,\varepsilon\}$ and $\{\gamma,\delta,\varepsilon\}$ be the λ lengths of their sides; then we have

$$\alpha+\beta > \gamma, \quad \alpha+\gamma > \beta, \quad \beta+\gamma > \alpha \ , \tag{4.29}$$

as well as

$$\begin{aligned} 0 \le\ & \mathrm{sgn}(\alpha^2+\beta^2-\varepsilon^2) K(\alpha,\beta,\varepsilon) \\ & + \mathrm{sgn}(\gamma^2+\delta^2-\varepsilon^2) K(\gamma,\delta,\varepsilon) \\ K(\alpha,\beta,\gamma) =\ & [(\alpha+\beta-\gamma)(\alpha+\gamma-\beta) \\ & \times (\beta+\gamma-\alpha)(\alpha+\beta+\gamma)]^{1/2} \ . \end{aligned} \tag{4.30}$$

With the help of this cell decomposition and the action of the mapping class group, Penner has succeeded in computing the orbifold Euler number of moduli space $\mathcal{M}_h$ and rederived the well-known formula of Harer and Zagier (1986),

$$\chi(\mathcal{M}_{h,1}) = \zeta(1-2h) \ , \tag{4.31}$$

where $\zeta(z)$ is the Riemann zeta function.

3. Integration over moduli space

There also exists a method for integrating forms invariant under the action of the mapping class group, in terms of integrations over cells whose edges satisfy the nonlinear inequalities above. The triangles are assembled into "fat graphs" G, which describe concisely the homotopy classes Δ and thus the cells $\mathcal{C}(\Delta)$ of $\mathcal{T}_{h,1}$. The formula for the integration of a top-dimensional form of ω of $\mathcal{M}_{h,1}$ is particularly elegant:

$$\int_{\mathcal{M}_{h,1}} \omega = \sum_{[G]} \frac{2^{\varepsilon[G]}}{\#\Gamma(G)} \int_{D_G} \phi^* \omega \ , \tag{4.32}$$

where D_G is the region of integration described by the nonlinear inequalities for the fat graph G, and the sum is over all G with trivalent vertices only. $\varepsilon[G]=1$ when G has two vertices and is hyperelliptic, and $\varepsilon[G]=0$ otherwise. $\Gamma(G)$ is the isotropy group of the cell corresponding to the fat graph G (i.e., the combinatorial factor familiar from Feynman diagrams).

The Weil-Petersson measure can be obtained from the Weil-Petersson Kähler form in the standard fashion, and the latter is given in terms of the real (positive) λ lengths $\{\alpha,\beta,\gamma\}$ assigned to each vertex $\{a,b,c\}$ of G by

$$\omega_{\mathrm{WP}} = -2\sum (d\ln\alpha \wedge d\ln\beta + d\ln\beta \wedge d\ln\gamma + d\ln\gamma \wedge d\ln\alpha) \ , \tag{4.33}$$

where the sum runs over all vertices of G. This formula results from Wolpert's explicit expression for the Weil-Petersson Kähler form as a function of geodesic lengths.

For the mathematical literature, we refer the reader to Epstein and Penner (1988), and Penner (1987a, 1987b). The original proof of Eq. (4.31) is in Harer and Zagier (1986). A recent survey is that of Harer (1982, 1985).

To conclude, we remark that the Penner decomposition method has been used in string theory applications in only one instance so far, even though it allows for some of the most explicit and calculable formulations of the integration measure on moduli space (see Gross and Periwal, 1988).

G. Mandelstam diagrams

The formulation of string theory in the light-cone gauge by Goddard, Goldstone, Rebbi, and Thorn (1973) has led Mandelstam (1973a, 1973b, 1974a, 1974b) to introduce the *interacting string picture.* In this picture, freely moving strings propagate as cylinders and split and join at definite light-cone interaction times τ_a, and the interaction vertex (at least for the bosonic string) is just the overlap integral between initial and final strings. In the light-cone picture, the radius of an intermediate cylinder corresponds to the p^+ component of the momentum of that string, so that the sum of all radii remains conserved as a function of light-cone time. We shall, however, abstract the diagram from the momentum conservation issue and simply keep the same geometry. Such diagrams will be referred to as *Mandelstam diagrams,* and a typical example is presented in Fig. 17.

As a Riemann surface, the number of internal slits corresponds to the genus h, whereas the number of cylinders running off the ∞ corresponds to the number of punctures n. Clearly, then, the Mandelstam diagram can only describe surfaces with at least two punctures. In view of the results in Sec. II.L, however, the Polyakov integrals may be reformulated on surfaces with punctures.

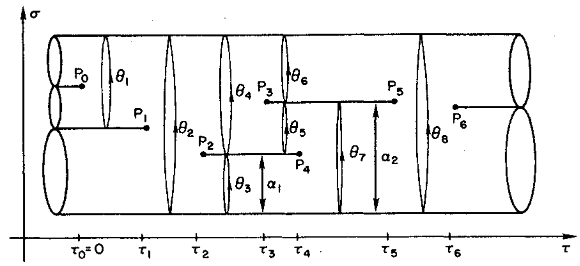


FIG. 17. A Mandelstam diagram for a surface with $n=5$ punctures, genus $h=2$, with the corresponding coordinates.

The coordinates labeling the diagram for genus h and n punctures are (see again Fig. 17)

$$\begin{aligned} &\tau_a, \quad a=1,2,\ldots,2h+n-3, \quad \text{interaction times} \ , \\ &\theta_\alpha, \quad \alpha=1,2,\ldots,3h+n-3, \quad \text{twist angles} \ , \\ &\alpha_I, \quad I=1,2,\ldots,h, \quad \text{internal } \textit{momenta} \ , \\ &\alpha_p, \quad p=1,2,\ldots,n, \quad \text{external } \textit{momenta} \ . \end{aligned} \tag{4.34}$$

The remarkable fact, implicit in the work of Mandelstam (1986a, 1986b) and proven by Giddings and Wolpert (1987), is that natural ranges for τ_a, θ_α, and α_I, as well as fixed α_p's, provide a single cover for the moduli space $\mathcal{M}_{h,n}$ of surfaces of genus h and with n punctures, with *residues* α_p prescribed at the punctures Q_i. The proof proceeds along the following lines.[28]

The quintessential property of a Mandelstam diagram (as is of course the case for the light-cone gauge in general) is that it admits a globally defined light-cone time τ, and so the differential $d\tau$ is exact. The σ direction, on the other hand, is not global, first because on each cylinder σ is not single valued, and second because there is a separate σ variable for each cylinder. However, the differential $d\sigma$ is well defined on each cylinder separately and is single valued there. It is thus natural to introduce the *light cone coordinate* $w=\tau+i\sigma$, whose differential is well defined,

$$\omega = dw = d\tau + i\, d\sigma \ . \tag{4.35}$$

Of course, w does not define a smooth coordinate in the neighborhood of an interaction point w_0. Smooth holomorphic coordinates z may be introduced in the neighborhood of w_0 by mapping the region into a planar region

$$w - w_0 = (z-z_0)^2 \ . \tag{4.36}$$

[28]Arguments along somewhat different lines were presented by Taylor (1987).

Actually, if N interaction times coincide,[29] one will have to introduce z through

$$w-w_0=(z-z_0)^N .$$

In any case, it is clear that ω vanishes at the interaction points and is nonvanishing everywhere else on the surface. Since the radii of the outgoing cylinders are α_p, we also have

$$\frac{1}{2\pi i}\oint_{\text{around } Q_p}\omega=\alpha_p, \quad \sum_{p=1}^{n}\alpha_p=0 , \tag{4.37}$$

so we may view ω as having a simple pole at the puncture Q_p with residue α_p.

Since ω is a (1,0) form, $\bar{\omega}\omega$ is a metric on the surface. It is flat Euclidean everywhere, except at the interaction points, where all the curvature of the diagram is concentrated:

$$\sqrt{g}R(z)=-2\pi\sum_{a=0}^{2h-3+n}\delta(z-P_a) , \tag{4.38}$$

where $g_{z\bar{z}}=\omega_z\bar{\omega}_z$.

Now notice that ω has purely imaginary periods around any homology cycle, since τ is single valued. Thus for every Mandelstam diagram there exists a unique holomorphic Abelian differential ω, with purely imaginary periods, and with residues α_p at the puncture Q_p. Conversely, every such holomorphic differntial ω determines a metric $\bar{\omega}\omega$ and hence a unique Riemann surface.

For a general Riemann surface (not necessarily viewed in the Mandelstam picture) there exists an analogous holomorphic differential. It is constructed out of the meromorphic Abelian differentials of the third kind $\omega_{PQ}(z)$ with simple poles at P and Q of residues 1 and -1, respectively. Such a differential is defined only up to holomorphic differentials ω_I, but a unique $\omega_{PQ}(z)$ emerges if one demands that all its periods be purely imaginary:

$$\mathrm{Re}\oint_{A_I}\omega_{PQ}=\mathrm{Re}\oint_{B_I}\omega_{PQ}=0 . \tag{4.39}$$

Then ω is given by

$$\omega(z)=\sum_{p=1}^{n-1}\beta_p\omega_{Q_pQ_{p+1}}(z) , \tag{4.40}$$

with the coefficients β_p expressed in terms of the residues α_p:

$$\beta_p-\beta_{p-1}=\alpha_p, \quad p=1,\ldots,n$$

and

$$\beta_0=\beta_n=0 .$$

As a differential on the surface with punctures $Q_1,\ldots,Q_n$, ω is of course holomorphic, and by construction all its periods are imaginary.

Thus on every Riemann surface there exists a unique holomorphic differential ω with residues α_p at the punctures Q_p and with purely imaginary periods. Conversely, such a differential defines a metric and hence specifies the Riemann surface uniquely.

The proof is completed by setting the two unique differentials ω equal to one another.

Actually, this proof also informs us immediately about the natural range of parameters mentioned above. The differential ω will be completely specified once the "geometry" of the Mandelstam diagram is given. But there are redundancies in the parametrization of the geometry which yield the same geometrical configuration of the diagram. Clearly, these are the only restrictions on the range of the parameters.

Another remarkable property of the Mandelstam diagram representation is that quadratic and $\frac{3}{2}$ holomorphic differentials admit an explicit representation in terms of the canonical differential ω and h holomorphic first-Abelian differentials ω_I, $I=1,\ldots,h$.

- Holomorphic quadratic differentials are given by

$$\begin{aligned}&\phi_I=\omega\omega_I, \quad I=1,\ldots,h ,\\ &\phi_a=\omega\omega_{P_0P_a}, \quad a=1,\ldots,2h+n-3 ,\end{aligned} \tag{4.41}$$

where $P_0,P_1,\ldots,P_{2h+n-3}$ are the interaction points of the diagram. The poles of the meromorphic (third) Abelian differentials ω_{PQ} are precisely canceled by the zeros of ω, so that ϕ_a and ϕ_I are holomorphic on the n-punctured surface.

- Holomorphic $\frac{3}{2}$ differentials for even-spin structure ν and a generic point on $\mathcal{M}_{h,n}$ are constructed as follows. There are no holomorphic $\frac{1}{2}$ differentials on the underlying compact surface (the analogs of ω_I), and there is a unique meromorphic $\frac{1}{2}$ differential [the so-called Szegö kernel of Eq. (3.202)] $\kappa_P(z)=S_\nu(z,P)$ with a single simple pole at P. We obtain the $2h+n-3$ holomorphic $\frac{3}{2}$ differentials as

$$\rho_a=\omega\kappa_{P_a}, \quad a=0,1,\ldots,2h+n-3 , \tag{4.42}$$

where again the pole of κ is canceled by the zeros of ω.

- Holomorphic $\frac{3}{2}$ differentials for odd-spin structure ν and a generic point in $\mathcal{M}_{h,n}$ are obtained as follows. There is now a unique holomorphic $\frac{1}{2}$ differential h_ν on the underlying compact surface (the analog of ω_I), and there is a unique meromorphic $\frac{1}{2}$ differential $\kappa_{PQ}(z)=S_\nu(z;P,Q)$, with simple poles of opposite residue $1,-1$ at P and Q (the analog of ω_{PQ}). We again obtain $2h+n-3$ holomorphic $\frac{3}{2}$ differentials as

$$\begin{aligned}&\rho_a=\omega\kappa_{P_0P_a}, \quad a=1,\ldots,2h+n-3 ,\\ &\rho_{2h+n-3}=\omega h_\nu .\end{aligned} \tag{4.43}$$

In Sec. V.G, we shall make use of these constructions to exhibit certain simple relations between ghost and matter determinants on Mandelstam diagrams.

[29]Only τ's coincide here; the interaction points may well be distinct.

H. Universal Teichmüller curve and compactification of moduli space

1. Teichmüller curve

We conclude this section by discussing the geometry of the (universal) Teichmüller curve. This is the fiber bundle over moduli space whose fiber above a given point in moduli space (a given complex structure m) is just the Riemann surface with this complex structure m. (See Fig. 18.) Its interest to us stems from the close connection between its curvature and the curvature of determinant line bundles over moduli space (Sec. VII) and, even more importantly, from the fact that it provides the proper setting for certain gauge-fixing procedures in the superstring.

The formal construction of the universal Teichmüller curve is the following. Let $\mathcal{M}$ be the space of all metrics g on a fixed topological surface M and consider the fibration

$$\begin{array}{ccc} (\xi,g) & \in & M\times\mathcal{M} \\ \downarrow & & \downarrow \\ g & \in & \mathcal{M} \end{array} \quad . \tag{4.44}$$

In the product space $M\times\mathcal{M}$ we shall view the pairs (ξ,g) and (ξ',g') as equivalent if there is a reparametrization of M sending simultaneously $\xi\to\xi'$ and $g\to g'$, or if ξ equals ξ' and g and g' differ only by a Weyl scaling. Denoting this equivalence by $\mathrm{Diff}(M)\times\mathrm{Weyl}(M)$, we can now define the universal Teichmüller curve as the fiber bundle

$$\begin{array}{ccc} \mathcal{C}_h & = & (M\times\mathcal{M})/\mathrm{Diff}(M)\times\mathrm{Weyl}(M) \\ \downarrow & & \downarrow \\ \mathcal{M}_h & = & \mathcal{M}/\mathrm{Diff}(M)\times\mathrm{Weyl}(M) \end{array} \quad . \tag{4.45}$$

We note that the original bundle (4.44) is trivial, but it follows, for example, from curvature computations (4.47) below that bundle (4.45) is not. In fact, the universal curve does not admit any global continuous sections.[30] Nevertheless, local sections exist and are important: for example, in a basis of $\frac{3}{2}$ differentials $\mu_a(z)=\delta(z-z_a)$, the points z_a should be reviewed as local sections of the Teichmüller curve. Another useful property that emerges out of the construction (4.44) and (4.45) is that vector fields on moduli space have a lifting to vector fields on the Teichmüller curve: the natural lift $g\to(\xi,g)$ of Eq. (4.44) is invariant under $\mathrm{Diff}(M)\times\mathrm{Weyl}(M)$, and hence makes sense as a lifting from $\mathcal{M}_h$ to $\mathcal{C}_h$. Liftings are needed to have a proper notion of derivatives of points of the surface M with respect to moduli parameters.

In the above construction the holomorphic structure of the Teichmüller curve is not manifest. However, if we represent a Riemann surface by a domain D_0 in the upper half plane $\mathbf{H}$ and extend the Beltrami differentials $\varepsilon\mu$ to be zero in the lower half plane, then D_0 will be deformed "holomorphically" to another domain $D_{\varepsilon\mu}=w(D_0)$ by quasiconformal solutions w of the Beltrami equation (4.16). As μ vary over a small neighborhood of 0 in $\mathbf{C}^{3h-3}$, this provides an embedding of a small neighborhood of the fiber corresponding to D_0 in the Teichmüller curve into $\mathbf{C}^{3h-3}\times\mathbf{C}$. We have just presented a very rough description of the Bers embedding, which endows the Teichmüller curve with a holomorphic structure. It is instructive to realize the lifting discussed in the preceding paragraph in the Bers embedding. If μ is a tangent vector to moduli space at $M=\mathbf{H}/\Gamma$, let $w_{\varepsilon\mu}$ again be the solution to Eq. (4.16), which fixes 0, 1, and ∞. The vector field

$$\left[\frac{d}{d\varepsilon}w_{\varepsilon\mu}\bigg|_{\varepsilon=0}\right]\frac{\partial}{\partial z}$$

is defined on the upper half plane, but cannot be viewed as a vector field on the surface M, since the arbitrary choices $0,1,\infty$ prevent it from transforming equivariantly under the group Γ. On the other hand, at each point z of M the vector field on the universal Teichmüller curve

$$\tau_\mu=\left[\frac{d}{d\varepsilon}w_{\varepsilon\mu}\bigg|_{\varepsilon=0}\right]\frac{\partial}{\partial z}+\mu=\dot{w}_\mu\frac{\partial}{\partial z}+\mu \tag{4.46}$$

will be equivariant and hence well defined. The correspondence $\mu\to\tau_\mu$ is the lifting we are looking for. This means that in the Bers realization, as we deform the complex structure along μ, the fundamental domain will be deformed as well. Each point in the fundamental domain describes, then, a path in $\mathbf{C}^{3h-3}\times\mathbf{C}$, whose direction is the vector field τ_μ of Eq. (4.46). We observe that τ_μ is a smooth vector field on the universal curve, while the related quasiconformal vector field V^z of Sec. II.I is a vector field on the surface M which cannot be smooth if μ is a nontrivial deformation.

It is evident that any choice of metrics on Riemann surfaces can be viewed as a metric on the vertical bundle above the Teichmüller curve, i.e., the bundle of tangent vectors to the fibers of the Teichmüller curve. In general

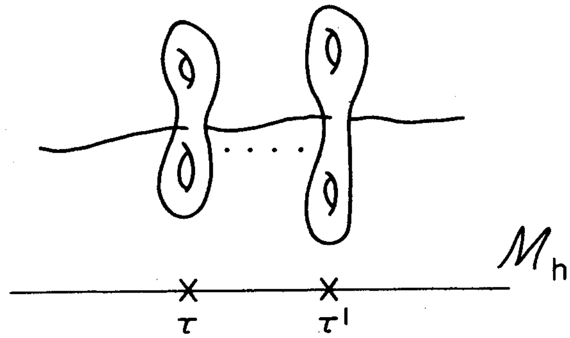


FIG. 18. The Teichmüller curve and its sections for the moduli space of surfaces of genus h.

[30]This follows from results of Johnson (1980), as pointed out by E. Miller.

a metric and a holomorphic structure will then determine a unique connection by the requirements of unitarity, hermiticity, and that its (0,1) component agree with the $\bar{\partial}$ operator (cf. Sec. VI.A). If we choose constant-curvature metrics on the surface M, the curvature of the corresponding connection on the Teichmüller curve can be computed explicitly. Since the curvature can be viewed as a Hermitian 2-form, it can be described by its values on pairs of vielbein vectors for the Teichmüller curve. Since a vielbein for moduli space is provided by a basis of Beltrami differentials μ_j, $j=1,\ldots,3h-3$, which lifts to vectors

$$\tau_j=\dot{w}_{\mu_j}\frac{\partial}{\partial z}+\frac{\partial}{\partial t_j}$$

on the universal Teichmüller curve, a vielbein for the latter will consist of $\{\tau_j\}$ and $\partial/\partial z$, this last vector field being viewed as a vector field along each fiber. The curvature Ω is now given by

$$\Omega\left[\frac{\partial}{\partial\bar{z}},\frac{\partial}{\partial z}\right]=-\frac{2}{(z-\bar{z})^2},\quad \Omega\left[\frac{\partial}{\partial\bar{z}},\tau_j\right]=0\ ,$$
$$\Omega(\bar{\tau}_j,\tau_k)=2(\Delta+2)^{-1}(\bar{\tau}_j\cdot\tau_k)\ . \tag{4.47}$$

Here Δ is the Laplacian on scalars. From this it is easy to deduce the higher powers of the curvature tensor and the characteristic classes

$$\kappa_n=\int_{\text{fiber}}\left[\frac{i\Omega}{2\pi}\right]^{n+1} \tag{4.48}$$

of the Teichmüller curve. In particular, one readily finds the first Chern class of the vertical line bundle

$$\kappa_1=-\frac{1}{4\pi^2}\int_{\text{fiber}}\Omega^2=\frac{1}{2\pi^2}\omega_{\text{WP}}\ , \tag{4.49}$$

where ω_{WP} is the Weil-Petersson Kähler form encountered in Eq. (4.17). This result will help us later (cf. Sec. VII.E) to identify the precise form of the holomorphic anomaly from the determinant line bundles formalism.

The Teichmüller curve is the natural setting for a careful treatment of derivatives of differentials with respect to moduli parameters. Indeed such differentials can be viewed as sections of tensor powers of the vertical line bundle over the Teichmüller curve. If metrics are chosen to represent conformal classes (as is usually done in gauge-fixing the superstring), these bundles will be endowed with a metric as well, and hence with a connection in the presence of a holomorphic structure. If we wanted to differentiate in the moduli direction, μ, it would suffice to again lift μ to a vector field along the fiber of the Teichmüller curve by the natural lift, and differentiate along that vector field using the connection we just discussed. This procedure can be applied, for example, to the $\frac{3}{2}$ differentials χ_a needed to absorb the superconformal ghost zero modes.

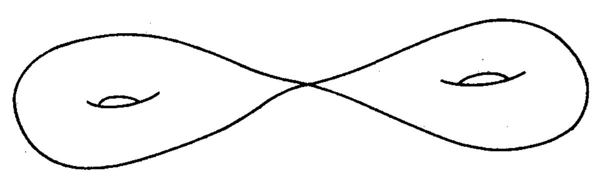
We also note that the Teichmüller curve can be viewed as the moduli space of surfaces with one puncture, already encountered in Secs. II.L and IV.F. Formulating string amplitudes as integrals over the Teichmüller curve and its generalizations is quite convenient in many respects.

2. Compactification of moduli space

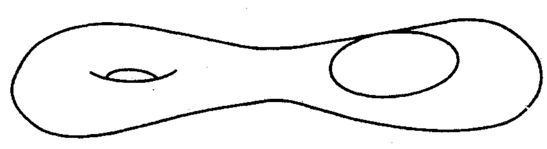
Finally, we come to the issue of the boundary of moduli space, of which mention has been made in connection with BRST invariance (Secs. III.J–III.P), and whose role will emerge more clearly in connection with finiteness of string amplitudes and supersymmetry breaking.

That moduli space $\mathcal{M}_h$ is not a space without boundary is not evident from the definition we adopted in Eq. (2.33). However, for genus $h=1$, we have an explicit representation of $\mathcal{M}_1$ as a fundamental domain for SL(2,**Z**) within the upper half space, which admits a natural one-point compactification. For higher genus $h\geq 2$, the Fenchel-Nielsen coordinates of Sec. IV.E for Teichmüller space provide a simple explanation: the boundary of Teichmüller space consists of the surfaces where one of the $3h-3$ geodesics has been pinched to a point (see Fig. 19). This is the basic geometric principle underlying the Deligne-Mumford compactification of moduli space, where one adjoins to the regular Riemann surfaces the divisor Δ of Riemann surfaces with nodes. A Riemann surface M_0 with nodes is a surface with special points p_i called nodes, around each of which the surface is conformally equivalent to two discs with their centers identified. If the coordinate of the node is 0, such a neighborhood of the node can be given by

$$\mathcal{U}_0=\{(z,w);\ zw=0,\ |z|,|w|<1\}\ . \tag{4.50}$$



(a)



(b)

FIG. 19. The boundary of moduli space—surfaces with nodes: (a) pinching a cycle homologous to zero: (b) pinching a cycle not homologous to zero.

This neighborhood can be viewed as the end product of a degeneration of regular holomorphic neighborhoods indexed by a parameter t, $|t|<1$ tending to 0:

$$\mathcal{U}_t=\{(z,w);\ zw=t,\ |t|<|z|<1,\ \ |t|<|w|<1\}\ . \tag{4.51}$$

For each fixed t, $\mathcal{U}_t$ can be viewed as an annulus, or equivalently a cylinder. The principle of the "plumbing fixture" is that the degeneracy family $\mathcal{U}_t$ of cylinders can be fitted in a family of regular Riemann surfaces M_t, which tend to M_0 as $t\to 0$. More precisely, let M_1 and M_2 be two regular Riemann surfaces of genus i and $h-i$ and p_1 and p_2 be points on M_1 and M_2, respectively. If z_i are holomorphic coordinates around p_i and D_i are the discs $\{|z_i|<1\}$, we can remove the smaller discs $\{|z_i|\le|t|^{1/2}\}$ and join the remaining points of M_1 and M_2 by attaching them both to the plumbing fixture $\mathcal{U}_t$ in the following manner:

z_1 in D_1 is identified with $(z=z_1,\ w=t/z_1)$ in $\mathcal{U}_t$ if $|t|^{1/2}<|z_1|<1$;

z_2 in D_2 is identified with $(z=t/z_2,\ w=z_2)$ in $\mathcal{U}_t$ if $|t|^{1/2}<|z_2|<1$.

This gives us a family of regular Riemann surfaces M_t of genus h which tend to a Riemann surface M_0 with node $p=p_1,p_2$. This type of degeneration corresponds to pinching to a point a cycle homologous to zero [Fig. 19(a)]. Holomorphic parameters for moduli space near such a moduli boundary point M_0 are the moduli parameters for M_1 and M_2, the points p_1 and p_2, and the parameter t characterizing the annulus $\mathcal{U}_i$. Of these, t should viewed as the parameter transversal to the boundary of moduli space, while the others parametrize the boundary itself. If we choose instead to pinch a cycle that is not homologous to zero [Fig. 19(b)], then we can repeat essentially the same plumbing fixture construction, starting this time with a regular Riemann surface M of genus $h-1$ with two marked points p_1 and p_2. Again coordinate discs $D_i=\{|z_i|<1\}$ around p_i are introduced, smaller discs $\{|z_i|<|t|^{1/2}\}$ are removed, and the above construction yields a "bridge" between the remaining parts of D_i. The resulting surfaces M_t now have genus h and tend to a surface M_0 with node $p=p_1,p_2$. Holomorphic coordinates for moduli space are the moduli parameters for M, the points p_1 and p_2, and the parameter t, which is again the transversal coordinate. The two types of degenerations can equivalently be distinguished by whether removal of the node at the end disconnects the surface or not. In either case, the plumbing fixture construction shows explicitly that the behavior of M_t outside of $\mathcal{U}_t$ is independent of the degeneration taking place within $\mathcal{U}_t$.

Viewing Δ as arising from regular Riemann surfaces by pinching closed curves to a point, it should be evident that there are nevertheless restrictions as to which and how many curves can be pinched simultaneously. For $h\ge 2$, in the Fenchel-Nielsen picture, it is any number of $3h-3$ defining geodesics. More formally, one requires that the surfaces with nodes have at most as many conformal Killing vectors as the regular surfaces that converge to them. The compactification $\overline{\mathcal{M}}_h$ of $\mathcal{M}_h$ obtained in this manner is called the moduli space of stable curves.

The compactification locus $\Delta=\overline{\mathcal{M}}_h-\mathcal{M}_h$ is a divisor with normal crossings. It is reducible and can be written as

$$\Delta=\Delta_0\cup\cdots\cup\Delta_{[h/2]}\ , \tag{4.52}$$

where the generic surface M in Δ_i has exactly one node separating it into two components of genus i and $h-i$, while Δ_0 consists of degenerations that do not disconnect the surface. The divisors Δ_k define cohomology classes in $H_{6h-8}(\overline{\mathcal{M}}_h)$. If $[\varpi_{\rm WP}]$ is the cohomology class obtained by Poincaré duality from the Weil-Petersson Kähler form, it is a remarkable theorem of Harer (1985) and Wolpert (1987) that $(\Delta_0,\ldots,\Delta_{[h/2]},[\varpi_{\rm WP}])$ is actually a basis for $H_{6h-8}(\overline{\mathcal{M}}_h)$.

A last fundamental feature of degenerations is that the universal Teichmüller curve extends to a holomorphic fibration over the compactified moduli space $\overline{\mathcal{M}}_h$ if the fibers above the locus Δ are the corresponding surfaces with nodes. That the total space of the fibration has no singularities (by opposition to the fiber) can be intuitively seen from the degeneration picture provided by Eq. (4.51): the total space there can be viewed as the perfectly regular two-dimensional complex variety

$$\{(z,w,t);\ zw=t,\ \ |z|,|w|,|t|<1\}\ , \tag{4.53}$$

whose projection by $(z,w,t)\to t$ just ceases to be a submersion at $t=0$. The compactified universal Teichmüller curve has been used to investigate the asymptotic behavior near the boundary of moduli space of the string integrand. It is likely that its potential applications to string theory have not been exhausted.

Differential geometric constructions of the universal Teichmüller curve seem to have started with Earle and Eells (1969). The Bers embedding is described in Bers (1973). The curvature of the (uncompactified) universal curve given in Eq. (4.47) is due to Wolpert (1986). A more recent treatment extending to the curvature of surfaces with nodes over the Diligne-Mumford compactification (see Sec. IV.H) is that of Wolpert (1988). The role of the universal Teichmüller curve in Grothendieck-Riemann-Roch constructions is explained in detail by Nelson (1987a). Although we do not need them here, it may be worth reporting that the curvature of moduli space with respect to the Weil-Petersson metric is completely known. Different methods of calculation are given in Royden (1985), Siu (1986), Tromba (1986), Wolpert (1986), and Wolf (1986).

V. EVALUATION OF DETERMINANTS

Determinants of Laplacians and $\bar\partial$ operators for the torus can be expressed in terms of the Dedekind eta function $\eta(\tau)$ and special values of the theta function $\vartheta(z,\tau)$. For higher loops, this can be generalized in a number of ways. If we choose to represent the conformal class by a constant-curvature metric, then the natural function is the Selberg zeta function $Z(s)$. We shall show in this section that all the determinants of Laplacians needed for quantization are given by special values of $Z(s)$. This will allow a simple analysis of the asymptotic behavior of the string integrand near the boundary of moduli space, confirming the presence of a tachyon in the bosonic string spectrum, and clarifying the respective roles of worldsheet and space-time supersymmetry in finiteness issues. It is difficult to extract the determinants of chiral $\bar\partial$ operators in this approach, if only because hyperbolic geometry and Selberg zeta functions are defined by real quantities. Actually, even an appropriate definition of a chiral determinant is problematic. The proper resolution of these issues requires a study of holomorphic anomalies and bosonization, and will lead to expressions for chiral determinants in terms of Riemann theta functions. A full account will be provided in Sec. VII.

Mandelstam diagrams are another convenient way of representing (punctured) Riemann surfaces. Although we shall not evaluate the determinants for Mandelstam diagrams individually, we shall show that remarkable relations hold between determinants of different U(1) weights. Such relations are usually required to relate the light-cone gauge to the Polyakov formulation.

A. One-loop formulas

We begin with the simpler case of one loop, which will serve as an introduction to the more complicated formulas required for multiloops. Recall that a complex structure for a torus is characterized by a lattice $\mathbf{Z}+\tau\mathbf{Z}$ in the complex plane, and that moduli space $\mathcal{M}_1$ is $\mathbf{H}/PSL(2,\mathbf{Z})$, with $\mathbf{H}=\{\tau=\tau_1+i\tau_2;\tau_2>0\}$. The key forms on moduli space are the Dedekind eta and the theta functions defined in Appendix E, which transform according to the Jacobi rule of Eqs. (E4) and (E10). For fermion determinants, a spin structure has to be prescribed. There are four spin structures $\nu=(\nu_1,\nu_2)$, $\nu_{1,2}=0,1$ corresponding to the boundary conditions

$$\begin{aligned}\varphi(\xi_1+1,\xi_2)&=-\varphi(\xi_1,\xi_2)e^{\pi i\nu_1}\,,\\ \varphi(\xi_1+\tau_1,\xi_2+\tau_2)&=-\varphi(\xi_1,\xi_2)e^{\pi i\nu_2}\,.\end{aligned}\tag{5.1}$$

The relevant operators become

$$\begin{aligned}&\Delta=-2\partial\bar\partial,\quad P_1^\dagger P_1=2\Delta\,,\\ &\not{D}=\begin{bmatrix}0&\bar\partial\\-\partial&0\end{bmatrix},\quad P_{1/2}=\not{D}\,.\end{aligned}\tag{5.2}$$

Introducing the basis

$$\varphi_{n_1,n_2}=\exp\left[2\pi i\left[(n_1+\tfrac12-\tfrac12\nu_1)\xi^1+\frac{1}{\tau_2}[n_2-(n_1+\tfrac12-\tfrac12\nu_1)\tau_1+\tfrac12-\tfrac12\nu_2]\xi^2\right]\right],$$

we find the eigenvalues of $\bar\partial$,

$$\lambda_{n_1n_2}=\frac{2\pi}{\tau_2}[(n_1+\tfrac12-\tfrac12\nu_1)\tau-(n_2+\tfrac12-\tfrac12\nu_2)]\,,$$

and zeta-function regularization[31] produces

$$\begin{aligned}&\ln\det{}_\nu\partial\bar\partial=-\zeta'(0)+2\ln\left[\frac{2\pi}{\tau_2}\right]\zeta(0)\,,\\ &\zeta(s)=\sum_{n_1,n_2}\{(n_1+\tfrac12-\tfrac12\nu_1)^2\tau_2^2\\ &\qquad+[(n_1+\tfrac12-\tfrac12\nu_1)\tau_1-n_2-\tfrac12+\tfrac12\nu_2]^2\}^{-s}\,,\end{aligned}\tag{5.3}$$

which is absolutely convergent for $\mathrm{Re}(s)>1$. When $\nu=(1,1)$, it is understood that the summation does not include $n_1=n_2=0$. The n_2 sum may be represented by a contour integral

$$\zeta(s)=\sum_{n_1}\oint_C dz\frac{1}{1-e^{-2\pi iz}}\{(n_1+\tfrac12-\tfrac12\nu_1)^2\tau_2^2+[(n_1+\tfrac12-\tfrac12\nu_1)\tau_1-z-\tfrac12+\tfrac12\nu_2]^2\}^{-s}\,,\tag{5.4}$$

where the contour surrounds the real axis once in the counterclockwise direction. The contour may be restricted to the line infinitesimally above the real axis, the other contribution being related to the complex conjugate. In turn, this contour can be deformed into an integration along both sides of a vertical cut in the upper half plane, starting at $(n_1+\frac12-\frac12\nu_1)\tau+\frac12\nu_2-\frac12$. When $\nu\neq(1,1)$ it is straightforward to see that $\zeta(0)=0$, whereas for $\nu=(1,1)$ we have $\zeta(0)=-1$. Furthermore

$$\zeta'(0)=\sum_{n_1}2\ln\{\exp[2\pi i(n_1+\tfrac12-\tfrac12\nu_1)\tau+\pi i(\nu_2-1)]-1\}+\text{c.c.}\,,$$

with the appropriate subtraction when $\nu=(1,1)$. Using the product representation (E6) from Appendix E for the theta function, and (E10) for the eta function, we find

$$\begin{aligned}&\det{}_\nu(-\bar\partial\partial)=\left|\frac{\vartheta_{\nu_1\nu_2}(0,\tau)}{\eta(\tau)}\right|^2,\quad\nu\neq(1,1)\,,\\ &\det{}'_\nu(-\bar\partial\partial)=\tau_2^2\,|\eta(\tau)|^4\,,\quad\nu=(1,1)\,,\end{aligned}$$

where the τ_2 factor above comes from the term proportional to $\zeta(0)$, which contributes only when there are zero modes. We observe that since the left-hand side is reparametrization invariant by construction, this result can be used to derive the usual transformation law for the η function under modular transformations. Separating out the holomorphic factors where appropriate gives

$$\begin{aligned}&\det{}'\Delta=\tau_2^2\,|\eta(\tau)|^4\,,\\ &\det{}'_\nu\bar\partial=\tau_2/2\pi\vartheta_1'(0,\tau)/\eta(\tau),\quad\nu=(1,1)\,,\\ &\det{}_\nu\bar\partial=\vartheta_{\nu_1\nu_2}(0,\tau)/\eta(\tau),\quad\nu\neq(1,1)\,,\end{aligned}\tag{5.5}$$

[31]See the opening remark of Sec. V.E.

The contributions of left-movers can be identified with those of $\bar{\partial}$ operators. The above formulas actually determine the functional determinants only up to global phases independent of the moduli parameter τ. Recall that spin structures can be divided into two groups, characterized by the parity of the number of zero modes of the Dirac operator. The mapping class group will permute all of them and hence determine the phases within each group. The relative phases of the two groups should be determined by factorization requirements. For one loop, the odd-spin structure $\nu=(1,1)$ does not contribute to the partition function because the zero mode of the Dirac operator decouples from the supermoduli modes, and modular invariance forces a combination of phases for the remaining three even-spin structures which produces 0 by the Jacobi identity. This vanishing of the cosmological constant can be viewed as a consequence of space-time supersymmetry (for details see Sec. III.M).

One-loop determinants for the bosonic string were evaluated by Polchinski (1986) and for the fermionic string by D'Hoker and Phong (1986b) and Namazie, Narain, and Sarmidi (1986).

B. Multiloop formulas and Selberg zeta functions

In this case a complex structure on the worldsheet M can be represented by a metric of constant negative curvature -1, and an analog of the theta function is provided by the Selberg zeta functions

$$Z(s)=\prod_{\gamma \text{ primitive}}\prod_{k=1}^{\infty}[1-\nu(\gamma)^n e^{-(s+k)l_\gamma}]\ ,\qquad n=0,1\ . \tag{5.6}$$

Here the primitive γ denotes simple closed geodesics on M, l_γ is the length of γ in the hyperbolic metric, and $\nu(\gamma)\in\{\pm 1\}$ is determined by the spin structure. In more algebraic terms, we view the worldsheet as $\mathbf{H}/\Gamma$, with Γ a Fuchsian group with compact quotient. A primitive γ is then an element of Γ that cannot be written as a power ≥ 2 of any element, l_γ is equal to $\cosh^{-1}(\mathrm{tr}\gamma/2)$, and $\nu(\gamma)$ is a $\mathbf{Z}_2$-valued character of the group $\tilde{\Gamma}\subset \mathrm{SL}(2,\mathbf{R})$ which projects onto $\Gamma\subset\mathrm{PSL}(2,\mathbf{R})$. If we recall that the complex parameter τ for the torus is just $\varphi+i/l$ in Fenchel-Nielsen coordinates, there is evidently a close similarity between Eqs. (5.10) and (5.6), with the difference, however, that $Z(s)$ is real and that there are many more geodesics on a hyperbolic surface. Although we use the same symbol $Z(s)$ for the various Selberg zeta functions, it should be clear that the definition (5.6) with $n=0$ is to be taken when dealing with bosons, while $n=1$ corresponds to fermions.

The function $Z(s)$ will converge for $\mathrm{Re}s>1$, admit a functional equation similar to the Riemann zeta function,

$$Z(1-s)=x(s)Z(s)\ ,$$
$$x(s)=\exp\left[4\pi(h-1)\int_0^{s-1/2}dy\, y\, tg\,\pi y\right]\ ,$$

and extend to an entire function in the s complex plane. In terms of $Z(s)$ the functional determinants appearing in the quantum superstring measure were evaluated by D'Hoker and Phong (1986d), Fried (1986b), and Sarnak (1987);

$$\begin{aligned}&\det\Delta_g=e^{-c_0\chi}Z'(1)\ ,\\ &\det P_1^\dagger P_1=e^{-c_1\chi}Z(2)\ ,\\ &\det P_{1/2}^\dagger P_{1/2}=e^{-c_{1/2}\chi}Z(\tfrac{3}{2})\ ,\\ &\det \not{D}\not{D}=e^{-c_{-1/2}\chi}Z^{(2N)}(\tfrac{1}{2})/(2N)!\ ,\end{aligned} \tag{5.7}$$

with

$$c_n=\sum_{0\leq m<n-1/2}(2n-2m-1)\ln(2n-m)-(n+\tfrac{1}{2})^2+2(n-[n])(n+\tfrac{1}{2})\ln 2\pi+2\zeta'(-1)\ .$$

Here N is the number of zero modes of the chiral Dirac operator. In general it depends on the spin structure as well as on the complex structure. This will be discussed at length in Sec. VI.F.

The Selberg zeta function was introduced by Selberg (1956) and appeared in evaluation of determinants in the work of Ray and Singer (1971, 1973). A good review of its properties can be found in Hejhal (1976a, 1976b). Special cases of Eq. (5.7) have been obtained by Baranov and Schwartz (1985), D'Hoker and Phong (1986a), Fried (1986a), Gilbert (1986), Kierlanczyk (1986), and Namazie and Rajeev (1986). Various relations between Selberg zeta functions were discussed by Beilinson and Manin (1986) and by Voros (1987). The case of worldsheets with boundary is treated in Blau and Clements (1987). Superdeterminants on super Riemann surfaces of constant supercurvature were related to super Selberg zeta functions by Aoki (1988) and by Baranov, Manin, Frolov, and Schwartz (1987). Character-valued generalizations of the Selberg function in which $\nu(\gamma)$ is the character of an Abelian group have been related to determinants on Riemann surfaces with an Abelian orbifold as target space-time by Periwal (1987). Finally, Mandelstam (1986a, 1986b) also used Selberg trace formula techniques to relate determinants to the partition function of the old dual models. The superstring case is discussed by Martinec (1987).

C. Hyperbolic geometry on a Riemann surface

In the remaining subsections of Sec. V we shall discuss the mathematical ingredients necessary to an understanding of Eq. (5.7). Some fundamental facts about Fuchsian groups $\Gamma\subset\mathrm{PSL}(2,\mathbf{R})$ with compact quotient are the following. Elements γ of Γ all have traces $|\mathrm{Tr}\gamma|$ greater than 2 and are conjugate within $\mathrm{SL}(2,\mathbf{R})$ to a dilation $z\mapsto e^{l_\gamma}z$. The dilation coefficient follows from the hyperbolic distance $d(z,z')$ defined in Eq. (4.13):

$$\min_z d(z,\gamma z)=l_\gamma\ . \tag{5.8}$$

The number of simple closed geodesics of length smaller than l is asymptotically given by $l^{-1}e^{l}[1+O(1)]$ as $l\to\infty$. For a fixed Γ, the set of lengths is bounded from below by some smallest length $l_0>0$.

Next a tensor $f(z)dz^n d\bar z^m$ on $H=\mathbf{H}/\Gamma$ may be identified with a function $f(z)$ on $\mathbf{H}$ transforming under Γ as

$$f(\gamma z)=(cz+d)^{2n}(c\bar z+d)^{2m}f(z) \quad \text{for } \gamma=\begin{bmatrix} a & b\\ c & d\end{bmatrix}\in \mathrm{PSL}(2,\mathbf{R}) . \tag{5.9}$$

There is no ambiguity for $n+m$ integer, but for n half-integer (which corresponds to spinor fields) the sign of the trace of γ matters, and we must introduce a multiplier $\nu(\tilde\gamma)\in\{\pm 1\}$ for $\tilde\gamma\subset \mathrm{SL}(2,\mathbf{R})$ projecting onto Γ. The condition (5.9) is then replaced by

$$f(\tilde\gamma z)=\nu(\tilde\gamma)^{2(n+m)}(cz+d)^{2n}(c\bar z+d)^{2m}f(z) , \quad \tilde\gamma=\begin{bmatrix} a & b\\ c & d\end{bmatrix}\in\tilde\Gamma . \tag{5.10}$$

In the number theory literature, functions $f(z)$ satisfying Eq. (5.10) are called automorphic forms. Their spaces are denoted by $\mathbf{T}^{n,m}$ and carry the natural inner product

$$\|f\|^2=\int_{H/\Gamma} d^2z\, |f|^2 y^{-2+n+m} , \tag{5.11}$$

which is just an inner product of the form (2.21) or (2.45) in terms of the constant-curvature metric $ds^2=2y^{-2}dz\,d\bar z$. Similarly, from Eqs. (2.42)–(2.44), the covariant complex derivatives $\nabla^n:\mathbf{T}^n\to\mathbf{T}^{n+1}$, $\nabla_n:\mathbf{T}^n\to\mathbf{T}^{n-1}$ become operators on automorphic forms:

$$\begin{aligned} &\nabla_n=y^2\bar\partial, \quad \nabla^n=y^{-2n}\partial(y^{2n}) ,\\ &\Delta_n^{(+)}=\nabla_n^\dagger\nabla_n, \quad \Delta_n^{(-)}=(\nabla^n)^\dagger\nabla^n . \end{aligned} \tag{5.12}$$

It will usually be simpler to work with the space $S(n)=\mathbf{T}^{n/2,-n/2}$ and the Maass operators K_n and L_n defined by

$$\begin{aligned} &K_n:S(n)\to S(n+1), \quad K_n=(z-\bar z)\partial/\partial z+n ,\\ &L_n:S(n)\to S(n-1), \quad L_n=(z-\bar z)\partial/\partial\bar z-n , \end{aligned} \tag{5.13}$$

which are isomorphic to $\mathbf{T}^n$, ∇^n, and ∇_n through the isometry $\mathbf{T}^n\ni f\mapsto y^{n/2}f\in S(n)$. The Laplacians $\Delta_n^{(\pm)}$ on $\mathbf{T}^n$ reduce to operators on $S(n)$,

$$\Delta_n^{(\pm)}=-D_{-n}+n(n\pm 1)$$

with

$$D_n=y^2\partial\bar\partial-2iny(\partial+\bar\partial) . \tag{5.14}$$

Note that D_n is real: $\bar D_n=D_{-n}$.

D. Poincaré series, heat kernels, Selberg trace formulas

To construct automorphic forms we rely on the method of images. More precisely, with any function $h(z)$ on Γ with suitable decrease at infinity we can associate the Poincaré series

$$\theta_h(z)=\sum_{\tilde\gamma\in\tilde\Gamma}[\nu(\tilde\gamma)]^{2(p+q)}(cz+d)^{2p}(c\bar z+d)^{2q}h(\gamma z) , \tag{5.15}$$

which will then be an automorphic form of weight (p,q). In particular, we may construct the heat kernels for operators on $\mathbf{H}/\Gamma$ from heat kernels for the corresponding operators on $\mathbf{H}$. In view of Eq. (5.14), the key ingredient is then the heat kernel $g_n{}^t(z,z')$ for $-D_{-n}$ on $\mathbf{H}$, which has actually been computed by McKean (1972), Hejhal (1976b), and Fay (1977):[32]

$$g_n{}^t(z,z')=\frac{\sqrt{2}e^{-t/4}}{(4\pi t)^{3/2}}\int_d^\infty db\frac{be^{-b^2/4t}}{\sqrt{\cosh b-\cosh d}} \times T_{2n}\left[\frac{\cosh b/2}{\cosh d/2}\right] , \tag{5.16}$$

where $T_{2n}(t)$ is the $2n$th Chebyshev polynomial and d is the hyperbolic distance between z and z'. As a consequence the heat kernel for $\Delta_n^{(\pm)}$ on $S(n)$ is given by $e^{-tn(n\pm1)}K_n{}^t(z,z')$ with

$$K_n{}^t(z,z')=\sum_{\gamma\in\tilde\Gamma}\nu(\gamma)^{2n}\left[\frac{c\bar z'+d}{cz'+d}\right]^n\left[\frac{z-\gamma\bar z'}{\gamma z'-\bar z}\right]^n \times g_n{}^t(z,\gamma z') . \tag{5.17}$$

(Observe that K_n^t is of weight n in z', but of weight $-n$ in z.)

The trace of the heat kernel is given by

$$\mathrm{Tr}(e^{-t\Delta_n^{(\pm)}})=e^{-tn(n\pm1)}\int_{\mathbf{H}/\Gamma}dx\,dy\,y^{-2}K_n^t(z,z) \tag{5.18}$$

and can be computed through Selberg trace formula techniques. The method roughly goes as follows. The heat kernel $K_n^t(z,z)$ is a sum over elements of Γ, which can be classified into conjugacy classes and elements in their centralizers. If Γ_β is the conjugacy class of β, then the integral over H/Γ of the sum over elements $k\in\Gamma/\Gamma_\beta$ can be viewed as the integral over a fundamental domain of Γ_β. But Γ_β is cyclic and of the form $\{\gamma^p,p\in Z\}$ for some primitive element γ. Since the heat kernel in the upper half space is invariant under SL(2,$\mathbf{R}$), we may now conjugate γ within SL(2,$\mathbf{R}$) to a dilation, choose a fundamental domain for Γ_γ to be of the form $\{-\infty<x<\infty, 1\le y<e^l\}$, and carry out the integrals explicitly. Using the generating function for Chebyshev polynomials, we obtain

[32] John Fay has kindly pointed out to us that the discrete series that occurs in his expression for g^t is erroneous and should be deleted, so that one indeed obtains Eq. (5.16).

$$\mathrm{Tr}(e^{-t\Delta_n^{(\pm)}})=e^{-tn(n\pm1)}I^n(t)+e^{-tn(n\pm1)}I_e^n(t)\ , \tag{5.19}$$

$$\begin{aligned}I^n(t)&=\sum_{\gamma\ \mathrm{primitive}}\sum_{p=1}^{\infty}\nu(\gamma)^{2np}\frac{1}{\sinh pl/2}\frac{e^{-t/4}}{4\sqrt{\pi t}}e^{-p^2l^2/4t}\ ,\\ I_e^n(t)&=|\chi(M)|\sum_{0\le m<|n|-1/2}(2|n|-2m-1)e^{(|n|-m)(|n|-m-1)t}\\ &\quad+|\chi(M)|\frac{e^{-t/4}}{2\sqrt{\pi}t^{3/2}}\int_0^\infty db\frac{be^{-b^2/4t}}{\sinh b/2}\cosh(|n|-[|n|])b\ .\end{aligned} \tag{5.20}$$

Here γ and γ^{-1} are counted as distinct primitive elements. We have singled out the contribution $I_e^n(t)$ of the identity element in the Poincaré series, which encodes all the short-time information of the heat kernel. Note that $I^n(t)$ depends only on whether the field is a tensor or a spinor, and otherwise not on its weight n.

From Eq. (5.20) we can obtain the number $N_n^{\pm}[=\lim_{t\to\infty}\mathrm{Tr}(e^{-t\Delta_n^{(\pm)}})]$ of zero modes of $\Delta_n^{(\pm)}$. Since $N_0^{\pm}$ is just 1 and $\mathrm{Tr}(e^{-t\Delta_{1/2}^{(-)}})$ is certainly bounded as $t\to\infty$, it follows that $I^0(t)\to1$ and $|I^{1/2}(t)|\le O(e^{-t/4})$ as $t\to\infty$. W can now combine this with asymptotics for $I_e^n(t)$ to deduce that $N_n^+=0$ for $n\ge\frac{1}{2}$, $N_1^-=2h$, $N_n^-=(2n-1)|\chi(M)|$ for $n\ge\frac{3}{2}$. The number $N_{1/2}^-=N_{-1/2}^{\pm}$ of zero modes of the Dirac operator satisfies no such simple formula, since it depends in general on both the spin and the complex structure [cf. Eq. (2.51) and Sec. VI.F].

Automorphic forms are discussed by Ford (1951). Fay's formula for the heat kernel appeared in Fay (1977). The Selberg trace formula was introduced by Selberg (1956) and applied in McKean (1972) and McKean and Singer (1967) for the scalar Laplacian. An extensive discussion is in Hejhal (1976b). The above generalization based on Maass operators and Fay's formula (5.16) appeared in D'Hoker and Phong (1986d). The generalization to the case of the superstring is discussed by Aoki (1988).

E. Zeta-function regularization

The determinants will be evaluated through zeta-function regularization,[33]

$$\begin{aligned}\det{}'\Delta_n^{(\pm)}&=\exp[-\zeta_n^{\pm\prime}(0)]\ ,\\ \zeta_n^{\pm}(s)&=\mathrm{Tr}'(\Delta_n^{(\pm)})^{-s}\\ &=\frac{1}{\Gamma(s)}\int_0^\infty dt\ t^{s-1}[\mathrm{Tr}(e^{-t\Delta_n^{\pm}})-N_n^{\pm}]\ .\end{aligned} \tag{5.21}$$

With the above formulas for the heat kernel, it is simplest to adapt to our context the elegant method of Fried (1986a). Set

$$\begin{aligned}M_n^{\pm}(s)&=\int_0^\infty dt\ t^{s-1}e^{-tn(n\pm1)}I(t)\ ,\\ M_{n,e}^{\pm}(s)&=\int_0^\infty dt\ t^{s-1}e^{-tn(n\pm1)}I_e^n(t)\ ,\end{aligned} \tag{5.22}$$

and choose $\alpha_n^{\pm}$ so that $M_{n,e}^{\pm}(s)-\alpha_n^{\pm}\Gamma(s)$ will be holomorphic at $s=0$. From Eq. (5.20) we see readily that α_n^+ should be taken as the constant term in the short-time expansion of $e^{-t(n\pm1)n}I_e^n(t)$ for $n\ge-\frac{1}{2}$, while α_n^- should be the constant term in the short-time expansion of $e^{-tn(n-1)}I_e^n(t)-(2n-1)|\chi(M)|$ for $n\ge1$. [This distinction is based on the asymptotic behavior for large t of $e^{-t(n\pm1)n}I_e^n(t)$, which is at most $\sim t^{-1/2}$ in the first case while it is $(2n-1)|\chi(M)|$ in the second case.] The exact value at $s=0$ of $M_{n,e}^{\pm}-\alpha_n^{\pm}\Gamma(s)\equiv c_n^{\pm}|\chi(M)|$ can actually be computed to be

$$c_n^+=c_n,\quad c_n^-=c_{n-1}\ , \tag{5.23}$$

with c_n as in Eq. (5.7). Returning to $\zeta_n^{\pm\prime}(0)$ we decompose it as

$$\begin{aligned}\zeta_n^{\pm\prime}(0)&=\lim_{s\to0^-}[M_n^{\pm}(s)-\Gamma(s)\zeta_n^{\pm}(0)+\alpha_n^{\pm}\Gamma(s)]\\ &\quad+\lim_{s\to0^-}[M_n^{\pm}(s)-\alpha_n^{\pm}\Gamma(s)]\ ,\end{aligned} \tag{5.24}$$

where the zeta function is known to be holomorphic near $s=0$ by the general theory of functional determinants. To express it in terms of number-theoretical zeta functions, rewrite t^{s-1} in Eq. (5.22) in terms of its Mellin transform, change the order of integration, and evaluate the integrals explicitly to obtain

$$\begin{aligned}M_n^{\pm}(s)&=\frac{1}{\Gamma(1-s)}\int_0^\infty d\lambda[\lambda(\lambda+2|n\pm\tfrac{1}{2}|)]^{-s}\sum_{\gamma,p}\nu(\gamma)^{2np}\frac{l}{2\sinh pl/2}e^{-pl(\lambda+|n\pm1/2|)}\\ &=\frac{1}{\Gamma(1-s)}\int_0^\infty d\lambda[\lambda(\lambda+2|n\pm\tfrac{1}{2}|)]^{-s}\frac{Z'(\lambda+|n\pm\frac{1}{2}|+\frac{1}{2})}{Z(\lambda+|n\pm\frac{1}{2}|+\frac{1}{2})}\ .\end{aligned} \tag{5.25}$$

[33]The result from zeta-function regularization differs from that of small-time cutoff by harmless factors involving the area and the Euler characteristic of the worldsheet, and the number of zero modes.

This is the key relation linking the heat kernel and Selberg zeta functions, allowing us to determine the poles at $s=0$ of $M_n^{\pm}(s)$ in terms of the order of vanishing $r_n^{\pm}$ of $Z(s)$ at $s=|n\pm\frac{1}{2}|+\frac{1}{2}$. In fact, for small $\varepsilon>0$ we can split the integral representing $M_n^{\pm}(s)$ into an integral over $\lambda>\varepsilon$ and an integral over $\lambda<\varepsilon$. The first is holomorphic at $s=0$ and behaves like $-r_n^{\pm}/s+r_n^{\pm}\ln|n\pm\frac{1}{2}|+r_n^{\pm}\ln\varepsilon+O(s)$ for $n\pm\frac{1}{2}\neq 0$, and like $-r_n^{\pm}/2s+r_n^{\pm}\ln\varepsilon$ for $n\pm\frac{1}{2}=0$. Since the pole of the second integral must cancel that of $\Gamma(s)[\zeta_n^{\pm}(0)-\alpha_n^{\pm}]$, it follows that $r_n^{\pm}=\alpha_n^{\pm}-\zeta_n^{\pm}(0)$ for $n\pm\frac{1}{2}\neq 0$, $r_n^{\pm}=2[\alpha_n^{\pm}-\zeta_n^{\pm}(0)]$ for $n\pm\frac{1}{2}=0$. Now it is easy to see that $\zeta_n^{\pm}(0)$ equals the difference between the constant term in the short-time expansion of $\mathrm{Tr}(e^{-t\Delta_n^{\pm}})$ and the number of zero modes $N_n^{\pm}$. Recalling the definition of $\alpha_n^{\pm}$ and the formulas for $N_n^{\pm}$ we obtain at once

$$
\begin{aligned}
&r^{+}_{-1/2}=2N^{+}_{-1/2}, \quad r^{-}_{1/2}=2N^{-}_{1/2} ,\\
&r_0^{+}=1, \quad r_1^{-}=1 ,\\
&r_n^{+}=0 \quad \text{for } n\geq\tfrac{1}{2}, \quad r_n^{-}=0 \quad \text{for } n\geq\tfrac{3}{2} .
\end{aligned} \tag{5.26}
$$

This cancellation of the poles leaves us with

$$
M_n^{\pm}(s)-\Gamma(s)\zeta_n^{\pm}(0)+\alpha_n^{\pm}\Gamma(s)\,|_{s=0} = -\ln\left[\frac{1}{r_n^{\pm}!}Z^{(r_n^{\pm})}(|n\pm\tfrac{1}{2}|+\tfrac{1}{2})\right] \tag{5.27}
$$

and thus

$$
\det{}'\Delta_n^{(\pm)}=e^{-c_n^{\pm}\chi(M)}\frac{1}{(r_n^{\pm})!}Z^{(r_n^{\pm})}(|n\pm\tfrac{1}{2}|+\tfrac{1}{2}) . \tag{5.28}
$$

This formula includes Eqs. (5.7) as special cases.

In the mathematical literature, zeta-function regularization of determinants goes back to Ray and Singer (1971). The above techniques were used to evaluate determinants by Fried (1986a, 1986b), D'Hoker and Phong (1986a, 1986d), and Sarnak (1987).

F. Asymptotics for determinants

Physical quantities are given in string theory by integrals over moduli space. The integrands have no singularity inside, so the only possible divergences must come from their asymptotic behavior near the boundary of moduli. The importance of boundary contributions has emerged before in Sec. II.K in our discussion of BRST invariance. As explained in Sec. IV.H, this boundary corresponds to the length of some closed geodesic on the worldsheet tending to 0, and our first task is to study the behavior of the determinants (5.7) in such a limit. As expected, the partition function for the bosonic string will diverge. For the fermionic string the evidence suggests that the contribution from some spin structures will diverge as well, so that finiteness of superstring amplitudes (if true) must result from delicate cancellations between various spin structures.

In view of Eqs. (5.7) the asymptotic behavior of the absolute values of the determinants of string theory reduces to that of special values of Selberg zeta functions. Let l_0 be the length of the closed geodesic γ_0 that is being pinched. We begin by noting that the asymptotics of its contribution to $Z(s)$ for $\nu(\gamma_0)=1$ can be determined from the Jacobi identity,

$$
\prod_{k=1}^{\infty}(1-e^{-(s+k)l_0})\sim l_0^{-s+1/2}\exp(-\pi^2/6l_0) . \tag{5.29}
$$

For values of $\mathrm{Re}(s)>1$ this actually gives already the asymptotics for the full $Z(s)$. The key to understanding this lies in the collar phenomenon when a geodesic tends to 0. As l_0 tends to 0, a collar, i.e., a region diffeomorphic to a cylinder, around the geodesic will stretch out with its length of the size of $\ln 1/l_0$. Outside the collar the area and diameters remain bounded independently of l_0 (see Fig. 20). (Note that this is consistent with the fact that the area must remain constant for hyperbolic metrics of fixed negative curvature.) It is also indicated by the plumbing fixture constructions of Sec. IV.H. This suggests dividing the remaining geodesics in the infinite product defining $Z(s)$ into two groups. The first group consists of the geodesics not intersecting γ_0. Their contributions will tend to the special value of the Selberg zeta function of the punctured surface obtained at $l_0=0$. These are well behaved and will merely contribute an asymptotic constant. The second group consists of geodesics intersecting γ_0. Since these geodesics must cross the collar, their lengths must go to ∞, and hence their contribution will tend to 1. (Strictly speaking, before these lengths increase they may first decrease due to the fact that they may wrap around γ_0 a large number of times.) Thus recalling that γ_0 and γ_0^{-1} are counted as distinct primitive geodesics in the Selberg zeta function we deduce that

$$
Z(s)\sim l_0^{-2s+1}\exp(-\pi^2/3l_0), \quad \mathrm{Re}(s)>1 . \tag{5.30}
$$

The cases of $d^kZ(s)/ds^k$ for $\mathrm{Re}(s)\leq 1$ are more difficult, since analytic continuation is needed to define these values. It is a good heuristic principle, however, that up to smaller factors the asymptotics of $d^kZ(s)/ds^k$ (k is the first integer for which the derivative does not vanish) are given by the same formula (5.29) coming from the asymptotics of the terms in $Z(s)$ involving the pinching geodesics. This heuristic principle is justified by the precise

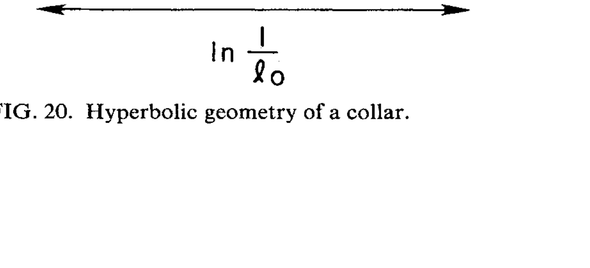


FIG. 20. Hyperbolic geometry of a collar.

formula for $Z'(1)$ obtained recently by Gava *et al.* (1986), Wolpert (1986), and Hejhal (1987):

$$Z'(1)\sim l_0^{-1}\exp(-\pi^2/3l_0)\prod_{0<\lambda_n<1/4}\lambda_n\ . \tag{5.31}$$

Here λ_n are the eigenvalues of the Laplacian on scalars on the Riemann surface. It is known from work by Schoen *et al.* (1980) and Dodziuk *et al.* (1986) that there are at most $4h-2$ eigenvalues that are less than $\frac{1}{4}$, and that the lowest eigenvalues λ_n are of the size of the sum of lengths of closed geodesics disconnecting the surface into $(n+1)$ components. Thus for a closed geodesic γ_0 of the type in Fig. 19, λ_1 is of the order of l_0. (Observe that this does not contradict the fact that the diameter of the surface is of the order of $\ln 1/l_0$. The reason is that, in the hyperbolic metric, the area element of a cylinder grows exponentially. More precisely, in the energy integral the contribution from the complement of the collar remains bounded, while by conformal invariance the collar contributions are the same as an energy integral for a Euclidean cylinder of radius 1 and length $\sim 1/l_0$.) Thus the additional factors involving λ_n in Eq. (5.31) are of lower order than the main terms.

In the presence of a spin structure $\nu(\gamma)$ with $\nu(\gamma_0)=-1$, the asymptotics of the contributions from γ_0 to $Z(s)$ become

$$\prod_{k=1}^{\infty}(1+e^{-(s+k)l_0})\sim e^{\pi^2/12l_0}\ , \tag{5.32}$$

and the above arguments apply when $\mathrm{Re}(s)>1$. No precise asymptotics such as those of Eq. (5.31) have been established rigorously for $d^{2N_{1/2}^-}Z(s)/ds^{2N_{1/2}^-}|_{s=1/2}$ at the present time. There is, nevertheless, a general method that should give good information in principle on any $Z^{(k)}(\rho)$ for $\mathrm{Re}(\rho)\leq 1$. This method, based on functional equations, goes back to Lavrik and was suggested in this context by Goldfeld. Roughly speaking $Z^{(k)}(\rho)$ can be obtained by integrating $Z(\rho+s)/s^k$ on a vertical line far to the right in the s plane, and shifting the line of integration far to the left, picking up the only pole in $Z(\rho+s)/s^k$ at $s=0$. The functional equation for Z allows us to rewrite the integral on the far left as an integral on the far right, where the infinite products for Z_ν converge absolutely and collar arguments are valid. Asymptotics follow in principle by expanding the zeta functions into Dirichlet series in l. Applied, for example, to $Z'(1)$, this method gives back Eq. (5.31) with the precise factor $\prod\lambda_n$ replaced by $O(l^{-\varepsilon})$ for any $\varepsilon>0$. For $Z^{(2N)}(\frac{1}{2})$ we expect it to confirm the heuristic principle stated above again, with a possible uncertainty of $O(l_0^{-\varepsilon})$.

In the above we have written down formulas for the pinching of only one geodesic. It should, however, be evident that they can be extended to the case of several pinching geodesics, and that the maximum number of geodesics that can be pinched independently is $3h-3$.

Before returning to string partition functions, we will need one more ingredient, namely, the asymptotic behavior of the Weil-Petersson measure. Recalling the correspondence between the complex coordinate t defining the divisor Δ of Riemann surfaces with nodes and the length of a pinching geodesic,

$$|t|\sim\exp(-2\pi^2/l_0)\ , \tag{5.33}$$

we note that the Weil-Petersson metric is described near Δ by Masur (1976),

$$d(\mathrm{WP})\sim\prod_j |dt_j^2|/|t_j|^2(\ln 1/|t_j|)^3\ . \tag{5.34}$$

For the bosonic string, the partition function is the integral over moduli of

$$\exp[c\chi(M)]Z'(1)^{-13}Z(2)d(\mathrm{WP})\ , \tag{5.35}$$

which in view of Eqs. (5.30)–(5.35) behaves up to smaller factors as

$$\prod_j |dt_j^2|\,|t_j|^{-4}\ . \tag{5.36}$$

This is the double-pole behavior obtained by Belavin and Knizhnik (1986) using essentially holomorphicity and characteristic classes arguments. A rigorous treatment along Selberg zeta-function lines as above may be found in Wolpert (1987).

Assuming the heuristic principle stated in the previous paragraph for asymptotics of Selberg zeta functions beyond $\mathrm{Re}(s)>1$ and neglecting the factors arising from supermoduli in the superstring functional integrals, we can derive similar asymptotics as well for fermionic partition functions. The importance of spin structures then becomes manifest, since the asymptotic behavior of fermionic determinants can change drastically if the sign of parallel transport around γ_0 is flipped. Thus for "wrong-spin structures" we cannot have cancellation between bosonic and fermionic determinants and must hope instead for a cancellation between the various spin structures. In physical terms, this means that finiteness of superstring theories must come here from space-time supersymmetry rather than worldsheet supersymmetry. Some of these issued have also been addressed by Iengo (1987) and Bonini and Iengo (1987a, 1987b).

G. Determinants on Mandelstam diagrams and unitarity

We have shown in Sec. II.L that Polyakov string amplitudes may be obtained in two different ways, yielding, however, the same answer. In the first approach, reparametrization-invariant vertex operators that satisfy certain Weyl invariance conditions are inserted on a compact surface, and their positions are integrated over. To obtain the full amplitude, one sums over all (inequivalent) compact surfaces of a given number of handles. In the second approach, one sums instead over surfaces of a given number of handles, with each vertex operator replaced by a puncture on the surface. Thus one sums over all surfaces of given genus h and given number of punctures n, corresponding to n vertex operators.

Now it must have become clear from Secs. IV.F and IV.G that very nice parametrizations are available for surfaces and their moduli spaces as soon as one allows for at least one puncture. Here, we shall only consider the case of Mandelstam diagrams, for which it was proven in Sec. IV.G that with their natural ranges of parameters, the Mandelstam diagrams parametrize the moduli space $\mathcal{M}_{h,n}$ of surfaces with n punctures precisely once.

This important result would be even more useful in string theory if the measure and the integrand for the scattering amplitudes were to assume a relatively simple and explicit form. We shall not analyze this question in full here, but restrict ourselves to showing the following important simplifications that occur when evaluating ghost determinants on the Mandelstam diagrams. The determinant for the spin-1 ghost, together with its finite-dimensional determinant involving holomorphic quadratic differentials, is simply given in terms of the determinant on spin-0 scalars. Similarly, the determinant for the spin-$\frac{1}{2}$ superghost is simply related to the Dirac determinant.

1. The spin-1 ghost determinant

Recall the quadratic holomorphic differentials on the Mandelstam diagram that were produced in Sec. IV.G:

$$\begin{aligned} &\phi_I=\omega\omega_I, \quad I=1,\ldots,h , \\ &\phi_a=\omega\omega_{P_0P_a}, \quad a=1,\ldots,2h+n-3 , \end{aligned} \tag{5.37}$$

where ω, ω_I, and $\omega_{P_0P_a}$ are the canonical differentials, Abelian of first and third kind, respectively. Using the Riemann bilinear relations of Appendix D, it is easy to compute the corresponding inner products,

$$\begin{aligned} &\langle \phi_I \mid \phi_J \rangle = 4\,\mathrm{Im}\Omega_{IJ}, \quad \langle \phi_I \mid \phi_a \rangle = 0 , \\ &\langle \phi_a \mid \phi_b \rangle = 2\,\mathrm{Re}\int_{P_0}^{P_b} \omega_{P_0P_a} \equiv 2G_{P_0}(P_a,P_b) , \end{aligned} \tag{5.38}$$

where Ω is the period matrix and G is symmetric in P_a and P_b. Actually, the function $G_{P_0}(P_a,P_b)$ is a Green's function for the scalar Laplacian on the Riemann surface, as can be seen by considering

$$\begin{aligned} \Delta_P G_{P_0}(P,Q) &= -2\frac{\partial}{\partial \bar P}\omega_{P_0Q}(P) \\ &= 2\pi[\delta(P,Q)-\delta(P,P_0)] . \end{aligned} \tag{5.39}$$

Having the inner products between ϕ's, it is straightforward to compute the finite-dimensional determinants

$$(\det\langle \phi_\alpha \mid \phi_\beta \rangle)^{1/2} = 2^h \det(\mathrm{Im}\Omega)\det G_{P_0}(P_a,P_b) , \tag{5.40}$$

where the latter determinant is taken of a $(2h+n-3)\times(2h+n-3)$ matrix. Similarly, the finite-dimensional determinant involving the Beltrami differentials μ_α can be evaluated by using the quasiconformal vector fields listed in Eq. (2.59). A somewhat lengthy argument, given in D'Hoker and Giddings (1987), allows one to find that

$$\det\langle \mu_\alpha \mid \phi_\beta \rangle = (8\pi)^h (4\pi)^{2h+n-3}\det(\mathrm{Im}\Omega) . \tag{5.41}$$

Next we evaluate the infinite-dimensional Faddeev-Popov ghost determinant $\det{}^* P_1^\dagger P_1$, considered on those reparametrization vector fields V^z and $V^{\bar z}$ that vanish at the punctures and are regular anywhere else. With the help of the canonical differential $\omega=\omega_z dz$, such vector fields may be rewritten in terms of scalar fields ϕ,

$$V^z = \frac{1}{\omega_z}\phi, \quad V^{\bar z} = \frac{1}{\omega_{\bar z}}\bar\phi , \tag{5.42}$$

provided the scalar field ϕ vanishes at the interaction points P_a where ω_z also vanishes. Provided ϕ is continuous at the punctures, V^z will automatically vanish there. Away from interaction points and punctures, the operators $P_1^\dagger P_1$ and Δ coincide, since the metric is Euclidean. Actually, the only reason they differ is that they act on vectors and scalars, respectively. But we have established above a correspondence between these two, and so the determinant of $P_1^\dagger P_1$ on vector fields vanishing at punctures equals the determinant of Δ vanishing at the interaction points (again indicated by an asterisk),

$$(\det{}^* P_1^\dagger P_1)^{1/2} = \det{}^*\Delta . \tag{5.43}$$

This is easily evaluated with the functional integral representation by inserting delta functions of the scalar field at the interaction points:

$$\begin{aligned} (\det{}^*\Delta)^{-1/2} = \int \mathcal{D}\phi\, e^{-\langle \bar\phi \mid \Delta\phi \rangle/8\pi}\delta(\phi(P_0)) \\ \times\delta(\phi(P_1))\cdots\delta(\phi(P_{2h+n-3})) . \end{aligned} \tag{5.44}$$

Delta functions can be represented by their Fourier transform, the constant ϕ mode can be integrated out, and the remaining Gaussian integral (with source at P_a) evaluated in terms of the Green's function $G_{P_0}(P_a,P_b)$ satisfying Eq. (5.39), and one finds

$$\begin{aligned} (\det{}^* P_1^\dagger P_1)^{1/2} &= \det{}^*\Delta \\ &= \left[\frac{8\pi^2}{\int_M d^2\xi\sqrt{g}}\det{}'\Delta\right]\det G_{P_0}(P_a,P_b) . \end{aligned} \tag{5.45}$$

It is instructive to combine this answer with that of the finite-dimensional determinants in Eq. (5.40):

$$\left[\frac{\det{}^* P_1^\dagger P_1}{\det\langle \phi_\alpha \mid \phi_\beta \rangle}\right]^{1/2} = 2^{-h}\frac{8\pi^2\det{}'\Delta}{\int_M d^2\xi\sqrt{g}\,\det\mathrm{Im}\Omega} . \tag{5.46}$$

The equivalence of the above determinants, together with the equivalence of the formulations with vertex operators and with punctures was obtained by D'Hoker and Giddings (1987) and was used to establish equivalence between the Polyakov approach and the interacting string picture. Since the latter is (formally) unitary by construction—recall it has a tachyon—this establishes the unitarity of the Polyakov approach. A direct comparison was also made by Sonoda (1987b).

2. The superghost determinant for even-spin structure

For even-spin structure and a generic point in moduli space, there is a unique meromorphic $\frac{1}{2}$ differential[34]—the Szegö kernel—with a single pole at P,

$$S_v(z,P)=\kappa_P(z)\sim\frac{(dz)^{1/2}}{z-P} ,$$

and on a Mandelstam diagram with canonical differential ω, the holomorphic $\frac{3}{2}$ forms are given by

$$\rho_a=\omega\kappa_{P_a},\quad a=1,2,\ldots,2h+n-2 , \tag{5.47}$$

as was shown in Sec. IV.G. The matrix of inner products of $\omega\kappa_P$ with P not necessarily at an interaction point is closely related to the Dirac Green's function

$$S(P,Q)=\langle\omega\kappa_Q\mid\omega\kappa_P\rangle=\int(\omega\overline{\omega})^{1/2}\kappa_P\overline{\kappa_Q} . \tag{5.48}$$

In conformal coordinates the Dirac operator is given by

$$\not{D}=\nabla^z_{1/2}\oplus\nabla_z^{-1/2} , \tag{5.49}$$

with

$$\begin{aligned}\nabla^z_{1/2}&=g^{z\bar z}\frac{\partial}{\partial\bar P} ,\\ (\nabla^z_{1/2})^\dagger&=-(g_{z\bar z})^{-1/2}\frac{\partial}{\partial P}(g_{z\bar z})^{1/2} ,\end{aligned} \tag{5.50}$$

so that

$$\begin{aligned}\not{D}\not{D}_P S(P,Q)&=-(\omega\overline{\omega})^{-1/2}(P)\frac{\partial}{\partial P}(\omega\overline{\omega})^{-1/2}(P)\\ &\quad\times\int(\omega\overline{\omega})^{1/2}(z)\frac{\partial\kappa_P(z)}{\partial\bar P}\overline{\kappa_Q(z)}\\ &=2\pi(\omega\overline{\omega})^{-1/2}(P)\frac{\partial}{\partial P}\overline{\kappa_Q(P)}\\ &=4\pi^2\delta^2(P,Q) .\end{aligned} \tag{5.51}$$

Thus we obtain

$$\det\langle\rho_a\mid\rho_b\rangle=(2\pi)^{2h+n-2}\det S(P_a,P_b) , \tag{5.52}$$

a formula very similar to Eq. (5.40).

The infinite-dimensional determinant for the superghost $\det^* P^\dagger_{1/2}P_{1/2}$, considered on spinor fields that vanish at the punctures, is also easily computed, since it can be related to the Dirac determinant $\det^*\not{D}$ over spinor fields that vanish at the interaction points. We shall not reproduce this calculation here and only quote the answer,

$$(\det^* P^\dagger_{1/2}P_{1/2})^{1/2}=(\det\not{D})\det S(P_a,P_b) . \tag{5.53}$$

[34]In Sec. VI.F we shall write down this differential explicitly in terms of ϑ functions.

Combining it with the expression for the finite-dimensional determinants, we find

$$\left[\frac{\det^* P^\dagger_{1/2}P_{1/2}}{\det\langle\rho_a\mid\rho_b\rangle}\right]^{1/2}=\det\not{D} . \tag{5.54}$$

3. The superghost determinant for odd-spin structure

For odd-spin structure δ and a generic point in moduli space, there is one holomorphic $\frac{1}{2}$ differential h_δ and a unique meromorphic $\frac{1}{2}$ differential[35] κ_{PQ} with simple poles at P and Q and unit residue at P. Holomorphic $\frac{3}{2}$ differentials on the Mandelstam diagram are given by (see Sec. IV.G)

$$\begin{aligned}\rho_a&=\omega\kappa_{P_aP_{a+1}},\quad a=1,2,\ldots,2h+n-3 ,\\ \rho_{2h-n-2}&=\omega h_\delta .\end{aligned} \tag{5.55}$$

Actually, the meromorphic differential $\kappa_{PP'}(z)$ depends both on an auxiliary point R, where it has a zero, and on the spin structure. We shall denote such a differential (in particular the one exhibited in Sec. VI.F) by $S_\delta(z,P,R,P')$, and reserve for κ_{PQ} the one that is orthogonal to h_δ:

$$\kappa_{PP'}(z)=S_\delta(z,P,R,P')-A(P,R,P')h_\delta(z)$$

and

$$A(P,R,P')=\frac{\int(\omega\overline{\omega})^{1/2}(z)\overline{h_\delta(z)}S_\delta(z,P,R,P')}{\langle h_\delta\mid h_\delta\rangle} . \tag{5.56}$$

Note that the normalization $\langle h_\delta\mid h_\delta\rangle$ is independent of P,R,P'. Now κ_{PQ} satisfies two essential equations (they can be established using the results of Sec. VI.F),

$$\begin{aligned}\frac{1}{2\pi}\frac{\partial}{\partial\bar z}\kappa_{PP'}(z)&=\delta^2(z-P)-\delta^2(z-P')\frac{h_\delta(P)}{h_\delta(P')} ,\\ \frac{1}{2\pi}\frac{\partial}{\partial\bar P}\kappa_{PP'}(z)&=-\delta^2(z-P)+\frac{h_\delta(z)\overline{h_\delta(P)}}{\langle h_\delta\mid h_\delta\rangle}(\omega\overline{\omega})^{1/2}(P) .\end{aligned} \tag{5.57}$$

It is now easy to show that the inner product between $\frac{3}{2}$ differentials produces a propagator for the square of the Dirac operator:

$$\begin{aligned}S'(P,Q)&=\langle\omega\kappa_{QQ'}\mid\omega\kappa_{PP'}\rangle\\ &=\int(\omega\overline{\omega})^{1/2}(z)\overline{\kappa_{QQ'}(z)}\kappa_{PP'}(z) .\end{aligned} \tag{5.58}$$

This is seen by applying a derivative in $\bar P$:

$$\frac{\partial}{\partial\bar P}S'(P,Q)=\int(\omega\overline{\omega})^{1/2}(z)\overline{\kappa_{QQ'}(z)}\frac{\partial}{\partial\bar P}\kappa_{PP'}(z) . \tag{5.59}$$

Using the second equation in (5.57) and the orthogonality of $\kappa_{QQ'}$ and h_δ, we get

[35]In Sec. VI.F we shall give explicit formulas for both these differentials in terms of ϑ functions.

$$\frac{\partial}{\partial \bar{P}} S'(P,Q) = -2\pi(\omega\bar{\omega})^{1/2}(P)\overline{\kappa_{QQ'}(P)} \ . \tag{5.60}$$

Applying now the Dirac operator as in Eq. (5.51), we find

$$\not{D}\not{D}_P S'(P,Q) = 2\pi(\omega\bar{\omega})^{-1/2}(P)\frac{\partial}{\partial P}\overline{\kappa_{QQ'}(P)} \ , \tag{5.61}$$

and using the first equation in (5.57), we get

$$\not{D}\not{D}_P S'(P,Q) = 4\pi^2\delta^2(P,Q) - 4\pi^2\delta^2(P,Q')\frac{h_\delta(Q)}{h_\delta(Q')} \ . \tag{5.62}$$

Thus the matrix of inner products of all holomorphic $\frac{3}{2}$ differentials is given by

$$\det\langle \rho_a \mid \rho_b \rangle = \langle h_\delta \mid h_\delta \rangle \det S'(P_a,P_b) \ , \tag{5.63}$$

where the last determinant is over a $(2h+n-3)\times(2h+n-3)$ matrix. It should be noted that the Green's function S' depends on the auxiliary points P' and Q' and that one has the property

$$S'(P',Q) = S'(P,Q') = 0 \tag{5.64}$$

for fixed P' and Q'. Computing the infinite-dimensional determinant

$$(\det{}^* P_{1/2}^\dagger P_{1/2})^{1/2} = \det{}^* \not{D} \tag{5.65}$$

is done by functional integral methods again, and it is a matter of patiently sorting out the zero-mode contribution and using Eq. (5.64) to obtain

$$(\det{}^* P_{1/2}^\dagger P_{1/2})^{1/2} = h_\delta(P')\overline{h_\delta(Q')}(\det{}'\not{D}) \times \det S'(P_a,P_b) \ . \tag{5.66}$$

Combining this result with the finite-dimensional determinant of the holomorphic $\frac{3}{2}$ differentials, we obtain the remarkable relation

$$\left[\frac{\det{}^* P_{1/2}^\dagger P_{1/2}}{\det\langle \rho_a \mid \rho_b\rangle}\right]^{1/2} = \frac{\det{}'\not{D}}{\langle h_\delta \mid h_\delta\rangle} h_\delta(P')\overline{h_\delta(Q')} \ . \tag{5.67}$$

A more detailed treatment with appropriate regularizations is in D'Hoker and Phong (1988b). Equations (5.54) and (5.67) are crucial ingredients in a unitarity proof of the fermionic string in component language (D'Hoker and Phong, 1988).

VI. COMPLEX GEOMETRY OF MODULI SPACE

In this section, we present the necessary mathematical background for the study in Sec. VII of the holomorphic structure of strings. A key ingredient is the topology and geometry of line bundles, so we begin with a short survey of their formalism. Accessible full treatments of the theory of line bundles over Riemann surfaces are provided by Gunning (1966, 1967), Hirzebruch (1966), and Griffiths and Harris (1978). The basic mathematical references for Secs. VI.D–VI.F are Fay (1973) and Mumford (1975, 1983).

A. Line bundles, Chern classes, and curvature

Let M be a smooth manifold. A line bundle L on M is an assignment of a one-dimensional complex vector space L_z to each point z of M. Sections of L are then functions assigning an element of L_z to each z. The vector spaces L_z should fit together smoothly, and we enforce this in two stages. First locally, i.e., for all z in small coordinate charts $\{B_\alpha\}$ for M, the set $\{L_z\}_{z\in B_\alpha}$ should just become isomorphic to a product $\mathbf{C}\times B_\alpha$, so that a section f of L on B_α should reduce to a smooth $\mathbf{C}$-valued function f_α on B_α. Second, there should exist smooth nowhere-vanishing complex functions $\phi_{\alpha\beta}$ defined on $B_\alpha \cap B_\beta$ so that the f_α's arising from a section f of L are characterized by the condition

$$f_\alpha = \phi_{\alpha\beta} f_\beta \quad \text{on } B_\alpha \cap B_\beta \ . \tag{6.1}$$

Clearly the $\phi_{\alpha\beta}$ themselves must satisfy the consistency condition

$$\phi_{\alpha\beta}\phi_{\beta\gamma} = \phi_{\alpha\gamma} \ . \tag{6.2}$$

They are called *transition functions* and describe L completely. Examples of line bundles are the space $\mathbf{T}^n$ on Riemann surfaces M, encountered earlier in Sec. II.E, where L_z is the $(-n)$th power of the tangent space to M of z, and transition functions are $(\partial z_\alpha/\partial z_\beta)^{-n}$ with z_α, z_β coordinate systems for the patches B_α and B_β. A more sophisticated example relevant to anomalies is that of determinant bundles. Here the manifold M is, for example, the space of metrics on a surface, and the vector space L_g at a metric g is

$$(\overset{\max}{\wedge}\,\mathrm{Ker}\bar{\partial}_n)^{-1} \otimes (\overset{\max}{\wedge}\,\mathrm{Ker}\bar{\partial}_n^\dagger) \ .$$

For chiral anomalies, M is instead the space of vector potentials A_μ, and L_{A_μ} at A_μ is similarly built out of zero modes for the Dirac operator coupled to A_μ and its adjoint. In these situations we note that the number of zero modes may jump, and it is a subtle issue to define properly the transition functions. This fact has important ramifications that will be discussed at length in Sec. VII.E.

Our next task is to introduce a topological classification of line bundles over a manifold. For this write the transition functions as $\phi_{\alpha\beta} = \exp(\psi_{\alpha\beta})$; then Eq. (6.2) is equivalent to $\psi_{\alpha\beta} + \psi_{\beta\gamma} - \psi_{\alpha\gamma} = 2\pi i(c_{\alpha\beta\gamma})$. The integers $c_{\alpha\beta\gamma}$ satisfy

$$c_{\alpha\beta\gamma} - c_{\beta\gamma\delta} + c_{\gamma\delta\alpha} - c_{\delta\alpha\beta} = 0 \ . \tag{6.3}$$

Since the $\psi_{\alpha\beta}$ are defined only modulo $2\pi i n_{\alpha\beta}$ with $n_{\alpha\beta}$ integers, we should identify two sets $c_{\alpha\beta\gamma}$ and $c'_{\alpha\beta\gamma}$ differing by elements of the form

$$n_{\alpha\beta} + n_{\beta\gamma} - n_{\alpha\gamma} \ . \tag{6.4}$$

Coefficients $(c_{\alpha\beta\gamma})$ satisfying (6.3) are called *closed cocycles*, while those that can be written in the form (6.4) are called *exact cocycles*. The space of closed cocycles modulo the exact ones is the second (Čech) cohomology group $H^2(M,\mathbf{Z})$ (with coefficients in $\mathbf{Z}$), and our discussion has shown that to each line bundle L on M corresponds an element of $H^2(M,\mathbf{Z})$, usually denoted $c_1(L)$ and called the first Chern class (or topological charge) of L.

The line bundles we need usually have more structure, whether it be under the form of a metric, a holomorphic structure, or a connection. A metric on L is a set of positive functions g_α satisfying $g_\alpha = |\phi_{\alpha\beta}|^{-2} g_\beta$. Thus a metric on L is a metric on each fiber L_z varying smoothly with z and does not involve any metric on the manifold M itself. Given a section f of L, $g_\alpha f_\alpha \bar{f}_\alpha$ is then a scalar on M which represents the modulus squared of f at each point. It will sometimes be denoted by $\|f\|^2$. The line bundle M is said to be a holomorphic line bundle if the manifold M is a complex manifold, and the transition functions $\phi_{\alpha\beta}$ are holomorphic functions on M. A connection is simply a U(1) gauge field on M, i.e., a collection $A_{\mu,\alpha}$ transforming as $A_{\mu,\alpha} = A_{\mu,\beta} - \partial_\mu \ln\phi_{\alpha\beta}$ under change of coordinate patches. There are of course many connections, but in the presence of a metric and a holomorphic structure on M there is a unique connection compatible with them both. To see this, let z^j be holomorphic coordinates on M and observe that, for any section f of L, $(\partial f_\alpha / \partial \bar{z}^j)$ satisfies Eq. (6.2), since $\phi_{\alpha\beta}$ is holomorphic. Thus

$$(\nabla_{\bar{j}} f)_\alpha = \frac{\partial}{\partial \bar{z}^j} f_\alpha \tag{6.5}$$

makes a well-defined section of L. The covariant derivatives $\nabla_j f$ are determined next by the requirement that

$$\frac{\partial}{\partial z^j}(g_\alpha f_\alpha \bar{f}'_\alpha) = g_\alpha \nabla_j f_\alpha \bar{f}'_\alpha + g_\alpha f_\alpha \overline{\nabla_{\bar{j}} f'_\alpha} \tag{6.6}$$

for any sections f and f', which implies that

$$\nabla_j f_\alpha = \left[\frac{\partial}{\partial z^j} + \frac{\partial}{\partial z^j} \ln g_\alpha \right] f_\alpha \,. \tag{6.7}$$

We have actually seen this process at work before when dealing with the spaces $\mathbf{T}^n$. They are holomorphic line bundles over the Riemann surface M, and Eqs. (6.5) and (6.7) are just extensions to this more general case of the constructions of covariant derivatives in Sec. II.E.

Since the connection is Abelian, the curvature $F_{\mu\nu}$ is a 2-form on the manifold M. It is immediate that the only nonvanishing components are

$$F_{j\bar{k}} = \partial^2(\ln g_\alpha)/\partial z_j \partial \bar{z}_{\bar{k}} \,. \tag{6.8}$$

A convenient way of phrasing this in completely intrinsic terms is the following: let f be any nonvanishing holomorphic section of L, i.e., a section for which the f_α's are holomorphic functions. Then the curvature of L is the (1,1) form given by

$$F = \partial\bar{\partial} \ln \|f\|^2 \,. \tag{6.9}$$

Here $d = \partial + \bar{\partial}$ is the splitting of the exterior derivative in the presence of a complex structure, and it is obvious that Eq. (6.9) is independent of the choice of f.

Finally, the Gauss-Bonnet theorem and the anomalous fermion number currents of Sec. II.I have taught us that there should be a direct link between topology and integrals of curvature. The proper generalization to the present context can be based on the DeRham theorem and formulated as follows. Recall that the kth DeRham cohomology group $H^k_{\mathrm{DR}}(M)$ is defined by

$$H^k_{\mathrm{DR}}(M) = \{\text{closed } k\text{-forms}\}/\{\text{exact } k\text{-forms}\} \,, \tag{6.10}$$

where a form ϕ is closed if $d\phi = 0$ and exact if $\phi = d\psi$ for some globally defined $(k-1)$ form ψ. The DeRham theorem asserts identity between these groups and real Čech cohomology groups $H^k(M,R)$ defined by real cocycles $c_{\alpha_1 \cdots \alpha_{k+1}}$ with conditions generalizing Eqs. (6.3) and (6.4). Since we shall need only the case $k=2$, we shall restrict ourselves to this case to simplify the discussion.

Let $[F]$ be an element of $H^2_{\mathrm{DR}}(M)$ with representative a closed 2-form F. On small patches B_α we can write $F = dA_\alpha$ for some one-forms A_α. On $B_\alpha \cap B_\beta$, $A_\alpha - A_\beta$ can in turn be written as

$$A_\alpha - A_\beta = d\lambda_{\alpha\beta} \tag{6.11}$$

for some functions $\lambda_{\alpha\beta}$. The class of the real cocycle $c_{\alpha\beta\gamma} = \lambda_{\alpha\beta} + \lambda_{\beta\gamma} - \lambda_{\alpha\gamma}$ can be checked to depend only on $[F]$ and defines an element of $H^2(M,R)$. According to the DeRham theorem this correspondence $[F] \leftrightarrow [c_{\alpha\beta\gamma}]$ is an isomorphism $H^2_{\mathrm{DR}}(M) \rightarrow H^2(M,R)$.

Returning now to any U(1) connections $A_{\mu,\alpha}$ on L [not necessarily the one singled out by metric and complex structures as in Eqs. (6.5) and (6.6)], we see that the curvature form $F_{\mu\nu}$ is clearly closed and thus defines a DeRham cohomology class $[(i/2\pi)F]$ in $H^2_{\mathrm{DR}}(M)$. Retracing the above steps, we see that the A_α in this case can be taken to be the connection forms $(i/2\pi) A_{\mu,\alpha} dx^\mu$, the $\psi_{\alpha\beta}$ become $(1/2\pi i)\ln\phi_{\alpha\beta}$ in view of Eq. (6.11), and thus the cohomology class $[(i/2\pi)F]$ actually coincides with the first Chern class $c_1(L)$.

To make contact with Gauss-Bonnet theorems, we observe that cohomology classes $[F]$ in $H^2_{\mathrm{DR}}(M)$ are characterized by their integrals $\int_C F$ over two-dimensional cycles C. When the dimension of the manifold M is two and M is compact and connected, M itself is the unique such cycle, and the integrals give topological numbers and hence multiples of the Euler characteristic,

$$c_1(L) = \frac{i}{2\pi} \int_M F \,. \tag{6.12}$$

That was essentially the content of equations such as (2.2), (2.27), and (2.50). In particular, we see that the Chern class of the canonical bundle K is $c_1(K) = 2h-2$, and more generally $c_1(\mathbf{T}^n) = n(2h-2)$.

B. The Jacobian variety of a Riemann surface

In this section we specialize to the case where the base manifold M is a Riemann surface and provide a classification of holomorphic line bundles on M.

Recall from Sec. VI.A that line bundles on M are distinguished already by their first Chern classes, which are elements of $H^2(M,\mathbf{Z})$. For compact two-dimensional surfaces, $H^2(M,\mathbf{Z})=\mathbf{Z}$, so that bundles are first indexed by integers. Next, bundles with the same Chern class may be topologically but not necessarily holomorphically equivalent, i.e., the smooth sections are in correspondence but not the holomorphic ones. Thus we introduce the *Picard varieties*

$$\mathrm{Pic}_d=\{\text{line bundles } L \text{ on } M \text{ with } c_1(L)=d\} \ . \tag{6.13}$$

It will turn out that the Picard varieties for various values of d are very similar, so we shall often concentrate on Pic_0, which is usually called the *Jacobian variety* of M and is denoted by $J(M)$. There are several ways of describing the Jacobian, each suitable for a different purpose, so we give them in turn.

First we observe that the space of all holomorphic line bundles on M can be conveniently viewed as a Čech cohomology group, albeit with coefficients that are not integers. More precisely, given a holomorphic line bundle L, recall that its transition functions $\phi_{\alpha\beta}$ satisfy Eq. (6.2) and note that bundles L' whose functions $\phi'_{\alpha\beta}$ are of the form

$$\phi'_{\alpha\beta}=\phi_{\alpha\beta}h_\alpha h_\beta^{-1} \tag{6.14}$$

for holomorphic nonvanishing functions h_α have their holomorphic sections f'_α in one-to-one correspondence with those of L: $f'_\alpha=h_\alpha f_\alpha$. We shall not distinguish between such bundles L and L'. If we introduce the Čech cohomology group $H^1(M,O^*)$ "with coefficients in O^*"[36] as the class of (multiplicative) holomorphic cocyles $\phi_{\alpha\beta}$ satisfying Eq. (6.2) modulo the exact ones $h_\alpha h_\beta^{-1}$, we see that $H^1(M,O^*)$ is just the space of holomorphic line bundles on M.

To single out the Jacobian variety from within $H^1(M,O^*)$, we begin by defining the first Čech cohomology group $H^1(M,O)$ with coefficients in[37] O in analogy with the previous discussion: $H^1(M,O)$ is the space of (additive) holomorphic cocycles $\psi_{\alpha\beta}$ satisfying

$$\psi_{\alpha\beta}+\psi_{\beta\gamma}-\psi_{\alpha\gamma}=0 \tag{6.15}$$

modulo exact cocycles, i.e., those of the form $\Theta_\alpha-\Theta_\beta$ for holomorphic $\Theta_\alpha,\Theta_\beta$.

There is then a natural mapping $H^1(M,O)\to H^1(M,O^*)$ given by $\phi_{\alpha\beta}=\exp(\psi_{\alpha\beta})$. In view of Eq. (6.15), the Chern class of bundles in $H^1(M,O^*)$ arising from this map must be 0. Furthermore the kernel of the map is evidently given by integer-valued cocycles $\psi_{\alpha\beta}$, naturally called the first Čech cohomology group $H^1(M,\mathbf{Z})$. The net outcome is the fundamental equation

$$J(M)=H^1(M,O)/H^1(M,\mathbf{Z}) \ . \tag{6.16}$$

[36] O^* usually stands for "germs of holomorphic nonvanishing functions."

[37] O denotes the space of germs of holomorphic functions.

In the mathematical literature, the above arguments are summarized by saying that the short exact sequence of sheaves

$$0\to 2\pi i Z\to O\xrightarrow{\exp} O^*\to 0 \tag{6.17}$$

leads to a long exact sequence in cohomology,

$$\cdots \to H^1(M,\mathbf{Z})\to H^1(M,O)\to H^1(M,O^*)\xrightarrow{c_1} H^2(M,\mathbf{Z}) \to H^2(M,O)\to \cdots$$

which terminates, since $H^2(M,O)=0$, as we shall see below. Equation (6.16) follows at once.

Our next description of the Jacobian variety is based on rewriting the Čech cohomology group $H^1(M,O)$ as the dual of the space of holomorphic one-forms on M. The key tool is the Dolbeault theorem, which is the version of the DeRham theorem that applies to the $\bar\partial$ operator instead of the exterior derivative. Let

$$H^{0,q}(M)=\frac{\{(0,q) \text{ forms } \omega \text{ on } M \text{ with } \bar\partial\omega=0\}}{\{\text{exact forms } \bar\partial s\}}$$

be the $(0,q)$ Dolbeault cohomology group. When $q=1$, we can associate with an element $[\Theta]$ of $H^{0,1}(M)$ an element of $H^1(M,O)$, very much as in the earlier DeRham discussion: on patches B_α write $\Theta=\bar\partial s_\alpha$. Clearly $\psi_{\alpha\beta}=s_\alpha-s_\beta$ is a holomorphic function on $B_\alpha\cap B_\beta$ satisfying the additive cocycle condition (6.15). Moreover, $\psi_{\alpha\beta}$ is an exact cocycle if and only if $\psi_{\alpha\beta}=h_\alpha-h_\beta$ for some holomorphic h_α, so that $s_\alpha-h_\alpha$ is then a globally defined function s on M still satisfying $\bar\partial s=\Theta$, which means that Θ is $\bar\partial$ exact. Thus we have a correspondence $H^{0,1}(M)\to H^1(M,O)$, which is an isomorphism by Dolbeault's theorem. Similarly the cohomology groups $H^{0,q}(M)$ and $H^q(M,O)$ can be shown to be isomorphic. Since for $q=2$ and M a Riemann surface there are no (0,2) forms, we conclude that $H^2(M,O)=0$, which is the statement we made in the previous paragraph.

It is actually more convenient to view $H^{0,1}(M)$ as the dual of $H^0(M,K)$, which is defined to be the space of holomorphic sections of the canonical bundle K, in other words, the space of Abelian differentials on M. This duality arises from the natural nondegenerate pairing

$$\begin{aligned} &H^{0,1}(M)\times H^0(M,K)\to \mathbf{C}\ , \\ &(\Theta,\omega)\to \int_M \Theta\wedge\omega\ , \end{aligned} \tag{6.18}$$

where the right-hand side makes sense intrinsically since $\Theta\wedge\omega$ is a (1,1) form. Thus another formula for the Jacobian of M is

$$J(M)=H^0(M,K)^+/H^1(M,\mathbf{Z}) \ . \tag{6.19}$$

We now pass to a description of the Jacobian in terms of curvature and holonomy. The key observation here is that a line bundle L has zero Chern class if and only if it admits a metric whose holomorphic connection has identically vanishing curvature. That the existence of such a metric implies that $c_1(L)=0$ is an immediate consequence of the Gauss-Bonnet formula (6.12). Conversely let $\hat{g}$ be any metric on L, and look for a factor $e^{2\sigma}$ so that the curvature of $g=e^{2\sigma}\hat{g}$ will be 0. This means that $2\partial\bar{\partial}\sigma=-$(curvature of $\hat{g}$), an equation that can be solved, since the right-hand side is orthogonal to constants again by the Gauss-Bonnet formula. We also note that a metric with zero curvature is unique up to constants. This follows from the simple fact that the ratio of two metrics with the same curvature must be the exponential of a harmonic function.

Now let L be a line bundle with $c_1(L)=0$, equipped with the unique flat metric as above. Let $A_\mu dx^\mu = A_z dz$ be the corresponding connection. Flatness is a local statement, and all such bundles L are locally the same. Globally, however, there are holonomy issues, and it is the values of parallel transport around closed cycles in M that completely determine the complex structure of L. More precisely, we introduce the following version of the familiar Wilson loop observable of gauge theories:

$$W(\gamma)=\frac{i}{2\pi}\oint_\gamma A_z dz \pmod{\mathbf{Z}} . \tag{6.20}$$

The function $W(\gamma)$ is real and can be interpreted as a phase shift. Indeed, if we parallel transport a vector around γ, it will return with a phase shift of $2\pi i W(\gamma)$. Since the curvature of $A_z dz$ is zero, it follows from Green's formula that $W(\gamma)$ depends only on the homology class of the cycle γ. In other words, W should be viewed as a real cohomology class, an element of $H^1(M;R)/H^1(M,\mathbf{Z})$.

We have thus associated an element W of $H^1(M;R)/H^1(M,\mathbf{Z})$ with each line bundle L with $c_1(L)=0$. Conversely, given W, we can construct L by taking the line bundle with constant transition functions $\exp[-2\pi i W(A_i)]$, $\exp[-2\pi i W(B_i)]$ across the cuts. Since W is trivial as a cohomology class if and only if L is trivial as a line bundle (if W is trivial, we can construct a covariantly constant section of L by parallel transport on the cut surface; the triviality of W guarantees that this section has no jumps across the cuts; the reverse statement is obvious), we have a third description of the Jacobian variety,

$$J(M)=H^1(M;R)/H^1(M,\mathbf{Z}) . \tag{6.21}$$

Another useful characterization of the Jacobian is in terms of *divisors*. The basic construction is the following. Given a point w of the surface M, we let z be a holomorphic coordinate centered at w, B_0 a small disk around w, and $B_\infty = M\setminus\{w\}$. A line bundle $[w]$ can now be defined by taking z as a transition function between B_0 and B_∞. Thus a holomorphic section f of $[w]$ is just a pair f_0, f_∞ of holomorphic functions on B_0, B_∞ satisfying $f_0 = z f_\infty$. In particular, the constant holomorphic function 1 on $M\setminus\{w\}$ gives rise to the holomorphic section $1_{[w]}$ defined by $f_\infty=1$ and $f_0=z$. Note that this section has a simple zero at w. Furthermore, the first Chern class $c_1[w]$ is equal to 1. This corresponds to the simple fact that the logarithm of the transition function is multiple valued in $B_0\cap B_\infty$ and changes by $(2\pi i)\times 1$ as we go around a small circle $|z|=$const in $B_0\cap B_\infty$. This argument can easily be made rigorous by taking a refinement of the covering $\{B_0, B_\infty\}$ and retracing the definition of Chern classes of line bundles.

More generally, given a formal expression of the form

$$D=\sum_{i=1}^N n_i w_i \quad \text{positive or negative } n_i \text{ integers} , \tag{6.22}$$

we can take holomorphic coordinates z_i centered at w_i, small disjoint disks B_i around w_i, and set $B_\infty = M\setminus\{w_1,\ldots,w_N\}$. The line bundle $[\sum_{i=1}^N n_i w_i]$ is defined by the covering $\{B_1,\ldots,B_N,B_\infty\}$ and the transition functions $z_i^{n_i}$ on the overlap $B_\infty\cap B_i$. Holomorphic sections of $[\Sigma n_i w_i]$ are now holomorphic functions $f_1,\ldots,f_N,f_\infty$ on $B_1,\ldots B_N,B_\infty$, respectively, satisfying $f_i=z_i^{n_i}f_\infty$. The holomorphic function 1 on $M\setminus\{w_1,\ldots,w_N\}$ thus extends to a meromorphic section of $[\sum_{i=1}^N n_i w_i]$ that has a pole of order n_i at w_i if n_i is negative. The multiple valuedness of the transition functions $z_i^{n_i}$ adds up to a net value of $\sum_{i=1}^N n_i$ for the Chern class of $[\sum_{i=1}^N n_i w_i]$.

The above construction will yield a trivial line bundle provided the expression $\sum_{i=1}^N n_i w_i$ is the set of zeros and poles of a meromorphic function ϕ, counted with their multiplicities. Indeed $\phi^{-1}1_{[\Sigma n_i w_i]}$ is then a holomorphic nowhere-vanishing section of $[\Sigma n_i w_i]$, and a line bundle with a nowhere-vanishing global section is evidently trivial. Similarly the line bundles arising from two formal expressions D and D' differing by the zeros and poles of a meromorphic function will be isomorphic. Thus we define a divisor $[D]$ to be a class of expressions (6.22) modulo such zeros and poles, and actually have a correspondence between divisors and line bundles. Every line bundle L does arise in this manner, since with L we can associate the divisor of zeros and poles of one of its meromorphic sections. It will be shown later from index theorems that such sections do exist, and it does not matter which one we choose, since the divisors of two sections will differ only by the zeros and poles of their quotient, which is a meromorphic function.

In this way we obtain another description of the Jacobian:

$$J(M)=\{\text{divisors } [\Sigma n_i w_i] \text{ with } \Sigma n_i=0\} . \tag{6.23}$$

As a by-product of the above discussion we have the useful fact that the difference between the number of zeros and the number of poles of a meromorphic function must be 0, and more generally, for a section of a line bundle L,

$$\#(\text{zeros})-\#(\text{poles})=c_1(L) . \tag{6.24}$$

C. Index and Riemann-Roch theorems

The basic operator on a line bundle L on a Riemann surface M is the Cauchy-Riemann operator $\bar{\partial}_L$:L

$\rightarrow L\otimes\bar{K}$. It will be important to determine the number of its zero modes, i.e., the number of holomorphic sections of L. When L is a bundle $\mathbf{T}^n$ of spinors, these are the zero modes of the Dirac operator (coupled to various vector potentials), and we determined them through index theorems and heat kernels in Secs. II.E and V.D. Here we discuss the version that applies for general L.

In the presence of a metric on L, $\bar{\partial}_L$ has an adjoint $\bar{\partial}_L^\dagger : L\otimes\bar{K}\rightarrow L$ which is just $\bar{\partial}_L^\dagger f = -\nabla_z f$ where ∇ is the holomorphic connection determined by the metric on L [cf. Eq. (6.7)]. The index theorem familiar from the study of chiral anomalies suggests that the index of $\bar{\partial}_L$ should be the integral over M of a polynomial in the curvature of L and M. Since the dimension of M is 2 and curvatures are 2-forms, we must have a linear function of $c_1(L)$ and $\chi(M)$. Comparing with Eq. (2.50) and recalling that $c_1(\mathbf{T}^n) = -n\chi(M)$, we arrive at

$$\dim \mathrm{Ker}\bar{\partial}_L - \dim \mathrm{Ker}\bar{\partial}_L^\dagger = c_1(L) + \tfrac{1}{2}\chi(M) \ . \tag{6.25}$$

It will be useful to reformulate this result independently of any metric g on L and just in terms of $\bar{\partial}$ operators. For this we appeal again to a duality statement known as Serre duality,

$$\begin{aligned}&(\mathrm{Ker}\bar{\partial}_L^\dagger)\times(\mathrm{Ker}\bar{\partial}_{L^{-1}\otimes K})\rightarrow \mathbf{C} \ ,\\ &(f\, d\bar{z}, e\, dz)\rightarrow \int fe\, d\bar{z}\, dz \ .\end{aligned} \tag{6.26}$$

The right-hand side is well defined for f and e in L and L^{-1}. This pairing is nondegenerate, since the vanishing of (6.26) means that $e\,dz$ is orthogonal in the Hilbert space sense to $h\bar{f}\,dz$. Since $f\,d\bar{z}$ is in the kernel of $\nabla_{L\otimes\bar{K}}$ if and only if $h\bar{f}\,dz$ is in the kernel of $\bar{\nabla}_{L^{-1}\otimes K} = \bar{\partial}_{L^{-1}\otimes K}$, our assertion follows. The index theorem (6.25) becomes

$$\dim \mathrm{Ker}\bar{\partial}_L - \dim \mathrm{Ker}\bar{\partial}_{L^{-1}\otimes K} = c_1(L) + \tfrac{1}{2}\chi(M) \tag{6.27}$$

and as such is known as the Riemann-Roch theorem.

We now illustrate the use of the Riemann-Roch theorem by deriving the existence of meromorphic sections of various line bundles. First recall that we claimed in Sec. VI.B that any holomorphic line bundle L admitted meromorphic sections. To see this, apply Eq. (6.27) with L replaced by $L\otimes[nw]$ where w is some fixed point and n is an integer taken so large that the right-hand side of Eq. (6.27) becomes positive. In particular, $L\otimes[nw]$ admits some holomorphic section s. But then $s1_{[-nw]}$ is a meromorphic section of L.

Next we investigate meromorphic differentials on M, i.e., sections of K. The case of Abelian (i.e., holomorphic) differentials has already been considered several times (Secs. II.E and V.D) and follows from Eq. (6.27) with L = trivial bundle. There are h Abelian differentials $\omega_1,\ldots,\omega_h$. Turning to the meromorphic ones, we shall establish the existence of meromorphic differentials with simple poles at exactly any two given points w_1 and w_2, or a double pole at any given w. We apply Eq. (6.27) with $L=[-w_1-w_2]$. Since L has Chern class $=-2$ and admits no holomorphic sections, it follows that $\dim\bar{\partial}_{[w_1+w_2]\otimes K} = h+1$. There must exist some section f of $[w_1+w_2]\otimes K$ that will complete $1_{[w_1+w_2]}\omega_1,\ldots, 1_{[w_1+w_2]}\omega_h$ into a basis for $\mathrm{Ker}\bar{\partial}_{[w_1+w_2]\otimes K}$. Evidently $1_{[-w_1-w_2]}f$ is then a section of K, with at most simple poles at w_1 and w_2, and in fact exactly simple poles at both points (if it had a pole at only one point the residue there would vanish, since we can integrate the differential on a closed contour and deform it away). It is now not difficult to see that by scaling and adding a suitable combination of Abelian differentials, we can produce a unique $\omega_{w_1w_2}$ having simple poles at w_1, w_2 with residues ± 1, and vanishing A periods. These $\omega_{w_1w_1}$ are the normalized differential of the third kind, encountered already in Sec. IV.G. Finally the above arguments can be modified to produce a meromorphic differential ω_w with exactly one double pole at w. These are called differentials of the second kind. It is not, however, true that there exist differentials with poles of any order at a given point. Certain orders $n_1,\ldots,n_k$ $(k\le h-1)$, which are called Weierstrass gaps, may be missing.

D. Period matrix and Abel map

The above concepts take a very concrete form if we choose a homology basis (A_J, B_J) satisfying the canonical relations (3.5). To a choice of homology basis corresponds a choice of basis $\omega_1,\ldots,\omega_h$ of Abelian differentials, defined unambiguously by the requirement

$$\oint_{A_J}\omega_K = \delta_{JK} \ . \tag{6.28}$$

The period matrix Ω is then the $h\times h$ complex matrix with entries

$$\oint_{B_J}\omega_K = \Omega_{JK} \ . \tag{6.29}$$

Two crucial properties of period matrices are the bilinear relations of Riemann (see Appendix D for a proof), which in particular imply that

$$\begin{aligned}&\Omega_{JK} = \Omega_{KJ} \ ,\\ &\mathrm{Im}\Omega \text{ is positive definite} \ .\end{aligned} \tag{6.30}$$

The space of all $h\times h$ matrices satisfying these conditions is called the Siegel upper half space. We note that it has complex dimension $h(h+1)/2$, while the subspace of all period matrices of Riemann surfaces has dimension at most $3h-3$ (for $h\ge 2$), in fact exactly $3h-3$, as period matrices actually characterize complex structures.

Next recall that the Jacobian variety of M can be viewed as the coset space $H^0(M,K)^+/H^1(M,\mathbf{Z})$. Since we have chosen a basis $\omega_1,\ldots,\omega_h$ of Abelian differentials, a cycle C in $H^1(M,\mathbf{Z})$ can be identified with its vector of periods $(\int_C\omega_1,\ldots,\int_C\omega_h)$ and hence with a point on the lattice $\mathbf{Z}^h+\Omega\mathbf{Z}^h$. Thus the Jacobian variety becomes

$$J(M) = \mathbf{C}^h/(\mathbf{Z}^h+\Omega\mathbf{Z}^h) \ , \tag{6.31}$$

which is evidently a complex torus of dimension h.

Observe that a change of homology bases preserving the intersection numbers (3.5) is effected by a modular transformation

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

in $\mathrm{Sp}(2h,\mathbf{Z})$. Under such a transformation the period matrix changes as

$$\Omega'=(A\Omega+B)(C\Omega+D)^{-1} . \tag{6.32}$$

It is evident that the lattice $\mathbf{Z}^h+\Omega\mathbf{Z}^h$ is then unchanged, which confirms the intrinsic meaning of the Jacobian (6.31).

We can now construct explicitly the correspondence between the divisors and elements of the Jacobian variety, and in particular embed the surface M itself within $J(M)$ (see Fig. 21). Since $J(M)$ is a torus, its function theory can be built on modular forms, and this fundamental embedding will allow us to study function theory on M through modular forms.

Fix a point z_0. Then for d points $z_1,\ldots,z_d$ in M, the *Abel map* is defined by

$$I(z_1+\cdots+z_d)=\int_{z_0}^{z_1}\omega+\cdots+\int_{z_0}^{z_d}\omega , \tag{6.33}$$

where the addition signs in the argument of I are understood in the divisor sense. The right-hand side represents an h-dimensional vector, with ω denoting the h vector of Abelian differentials $(\omega_1,\ldots,\omega_h)$. Evidently there is an arbitrariness in the choice of integration paths, but this leads only to an ambiguity of the form of a lattice point in $\mathbf{Z}^h+\Omega\mathbf{Z}^h$, so that the Abel map I is single valued in the Jacobian variety. Actually I is naturally defined on divisor classes in the sense that $I(\Sigma z_j-\Sigma w_l)\equiv 0$ if and only if z_j, w_l are the zeros and poles of a meromorphic function. This statement is usually known as Abel's theorem, for which we refer the reader to Appendix D. The Abel map I viewed as a map from divisor classes to the Jacobian variety becomes one-to-one and onto when restricted to the space of divisors with zero Chern class. This is the explicit correspondence between such divisors and the elements of a complex torus that we are looking for, although, strictly speaking, we have not as yet checked that under this identification $-I(D)$ does go over to the line bundle admitting D for divisor. This will follow most easily from theta-function constructions to be outlined in the next section.

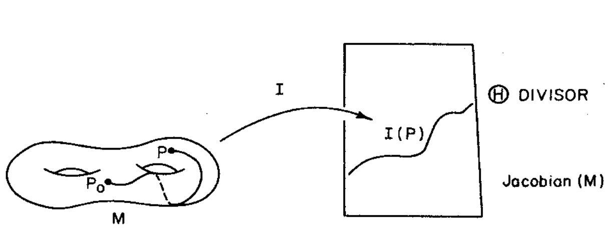


FIG. 21. The embedding of the Riemann surface into the Jacobian by the Abel map.

We note that the Abel map I can be viewed as well as a map from the space of the bundles, since with each line bundle L we can associate $I(D)$ where D is the divisor class of L. We shall use indiscriminately both symbols $I(L)$ and $I(D)$ to denote this same point on the Jacobian variety.

If we restrict I to single points $I(z)=\int_{z_0}^{z}\omega$, we obtain a one-to-one map of the surface M. This embedding is not completely free of any choices, since it depends on the base point z_0. This results in an arbitrary translation within $J(M)$, which may serve as justification for some formulas we shall encounter later.

Finally, it is not difficult to establish the following useful formula for the variation of the period matrix as we deform the complex structure by a Beltrami differential:

$$\delta\Omega_{IJ}=-i\int d^2z\,\mu_{\bar z}{}^{z}\omega_I\omega_J . \tag{6.34}$$

E. Theta functions

The Jacobi theta function is defined by

$$\vartheta(\zeta,\Omega)=\sum_{n\in\mathbf{Z}^h}\exp(\pi i n^t\Omega n+2\pi i n^t\zeta),\quad \zeta\in\mathbf{C}^h . \tag{6.35}$$

It satisfies the heat equation

$$\left[4\pi i\frac{\partial}{\partial\Omega_{IJ}}+\frac{\partial^2}{\partial\zeta_I\partial\zeta_J}\right]\vartheta(\zeta,\Omega)=0$$

and the key transformation laws

$$\vartheta(\zeta+M+\Omega N,\Omega)=\exp(-\pi i N^t\Omega N-2\pi i N^t\zeta)\vartheta(\zeta,\Omega) \tag{6.36}$$

for M and N vectors of integers. This periodicity up to a factor with respect to the lattice $Z^h+\Omega Z^h$ shows that $\vartheta(\zeta,\Omega)$ should be viewed as a holomorphic section of a line bundle over the Jacobian variety, a line bundle whose holonomy around the cycles of $J(M)$ is defined by the factors in Eq. (6.36). It can be shown that this bundle—called the ϑ line bundle—admits in fact (up to multiplicative constants) only one holomorphic section, represented by the theta function.

Although theta functions are strictly speaking sections of line bundles, we can easily manufacture meromorphic functions on $J(M)$ out of them. For example, it is easy to check that

$$\frac{\prod_{i=1}^{n}\vartheta(\zeta+a_i,\Omega)}{\prod_{i=1}^{n}\vartheta(\zeta+b_i,\Omega)}\quad\text{with }\sum_i a_i-\sum b_i\equiv 0\ (\mathrm{mod}\ \mathbf{Z}^h) ,$$

$$\frac{\partial}{\partial\zeta^I}\ln\frac{\vartheta(\zeta+a,\Omega)}{\vartheta(\zeta+b,\Omega)} ,$$

$$\frac{\partial^2}{\partial\zeta^I\partial\zeta^J}\ln\vartheta(\zeta,\Omega) ,$$

are periodic and hence functions on $J(M)$. To go further we need a detailed knowledge of the zero set of $\vartheta(\zeta,\Omega)$, and more precisely of its intersection with the image of the Abel map. Such information is provided by the *Riemann vanishing theorem.*

Let Δ, the "vector of Riemann constants," be defined by

$$\Delta_J=\tfrac{1}{2}-\tfrac{1}{2}\Omega_{JJ}+\sum_{K\neq J}\oint_{A_K}\omega_K(z)\int_{z_0}^{z}\omega_J \ . \qquad (6.37)$$

Then

- $2\Delta=I(K), K=$canonical bundle ;
- $\vartheta(\zeta,\Omega)=0$ if and only if $\zeta=\Delta-I(z_1+\cdots+z_{h-1})$ for any $h-1$ points $z_1,\ldots,z_{h-1}$ in M;
- $\vartheta(\zeta+I(z),\Omega)$ either vanishes identically as a function of z, or else has exactly h zeros $z_1,\ldots,z_h$ characterized by the equation

$$I(z_1+\cdots+z_h)=-\zeta+\Delta \ . \qquad (6.38)$$

The zero set of $\vartheta(\zeta,\Omega)$ is called the Θ divisor. Note that it is well defined as a subset of the Jacobian, thanks to its periodicity.

We pause to discuss briefly some ramifications of the Riemann vanishing theorem. The points ζ for which $\vartheta(\zeta+I(z),\Omega)=0$ as a function of z are rather special, and can be shown to coincide with points of the form

$$\zeta=\Delta-I(z_1+\cdots+z_h) \ ,$$

where $z_1+\cdots+z_h$ is a so-called *special divisor,* i.e., must contain all the poles of some nonconstant meromorphic function. A general divisor $w_1+\cdots+w_h$ in general will not satisfy this property, and the set of special ζ's above (6.38) is a strict analytic subvariety of the Jacobian. Points w for which hw is special are called the *Weierstrass points.* From the Riemann-Roch theorem, w is a Weierstrass point if and only if there exists a holomorphic Abelian differential vanishing to order h at w. Weierstrass points carry a lot of information about the complex structure of M. It is known that there are none in genus $h\leq 1$ and exactly $2h+2$ when M is hyperelliptic. In this case they can be viewed as the branch points of M, when represented as a double covering of the sphere. More generally a theorem of Hurwitz asserts that the number of Weierstrass points is between $2h+2$ and $h(h^2-1)$.

The last statement in the Riemann vanishing theorem provides an explicit answer to a question of Jacobi, namely, given ζ in $J(M)$, find h points $z_1,\ldots,z_h$ so that

$$I(z_1+\cdots+z_h)=\zeta \ . \qquad (6.39)$$

For generic ζ the desired points $z_1,\ldots,z_h$ are obtained simply by translating by $-\zeta+\Delta$ the image by the Abel map I of the Riemann surface M, and taking its intersections with the zero set Θ of the theta function. This invertibility of Eq. (6.39) for generic ζ is usually referred to as the Jacobi inversion theorem and will play a key role in the study of Bose-Fermi correspondence in Sec. VII.

Finally, the equation $2\Delta=I(K)$ suggests that Δ is intimately linked with the square root of K, in other words, bundles of spinors. We shall discuss this aspect in greater detail in the next section.

It is now easy to see why functions such as

$$\vartheta\left[\zeta+\int_w^z\omega,\Omega\right] \qquad (6.40)$$

will be the main ingredient in the construction of propagators. Indeed, if ζ is in the zero set of $\vartheta(\zeta,\Omega)$, then w must be among the h zeros $z_1,\ldots,z_h$ of this function. If, say, $w=z_h$, Eq. (6.38) will reduce to

$$\sum_{i=1}^{h-1}\int_{z_0}^{z_i}\omega_J\equiv-\zeta_J+\Delta_J \ , \qquad (6.41)$$

which shows that the points $z_1,\ldots,z_{h-1}$ are actually independent of w and depend on ζ alone. Since the function $\vartheta(\zeta,\Omega)$ is even, we can interchange the roles of z and w and conclude that there exist points $z_1,\ldots,z_{h-1},w_1,\ldots,w_{h-1}$ depending on ζ such that

$$\vartheta\left[\zeta+\int_w^z\omega,\Omega\right]=0\leftrightarrow\begin{cases}z=w,\\ \text{or } z \text{ is among } z_1,\ldots,z_{h-1},\\ \text{or } w \text{ is among } w_1,\ldots,w_{h-1}.\end{cases} \qquad (6.42)$$

Thus $\vartheta(\zeta+\int_w^z\omega,\Omega)$ has the key property of essentially vanishing only along the diagonal. It is still multiple valued as a function of z and w, but this difficulty can often be overcome as before, by taking suitable ratios.

As a simple illustration we can produce explicitly the meromorphic function with given divisor $z_1+\cdots+z_d-(w_1+\cdots+w_d)$ under the condition of Abel's theorem, i.e., that $I(z_1+\cdots+z_d)=I(w_1+\cdots+w_d)$. A candidate is

$$f(z)=\frac{\prod_{i=1}^{d}\vartheta\left[\zeta+\int_{z_i}^{z}\omega,\Omega\right]}{\prod_{i=1}^{d}\vartheta\left[\zeta+\int_{w_i}^{z}\omega,\Omega\right]} \ , \qquad (6.43)$$

where ζ is chosen so as not to have the functions involved vanish identically, $\vartheta(\zeta)=0$, and the paths of integration are yet to be described. A natural way of prescribing the paths is to choose one same path from z_0 to z, and link it to fixed paths α_i and β_i from z_i to z_0, and from w_i to z_0, respectively. Under changes of the path from z_0 to z the transformation laws (6.36) show that $f(z)$ may change by integral powers of

$$\exp\left[-2\pi i\sum_{i=1}^{d}\left[\int_{\alpha_i}\omega_J-\int_{\beta_i}\omega_J\right]\right] \ . \qquad (6.44)$$

By hypothesis the expression in the exponential is a lattice point in $\mathbf{Z}^h+\Omega\mathbf{Z}^h$ for any paths α_i,β_i from z_i and z_i to z_0. By adding, if necessary, appropriate multiples of the homology basis cycles A_J,B_J, we may make sure that it actually vanishes. Thus $f(z)$ is a single-valued meromorphic function on M and has exactly the desired zeros and poles.

Next, we should like to construct explicitly sections of any line bundle in the Jacobian of M. For this it is most convenient to introduce the *theta function with characteristics,*

$$\begin{aligned}\vartheta[\delta](\zeta,\Omega)=&\sum_{n\in Z^h}\exp[\pi i(n+\delta')\Omega(n+\delta')\\&+2\pi i(n+\delta')(\zeta+\delta'')]\\=&\exp[\pi i\delta'\Omega\delta'+2\pi i\delta'(\zeta+\delta'')]\\&\times\vartheta(\zeta+\Omega\delta'+\delta'',\Omega)\end{aligned}\tag{6.45}$$

for any *characteristics* $\delta=[{\delta' \atop \delta''}]$ in $(0,1)^{2h}$. It is readily seen that the transformation laws are

$$\begin{aligned}&\vartheta[\delta](\zeta+M+\Omega N,\Omega)=\exp[-\pi i N\Omega N-2\pi i N(\zeta+\delta'')\\&\qquad+2\pi i\delta' M]\vartheta[\delta](\zeta,\Omega)\ ,\\&\vartheta\begin{bmatrix}\delta'+M\\\delta''+N\end{bmatrix}(\zeta,\Omega)=\exp[2\pi i\delta' N]\vartheta\begin{bmatrix}\delta'\\\delta''\end{bmatrix}(\zeta,\Omega)\ .\end{aligned}\tag{6.46}$$

Given $[\delta]$, we can now construct three different objects that describe in different ways the same line bundle with zero Chern class: $-(\delta''+\Omega\delta')\in\mathbf{C}^h/\mathbf{Z}^h+\Omega\mathbf{Z}^h$; sections f defined by holonomy conditions $(\delta',\delta'')\in H^1(M,\mathbf{R})/H^1(M,\mathbf{Z})$ around the A_I,B_I cycles

$$\begin{aligned}&f\left[z+\int_{A_I}\right]=e^{2\pi i\delta'_I}f(z)\ ,\\&f\left[z+\int_{B_I}\right]=e^{-2\pi i\delta''_I}f(z)\ ;\end{aligned}\tag{6.47}$$

and

$$\frac{\vartheta[\delta]\left[\zeta+\int_{z_0}^z\omega,\Omega\right]}{\vartheta[0]\left[\zeta+\int_{z_0}^z\omega,\Omega\right]}\ .\tag{6.48}$$

To see that they indeed correspond to the same line bundle, it suffices to observe that the expressions in Eq. (6.48) transform as (6.47) when z is transported around each cycle A_I or B_I, so that (6.48) is a section of the line bundle with holonomy (δ',δ''). Furthermore, its divisor can be read off from the Riemann vanishing theorem: the zeros $z_1,\dots,z_h$ and poles $w_1,\dots,w_h$ must satisfy

$$\begin{aligned}&I(z_1+\cdots+z_h)=-\zeta-(\delta''+\Omega\delta')+\Delta\ ,\\&I(w_1+\cdots+w_h)=-\zeta+\Delta\ ,\end{aligned}\tag{6.49}$$

and hence

$$I(z_1+\cdots+z_h-w_1-\cdots-w_h)=-(\delta''+\Omega\delta')$$

as predicted.

In Sec. VI.B, we gave several equivalent descriptions of the Jacobian variety of M as a *set*, but it was not so clear exactly how to pass from one description to another for a specific element L of $J(M)$. Equations (6.46)–(6.48) provide a satisfying answer to this question, and allow us as well to characterize a line bundle in $J(M)$ by its characteristics $[\delta]$.

F. Spin structures, Dirac zero modes, and the prime form

The previous sections have provided a thorough investigation of line bundles with zero Chern class. Fermionic strings, however, involve spinors on the worldsheet M, i.e., sections of square roots of the canonical bundle K. Such square roots are called spin bundles and must have Chern class

$$c_1=\tfrac{1}{2}c_1(K)=h-1\ .$$

We have argued elsewhere (Sec. III.A) that there are 2^{2h} distinct spin bundles. They form a finite set inside the $(2h)$-dimensional Picard variety Pic_{h-1} of line bundles of Chern class $h-1$.

Now the Picard varieties Pic_d for different values of d are very similar in structure, but there is no natural correspondence between them without making some choices. One way of obtaining a correspondence is to single out a specific element within Pic_d, so that other elements of Pic_d can be identified by their differences from the chosen element. Since these differences must have vanishing Chern classes, this provides us with an isomorphism between Pic_d and the Jacobian variety.

It is remarkable that once a homology basis A_I,B_I has been chosen, we have in fact a particular spin structure S_0 determined by the basis. The key to this phenomenon lies in the fact that there is a natural correspondence between spin bundles and symmetric translates of the Θ divisor:

$$\begin{aligned}&S\text{ spin bundle}\leftrightarrow\text{translate of }\Theta\text{ divisor by }I(S)-\Delta\ ,\\&\{\text{spin bundles}\}\leftrightarrow\{\text{symmetric translates of }\Theta\}\ .\end{aligned}\tag{6.50}$$

Here by a symmetric subset of the Jacobian variety, we mean a subset invariant under $\zeta\to-\zeta$. To establish Eq. (6.50), we begin by noting that a line bundle admits holomorphic sections if its divisor is positive and, in particular for line bundles L of Chern class $h-1$, if its divisor is of the form $z_1+\cdots+z_{h-1}$. In view of the Riemann-Roch theorem, L will admit holomorphic sections if and only if $L^{-1}\otimes K$ does. In other words,

$$[L]=z_1+\cdots+z_{h-1}\Longrightarrow[L^{-1}\otimes K]=w_1+\cdots+w_{h-1}\ .$$

In particular, for each $z_1,\dots,z_{h-1}$, there exist $w_1,\dots,w_{h-1}$ so that

$$I(K)-I(z_1+\cdots+z_{h-1})=I(w_1+\cdots+w_{h-1})\ .$$

For a spin bundle S, $I(K)=2I(S)$, so this equation becomes

$$\begin{aligned}&I(S)-I(z_1+\cdots+z_{h-1})\\&\qquad=-[I(S)-I(w_1+\cdots+w_{h-1})]\ ,\end{aligned}$$

which just means that $\Theta+I(S)-\Delta$ is symmetric. Since $\vartheta(\zeta,\Omega)$ is an even function, the Θ divisor itself is symmetric. Furthermore, it is not difficult to show that the only way of obtaining a symmetric translate of Θ is actually to translate it by half-lattice points, i.e., points of the form $-(\delta''+\Omega\delta')$ where δ',δ'' are half-integers. There are thus exactly 2^{2h} symmetric translates, so the above correspondence is one-to-one and onto. In particular, to the Θ divisor itself must correspond some specific spin bundle S_0, and this is the one we are looking for. Note that it satisfies $I(S_0)=\Delta$, but depends only on the homology basis, not on the choice of base point P_0.

With the choice of the spin bundle S_0 we can identify the Jacobian variety and the Picard variety Pic_{h-1} via

$$\begin{bmatrix}\delta'\\ \delta''\end{bmatrix}\in J(M)\leftrightarrow S_0\otimes\begin{bmatrix}\delta'\\ \delta''\end{bmatrix}\in \mathrm{Pic}_{h-1}\ ,$$

while spin bundles $S_0\otimes[^{\delta'}_{\delta''}]$ within the Jacobian are given by $(S_0\otimes[^{\delta'}_{\delta''}])^2=K$. This means that $[^{\delta'}_{\delta''}]$ must be a half-integer point, in agreement with the discussion based on symmetric translates of the Θ divisor. Observe that, for spin bundles, the theta function with characteristics satisfies

$$\vartheta[\delta](-\zeta,\Omega)=(-1)^{4\delta'\delta''}\vartheta[\delta](\zeta,\Omega)\ , \tag{6.51}$$

which shows that spin structures can be divided into two groups, the even and the odd ones, depending on whether $4\delta'\delta''$ is an even or an odd integer. Simple counting yields $2^{h-1}(2^h+1)$ even-spin and $2^{h-1}(2^h-1)$ odd-spin structures. This parity will reflect itself in the number of zero modes in the Dirac operator.

We can now characterize within Pic_{h-1} those line bundles which admit holomorphic sections. Of course, when we have a spin bundle, the Dirac operator reduces to the $\bar\partial$ operator, and holomorphic sections are simply Dirac zero modes. The description is actually very simple: a line bundle $S_0\otimes[\delta]$ in Pic_{h-1} admits zero modes if and only if its divisor is of the form $z_1+\cdots+z_{h-1}$ for some points $z_1,\ldots,z_{h-1}$. Taking the Abel map and recalling that $I(S_0)=\Delta$, $I([\delta])=-(\delta''+\Omega\delta')$ [cf. Eqs. (6.47) and (6.48)], we obtain

$$-(\delta''+\Omega\delta')+\Delta-I(z_1+\cdots+z_{h-1})=0\ .$$

In other words, $\delta''+\Omega\delta'$ is in the Θ divisor,

$$\vartheta[\delta](0,\Omega)=0\ ,$$

in view of the Riemann vanishing theorem.

This characterization suggests strongly that the number of zero modes is just the order of vanishing of the theta function. This is, for example, in the same spirit as Selberg zeta-function-type formulas derived earlier for regularized determinants and can actually be proved with further work. In particular, for spin bundles it confirms that the parity of the number of zero modes is the same as the parity of the spin structure, and that generically there is no zero mode for the even-spin structures and exactly one for the odd ones.

We turn next to the remaining fundamental ingredient in the construction of chiral fields on a Riemann surface, namely the prime form.

Let $[\delta]$ be an odd-spin structure and assume that we are in the generic case where $\vartheta[\delta](\zeta,\Omega)$ vanishes exactly to first order at $\zeta=0$. This means that the Dirac operator has exactly one zero mode, which we can construct explicitly. For this consider the holomorphic Abelian differential

$$\omega_\delta(w)=\sum_{I=1}^{h}\frac{\partial\vartheta}{\partial\zeta^I}[\delta](0,\Omega)\omega_I(w)\ . \tag{6.52}$$

We claim that it vanishes to second order at $(h-1)$ points $z_1,\ldots,z_{h-1}$ and that these points are determined by $I(z_1+\cdots+z_{h-1})=\Delta-\delta$. To see this, let $w,z_1,\ldots,z_{h-1}$ be the h zeros of the function $\vartheta[\delta](\int_w^z\omega,\Omega)$. Thus Riemann's theorem implies that the stated relation holds, and in particular $z_1,\ldots,z_{h-1}$ are independent of w. Taking the differential of $\vartheta[\delta](\int_w^z\omega,\Omega)$ with respect to w at $z=w$ yields $\omega_\delta(w)$, which must then vanish at $z_1,\ldots,z_{h-1}$. Since ω_δ is an Abelian differential, its divisor is the divisor of the canonical bundle. This fact together with Eq. (6.24) readily implies that the missing $(h-1)$ zeros of ω_δ are again $z_1,\ldots,z_{h-1}$, which is the desired statement. Since the zeros of ω_δ are double, the spinor

$$h_\delta(w)=\sqrt{\omega_\delta(w)} \tag{6.53}$$

is well defined and holomorphic, and in fact is a section of the spin bundle corresponding to δ. We may now generalize the construction (6.40) to obtain the prime form

$$E(z,w)=\frac{\vartheta[\delta]\left[\int_w^z\omega,\Omega\right]}{h_\delta(z)h_\delta(w)}\ . \tag{6.54}$$

The prime form $E(z,w)$ can be viewed as a $(-\frac{1}{2},0)$ form in each variable on the universal covering of the surface M, whose transformation laws can be easily read off from Eq. (6.46). The introduction of the factors $h_\delta(z)h_\delta(w)$ in Eq. (6.54) has several beneficial effects: $E(z,w)$ has the correct U(1) weight for inverses of fermion propagators, is actually independent of the spin structure δ we selected originally, and vanishes only when $z=w$.

When the point z is moved around an A_I cycle once, E is left invariant up to a $\pm$ sign, whereas when it is moved around a B_I cycle one, it transforms as

$$E(z,w)\to -\exp\left[-i\pi\Omega_{II}-2\pi i\int_z^w\omega_I\right]E(z,w)\ . \tag{6.55a}$$

It is important to note that the prime form depends on the choice of homology basis and will transform under modular transformations as

$$E(z,w)\to\exp\left[\pi i\int_z^w\omega(C\Omega+D)^{-1}C\int_z^w\omega\right]E(z,w)\ . \tag{6.55b}$$

Meromorphic differentials, whose existence was established in Sec. VI.C through indirect index theorems arguments, can be written very simply in terms of the prime form. In fact,

$$\omega_{w_1 w_2}(z)=d_z \ln \frac{E(z,w_1)}{E(z,w_2)} \tag{6.56}$$

is a differential of the third kind with zero A periods and residues ± 1 at w_1 and w_2, while

$$\omega_{w_1}(z)=d_z \frac{\partial}{\partial w}\bigg|_{w=w_1} \ln E(z,w) \tag{6.57}$$

is a differential of the second kind with zero A periods and a double pole at w_1. Similarly, propagators can be constructed out of the prime form, but we shall return to this issue later.

All variations with respect to moduli parameters can be deduced from the following variational formulas for the Abelian differentials and the prime form:

$$\begin{aligned} &\frac{4\pi}{\sqrt{g}}\frac{\delta}{\delta g^{ww}}\omega_I(z)=\omega_I(w)\partial_z\partial_w \ln E(z,w)\ , \\ &\frac{4\pi}{\sqrt{g}}\frac{\delta}{\delta g^{ww}}E(z,x)=-\tfrac{1}{2}[\omega_{zx}(w)]^2\ . \end{aligned} \tag{6.58}$$

Finally, we discuss degenerations of Riemann surfaces in terms of plumbing fixtures (see Sec. IV.H), period matrices, and prime forms. Recall that there are two cases, distinguished by whether the plumbing fixture $\mathcal{U}_t$ used to model the degeneration process disconnects the Riemann surfaces M_t or not. In the case where it does, let M_1' and M_2' be the components of the complement in M_t of the plumbing fixture $\mathcal{U}_t$, and let M_1 and M_2 be the components of M_0 in the degeneration limit. Then the normalized basis of Abelian differentials $\omega_I^t(z)$ will approach the combined bases of Abelian differentials $\omega_{I_1}^t(z_1)$, $1\le I_1 \le i$, $\omega_{I_2}(z_2)$, $i+1\le I_2\le h$, of the surfaces M_1 and M_2. More precisely, we have

$$\omega_{I_1}^t(z)=\begin{cases}\omega_{I_1}(z)+\frac{1}{4}t\omega_{I_1}(p_1)\omega_{p_1}^1(z)+O(t^2) & \text{for } z\in M_1',\\ \frac{1}{4}t\omega_{I_1}(p_1)\omega_{p_2}^2(z)+O(t^2) & \text{for } z\in M_2',\end{cases} \tag{6.59}$$

and similarly for $\omega_{I_2}(z,t)$ with the roles of M_1 and M_2 interchanged. Here $\omega_w^i(z)$ are the Abelian differentials of the second kind on M_i, with double pole at w [cf. Eq. (6.57)]. The terms $O(t^2)$ are holomorphic differentials whose limits $t^{-2}O(t^2)$ may have a pole of order at most 4 at p_1 and p_2. Integrating over basis cycles gives the asymptotic behavior of the period matrix of M_t,

$$\Omega(t)=\begin{bmatrix}\Omega_1 & 0\\ 0 & \Omega_2\end{bmatrix}+O(t)\ , \tag{6.60}$$

where Ω_i are the period matrices of M_i. Next, if D_i is a divisor on M_t with zero Chern class, which decomposes into $D_t=D_1+D_2+D_{\mathcal{U}}$ with D_1, D_2, and $D_{\mathcal{U}}$ divisors on M_1', M_2', and $\mathcal{U}_t$ of degrees d_1, d_2, and $d_{\mathcal{U}}$, respectively, then the theta function will factorize as

$$\vartheta(D_t,\Omega)\to\vartheta(D_1-d_1p_1,\Omega_1)\vartheta(D_2-d_2p_2,\Omega_2)\ . \tag{6.61}$$

In particular, the Riemann class factorizes as $\Delta(t)\to\Delta_1+\Delta_2+p_{1,2}$. As for the prime form $E(z,w)$, it will behave as $z-w$ when both z and w are in the plumbing fixture, and otherwise

$$\begin{aligned} &E(z_1,w)\to E_1(z_1,p_1)wt^{-3/4}\ ,\\ &E(z_2,w)\to E_2(z_2,p_2)t^{-1/4}\ ,\\ &E(z_1,z_2)\to E_1(z_1,p_1)E_2(p_2,z_2)t^{-1/2}\ ,\end{aligned} \tag{6.62}$$

for $z_i\in M_i'$ and w in the plumbing fixture. Here of course the E_i's denote the prime forms of the surfaces M_i.

In the case in which the complement M_t' of the fixture remains connected [Fig. 19(b)], the normalized basis of h Abelian differentials $\omega_I^t(z)$ for the degenerating surface M_t will approach the normalized basis ω_I, $I=1,\ldots,h-1$ for the limiting surface M, while $\omega_h^t(z)$ will tend to the Abelian differential of the third kind $\omega_{p_1p_2}(z)$ [Eq. (6.56)], with poles at $z=p_1$ and p_2. On M_t' one can give precise asymptotes,

$$\begin{aligned} \omega_I^t(z)&=\omega_I(z)+\tfrac{1}{4}t[\omega_I(p_1)-\omega_I(p_2)]\\ &\qquad\times[\omega_{p_1}^1(z)-\omega_{p_2}^2(z)]+O(t^2)\ ,\\ &\qquad\qquad I=1,\ldots,h-1\ ,\\ \omega_h^t(z)&=\omega_{p_1p_2}(z)+t\tilde{\omega}_h(z)+O(t^2)\ , \end{aligned} \tag{6.63}$$

where $\tilde{\omega}_h$ is a meromorphic differential with poles at p_1 and p_2 of order three. It follows that the period matrix of M_t can be written as

$$\Omega_t=\begin{bmatrix}\Omega_{IJ}(M) & \int_{p_1}^{p_2}\omega_I\\ \int_{p_1}^{p_2}\omega_J & \ln t+\text{const}\end{bmatrix}+O(t)\ , \tag{6.64}$$

with $1\le I,J\le h-1$. The asymptotics of theta functions and the prime form can now be derived in analogy with the previous case.

A detailed discussion of theta functions, the prime form, and their degeneration is to be found in Fay (1973). In the physics literature, the characterization of line bundles with holomorphic sections by the theta divisor appears in Alvarez-Gaumé, Moore, and Vafa (1986).

VII. HOLOMORPHIC STRUCTURE OF STRINGS

A fundamental principle underlying theories of closed oriented strings is that massless fields in two dimensions decompose into independent left- and right-movers. The independence is maintained at the interacting level, since the action is still the free action, and the presence of interactions is only indicated by the topology of the worldsheet. This principle is crucial in the construction of fermionic strings: in the type-II string we have to separate the contributions of left- and right-movers to as-sign them independent spin structures, while in the heterotic string we have to amalgamate the left-movers of the fermionic string with the right-movers of the bosonic string. For bosonic strings, separation of the left- and right-movers is not required, but it should remain a useful property of the partition function. A careful treatment of this chiral splitting and of the related issue of internal loop momenta has been provided in Secs. III.K and III.O.

A key observation due to Belavin and Knizhnik (1986) is that separation of left- and right-movers on the worldsheet can be translated into holomorphicity of the string integrand on moduli space. In fact, if Z is the partition function of a conformal field theory with respect to a background metric $ds^2=\rho\, dz\, d\bar{z}$ and we deform the background metric by a Beltrami differential μ, then

$$\delta_\mu \delta_{\bar{\mu}} \ln Z = \left[\frac{1}{4\pi}\right]^2 \int d^2z \sqrt{g}\, g^{z\bar{z}} \times \int d^2w \sqrt{g}\, g^{w\bar{w}} \mu_{\bar{z}}{}^{z} \bar{\mu}_{w}{}^{\bar{w}} \langle T_{zz} T_{\bar{w}\bar{w}} \rangle_{\text{conn}} .$$

Thus the vanishing of $\langle T_{zz} T_{\bar{w}\bar{w}} \rangle_{\text{conn}}$ would imply that Z is the absolute value squared of a holomorphic function on moduli space.

Anomalies in principle could spoil this picture. Recall (Secs. II.I and II.J) that the bosonic string is built out of the conformal systems of the matter fields x^μ, $\mu=1,\ldots,d$ and ghost fields b,c. If we consider, say, the x^μ fields alone, reparametrization invariance and separation of left- and right-movers (in Euclidean signature, holomorphic and antiholomorphic) cannot be achieved simultaneously, the obstruction being the nonvanishing central charge in the Virasoro algebra. This means that $\langle T_{zz} T_{\bar{w}\bar{w}} \rangle_{\text{conn}}$ develops a Schwinger term that prevents the vanishing of $\delta\bar{\delta} \ln Z$. The same is true for the isolated b,c system. For the combined x^μ,b,c system, however, the anomalies should cancel in $d=26$, and it is after cancellation that the string partition function should split into a holomorphic factor times its antiholomorphic conjugate on moduli space. Earlier expressions for the bosonic string such as (2.145) should be understood in this sense.

A complete analysis of the second variation $\delta\bar{\delta} \ln Z$ was carried out by Belavin and Knizhnik (1986). Besides justifying the above principles, their results also provide a basis for investigating the holomorphic structure on moduli space of the conformal field theories encountered earlier in Secs. II.I and II.J. In particular, they can be a starting point for a detailed study on higher-genus surfaces of the Bose-Fermi correspondence of two-dimensional field theory.

In Sec. VII.A we provide an exposition of the holomorphic anomaly formula of Belavin and Knizhnik, based on heat-kernel regularization. This leads to their characterization of the string partition function as the unique (up to constants) holomorphic nonvanishing section of a line bundle over moduli space. Sections VII.C and VII.D are devoted to bosonization, following Verlinde and Verlinde. They culminate in complete expressions for correlation functions of bosons and chiral fermions in terms of the prime form. In Sec. VII.E a geometric interpretation of the holomorphic anomaly is given in terms of curvature of determinant line bundles. This refines the Atiyah-Singer (1984) interpretation of chiral anomalies as nontriviality of these bundles. There the bundle was that of Dirac operators over the space of vector potentials modulo gauge transformations. Here it is the bundle of $\bar{\partial}$ operators over moduli space. The key new feature is the existence of a new metric built out of regularized determinants, the Quillen (1984) metric, so that nontriviality of the line bundle can be measured at the level of differential forms (rather than Chern classes) by the curvature and holonomy of its connection.

The determinants for Dirac and gauge-fixing operators obtained this way in terms of theta functions encode nicely their dependence on spin structures. They also allow a simple study of degeneration behavior. However, the resulting expressions for string scattering amplitudes are still somewhat formal, since they require a convenient parametrization of period matrices within the Siegel upper half space. It is possible that recent solutions of the Schottky problem based on the KP hierarchy may be helpful in this context, but the issue has not been fully explored as yet.

In Sec. VII.F we investigate the superholomorphic structure of superstrings, following D'Hoker and Phong (1987b). There is indeed a superholomorphic anomaly, which cancels in the critical dimension $d=10$ and for the heterotic string with rank 16 gauge groups. Thus we may hope that the superholomorphic structure of supermoduli space will impose powerful constraints on the superstring. The full consequences will require a better understanding of superalgebraic geometry, which is being developed by many authors. Finally, in Sec. VII.G we provide a detailed comparison of chiral splitting, holomorphic splitting, and holomorphic splitting at fixed internal momenta. A crucial ingredient in this comparison is a supersymmetric extension $\hat{\Omega}$ of the period matrix Ω. One of the major difficulties encountered in multiloop amplitudes has been the fact that supermoduli space does not seem to have a natural projection onto moduli space. The existence of $\hat{\Omega}$ indicates that such a projection exists if we represent supermoduli by $\{\hat{\Omega},\chi\}$ and moduli by $\{\hat{\Omega}\}$, though it need not coincide with the standard idea of split supermanifolds. The matrix $\hat{\Omega}$ may ultimately be the way to express superstring amplitudes in terms of modular forms.

A. Holomorphic anomalies

There is a simple way of viewing the holomorphic anomalies we shall discuss in this section as chiral anomalies. In fact, if we wish to consider the chiral version of the fermionic theories $b(dz)^n$, $c(dz)^{1-n}$ of Sec. II.J, quantization will demand a suitable notion of a determinant for the chiral operator ∇_n^z. Now determinants of chiral operators make sense only as sections of line bundles, as we shall see in Sec. VII.E. To obtain a scalar we could try instead to construct an appropriate square root for the nonchiral determinant of $\nabla_n^{z\dagger}\nabla_n^z=\Delta_n^{(-)}$. The phases of such square roots are arbitrary, however, and can only be determined by requiring further that the dependence on moduli of $\det\nabla_n^z$ mimic that of ∇_n^z itself. To understand this dependence, let

$$ds^2=\rho\,|\,dz+\mu\,d\bar{z}\,|^2 \tag{7.1}$$

parametrize deformations of a fixed conformal structure $ds^2=\rho\,|\,dz\,|^2$. The corresponding deformation of ∇_n^z is

$$\delta\nabla_n^z=\mu_{\bar{z}}{}^z\nabla^{\bar{z}}+n\,\nabla^{\bar{z}}\mu_{\bar{z}}{}^z\ . \tag{7.2}$$

Since $\mu_{\bar{z}}{}^z$ constitutes the holomorphic coordinates for moduli space near $\rho\,|\,dz\,|^2$, the fact that $\delta\nabla_n^z$ depends only on $\mu_{\bar{z}}{}^z$ and not $\bar{\mu}_z{}^{\bar{z}}$ means that ∇_n^z depends holomorphically on moduli parameters. Thus a chiral theory of b,c fermions requires a reparametrization-invariant, holomorphic square root of $\det'\Delta_n^{(-)}$, with suitable modifications necessitated by absorption of zero modes.

If we choose to maintain manifest reparametrization invariance, say by a heat-kernel regularization, we shall see that we cannot extract a holomorphic square root on moduli space, as may naively have been expected from the previous discussion. A local "holomorphic anomaly" is measured by

$$\delta_\mu\delta_{\bar{\mu}}\ln\left[\frac{\det'\Delta_n^{(\pm)}}{\det\langle\phi_a\,|\,\phi_b\rangle\det\langle\psi_a\,|\,\psi_b\rangle}\right] \tag{7.3}$$

and has a very similar structure to the conformal anomaly. We turn now to its evaluation. We shall work with Lorentz-covariant derivatives D_z^n, $D_{\bar{z}}^n$ on tensors of weight n,

$$D_z^n=e_z^m(\partial_m+in\,\omega_m)\ ,$$
$$\Delta_n^{(+)}=-2D_{\bar{z}}^{n+1}D_z^n\ ,$$
$$\Delta_n^{(-)}=-2D_z^{n-1}D_{\bar{z}}^n\ ,$$

instead of the covariant derivatives ∇_n^z, since this setup is more convenient for the generalization to superholomorphic anomalies in Sec. VII.F. Recall that determinants are defined by

$$\ln\det'\Delta_n^{(+)}=-\int_\varepsilon^\infty\frac{dt}{t}(\mathrm{tr}e^{-t\Delta_n^{(+)}}-N_n^+)\ , \tag{7.4}$$

and a change with respect to μ produces

$$\delta_\mu\ln\det'\Delta_n^{(+)}=\int_\epsilon^\infty dt\ \mathrm{tr}(\delta_\mu\Delta_n^{(+)}e^{-t\Delta_n^{(+)}})\ . \tag{7.5}$$

Now the operators $\Delta_n^{(+)}$ and $\Delta_{n+1}^{(-)}$ are not in general invertible on the entire function spaces of rank n and $n+1$ tensors, so it is appropriate to single out their kernels by introducing the projection operators

$$\begin{aligned}\Pi_n^+&\equiv 1+2D_{\bar{z}}^{n+1}(\Delta_{n+1}^{(-)})^{-1}D_z^n\quad\text{onto }\mathrm{Ker}D_z^n\ ,\\ \Pi_{n+1}^-&\equiv 1+2D_z^n(\Delta_n^{(+)})^{-1}D_{\bar{z}}^{n+1}\quad\text{onto }\mathrm{Ker}D_{\bar{z}}^{n+1}\ .\end{aligned} \tag{7.6}$$

Hence we decompose the trace as follows:[38]

$$\begin{aligned}\mathrm{tr}\delta_\mu\Delta_n^{(+)}e^{-t\Delta_n^{(+)}}&=-2\,\mathrm{tr}\delta_\mu D_{\bar{z}}D_z e^{-t\Delta_n^{(+)}}\\ &\quad-2\,\mathrm{tr}\delta_\mu D_z D_{\bar{z}}e^{-t\Delta_{n+1}^{(-)}}\\ &=-2\,\mathrm{tr}\delta_\mu D_{\bar{z}}D_z e^{-t\Delta_n^{(+)}}(1-\Pi_n^+)\\ &\quad-2\,\mathrm{tr}\delta_\mu D_z D_{\bar{z}}e^{-t\Delta_{n+1}^{(-)}}(1-\Pi_{n+1}^-)\ .\end{aligned} \tag{7.7}$$

Since, on the space complementary to their kernel, the Laplace operators are now invertible, this equals

$$\begin{aligned}&2\frac{\partial}{\partial t}\mathrm{tr}\delta_\mu D_{\bar{z}}D_z(\Delta_n^{(+)})^{-1}e^{-t\Delta_n^{(+)}}(1-\Pi_n^+)\\ &\qquad+2\frac{\partial}{\partial t}\mathrm{tr}\delta_\mu D_z D_{\bar{z}}(\Delta_{n+1}^{(-)})^{-1}e^{-t\Delta_{n+1}^{(-)}}(1-\Pi_{n+1}^-)\ .\end{aligned} \tag{7.8}$$

Thus

$$\begin{aligned}\delta_\mu\ln\det'\Delta_n^{(+)}&=-2\,\mathrm{tr}\delta_\mu D_{\bar{z}}D_z(\Delta_n^{(+)})^{-1}\\ &\qquad\times e^{-\varepsilon\Delta_n^{(+)}}(1-\Pi_n^+)\\ &\quad-2\,\mathrm{tr}\delta_\mu D_z D_{\bar{z}}(\Delta_{n+1}^{(-)})^{-1}\\ &\qquad\times e^{-\varepsilon\Delta_{n+1}^{(-)}}(1-\Pi_{n+1}^-)\ .\end{aligned} \tag{7.9}$$

Then we evaluate the second derivative with respect to $\bar{\mu}$. Two formulas come in handy:

$$\delta_{\bar{\mu}}(1-\Pi_n^+)=-2D_{\bar{z}}(\Delta_{n+1}^{(-)})^{-1}\delta_{\bar{\mu}}D_z\Pi_n^++O(\delta_{\bar{\mu}}D_{\bar{z}})\ , \tag{7.10}$$

$$\delta_{\bar{\mu}}(1-\Pi_{n+1}^-)=-2\Pi_{n+1}^-\delta_{\bar{\mu}}D_z(\Delta_n^{(+)})^{-1}D_{\bar{z}}+O(\delta_{\bar{\mu}}D_{\bar{z}})\ ,$$

as well as

$$\delta e^A=\int_0^1 du\ e^{uA}\delta A e^{(1-u)A}\ . \tag{7.11}$$

With the help of these and some straightforward algebra, one finds

$$\begin{aligned}\delta_{\bar{\mu}}\delta_\mu\ln\det'\Delta_n^{(+)}=&-2\,\mathrm{tr}(\delta_{\bar{\mu}}D_{\bar{z}})D_z(\Delta_n^{(+)})^{-1}e^{-\varepsilon\Delta_n^{(+)}}(1-\Pi_n^+)-2\,\mathrm{tr}(\delta_{\bar{\mu}}\delta_\mu D_z)D_{\bar{z}}(\Delta_{n+1}^{(-)})^{-1}e^{-\varepsilon\Delta_{n+1}^{(-)}}(1-\Pi_{n+1}^-)\\ &-2\,\mathrm{tr}\delta_\mu D_{\bar{z}}\Pi_{n+1}^-\delta_{\bar{\mu}}D_z(\Delta_n^{(+)})^{-1}(1-\Pi_n^+)-2\,\mathrm{tr}\delta_\mu D_{\bar{z}}(1-\Pi_{n+1}^-)(\Delta_{n+1}^{(-)})^{-1}\delta_{\bar{\mu}}D_z\Pi_n^+\\ &+2\varepsilon\int_0^1 du\ \mathrm{tr}\delta_\mu D_{\bar{z}}e^{-\varepsilon u\Delta_{n+1}^{(-)}}\delta_{\bar{\mu}}D_z e^{-\varepsilon(1-u)\Delta_n^{(+)}}\ .\end{aligned} \tag{7.12}$$

[38]For brevity, we shall denote $D_z=D_z^n$ and $D_{\bar{z}}=D_{\bar{z}}^{n+1}$.

Here we have made use of the fact that within finite-dimensional traces, the heat kernel for short time ε reduces to the identity operator. The presence of nonlocal contributions in the first four terms on the right-hand side reminds us of the fact that it is natural to work with determinants divided by normalizations of zero modes, as in the case of the Weyl anomaly. The changes of the finite-dimensional determinants under δ_μ and $\delta_{\bar\mu}$ are obtained as follows. Let ϕ_j span a basis for $\mathrm{Ker} D_{\bar z}^{n+1}$ and ψ_a a basis for $\mathrm{Ker} D_z^n$. It is easy to show that

$$\delta_{\bar\mu}\delta_\mu \ln\det\langle\phi_j\,|\,\phi_k\rangle = 2\langle\delta_{\bar\mu}\delta_\mu\phi_j\,|\,\phi_j\rangle + \langle\delta_{\bar\mu}\phi_j\,|\,(1-\Pi_{n+1}^-)\,|\,\delta_\mu\phi_j\rangle\ .$$

From $D_{\bar z}^{n+1}\phi_j=0$, we deduce

$$\delta_\mu D_{\bar z}^{n+1}\phi_j + D_{\bar z}^{n+1}\delta_\mu\phi_j = 0 \tag{7.13}$$

and hence

$$(1-\Pi_{n+1}^-)\,|\,\delta_\mu\phi_j\rangle = 2D_z(\Delta_n^{(+)})^{-1}\delta_\mu D_{\bar z}^{n+1}\,|\,\phi_j\rangle\ , \tag{7.14}$$

so that

$$\delta_{\bar\mu}\delta_\mu\ln\det\langle\phi_j\,|\,\phi_k\rangle = 2\langle\delta_{\bar\mu}\delta_\mu\phi_j\,|\,\phi_j\rangle - 2\,\mathrm{tr}\Pi_{n+1}^-\delta_{\bar\mu}D_z^n(1-\Pi_n^+)(\Delta_n^{(+)})^{-1}\delta_\mu D_{\bar z}^{n+1} \tag{7.15}$$

and similarly

$$\delta_{\bar\mu}\delta_\mu\ln\det\langle\psi_a\,|\,\psi_b\rangle = 2\langle\delta_{\bar\mu}\delta_\mu\psi_a\,|\,\psi_a\rangle - 2\,\mathrm{tr}\Pi_n^+\delta_\mu D_{\bar z}^{n+1}(1-\Pi_{n+1}^-)(\Delta_{n+1}^{(-)})^{-1}\delta_{\bar\mu}D_z^n\ . \tag{7.16}$$

We may recast Eq. (7.12) in the form

$$\begin{aligned}\delta_{\bar\mu}\delta_\mu\ln\frac{\det'\Delta_n^{(+)}}{\det\langle\phi_j\,|\,\phi_k\rangle\det\langle\psi_a\,|\,\psi_b\rangle} &= -2\,\mathrm{tr}(\delta_{\bar\mu}\delta_\mu D_{\bar z}^{n+1})D_z^n(\Delta_n^{(+)})^{-1}e^{-\varepsilon\Delta_n^{(+)}}(1-\Pi_n^+)\\ &\quad -2\,\mathrm{tr}(\delta_{\bar\mu}\delta_\mu D_z^n)D_{\bar z}^{n+1}(\Delta_{n+1}^{(-)})^{-1}e^{-\varepsilon\Delta_{n+1}^{(-)}}(1-\Pi_{n+1}^-)\\ &\quad -2\langle\delta_{\bar\mu}\delta_\mu\phi_j\,|\,\phi_j\rangle - 2\langle\delta_{\bar\mu}\delta_\mu\psi_a\,|\,\psi_a\rangle\\ &\quad +2\varepsilon\int_0^1 du\ \mathrm{tr}\delta_\mu D_{\bar z}^{n+1}e^{-\varepsilon u\Delta_{n+1}^{(-)}}\delta_{\bar\mu}D_z^n e^{-\varepsilon(1-u)\Delta_n^{(+)}}\ .\end{aligned} \tag{7.17}$$

The next crucial observation is that the only way the operators and zero modes can depend on μ and $\bar\mu$ simultaneously is through a conformal change, as we indicated when we first wrote down the corresponding differential operators. Denoting this Weyl scaling by $\delta\sigma$, we have

$$\begin{aligned}\delta_{\bar\mu}\delta_\mu D_{\bar z}^{n+1} &= -(n+2)\delta\sigma D_{\bar z}^{n+1} + (n+1)D_{\bar z}^{n+1}\delta\sigma\ ,\\ \delta_{\bar\mu}\delta_\mu D_z^n &= (n-1)\delta\sigma D_z^n - nD_z^n\delta\sigma\ ,\end{aligned} \tag{7.18}$$

and correspondingly

$$\begin{aligned}\delta_{\bar\mu}\delta_\mu\phi_j &= -(n+1)\delta\sigma\phi_j\ ,\\ \delta_{\bar\mu}\delta_\mu\psi_a &= n\,\delta\sigma\psi_a\ .\end{aligned} \tag{7.19}$$

With the help of these, we see that the first four terms on the right-hand side of Eq. (7.17) reduce to

$$\begin{aligned} n\,\mathrm{tr}\delta\sigma e^{-\varepsilon\Delta_{n+1}^{(-)}} &- (n+1)\mathrm{tr}\delta\sigma e^{-\varepsilon\Delta_n^{(+)}}\\ &= -\frac{1}{4\pi\varepsilon}\int d^2\xi\sqrt{g}\,\delta\sigma\\ &\quad -\frac{6n^2+6n+1}{12\pi}\int d^2\xi\sqrt{g}\,R\,\delta\sigma + O(\varepsilon)\ ,\end{aligned} \tag{7.20}$$

which is precisely the effect of the conformal anomaly on the determinant of $\Delta_n^{(+)}$ [see Eq. (2.69)]. For a deformation of the form (7.1) $\delta\sigma=\bar\mu\mu$. Thus the only term that remains is the integral over u in Eq. (7.17), and this term is both local on the worldsheet and, as can be seen from its definition, reparametrization invariant. Thus we may evaluate it locally on the Riemann surface, its reparametrization invariance guaranteeing that these local contributions will fit together consistently. Since we work only up to a Weyl anomaly, we can in fact work around flat space, and the calculation is then easily performed. Putting all together, one finds

$$\begin{aligned}\delta_{\bar\mu}\delta_\mu\ln\frac{\det'\Delta_n^{(\pm)}}{\det\langle\phi_j\,|\,\phi_k\rangle\det\langle\psi_a\,|\,\psi_b\rangle}\\ = -\frac{6n^2\pm 6n+1}{12\pi}\int d^2\xi\sqrt{g}\,(\nabla_z\bar\mu\nabla^z\mu + 2R\mu\bar\mu)\ .\end{aligned} \tag{7.21}$$

We note that this second variation corresponds to the particular choice of variations $\rho\,|\,dz\,|^2\to\rho\,|\,dz+\mu\,d\bar z\,|^2$. Clearly we can accompany this variation with any additional Weyl scaling without changing the complex structure, so strictly speaking the holomorphic anomaly is not intrinsic and must be considered modulo the conformal anomaly.

We can now give a complex analytic characterization of the bosonic string partition function. In view of Eq. (7.21), the function

$$F=\left[\frac{\det'\Delta_0}{\int d^2\xi\sqrt{g}\,\det\langle\omega_I\mid\omega_J\rangle}\right]^{-13}\left[\frac{\det'\Delta_2^{(-)}}{\det\langle\phi_a\mid\phi_b\rangle}\right] \tag{7.22a}$$

is the square modulus of a holomorphic function on moduli space, as long as the Abelian differentials ω_I and quadratic differentials ϕ_a are chosen to depend holomorphically on moduli parameters. Stated as generally as that, such choices are not possible globally on moduli space. A weaker choice is, however, possible, which is dictated by the structure of Eq. (7.22) and suffices for our purposes. Let K and Λ be the maximum wedge powers of the spaces of quadratic differentials and Abelian differentials, respectively. Since these spaces vary holomorphically with moduli, they should be viewed as making two holomorphic line bundles over moduli space. Given a holomorphic section s of $K\otimes\Lambda^{-13}$, we can write it locally as

$$s=(\phi_1\wedge\cdots\wedge\phi_{3h-3})\otimes(\omega_1\wedge\cdots\wedge\omega_h)^{-13}\ , \tag{7.22b}$$

and the function F in Eq. (7.22a) depends only on s and not on the particular factorization into ϕ_a and ω_I. We can now apply a theorem of Mumford (1977) which guarantees the existence of the weaker choice we referred to earlier, namely, that of a global nowhere-vanishing holomorphic section s of $K\otimes\Lambda^{-13}$. In other words, neither line bundle K nor Λ is trivial over moduli space, but $K\otimes\Lambda^{-13}$ is. If s is a global section of $K\otimes\Lambda^{-13}$, the function F will be globally defined on moduli space and hence must be constant. Writing s as in Eq. (7.22b), we note that $\det\langle\omega_I\mid\omega_J\rangle^{-13}\ \phi_1\wedge\cdots\wedge\bar{\phi}_{3h-3}$ is now a well-defined global $(6h-6)$ volume form over moduli space, which coincides in local coordinates with the measure $[dm]\det\langle\mu_j\mid\phi_k\rangle\det\langle\omega_I\mid\omega_J\rangle^{-13}$ of Sec. II.G. Thus, up to a multiplicative constant c, the bosonic string partition function can be rewritten as

$$Z=c\int_{\mathcal{M}_h}\phi_1\wedge\cdots\wedge\bar{\phi}_{3h-3}\det\langle\omega_I\mid\omega_J\rangle^{-13}\ , \tag{7.23}$$

a formula that is manifestly conformally invariant.

The line bundles Λ and K are usually called, respectively, the Hodge bundle and the canonical bundle of moduli space.

In the Deligne-Mumford compactification $\overline{\mathcal{M}}_h$, one adjoins to moduli space the divisor $[\Delta]$ of Riemann surfaces with nodes. Both the canonical bundle K and the Hodge bundle admit natural extensions to $\overline{\mathcal{M}}_h$, the first as the canonical bundle of $\overline{\mathcal{M}}_h$, and the second as the line bundle of dualizing differentials. A characteristic class computation then shows that $K\otimes\Lambda^{-13}$ over $\overline{\mathcal{M}}_h$ is actually not trivial and admits $[-2\Delta]$ as divisor. Since the components of Δ are independent and F is nowhere vanishing on the interior of $\mathcal{M}_h$, it follows that F must have a second-order pole along Δ. Physically, this pole corresponds to the presence in the string mass spectrum of the tachyon.

Holomorphic anomalies in string theory were discovered by Belavin and Knizhnik (1986). In retrospect, related issues had occurred earlier in the work of Schwinger (1951), Coleman, Gross, and Jackiw (1969), and Quillen (1984) on two-dimensional Dirac operators coupled to vector potentials. Equation (7.23), which makes no reference to regularized determinants, appeared in Belavin and Knizhnik (1986) and also in Bost and Jolicoeur (1986) and Catenacci *et al.* (1986). It is the starting point for several expressions of the string partition function in terms of modular forms and theta functions, e.g., Beilinson and Manin (1986), Belavin *et al.* (1986), Manin (1986), Moore (1986), Dugan (1987), Morozov (1987a, 1987b). Other expressions in terms of theta functions can be derived from chiral bosonization formulas below, as indicated in Sec. VII.D. Applications to chiral determinants are considered in Knizhnik (1986a, 1986b, 1987). A careful discussion of the extensions of the Hodge and canonical bundles to the compactified moduli space $\overline{\mathcal{M}}_h$ is provided in the review of Nelson (1987a).

B. The free scalar field

We now begin a detailed study of the conformal fields introduced in Sec. II.J. The simplest field is a free scalar boson x, with action

$$I_x(x)=\frac{1}{4\pi}\int d^2z\,\partial_z x\partial_{\bar z}x\ .$$

Its two-point function $G(z,w)=\langle x(z)x(w)\rangle$ is familiar from Sec. II.G. Recall that it is not Weyl invariant, so that $x(z)$ does not have a well-defined conformal dimension. However, both $\partial_z x$ and the vertex operator

$$V_q(z)=\rho^{q^2/2}e^{iqx(z)}$$

are well-behaved conformal fields in view of Eqs. (2.87) and (2.90), and have conformal dimensions (1,0) and $(q^2/2,q^2/2)$, respectively.

Finally, we can now address the issue of chiral scalar fields. In the presence of the holomorphic anomaly discussed in Sec. VII.A, the partition function of x is not the absolute value squared of a holomorphic function on moduli space. Nevertheless we can define the partition function Z_Δ^{-1} of a chiral scalar field by

$$\int Dx\,\exp\left[-\frac{1}{4\pi}\int d^2z\,\partial_z x\partial_{\bar z}x\right]=|Z_\Delta^{-1}|^2 e^{S_L(\rho)}(\det \mathrm{Im}\Omega)^{-1/2}\ , \tag{7.24}$$

where $S_L(\rho)$ is the Liouville action in conformal gauge with $ds^2=\rho\,dz\,d\bar z$,

$$S_L(\rho)=\frac{1}{48\pi}\int d^2z\,\partial_z\ln\rho\partial_{\bar z}\ln\rho\ .$$

In this way, Z_Δ will be holomorphic on moduli space, although it has both local and global gravitational anomalies, as indicated by the presence of the Liouville action and $\det \operatorname{Im}\Omega$.

It is easy to determine the variation of Z_Δ with respect to moduli, since it reduces to the expectation value of the chiral stress tensor T_{zz} of Eq. (2.178),

$$\frac{4\pi}{\sqrt{g}}\frac{\delta}{\delta g^{zz}}\ln Z_\Delta = -T_{zz} . \tag{7.25}$$

Equation (2.90) for the propagator and chiral renormalization procedure yields

$$T_{zz} = -\tfrac{1}{2}\lim_{z\to w}\left[\partial_z\partial_w \ln E(z,w) - \frac{1}{(z-w)^2}\right] . \tag{7.26}$$

This means that the stress tensor is the third Taylor expansion coefficient of the expansion of E for z near w,

$$E(z,w) = z - w + (z-w)^3 T_{zz} + O(z-w)^5) . \tag{7.27}$$

C. Spin-$\frac{1}{2}$ bosonization

In this section and those that follow, we shall solve completely the theory of circle-valued bosonic fields. The formulas we shall derive for correlation functions will be explicit enough to allow us to identify them with the corresponding correlation functions for chiral fermions. We begin with the simplest case of no background charge Q, where the action reduces to

$$I_x(\varphi) = \frac{1}{4\pi}\int d^2z\, \partial_z\varphi\partial_{\bar z}\varphi . \tag{7.28}$$

The first task is a suitable indexing of the soliton sector. Recall that φ is to be thought of as circle valued, i.e., $d\varphi$ is a closed 1-form that is not necessarily exact [note, however, that the action $I_x(\varphi)$ in Eq. (7.28) is unambiguous, since it can as well be written as the integral of the (1,1) form $\partial\varphi\wedge\bar\partial\varphi$, after splitting $d\varphi$ as $\partial\varphi+\bar\partial\varphi$]. Up to exact forms, a closed 1-form is characterized by its winding numbers along cycles of the homology basis

$$2\pi n_I = \oint_{A_I} d\varphi,\quad 2\pi m_I = \oint_{B_I} d\varphi .$$

If we fix once and for all a set of closed 1-forms ϕ_{mn} with precisely winding numbers m_I and n_I, $d\varphi$ can be unambiguously written as

$$d\varphi = \phi_{mn} + dx , \tag{7.29}$$

where x is a genuine single-valued scalar, completely determined by the familiar normalization requirement $\int d^2z\sqrt{g}\,x = 0$, needed to remove the zero mode of the scalar Laplacian. Since the action $I_x(\varphi)$ then splits completely as

$$I_x(\varphi) = I_x(\phi_{mn}) + I_x(x) = I_{mn} + I_x(x) ,$$

the path integral becomes

$$\sum_{m,n} e^{-I_{mn}}\int Dx\, e^{-I_x(x)} ,$$

which is tractable. With the canonical basis choice for Abelian differentials, $\omega_1,\ldots,\omega_h$, it is easy to write down such forms ϕ_{mn}:

$$\phi_{mn} = -i\pi(m+\bar\Omega n)_I(\operatorname{Im}\Omega)^{-1}_{IJ}\omega_J + \text{c.c.} , \tag{7.30}$$

and the soliton contribution to the action is

$$I_{mn} = \frac{\pi}{2}(m+\bar\Omega n)(\operatorname{Im}\Omega)^{-1}(m+\Omega n) . \tag{7.31}$$

We shall illustrate the procedure with an explicit calculation of the partition function. In this case the contribution to the Dx integral is $(8\pi^2\det'\Delta_g/\int d^2z\sqrt{g})^{-1/2}$, while the sum over soliton sectors produces the factor

$$\sum_{m,n}\exp\left[-\frac{\pi}{2}(m+\bar\Omega n)(\operatorname{Im}\Omega)^{-1}(m+\Omega n)\right] . \tag{7.32}$$

This actually is a sum over all spin structures of theta functions evaluated at 0. To see this, we rewrite the summation index m as $2(k+\delta'')$ with δ'' half-integer valued and apply the Poisson summation formula to get

$$(\det\operatorname{Im}\Omega)^{1/2}\sum_{\delta''}\sum_{n,k} e^{-\pi n(\operatorname{Im}\Omega)n/2}e^{-\pi k(\operatorname{Im}\Omega)k/2} \times e^{i(2\pi\delta''+\pi n\operatorname{Re}\Omega)k} . \tag{7.33}$$

If we now rewrite the summation over n,k as the summation over integers p,q and half-integers δ' with $p+q+2\delta'=n$, $p-q=k$, we recognize the sum over n,k as

$$\sum_{\delta'} |\vartheta[\delta](0,\Omega)|^2 .$$

Thus the final formula for the circle-valued bosonic field with vanishing background charge is

$$Z_B = \sum_\delta Z_B^\delta ,$$

where

$$Z_B^\delta = \left[\frac{8\pi^2\det'\Delta_g}{\int d^2z\sqrt{g}\,\det\operatorname{Im}\Omega}\right]^{-1/2} |\vartheta[\delta](0,\Omega)|^2 . \tag{7.34}$$

We can compare this expression with the partition function of the chirally symmetric fermion theory with spin $\frac{1}{2}$ (cf. Sec. II.J):

$$Z_F^\delta = \int e^{-I(b,c)+\text{c.c.}} .$$

Clearly, both Z_B^δ and Z_F^δ vanish when there is a Dirac zero mode, so we discuss only generic even-spin structures, in which case

$$Z_F^\delta = (\det\Delta_{1/2}^{(-)}) . \tag{7.35}$$

To compare Eq. (7.34) with (7.35) it suffices to compare their variations with respect to the background metric. Since determinants are regularized by heat kernels, they are manifestly reparametrization invariant. Modular anomalies could come from changing the basis of Abelian differentials and hence changing Ω, but this is compen-