The geometry of string perturbation theory

Eric D'Hoker*

Department of Physics, Princeton University, Princeton, New Jersey 08544

D. H. Phong

Department of Mathematics, Columbia University, New York, New York 10027

This paper is devoted to recent progress made towards the understanding of closed bosonic and fermionic string perturbation theory, formulated in a Lorentz-covariant way on Euclidean space-time. Special emphasis is put on the fundamental role of Riemann surfaces and supersurfaces. The differential and complex geometry of their moduli space is developed as needed. New results for the superstring presented here include the supergeometric construction of amplitudes, their chiral and superholomorphic splitting and a global formulation of supermoduli space and amplitudes.

CONTENTS

*Present address: Department of Physics, University of California, Los Angeles, CA 90024.

I. INTRODUCTION

Local quantum field theory offers a remarkably successful description of the electromagnetic, weak, and strong interactions of the particles thus far observed. The standard electroweak theory of Glashow, Weinberg, and Salam, together with quantum chromodynamics, accounts extremely well for the vast amounts of high-energy particle accelerator data that have accumulated over the past forty years. Unification of quarks and leptons and of these three fundamental forces has been proposed by Georgi, Quinn, and Weinberg (1974), and several models have been constructed, amongst them the unified theories of Georgi and Glashow (1974) and Pati and Salam (1974). Though experimental evidence for the predicted decay of baryons in these theories is still lacking, there is a widespread belief that some type of unification should take place. As far as the physics of elementary particles is concerned, local quantum field theory thus provides a consistent and predictable framework.

Nature has provided us with one more force, however, that of gravitational attraction. The theory of general relativity accounts for this force, at a non-quantum-mechanical level, as a manifestation of the curved geometry of space-time. General relativity has been well tested on the cosmic scale, but has not yet been incorporated in a consistent scheme based on local quantum field theory. An overview of some of the attempts at the quantization of gravity may be found in Hawking and Israel (1979). Better yet, a natural goal would be to unify the four fundamental forces of nature into a single consistent and predictive quantum theory. Though supergravity theories pioneered by Freedman, van Niewenhuizen, and Ferrara (1976) and by Deser and Zumino (1976a) seemed at one time good candidates for such a unification within the framework of conventional quantum field theory, there are still problems with their consistent quantization.

In a key development, Neveu and Scherk (1972) found an effective Yang-Mills theory present in *dual models,* and Yoneya (1973) and Scherk and Schwarz (1974) argued that dual models with their "string" interpretation automatically contained a massless spin-2 particle, coupling precisely as the graviton couples in general relativity. The picture of elementary particles, and in particular the graviton, as pointlike objects with no internal structure could then be traded in for a theory in which elementary particles are thought of as one-dimensional curves with infinitesimal thickness, or so-called *strings.* Strings interact by joining and splitting. A unification of all forces along these lines was proposed by Scherk and Schwarz (1975). The length scale of such strings is set by the only scale characteristic of quantum gravity: the Planck length, which is on the order of 10^{-33} cm. It was also discovered that standard fermions and gauge bosons are automatically present in fermionic versions of the dual string models such as those found by Ramond (1971) and Neveu and Schwarz (1971). Furthermore, certain truncations of this model were shown to exhibit a supersymmetric spectrum by Gliozzi, Scherk, and Olive (1975, 1976), and the full supersymmetry was subsequently proven in one of the first papers on modern string theory by Green and Schwarz (1981). Soon thereafter the famous type-I theory of open and closed superstrings and the type-II A and B theories of closed superstrings only were identified by Green and Schwarz (1982). The type-I string possesses gauge symmetry from the outset, but the type-II string does not. For the type-II string, Witten (1983, 1985c) argued that serious problems arise if one wants to keep chiral fermion multiplets after compactification to four dimensions. In 1983, Alvarez-Gaumé and Witten showed that rather generic anomalies in gauge and gravitational symmetries cancel for the type-II superstring. The discovery of the absence of anomalies in the type-I superstring with gauge group O(32) by Green and Schwarz (1984) sparked a great deal of excitement about the phenomenological possibilities of that theory. The anomaly cancellation mechanism also allowed a gauge group $E_8 \times E_8$, and a theory with this symmetry seemed to be even more promising phenomenologically. A new type of string theory that encompasses this possibility—called the heterotic string—was soon discovered by Gross, Harvey, Martinec, and Rohm (1985a, 1985b). As these string models only seem to make sense in higher dimensions, it is usually assumed that the ground state rolls up in a tiny compact space in all but four dimensions, an idea going back to Kaluza (1921) and Klein (1926) and revived more recently in Cremmer and Scherk (1977). Promising compactifications and their phenomenological implications were discussed early on by Candelas, Horowitz, Strominger, and Witten (1985). Not only may superstrings contain the right particles, they also present strong evidence for being consistent, unitary, and predictable quantum theories of all particles and forces in nature. In a sense these string theories appear even healthier than quantum field theory itself, as calculations of scattering amplitudes do not seem to require renormalization, they are just finite. At least those are the indications gotten from analyses to tree level and sometimes to one-loop order in string perturbation theory.

It is obviously an important question whether the indications of one-loop finiteness and unitarity persist to all orders in perturbation theory. Most of the present review will be explaining the general framework for perturbatively calculating scattering amplitudes in string theories as we understand it today. One might compare this program with the derivation of the Feynman rules in a quantum field theory, to any order in perturbation theory. It has become clear that string theory offers a challenge with sometimes intricate but generally beautiful mathematical concepts, and we shall acquaint the reader gradually with the geometry that enters the perturbative methods, along with the physical ideas involved. Perhaps string perturbation theory does not provide us with sufficient flexibility and insight into questions of compactification and symmetry breaking, and a more general scheme is needed. Many attempts in this direction have been undertaken, and though their discussion would take us too far away from the mainstream of this review, we shall periodically indicate connections with such investigations.

Because of the almost unique nature of consistent string theories and the occurrence of surprising anomaly cancellation mechanisms we may expect a very simple but fundamental principle to underlie their existence. Such a principle remains to be fully uncovered. There is, however, a recurrent theme that sharply distinguishes strings and pointlike particle theories. With pointlike particles, there is a geometrical distinction between free propagation of particles and their interaction. The dynamics of the freely moving particle and of the interaction of several particles are separate components of the theory: in particular, the smooth world lines of free propagation experience a "singular" joining at the interaction point. The nature of the interaction is an additional input in the theory. An interaction occurs at a geometric point, and if it were observed from a different Lorentz frame, geometrically speaking the point of interaction would be unaltered [Fig. 1(a)]. In a theory of say, closed strings, formulated in a Lorentz-covariant way, two strings may touch at one instant and merge into one string, but the interaction point is not "geometric," as observation from different Lorentz frames will lead to different geometric locations of the interaction point [Fig. 1(b)]. The local dynamics of the string does not depend on whether there are interactions or not. In a Lorentz-covariant formulation, the action of the interacting string is the same as that of the free string. The topology of the worldsheet swept out by the strings alone is able to inform us that the strings interact. Thus the interaction appears global and "smeared out." This was known already in the days of the old dual models; there one noticed that the form factor of a string indicated nothing hard to scatter off, and this is clearly important for its nice short-distance properties.

From the point of view presented above, string theories describe surfaces moving in a target space-time, with no local interaction on the worldsheet. String interactions result from nontrivial topology of the surface; in particular, connectedness is related to the degree of interaction, boundary curves to initial and final strings, and the number of handles to the number of loops in an analogous dual or Feynman diagram representation. The formulation in which this topological and geometrical character of string amplitudes is manifest is that of Polyakov (1981a, 1981b), originally proposed mainly as a model of random surfaces. It provides a natural framework for maintaining reparametrization and conformal invariance, which are crucial symmetries of string theories, and further elucidates the role of the critical dimension. Actually, the conformal invariance properties in two dimensions are very restrictive, as was realized by Kadanoff (1969) and Polyakov (1969), who used it to de-

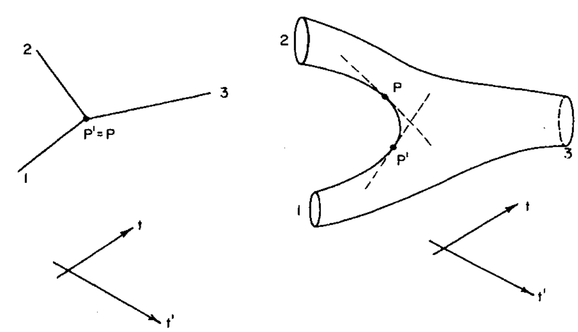

FIG. 1. Interactions of particles and of strings: (a) The point of interaction to two pointlike particles is geometrical and independent of the Lorentz frame of observation; (b) the point of interaction of two strings is not geometrical and depends on the Lorentz frame of observation.

velop the conformal bootstrap program. More recently, conformal field theory in two dimensions has been the scene of an intense independent development, sparked by the work of Belavin, Polyakov, and Zamolodchikov (1984) and its unitary restriction discovered by Friedan, Qiu, and Shenker (1984) and constructed explicitly by Goddard, Kent, and Olive (1986).

Though the table of contents should facilitate the reader's access to this review, we should like to sketch the broad outlines of our approach. We shall be considering only closed-string theories, first the closed oriented bosonic model of Virasoro (1969) and Shapiro (1970), generalizing the original open-string model of Veneziano (1968), then the type-II and heterotic superstrings. In many ways, the modifications required for open strings are of a purely technical nature, and we shall provide a few key references to the work on open strings when appropriate.

In keeping with manifest Lorentz invariance, we use the Polyakov formulation. Scattering amplitudes are evaluated perturbatively in the loop expansion. To order of h loops, the answer reduces to an integral over moduli space (or supermoduli space for superstrings) of an integrand consisting of determinants of certain operators and correlation functions. Great effort is devoted to evaluating these quantities and the moduli measure explicitly, first with the help of real geometry of moduli space, then with the help of complex geometry, leading us to make contact with the more algebraic conformal field theory formulation. Scattering amplitudes in the light-cone gauge have been investigated by Mandelstam (1973a, 1973b, 1974a, 1974b, 1974c, 1986a, 1986b).

A word about references may be in order. Within the field of string perturbation theory about flat Minkowski or Euclidean space-time, we have attempted to include many published references and preprints that are of direct relevance to the approach adopted here. Unfortunately, however, it has become exceedingly difficult to keep track of all the literature, and we present our apologies to those authors who feel their work has not been appropriately referenced. At places where we discuss connections with separate fields of investigation such as string field theory, propagation in nontrivial backgrounds and compactifications, open strings, universal moduli space, Grassmannians, etc., we shall quote only some of the earliest papers.

Finally, a number of reviews have been published over the past years. Some of the earlier work appears in Alessandrini, Amati, Le Bellac, and Olive (1971), Schwarz (1973), Frampton (1974), Mandelstam (1974a), Rebbi (1974), Veneziano (1974), and, perhaps the most accessible, Scherk (1975).

More recent reviews are those of Schwarz (1982), Green (1983), a reprint collection by Schwarz (1986), a book of Green, Schwarz, and Witten (1987), and a number of conference proceedings, including conferences held at Argonne (edited by Bardeen and White, 1985), Santa Barbara (edited by Green and Gross, 1986), and San Diego (edited by Yau, 1987). Polyakov's viewpoint and strings in other contexts than grand unification are in his book: Polyakov (1987b).

II. THE CLOSED ORIENTED BOSONIC STRING

The evolution of a closed string sweeps out a worldsheet, which is a two-dimensional surface embedded in a target space-time. The worldsheet is bounded by the position curves of the initial and final strings, and its handles indicate the creation and annihilation of virtual pairs. Thus the worldsheet is similar to a Feynman diagram in which propagator lines are replaced by cylinders and a loop now corresponds to a handle [Fig. 2(a)]. In this review we shall consider only S matrix elements, for which the initial and final strings are on shell and set at infinity. Under conformal transformations, which are the crucial symmetries of the theory, such a worldsheet can be transformed to a compact surface with a number n of points removed corresponding to the external string states. Such points are called punctures [Fig. 2(b)]. In

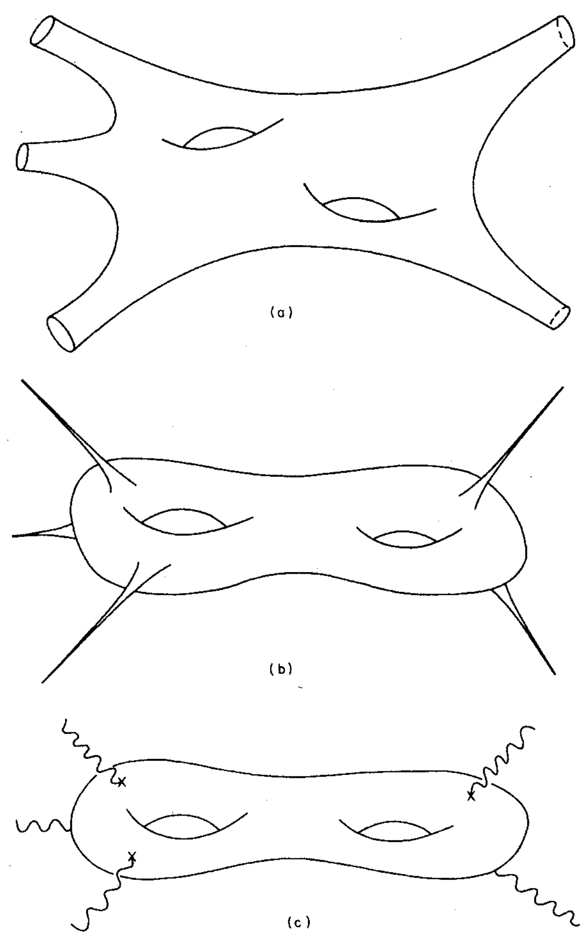

FIG. 2. The five-point function to two-loop ($h=2$) order, with incoming and outgoing strings represented as (a) full boundary curves; (b) punctures; (c) vertex operators.

the path-integral quantization procedure, the scattering amplitude is obtained by summing over all surfaces with n punctures and integrating at the punctures against the wave functions of the string states. Alternatively, we can rely on a string analog of the Lehmann-Symanzik-Zimmermann (LSZ) reduction formalism of quantum field theory, which gives on-shell scattering amplitudes in terms of vacuum expectation values of a time-ordered product of fields. The worldsheet is viewed then as a compact surface without punctures, but with insertions of local operators with the quantum numbers of the external string states. These operators are called vertex operators. In this formulation, the amplitude is obtained by summing over all compact surfaces and over all possible locations of the vertex operators [Fig. 2(c)]. The equivalence between the two formulations, together with the relation between the wave functions and vertex operators, will be discussed in detail later in Sec. II.L, and for the time being we shall adopt the vertex operator approach.

For closed oriented strings, the worldsheet is a compact orientable surface. At the h-loop level, there is topologically speaking only one such surface, which is a sphere with h handles. The number h is often referred to as the genus of the surface. Equivalently, we can classify the topology of the surface M by its Euler characteristic $\chi(M)$, defined as

$$\chi(M)=f-e+v \ ,$$

where f, e, and v are, respectively, the number of faces, edges, and vertices of any triangulation of M. The relation between $\chi(M)$ and h is readily seen to be

$$\chi(M)=2-2h \ . \tag{2.1}$$

In the presence of a metric g_{mn} on the surface M, the Gauss-Bonnet theorem asserts that $\chi(M)$ can be evaluated from the Gaussian scalar curvature R,

$$\chi(M)=\frac{1}{2\pi}\int_M d^2\xi\sqrt{g}R \ . \tag{2.2}$$

This formula can also be viewed as a topological constraint on the curvature of a surface of given genus.

We have already mentioned that conformal invariance plays a key role in string theories, and this issue will be discussed in detail as it emerges again and again in the review. It may be helpful to note at this point that two surfaces that are topologically equivalent may still not be equivalent as surfaces with complex structures. It is the space of complex structures on a given topological surface—the moduli space—which lies at the center of string perturbation theory.

As we progress, more facts about geometry of surfaces will be introduced as we need them.

A. Classical strings

A natural reparametrization-invariant action is the geometrical area, as proposed by Nambu (1970) and Goto (1971):

$$I_{\rm NG}(x^\mu)=T\int_M d^2\xi\sqrt{h} \ . \tag{2.3}$$

Here $\xi^m=(\xi^1,\xi^2)$ are coordinates on M, and $x^\mu(\xi)$, $\mu=1,\dots,d$ describe the propagation of a string in a space-time of dimension d. The metric $G_{\mu\nu}(x)$ in space-time should ultimately arise dynamically as excitations of the $x^\mu(\xi)$, but in string perturbation theory it is taken just as a background metric which satisfies the string equations of motion. The embedding x^μ then induces a metric h_{mn} on the worldsheet given by

$$h_{mn}=\partial_m x^\mu \partial_n x^\nu G_{\mu\nu}(x) \ , \tag{2.4}$$

and $h=\det(h_{mn})$ (see Fig. 3). Finally T has dimensions of inverse length squared (or equivalently mass squared) and is called the string tension. It is simply related to the Regge slope parameter α' of dual-model theory by $T=1/2\alpha'$.

The field equations for x^μ implied by the Nambu-Goto action have two constraints expressing the vanishing of the worldsheet stress tensor. These constraints may be obtained as field equations directly if an intrinsic metric g_{mn} independent of h_{mn} is introduced. This leads to the formulation of Polyakov (1981a). Its action is that of a σ model with space-time as the target Riemannian manifold and the key property of reparametrization invariance on the worldsheet:

$$I_0(x^\mu,g_{mn})=\frac{T}{8\pi}\int_M d^2\xi\sqrt{g}g^{mn}\partial_m x^\mu\partial_n x^\nu G_{\mu\nu}(x) \ . \tag{2.5}$$

Classically the Nambu-Goto and the Polyakov actions lead to identical dynamics. Quantum mechanically it is not known whether the corresponding theories are equivalent, mainly because the string theory obtained from $I_{\rm NG}$ is hard to quantize unambiguously. The main advantages of the Polyakov action are that there is a clear distinction between the intrinsic geometry g_{mn} of the worldsheet and its embedding in space-time, and that the action is quadratic in x^μ's if $G_{\mu\nu}$ is the flat Euclidean metric. This is the case we shall study in detail. Henceforth, we shall set the string tension to unity: $T=1$.

The classical symmetries of Eq. (2.5) are as follows.

(i) The group Diff(M) of differentiable reparametriza-

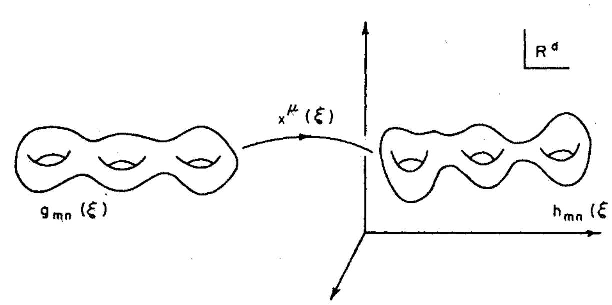

FIG. 3. The worldsheet: with intrinsic metric g_{mn} (left); embedded in the target space-time with the induced metric h_{mn} (right).

tions, or diffeomorphisms of M. Their action on the coordinates of the surface is given by $\xi^m \to \xi'^m(\xi)$, and the action on the metric is also familiar from general relativity:

$$g_{mn}(\xi) \to g'_{mn}(\xi') = \frac{\partial \xi^p}{\partial \xi'^m} \frac{\partial \xi^q}{\partial \xi'^n} g_{pq}(\xi) . \tag{2.6}$$

Diffeomorphisms connected to the identity form the smaller group $\mathrm{Diff}_0(M)$ and are generated by continuous vector fields $\delta v^m = \xi'^m - \xi^m$. The corresponding infinitesimal changes in the fields are

$$\delta g_{mn} = \nabla_m(\delta v_n) + \nabla_n(\delta v_m), \quad \delta x^\mu = \delta v^m \partial_m x^\mu . \tag{2.7}$$

(ii) The group Weyl(M) of all rescalings of the metric by (M-dependent) positive real functions. These transformations do not move the points of M and act infinitesimally as

$$\delta g_{mn} = 2\delta\sigma g_{mn}, \quad \delta x^\mu = 0 . \tag{2.8}$$

(iii) For flat target space-time, $G_{\mu\nu}$ is the Minkowski metric $\eta_{\mu\nu}$. The group of Poincaré transformations in the target space-time is

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu + x_0^\mu, \quad \eta_{\mu\nu}\Lambda^\mu{}_\kappa \Lambda^\nu{}_\lambda = \eta_{\kappa\lambda} . \tag{2.9}$$

Most of the time, we shall assume analytic continuation to imaginary time, so that the metric of the target space-time is Euclidean; it would then be more appropriate to call this the group of isometries of flat d space.

It is useful to remark that some diffeomorphisms preserve the angles and are thus conformal reparametrizations. On the other hand, as the Weyl transformations merely rescale the metric, all Weyl transformations are conformal.

The action (2.5) had actually been considered before Polyakov in the context of two-dimensional supergravity by Brink *et al.* (1976) and by Deser and Zumino (1976b).

Notice that both the actions I_{NG} and I_0 involve only the intrinsic geometry of the string, with no reference to the extrinsic curvature experienced by the string. This appears to be the appropriate setting for a string model of elementary particles. However, if string theory is to be viewed as an effective theory of flux tubes in QCD or of the Ising model, the extrinsic curvature corrections should be taken into account. Such a model has recently been proposed by Polyakov (1986), but we shall not discuss it here.

We shall also constantly make use of standard differential and Riemannian geometry. Some key formulas are collected in Appendix A. Fuller accounts can be found in Spivak for differential geometry, Weinberg (1972) for general relativity, and Bott and Tu (1983) for topological aspects.

B. Quantization

Quantization may be performed by summing in the functional integral over all closed compact surfaces, as originally proposed by Hsue, Sakita, and Virasoro (1970) and by Gervais and Sakita (1971b). In the Polyakov formulation, this corresponds to treating both the x^μ and the worldsheet metrics g_{mn} as two-dimensional quantum fields. For flat Euclidean space-time, the x^μ are free fields and their path integrals Gaussian, so the crucial part will be the path integral over metrics g_{mn}.

The functional integral approach requires the addition to the Polyakov action of all possible renormalization counterterms consistent with the symmetries of the theory. In general, Weyl invariance is broken upon quantization in view of the conformal anomaly,[1] so we must include Weyl-noninvariant counterterms as well, and the most general local action compatible with reparametrization invariance is

$$I(x^\mu, g_{mn}) = I_0(x^\mu, g_{mn}) + \lambda\chi(M) + \mu_0^2 \int_M d^2\xi \sqrt{g} . \tag{2.10}$$

The Weyl invariance lost in the action because of μ_0^2 can actually be restored in the critical dimension $d=26$. We shall derive this crucial fact in detail later, and for the moment restrict our discussion to why we should have Weyl invariance at all. The standard philosophy is roughly as follows. Weyl and reparametrization invariance make up for three degrees of freedom, exactly the number in the metric g_{mn}. The only true degrees of freedom are then those of the d x^μ fields, with of course the two constraints implied by the equations of motion of g_{mn}. Thus with Weyl invariance the quantum string has $(d-2)$ degrees of freedom, precisely the number of the classical string. Note that the requirement of having the same number of quantum and classical degrees of freedom is assumed from the start in the light-cone formulation of Goddard *et al.* (1973) and Mandelstam (1973a, 1973b, 1974). Actually Polyakov (1981a) originally proposed his model precisely with the objective of obtaining a consistent quantum theory without Weyl invariance. The scale factor develops, then, an effective dynamics that is described by the Liouville theory. Ultimately, the constraint of Weyl invariance must be analyzed in the light of the unitarity of the theory, as we shall discuss in Sec. II.G.

Now the physical quantities of interest are the partition function

$$Z = \sum_{h=0}^{\infty} \int \frac{Dg_{mn} Dx^\mu}{\mathcal{N}} e^{-I(x,g)} \tag{2.11}$$

(which can be identified with the space-time integral of the target space-time cosmological constant) and on-shell scattering amplitudes, obtained by inserting vertex operators V_i

$$\langle V_{i_1}(k_1^\mu) \cdots V_{i_p}(k_p^\mu) \rangle = \sum_{h=0}^{\infty} \int \frac{Dg_{mn} Dx^\mu}{\mathcal{N}} e^{-I(x,g)} \times V_{i_1}(k_1^\mu) \cdots V_{i_p}(k_p^\mu) . \tag{2.12a}$$

[1]With dimensional regularization, for example, Weyl symmetry would be destroyed away from two dimensions.

Here $\mathcal{N}$ denotes a normalization factor to be specified later on. In Sec. VIII a detailed discussion of vertex operators for on-shell physical particles will be presented. For the time being, it may be sufficient to say that they are typically of the form

$$V(k^\mu, x^\mu(\xi)) = P(\varepsilon, Dx^\mu(\xi)) e^{ik\cdot x(\xi)} , \qquad (2.12b)$$

where $P(\varepsilon, Dx^\mu)$ is a polynomial expression in the derivatives of x and ε is a polarization tensor. This form is dictated by the symmetries of the action, and a key requirement is Weyl invariance after inclusion of anomalies. The lowest mass levels will turn out to be

$$\begin{aligned} &k^2=2, \quad V(k^\mu)=e^{ik\cdot x} , \\ &k^2=0, \quad V(k^\mu)=\varepsilon_{\mu\nu} g^{mn}\partial_m x^\mu \partial_n x^\nu e^{ik\cdot x} , \end{aligned} \qquad (2.12c)$$

where the first corresponds to the tachyon, and the second corresponds to $\varepsilon_{\mu\nu}$ symmetric traceless, the graviton; $\varepsilon_{\mu\nu}$ antisymmetric, the antisymmetric tensor field; $\varepsilon_{\mu\nu}$ pure trace part, the dilaton.

The integration measures Dg_{mn} and Dx^μ are determined by requirements of symmetry and locality. The construction of Dg_{mn} will be discussed in the next section. For Dx^μ, the measure is completely determined once one has a metric function on the space of small variations δx^μ, so that one can measure lengths and angles and hence volumes. This metric on the space of embeddings x^μ is unique due to Poincaré and reparametrization invariance,

$$\|\delta x^\mu\|^2 = \int_M d^2\xi \sqrt{g}\, \delta x^\mu \delta x^\mu . \qquad (2.13)$$

It induces an inner product, which we denote by $\langle \delta x_1 \mid \delta x_2 \rangle$. Note that it is not Weyl invariant, a property providing another explanation for the Weyl anomaly. Since, however, the measure involves a product over an infinite number of variables, there may be some ambiguity in defining it from Eq. (2.13). This is resolved by the principle of ultralocality as stated by Polchinski (1986), which asserts that since the measure is a pointwise, reparametrization-invariant product over the worldsheet, any ambiguity must also be a reparametrization-invariant pointwise product. In particular, no derivatives should occur, and the only ambiguity can reside in a factor of the form

$$\exp\left[-\mu_1^2 \int_M d^2\xi \sqrt{g}\right] \qquad (2.14)$$

for some constant μ_1^2. In particular, no constant other than 1 in front of the exponential is allowed, since this could not be written as a pointwise reparametrization-invariant product over the surface. Upon substitution into functional integrals, Eq. (2.14) results in just a shift in the counterterm μ_0^2 in Eq. (2.10). Ultimately, in the critical dimension $d=26$, the net counterterm will be fixed by requiring Weyl invariance, so the measure associated with Eq. (2.13) will in effect be unique. Henceforth we shall assume that such a counterterm has been fixed, and we shall not exhibit the area term any longer. This argument applies equally well to the case of any space-time metric $G_{\mu\nu}(x)$ as long as it is independent of the derivatives of x^μ.

The main issue, then, is to evaluate the functional integrals in Eqs. (2.11) and (2.12a). The integral in x^μ is easily performed once we have specified the measure arising from Eq. (2.13). First, we write the action I_0 using the scalar Laplacian Δ_g on the Riemann surface:

$$I_0(x^\mu, g_{mn}) = \frac{1}{8\pi} \langle x \mid \Delta_g x \rangle , \qquad (2.15)$$

where

$$\Delta_g = -\frac{1}{\sqrt{g}} \partial_m \sqrt{g}\, g^{mn} \partial_n . \qquad (2.16)$$

Next, the x variable is divided into the constant zero mode x_0^μ of the Laplacian and all other modes x'^μ orthogonal to it: $x^\mu = x_0^\mu + x'^\mu$ with $\langle x_0 \mid x' \rangle = 0$. Upon splitting the functional integral accordingly, we have

$$\begin{aligned} \int Dx^\mu e^{-I_0(x,g)} &= \int Dx_0^\mu \int Dx'^\mu e^{-\langle x' \mid \Delta_g x' \rangle / 8\pi} \\ &= (\det{}'\Delta_g)^{-d/2} \\ &\quad \times \int dX_0^\mu \int Dx'^\mu e^{-\|x'\|^2/8\pi} . \end{aligned} \qquad (2.17)$$

With the principle of ultralocality, one deduces that the Gaussian integral

$$\int Dx^\mu e^{-\|x\|^2/8\pi} \qquad (2.18)$$

is a local product over the worldsheet, with no derivatives of the metric entering. Consequently it must be the exponential of the worldsheet area, which may be absorbed into the μ_0^2 coefficient in Eq. (2.10). Let us now split up this integral as we did before:

$$\begin{aligned} 1 &= \int Dx^\mu e^{-\|x\|^2/8\pi} \\ &= \int Dx_0^\mu \int Dx'^\mu e^{-\|x_0\|^2/8\pi - \|x'\|^2/8\pi} \\ &= \left[\frac{8\pi^2}{\int_M d^2\xi \sqrt{g}}\right]^{d/2} \int Dx'^\mu e^{-\|x'\|^2/8\pi} . \end{aligned} \qquad (2.19)$$

If, in addition, we note that $\int Dx_0^\mu = \Omega$, the volume of space-time, then we finally obtain

$$\int Dx^\mu e^{-I_0(x,g)} = \Omega \left[\frac{8\pi^2}{\int_M d^2\xi\sqrt{g}} \det{}'\Delta_g\right]^{-d/2} . \qquad (2.20)$$

If, instead of considering the contribution to the partition function as we did above, we also have a sequence of vertex operators, then the space-time volume element should be replaced by the total momentum conservation δ function, resulting from the x_0^μ integral over the $\exp(ik\cdot x)$ factors in the vertex operators (2.12b).

We turn to the more difficult task of integrating over metrics in Eqs. (2.11) and (2.12a) in the next section.

C. Worldsheet metrics and deformations of conformal classes

We now fix the topology, i.e., the number of handles of the worldsheet M. Let $\mathcal{M}=\{g_{mn} \text{ on } M\}$ be the space of positive worldsheet metrics. An infinitesimal deformation δg_{mn} of a metric is a symmetric two-tensor, and the natural norm for δg_{mn} is

$$\|\delta g_{mn}\|^2=\int_M d^2\xi\sqrt{g}\,(cg^{mn}g^{pq}+g^{mp}g^{nq})\delta g_{mn}\delta g_{pq} \ . \tag{2.21}$$

The arbitrary constant c will not appear in any physical answer, so we shall set it to zero. Associated with the norm is an inner product on the space of metric deformations, which will be denoted by $\langle \delta g_1 \mid \delta g_2 \rangle$. The measure Dg_{mn} will be the one associated with Eq. (2.21) (for $c=0$), with the usual harmless ambiguity which ultralocality fixes to be of the form (2.14). To determine the form of the final answer for Eqs. (2.11) and (2.12), we assume momentarily that all possible Weyl and reparametrization anomalies will cancel and ask whether all modes of g_{mn} can be gauged away with the help of reparametrizations and Weyl transformations. The number of degrees of freedom is the same (3 for g_{mn}, 1 for Weyl, and 2 for reparametrizations), and it is a classic theorem of Gauss that in any simply connected patch on the surface the metric can indeed be made Euclidean by such transformations. Whenever the topology is nontrivial, however, the reparametrizations of different patches need not match and there may be topological obstructions. To see this, we note that the joint action of reparametrizations and Weyl transformations on the metric is given by Eqs. (2.7) and (2.8),

$$\delta g_{mn}=(2\delta\sigma+\nabla^p\delta v_p)g_{mn}+(P_1\delta v)_{mn} \ , \tag{2.22}$$

where the operator P_1 sends vectors into symmetric traceless two-tensors,

$$(P_1\delta V)_{mn}=\nabla_m\delta v_n+\nabla_n\delta v_m-g_{mn}\nabla^p\delta v_p \ , \tag{2.23}$$

and describes the traceless piece of the deformation coming from reparametrization by the vector field δv^m. Clearly the total trace piece can always be eliminated without topological obstruction by a Weyl transformation alone. Thus the only metric deformations δg_{mn} that are not gotten by reparametrization and Weyl transformation are in $(\text{Range}P_1)^\perp$. This means that any deformation δg_{mn} is given by the decomposition orthogonal under Eq. (2.21):

$$\{\delta g_{mn}\}=\{\delta\sigma g_{mn}\}\oplus\text{Range}P_1\oplus\text{Ker}P_1^\dagger \ , \tag{2.24}$$

where the action of $P_1^\dagger$ on symmetric traceless two-tensors is given by

$$(P_1^\dagger\delta g)_m=-2\nabla^n\delta g_{mn} \tag{2.25}$$

and we have used the result that, under the inner product $\langle\ |\ \rangle$, we have the identification

$$(\text{Range}P_1)^\perp=\text{Ker}P_1^\dagger \ . \tag{2.26}$$

The first two spaces on the right-hand side of Eq. (2.24) consist of modes that can be gauged away by combined Weyl and reparametrization symmetries. The dimension of the remaining space is finite, and we shall now determine it.

The way to determine the number of zero modes of these operators is to appeal to an index theorem, which gives the difference between the number of zero modes of the operator and its adjoint in terms of a topological invariant. The problem is then reduced to a similar one for the adjoint, which often may be solved by independent methods such as vanishing theorems. In the present case, zero modes of P_1 are just reparametrizations inducing only trace changes in the metric, in other words, conformal Killing vectors; the topological invariant is the Euler characteristic, and the index theorem reduces to the following version of the Riemann-Roch theorem:

$$\dim\text{Ker}P_1-\dim\text{Ker}P_1^\dagger=3\chi(M) \ . \tag{2.27}$$

This relation will actually follow from the short-time heat-kernel expansion in Sec. II.F, as we shall see later. For the sphere, the conformal Killing transformations form the Möbius group

$$z\mapsto(az+b)/(cz+d), \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}\in SL(2,C) \ ,$$

so that $\dim\text{Ker}P_1=6$; for the torus, it is the group of translations that has dimension 2. For genus ≥ 2, there are no conformal Killing vectors on a surface without boundary. It is easy to provide a proof for the case of metrics of constant negative curvature R. As we shall indicate in the next section, any metric on a surface of genus ≥ 2 can be brought back to this case by a Weyl transformation, and the dimensions of these kernels are not changed. A conformal Killing vector δv^m satisfies $(P_1\delta v)_{mn}=0$, and upon differentiation one gets

$$\nabla^q\nabla_q\delta v^p=-R\,\delta v^p \ . \tag{2.28}$$

Integrating versus δv_p over the surface, one finds

$$\|\nabla^q\delta v^p\|^2-R\,\|\delta v^p\|^2=0 \ , \tag{2.29}$$

so that $\delta v^p=0$ for $R<0$. Using the index theorem (2.27) and the above counting of conformal Killing vectors, we conclude that

$$\dim\text{Ker}P_1^\dagger=\begin{cases} 0, & h=0 \ , \\ 2, & h=1 \ , \\ 6h-6, & h\geq 2 \ . \end{cases} \tag{2.30}$$

Thus we expect the partition function and scattering amplitudes to reduce to finite-dimensional integrals over spaces of the corresponding dimensions. Elements of $\text{Ker}P_1^\dagger$ are called real *quadratic differentials* or *moduli deformations,* and parametrize infinitesimal deformations of conformal classes of metrics.

The space of conformal classes of metrics is a vast subject in the mathematics literature, going back as far as Riemann. For modern texts, we shall refer systematically to the books by Schiffer and Spencer (1954), Ahlfors (1966), Siegel (1970, 1971, 1973), Griffiths and Harris (1978), Abikoff (1980), Farkas and Kra (1980), Beardon (1983), and the survey articles of Bers (1972, 1981). In the physics literature, moduli parameters appear implicitly in dual-model diagrams. A lucid geometric account in this early phase is that of Alessandrini (1971) and Alessandrini and Amati (1971). For light-cone diagrams, moduli parameters were essentially introduced by Mandelstam (1973a, 1973b). The above approach to quadratic differentials appears in Alvarez (1983); the detailed mathematical treatment is given by Fischer and Tromba (1984a, 1984b, 1984c).

D. Teichmüller and moduli spaces

In the absence of anomalies, the string path integrals in Eqs. (2.11) and (2.12) should reduce to integrals over the space of inequivalent metrics under the combined Weyl and reparametrization symmetries. The discussion in the preceding section has shown that this space is locally a finite-dimensional manifold of dimension 0, 2, and $6h-6$ when h is 0, 1, and ≥ 2, respectively. We still need a global description, taking into account the fact that Diff(M) acts on the space of metrics by isometries but Weyl(M) does not.

A natural way to do this is to make use of the key fact that for any metric g_{mn} on M there exists a unique scaling factor $e^{2\sigma}$ such that $\hat{g}_{mn}=e^{-2\sigma}g_{mm}$ has constant curvature $R_{\hat{g}}=1$ when $h=0$, $R_{\hat{g}}=0$ and Area$(\hat{g})=1$ when $h=1$, and $R_{\hat{g}}=-1$ when $h\geq 2$. [Note that the sign of the curvature must be consistent with the Gauss-Bonnet theorem of Eq. (2.2).] This is equivalent to the fact that the Liouville equation,

$$\Delta_g \sigma = R_g - R_{\hat{g}} e^{-2\sigma} \ , \qquad (2.31)$$

admits a unique solution. Thus

$$\mathcal{M}_{\text{const}} = \{\hat{g};\ R_{\hat{g}} = \text{const as above}\} \qquad (2.32)$$

is a well-defined global slice for the Weyl group without any complication of the type discussed by Gribov (1978). It naturally carries the metric (2.21), since $\mathcal{M}_{\text{const}}$ is a subspace of $\mathcal{M}$. We may now define Teichmüller $\mathcal{T}_h$ and the moduli space $\mathcal{M}_h$ of Riemann surfaces of genus h by

$$\begin{aligned} \mathcal{T}_h &= \mathcal{M}_{\text{const}}/\text{Diff}_0(M) \ , \\ \mathcal{M}_h &= \mathcal{M}_{\text{const}}/\text{Diff}(M) = \mathcal{T}_h/\text{MCG}_h \ , \end{aligned} \qquad (2.33)$$

with

$$\text{MCG}_h = \text{Diff}(M)/\text{Diff}_0(M) \ .$$

Teichmüller space $\mathcal{T}_h$ will turn out to be a complex manifold topologically equivalent to $(\mathbf{R}_+ \times \mathbf{R})^{3h-3}$. The mapping class group MCG_h is a discrete group, which acts holomorphically with fixed points, however. Thus moduli space will have the structure of an orbifold. A more detailed discussion of some of these issues will be taken up in Sec. IV. As defined above, both Teichmüller and moduli space come equipped with a natural metric given by Eq. (2.21). The reason is that Diff(M) acts isometrically on $\mathcal{M}_{\text{const}}$ so that the natural metric on $\mathcal{M}_{\text{const}}$ can be pulled back to either space under the action of this isometry. This metric on Teichmüller and moduli space is called the Weil-Petersson metric.

To conclude this section we now verify that the tangent space to Teichmüller and moduli space is the space of quadratic differentials $\text{Ker}P_1^\dagger$, as may be expected. In fact under any deformation δg_{mn} of metrics (see Appendix A) the curvature changes by

$$\delta R = -\tfrac{1}{2} g^{mn}\delta g_{mn} R - \tfrac{1}{2}\nabla^p \nabla_p (g^{mn}\delta g_{mn}) + \tfrac{1}{2}\nabla^m \nabla^n (\delta g_{mn}) \ . \qquad (2.34)$$

This shows that a deformation δg_{mn} in $\text{Ker}P_1^\dagger$ does not change the curvature and hence is tangent to $\mathcal{M}_{\text{const}}$. Combining this with Eq. (2.24) yields (see Fig. 4)

$$T_{\hat{g}}(\mathcal{M}_{\text{const}}) = \{\nabla_m(\delta v_n) + \nabla_n(\delta v_m)\} \oplus \text{Ker}P_1^\dagger$$

and, in particular,

$$T_{\hat{g}}(\mathcal{M}_h) = \text{Ker}P_1^\dagger \ . \qquad (2.35)$$

Thus the Weil-Petersson metric can be described as follows: to determine the norm of a tangent vector to Teichmüller or moduli space, represent it by a quadratic differential δg_{mn}. Then its norm is given by Eq. (2.21), taken with respect to a metric $\hat{g}$ of constant curvature.

Other slices for $\mathcal{M}_h$ will also prove useful in treatments of string path integrals. For example, one can choose instead metrics that are flat everywhere with Dirac singularities for the curvature at a finite number of points, so that the Gauss-Bonnet relation is satisfied. From the point of view of complex analysis, one can parametrize

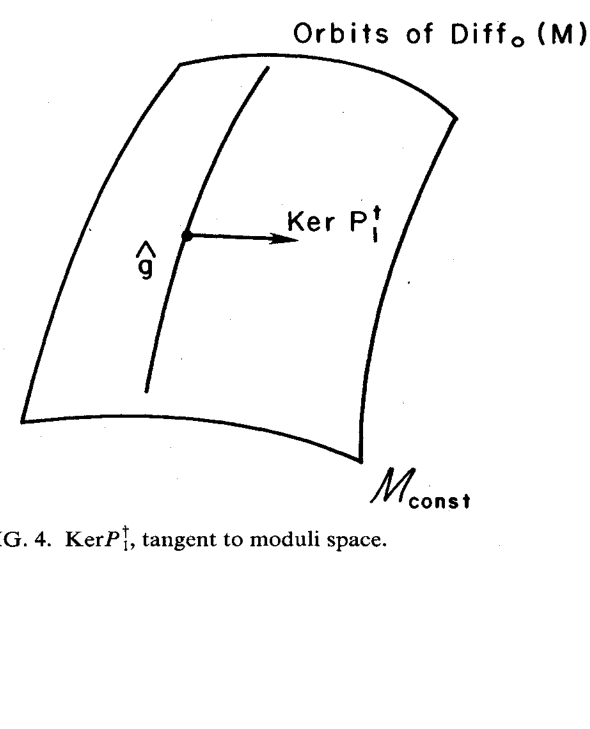

FIG. 4. $\text{Ker}P_1^\dagger$, tangent to moduli space.

moduli by period matrices (see Sec. VI.D). Different explicit parametrizations of moduli space will be discussed in Sec. IV.

It may be helpful to discuss at this point the distinction between Teichmüller and moduli spaces as far as string amplitudes are concerned. String amplitudes will reduce to integrals over Teichmüller space in the absence of $\mathrm{Diff}_0(M)$ anomalies, which are just perturbative gravitational anomalies. Elements in the mapping class group can be viewed as "large" reparametrizations. Only in the absence of global gravitational anomalies will string amplitudes reduce to integrals over moduli space.

For the bosonic string we can adopt throughout manifestly reparametrization-invariant methods of regularization such as ζ-function and short-time cutoff heat-kernel regularization. Thus neither perturbative nor global gravitational anomalies will occur, and physical quantities can all be expressed as integrals over moduli space. For fermonic strings, absence or cancellation of such anomalies is a highly nontrivial matter, and the famous constraints of modular invariance are just the requirement of absence of global gravitational anomalies when large reparametrizations act on various choices of homology bases. We shall discuss these issues in greater detail when they arise later.

E. Complex structures, tensors, covariant derivatives, and differentials

With Weyl and reparametrization invariance, the key geometric object on the worldsheet is not really the metric g_{mn}, but rather the complex structure $J_m{}^n$ it defines,

$$J_m{}^n=\sqrt{g}\,\varepsilon_{mp}g^{pn}\ , \tag{2.36}$$

with $\varepsilon_{12}=-\varepsilon_{21}=1$. We can readily check that

$$J_m{}^nJ_n{}^p=-\delta^p_m\,,\quad \nabla_pJ_m{}^n=0\ . \tag{2.37}$$

We see that the above definition of moduli space [Eq. (2.33)] is equivalent to the definition of moduli space as the set of equivalence classes under $\mathrm{Diff}(M)$ of the space of complex structures:

$$\mathcal{M}_h=\{J_m{}^n;J_m{}^pJ_p{}^n=-\delta_m{}^n\}/\mathrm{Diff}(M)\ .$$

We may also define holomorphic and antiholomorphic functions on M by the Cauchy-Riemann equations

$$J_m{}^n\partial_nf=i\partial_mf,\quad J_m{}^n\partial_n\bar f=-i\partial_m\bar f\ . \tag{2.38}$$

In any local coordinate patch, one can render the metric conformally flat by a reparametrization, so that $ds^2=2g_{z\bar z}dz\,d\bar z$, at least locally. This choice of coordinates exhibits the residual invariance under analytic reparametrizations $z\to z'(z)$, where z' is an analytic function of z. Thus M can be covered by coordinate charts with holomorphic transition functions and is consequently a Riemann surface. Since $J_m{}^n$ is Weyl invariant, and since it transforms as a tensor under general coordinate transformations, it characterizes the metric up to Weyl transformations and reparametrizations. This produces a one-to-one correspondence between points in moduli space and complex structures on M.

In the presence of a complex structure, tensors on M can be decomposed into tensors of weight (m,n), with m lower z and n lower $\bar z$ indices. The behavior of a tensor under an analytic coordinate transformation $z\to z'(z)$ is given by

$$T_{\underbrace{zz\cdots z}_{m}\ \underbrace{\bar z\bar z\cdots\bar z}_{n}}\to\left[\frac{\partial z}{\partial z'}\right]^m\left[\frac{\partial\bar z}{\partial\bar z'}\right]^nT_{\underbrace{zz\cdots z}_{m}\ \underbrace{\bar z\bar z\cdots\bar z}_{n}}\ . \tag{2.39}$$

Thus the invariant quantity characterizing such a tensor is

$$T_{\underbrace{zz\cdots z}_{m}\ \underbrace{\bar z\bar z\cdots\bar z}_{n}}(dz)^m(d\bar z)^n\ ; \tag{2.40}$$

this will also be called a tensor of weight (m,n). In particular, the space of tensors of weight $(m,0)$ will be denoted $\mathbf{T}^m$, and the space $\mathbf{T}^1$ is often referred to as (sections of) the canonical bundle K of the Riemann surface M. On a tensor $T(dz)^n$ in $\mathbf{T}^n$ the covariant derivative decomposes into

$$\nabla_pT(dz)^nd\xi^p=\nabla_z^nT(dz)^{n+1}+\nabla_{\bar z}^nT(dz)^nd\bar z\ . \tag{2.41}$$

The first term above defines an operator

$$\nabla_z^n:\ \mathbf{T}^n\mapsto\mathbf{T}^{n+1}\ ,$$
$$\nabla_z^n(T(dz)^n)=(g_{z\bar z})^n\frac{\partial}{\partial z}((g^{z\bar z})^nT)(dz)^{n+1}\ . \tag{2.42}$$

The second term, on the other hand, depends only on the conformal class of the metric and defines the key operator of the theory, namely the Cauchy-Riemann operator $\bar\partial_n=\nabla_{\bar z}^n$,

$$\bar\partial_n:\ \mathbf{T}^n\mapsto\{(n,1)\ \text{tensors}\}\ ,$$
$$\bar\partial_n(T(dz)^n)=\frac{\partial T}{\partial\bar z}(dz)^nd\bar z\ . \tag{2.43}$$

The $\bar\partial_n$ operators are intrinsically associated with moduli parameters and, in fact, holomorphically depend on them. Consequences for string amplitudes of this crucial property will be discussed at length in Sec. VII.

When we use the metric to change tensor weights from (m,n) to $(m-n,0)$, the operator $\bar\partial_n$ goes into an operator[2]

$$\nabla_n^z:\ \mathbf{T}^n\mapsto\mathbf{T}^{n-1},\quad \nabla_n^z(T(dz)^n)=g^{z\bar z}\frac{\partial}{\partial\bar z}T(dz)^{n-1}\ . \tag{2.44}$$

On each tensor space, there exists a unique inner prod-

[2]The operators ∇_z^n and ∇_n^z differ from those introduced by Alvarez's (1983) only because our $\mathbf{T}^n$ is Alvarez's $\mathcal{T}^{-n}$.

uct of tensor fields $T_1, T_2 \in \mathbf{T}^n$,

$$\langle T_1 \mid T_2 \rangle = \int d^2z \sqrt{g}\, (g^{z\bar{z}})^n T_1^* T_2 \ , \qquad (2.45)$$

and one can obtain the adjoint operator in the usual way:

$$(\nabla_n^z)^\dagger = -\nabla_z^{n-1} \ . \qquad (2.46)$$

We shall also make use of the Laplace operators

$$\Delta_n^{(+)} = -2\nabla_{n+1}^z \nabla_z^n, \quad \Delta_n^{(-)} = -2\nabla_z^{n-1}\nabla_n^z \ . \qquad (2.47)$$

The operator $\Delta_n^{(-)}$ is exactly $2\bar{\partial}_n^\dagger \bar{\partial}_n$, while $\Delta_n^{(+)}$ will correspond to $2\bar{\partial}_{n+1}\bar{\partial}_{n+1}^\dagger$ after identification of $(n,0)$ forms with $(n+1,1)$ forms.

To make contact with the space of "real" two-component tensors, we set $\mathbf{T}^n \oplus \mathbf{T}^{-n} = \mathbf{S}^n$, and the covariant derivatives on this space act by

$$\begin{aligned} P_n &: \ \mathbf{S}^n \to \mathbf{S}^{n+1}, \quad P_n = \nabla_z^n \oplus \nabla_{-n}^z \ , \\ P_n^\dagger &: \ \mathbf{S}^{n+1} \to \mathbf{S}^n, \quad P_n^\dagger = -(\nabla_{n+1}^z \oplus \nabla_z^{-n-1}) \ . \end{aligned} \qquad (2.48)$$

It is easy to see that P_1 is the operator of Eq. (2.23). Similarly, the Laplacian on scalars introduced in Eq. (2.16) is given by

$$\Delta_g = \Delta_0^{(+)} = \Delta_0^{(-)} \ . \qquad (2.49)$$

It should be borne in mind that even though these operators were first defined on tensor fields with n integer, one may in fact generalize this construction so as to allow for spinors and spinor tensors for which n is a half-integer. A proper definition of some sign ambiguities requires the notion of spin structure, which will be introduced in Secs. III.A and VI.F.

Zero modes of these operators are of great interest in string theory, since they are potential sources of anomalies. First of all, we have the following generalization of Eq. (2.27):[3]

$$\dim \mathrm{Ker}\nabla_z^n - \dim \mathrm{Ker}\nabla_{n+1}^z = \tfrac{1}{2}(2n+1)\chi(M) \ . \qquad (2.50)$$

This formula will be proven at the end of Sec. II.F. Since ∇_{-n}^z is the complex conjugate of ∇_z^n, we may restrict our attention to the case $n \geq -\frac{1}{2}$. By an argument similar to that given in Sec. II.C to show that $\mathrm{Ker}P_1 = 0$ for $h \geq 2$, one shows that for $h \geq 2$

$$\begin{aligned} &\dim \mathrm{Ker}\nabla_z^n = 0, \quad \text{for } n \geq 1 \ , \\ &\dim \mathrm{Ker}\nabla_z^0 = 1 \ , \\ &\dim \mathrm{Ker}\nabla_n^z = \dim \mathrm{Ker}\bar{\partial}_n \\ &\qquad = (2n-1)(h-1) \quad \text{for } n \geq \tfrac{3}{2} \ , \\ &\dim \mathrm{Ker}\nabla_1^z = \dim \mathrm{Ker}\bar{\partial}_1 = h \ . \end{aligned} \qquad (2.51)$$

For the torus, the dimension of every kernel is exactly one (except for spinors where there are no zero modes for even-spin structures):

$$\dim \mathrm{Ker}\nabla_z^n = 1, \quad \dim \mathrm{Ker}\nabla_n^z = 1 \ . \qquad (2.52)$$

For the sphere, there are no holomorphic forms, so that

$$\begin{aligned} &\dim \mathrm{Ker}\nabla_z^n = 2n+1 \quad \text{for } n \geq -\tfrac{1}{2} \ , \\ &\dim \mathrm{Ker}\nabla_n^z = 0 \quad \text{for } n \geq \tfrac{1}{2} \ . \end{aligned} \qquad (2.53)$$

As one can see, these dimensions involve only topological information. Note that, in the case of differentials of weight $\frac{1}{2}$, the index theorem yields no information. To obtain the dimensions for the torus we have just used Liouville's theorem, whereas for the sphere we used Lichnerowicz's theorem on the absence of harmonic spinors on the sphere. For genus $h \geq 2$, topological information is insufficient to determine the number of holomorphic $\frac{1}{2}$ differentials. In Sec. VI we shall see that indeed the number of holomorphic $\frac{1}{2}$ differentials depends on the spin structure for $h \geq 1$ and also on the moduli for $h \geq 3$.

It is convenient to single out now the differentials of special significance in string theory. We shall encounter holomorphic 1-forms or Abelian differentials of the first kind ω_I belonging to $\mathrm{Ker}\nabla_1^z$, whose integrals are the standard Abelian integrals of the first kind. There are h of these, and they generate the first cohomology group of the Riemann surface. One also has meromorphic 1-forms with one double pole (Abelian differentials of the second kind) or with two simple poles of opposite residues (Abelian differentials of the third kind). There is a differential in each case, and it is unique up to addition of holomorphic differentials. These facts about meromorphic forms require the full version of the Riemann-Roch theorem (cf. Sec. VI.C), and their explicit constructions in terms of the prime form will be given in Sec. VI.F. Next we have the holomorphic quadratic differentials belonging to $\mathrm{Ker}\nabla_2^z = \mathrm{Ker}\bar{\partial}_2$, of which there are $3h-3$ (complex ones). Together with their complex conjugates, they span the space of Teichmüller deformations, introduced in Sec. II.C and shown to be identical to the tangent space to moduli space $\mathcal{M}_h$. Finally, for fermionic strings, of special importance will be the holomorphic $\frac{1}{2}$-differentials, which are just zero modes of the Dirac operator, the meromorphic $\frac{1}{2}$-differentials, which will be used to construct fermion propagators, and the $2h-2$ complex $\frac{3}{2}$-differentials, which will make up the odd variables of supermoduli space (cf. Secs. III.E, III.F, and III.G).

To conclude this subsection, we introduce the concepts of Beltrami differentials and quasiconformal vector fields. Beltrami differentials span the space dual to holomorphic quadratic differentials. Thus they are differential forms of weight $(-1,1)$, of the form $\mu = \mu_{\bar{z}}{}^z d\bar{z}(dz)^{-1}$ and $\mu_z{}^z = 0$, and they can be integrated versus quadratic differentials $\phi = \phi_{zz}(dz)^2$:

$$\langle \mu \mid \phi \rangle = \int_M d^2z\, \mu_{\bar{z}}{}^z \phi_{zz} \ . \qquad (2.54)$$

[3]Dimensions of kernels of operators with complex indices are understood to be complex dimensions, whereas those of kernels of real operators are understood to be real.

Note that the pairing depends only on the conformal class and not on a particular choice of metric.

Beltrami differentials provide a natural parametrization of the metrics on the Riemann surface. If $d\hat{s}^2=\rho(z)\,|\,dz\,|^2$ is a metric on M, any other metric can be written as

$$ds^2=g_{mn}d\xi^m d\xi^n=\tilde{\rho}(z)\,|\,dz+\mu\,d\bar{z}\,|^2 , \tag{2.55}$$

with $\tilde{\rho}$ real and positive and μ a suitable Beltrami differential. The role of Beltrami differentials is best explained in terms of the associated Beltrami equation

$$\partial_{\bar{z}}w=\mu_{\bar{z}}{}^{z}\partial_z w . \tag{2.56}$$

At least for sufficiently small μ, this equation may be solved perturbatively in μ, and a solution is known always to exist locally. It can be written as

$$w(z,\bar{z})=z+v^z+O(v^z)^2 , \tag{2.57}$$

where the vector field v^z is defined locally by

$$\nabla_{\bar{z}}v^z=\mu_{\bar{z}}{}^z . \tag{2.58}$$

Since $ds^2=\tilde{\rho}(z)\,|\,dw/dz\,|^{-2}\,|\,dw\,|^2$, Eq. (2.57) means that the metric ds^2 comes from the metric $d\hat{s}^2$ by a Weyl transformation and a local reparametrization, and we have just restated the familiar fact that locally all conformal structures are the same. The Beltrami equation (2.56) takes real meaning only when we consider it in a global context. Indeed, whether it admits a global solution would tell us whether ds^2 and $d\hat{s}^2$ define the same conformal structure. There are several ways of expressing this more concretely: we could view Eq. (2.56) as defining a family of reparametrizations v^z on local coordinate patches, which, however, may not match. Whether they do can be measured by a vector-valued Cech cohomology class (see, for example, Sec. VI.A). This is the point of view of Kodaira-Spencer deformation theory; or, using uniformization (Sec. IV.A), we can represent M as cosets $\tilde{M}/\Gamma$ and $\tilde{M}/\hat{\Gamma}$ and solve Eq. (2.56) for a solution v^z which may not transform equivariantly; or, finally, consider solutions of the Beltrami equation which may have discontinuities. The last approach is the one we shall often adopt, with the vector fields v^z admitting jump discontinuities along closed curves on the surface M when μ deforms the complex structure. Vector fields of this type can be chosen to induce shifts, stretches, and twists. The corresponding transformations $z\to w(z,\bar{z})$ are called *quasiconformal transformations,* and we can in this way parametrize all deformations of complex structures by Beltrami differentials.

As an example of such vector fields, we may consider a piece of a surface that is a cylinder with Euclidean metric. Such configurations occur in Fenchel-Nielsen coordinates, in the light-cone diagrams of Mandelstam (1974a, 1974b, 1974c), or in the closed-string sector of the open-string field theory of Witten (1986a, 1986b), and they were explicitly given in D'Hoker and Giddings (1987). In Fig. 5 we illustrate the contours of discontinuity for the quasiconformal vector fields generating stretches (a), twists (b), or shifts (c) of the cylinder. Their analytic expressions are

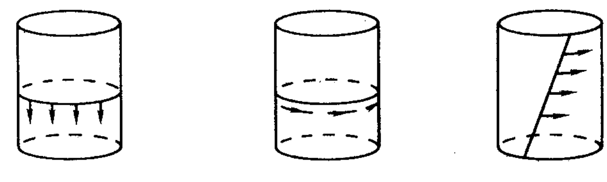

FIG. 5. Quasiconformal vector fields v: (a) generating stretches; (b) generating twists; (c) generating shifts of the cylinder.

$$\begin{aligned} v^z_{(a)}&=\frac{z+\bar{z}}{2a}H[a-\mathrm{Re}(z)] ,\\ v^z_{(b)}&=i\frac{z+\bar{z}}{2a}H[a-\mathrm{Re}(z)] ,\\ v^z_{(c)}&=\frac{z-\bar{z}}{2\alpha}+i\frac{\theta(z+\bar{z})}{2\Delta\tau} , \end{aligned} \tag{2.59}$$

where α is the radius of the cylinder, $\Delta\tau$ is its height, a is a parameter specifying the location of the discontinuity, and H is the Heaviside function.

It may be helpful at this point to clarify the tensor structure of moduli space. If we represent a conformal structure m by choosing a representative metric $g_{z\bar{z}}$, there is very little difference between Beltrami differentials and quadratic differentials, since we can raise and lower indices using $g_{z\bar{z}}$ to pass back and forth between the two notions. However, if we do not make such a choice, it is the Beltrami differentials that should be viewed as tangent vectors to moduli. In fact, Eq. (2.56) shows how to deform holomorphic structures without any choice of metrics. Since we still have to take into account reparametrization invariance, we see that

$$T(\mathcal{M}_h)=\frac{\{\text{Beltrami differentials}\}}{\{\text{Range}\,\partial_{\bar{z}}\text{ on vectors}\}} .$$

The pairing (2.54) exhibits, then, the quadratic differentials as cotangent vectors to moduli space. We now have a different way of explaining why this distinction disappears when a metric $g_{z\bar{z}}$ on the worldsheet is chosen. Such a metric provides a pairing on tensors, and hence on the tangent space to moduli at m. With this pairing, covariant and contravariant tensors on moduli space can be identified. This is why quadratic differentials appeared earlier [cf. Eq. (2.35)] as tangents to moduli.

F. Determinants and Weyl anomalies

We now study the behavior of the determinants of the Laplacians $\Delta_n^{(\pm)}$ under a Weyl scaling. These operators in general can have zero modes, which require special care. Let

$$N_n^+ = \dim \operatorname{Ker}(\nabla^z_{n+1})^\dagger = \dim \operatorname{Ker}\Delta_n^{(+)} ,$$
$$N_n^- = \dim \operatorname{Ker}\nabla^z_n = \dim \operatorname{Ker}\Delta_n^{(-)} , \tag{2.60}$$

and let ϕ_j be a basis for $\operatorname{Ker}\nabla^z_n$ and ψ_α a basis for $\operatorname{Ker}(\nabla^z_{n+1})^\dagger$. From Eqs. (2.42) and (2.44), it is evident that $N_n^\pm$ do not change under Weyl transformations. In fact, when one changes the metric from $\hat{g}_{mn}$ to $g_{mn} = e^{2\sigma}\hat{g}_{mn}$, the ϕ_j's do not change, whereas $\psi_\alpha = e^{2n\sigma}\hat{\psi}_\alpha$.

The regularized determinants can be defined by the heat-kernel, short-time cutoff procedure:

$$\ln \det{}'\Delta_n^{(\pm)} = -\int_\varepsilon^\infty \frac{dt}{t}(\operatorname{Tr}e^{-t\Delta_n^{(\pm)}} - N_n^\pm) . \tag{2.61}$$

Deletion of the zero modes from the determinant is indicated by a prime and requires the subtraction of the constants $N_n^\pm$, which makes the integral converge at $t=\infty$. Upon performing an infinitesimal Weyl transformation $\delta\sigma$, we have

$$\delta \ln \det{}'\Delta_n^{(\pm)} = \int_\varepsilon^\infty dt \operatorname{Tr}(\delta\Delta_n^{(\pm)} e^{-t\Delta_n^{(\pm)}}) .$$

The changes in the covariant derivatives and Laplacians follow from

$$\begin{aligned}
&\delta\nabla^z_n = -2\delta\sigma\nabla^z_n ,\\
&\delta\nabla_z^n = 2n\,\delta\sigma\nabla_z^n - 2n\nabla_z^n\delta\sigma ,\\
&\delta\Delta_n^{(-)} = 2(n-1)\delta\sigma\Delta_n^{(-)} + 4n\nabla_z^{n-1}\delta\sigma\nabla^z_n ,\\
&\delta\Delta_n^{(+)} = -2\delta\sigma\Delta_n^{(+)} - 4n\nabla^z_{n+1}\delta\sigma\nabla_z^n - 2n\Delta_n^{(+)}\delta\sigma ,
\end{aligned} \tag{2.62}$$

so that, for example,

$$\operatorname{Tr}(\delta\Delta_n^{(-)}e^{-t\Delta_n^{(-)}}) = \operatorname{Tr}[2(n-1)\delta\sigma\Delta_n^{(-)}e^{-t\Delta_n^{(-)}} - 2n\,\delta\sigma\Delta_{n-1}^{(+)}e^{-t\Delta_{n-1}^{(+)}}] . \tag{2.63}$$

Here we have used the rearrangement formula $e^{-AB}A = Ae^{-BA}$. Thus we get

$$\delta \ln \det{}'\Delta_n^{(-)} = \int_\varepsilon^\infty dt \operatorname{Tr}[2(n-1)\delta\sigma\Delta_n^{(-)}e^{-t\Delta_n^{(-)}} - 2n\,\delta\sigma\Delta_{n-1}^{(+)}e^{-t\Delta_{n-1}^{(+)}}] .$$

The t integral is easily carried out, and one obtains

$$\delta \ln \det{}'\Delta_n^{(-)} = -2(n-1)\operatorname{Tr}\delta\sigma e^{-t\Delta_n^{(-)}}\Big|_\varepsilon^\infty + 2n \operatorname{Tr}\delta\sigma e^{-t\Delta_{n-1}^{(+)}}\Big|_\varepsilon^\infty . \tag{2.64}$$

As $t\to\infty$, the heat kernels reduce to the projection operators onto $\operatorname{Ker}\nabla^z_n$ and $\operatorname{Ker}(\nabla^z_n)^\dagger$, respectively. Thus we have

$$\lim_{t\to\infty} \operatorname{Tr}\delta\sigma e^{-t\Delta_n^{(-)}} = \sum_{j=1}^{N_n^-} \langle \phi_j \mid \delta\sigma\phi_j \rangle . \tag{2.65}$$

On the other hand, the change in the finite-dimensional determinants of products of zero modes is given by

$$\begin{aligned}
\delta \ln \det\langle \phi_j \mid \phi_k \rangle &= \delta \operatorname{tr} \ln\langle \phi_j \mid \phi_k \rangle\\
&= \operatorname{tr}\delta\langle \phi_j \mid \phi_k \rangle\\
&= \sum_{j=1}^{N_n^-} (2-2n)\langle \phi_j \mid \delta\sigma\phi_j \rangle ,
\end{aligned} \tag{2.66}$$

where we have used the fact that the basis ϕ_j was chosen orthonormal. With the analogous result for the zero modes ψ_α, we find

$$\delta \ln \frac{\det{}'\Delta_n^{(-)}}{\det\langle \phi_j \mid \phi_k \rangle \det\langle \psi_\alpha \mid \psi_\beta \rangle} = 2(n-1)\operatorname{Tr}\delta\sigma e^{-\varepsilon\Delta_n^{(-)}} - 2n \operatorname{Tr}\delta\sigma e^{-\varepsilon\Delta_{n-1}^{(+)}} . \tag{2.67}$$

From the short-time expansions of the heat kernels, derived in Appendix B.

$$\begin{aligned}
\operatorname{Tr}\delta\sigma e^{-\varepsilon\Delta_n^{(+)}} &= \frac{1}{4\pi\varepsilon}\int_M d^2\xi\sqrt{g}\,\delta\sigma + \frac{1+3n}{12\pi}\int_M d^2\xi\sqrt{g}\,R\,\delta\sigma + O(\varepsilon) ,\\
\operatorname{Tr}\delta\sigma e^{-\varepsilon\Delta_n^{(-)}} &= \frac{1}{4\pi\varepsilon}\int_M d^2\xi\sqrt{g}\,\delta\sigma + \frac{1-3n}{12\pi}\int_M d^2\xi\sqrt{g}\,R\,\delta\sigma + O(\varepsilon) ,
\end{aligned} \tag{2.68}$$

one finally obtains

$$\delta \ln \frac{\det{}'\Delta_n^{(-)}}{\det\langle \phi_j \mid \phi_k \rangle \det\langle \psi_\alpha \mid \psi_\beta \rangle} = -\frac{1}{2\pi\varepsilon}\int_M d^2\xi\sqrt{g}\,\delta\sigma - \frac{6n^2-6n+1}{6\pi}\int_M d^2\xi\sqrt{g}\,R\,\delta\sigma . \tag{2.69}$$

Putting all together and integrating the infinitesimal Weyl transformation using Eq. (2.31), one finds

$$\frac{\det{}'\Delta_n^{(\pm)}}{\det\langle \phi_j \mid \phi_k \rangle_g \det\langle \psi_\alpha \mid \psi_\beta \rangle_g} = \frac{\det{}'\hat{\Delta}_n^{(\pm)}}{\det\langle \hat{\phi}_j \mid \hat{\phi}_k \rangle_{\hat{g}} \det\langle \hat{\psi}_\alpha \mid \hat{\psi}_\beta \rangle_{\hat{g}}} \times e^{-2(6n^2\pm 6n+1)S_L(\sigma)} . \tag{2.70}$$

Here the Liouville action is

$$S_L(\sigma) = \frac{1}{12\pi}\int_M d^2\xi\sqrt{\hat{g}}\,[\tfrac{1}{2}\hat{g}^{mn}\partial_m\sigma\partial_n\sigma + \mu^2(e^{2\sigma}-1) + R_{\hat{g}}\sigma] . \tag{2.71}$$

In particular, the formulas needed for bosonic string theory are given by

$$\begin{aligned}
\left[\frac{\det P_1^\dagger P_1}{\det\langle \phi_j \mid \phi_k \rangle_g}\right]^{1/2} &= \left[\frac{\det \hat{P}_1^\dagger \hat{P}_1}{\det\langle \hat{\phi}_j \mid \hat{\phi}_k \rangle_{\hat{g}}}\right]^{1/2} e^{-26 S_L(\sigma)} ,\\
\left[\frac{8\pi^2 \det{}'\Delta_g}{\int_M d^2\xi\sqrt{g}}\right]^{1/2} &= \left[\frac{8\pi^2 \det{}'\Delta_{\hat{g}}}{\int_M d^2\xi\sqrt{\hat{g}}}\right]^{1/2} e^{-S_L(\sigma)} .
\end{aligned} \tag{2.72}$$

In the expression for the scalar Laplacian, we have deleted the determinant of holomorphic Abelian differentials ω_I, because it is Weyl invariant all by itself, as can be seen from its explicit expression:

$$\det\langle \omega_I \mid \omega_J \rangle_g = \det \int_M d^2\xi \sqrt{g}\, g^{z\bar z}\{(\omega_I^*)_{\bar z}(\omega_J)_z + \text{c.c.}\} = \det\langle \omega_I \mid \omega_J \rangle_{\hat g} . \tag{2.73}$$

We are now also in a position to prove the Riemann-Roch theorem [Eq. (2.50)]. First we have

$$\dim \operatorname{Ker}\nabla_z^n - \dim \operatorname{Ker}\nabla^z_{n+1} = \dim \operatorname{Ker}\Delta_n^{(+)} - \dim \operatorname{Ker}\Delta_{n+1}^{(-)} = \operatorname{Tr} e^{-\varepsilon\Delta_n^{(+)}} - \operatorname{Tr} e^{-\varepsilon\Delta_{n+1}^{(-)}} . \tag{2.74}$$

Letting $\varepsilon \to 0$, we recover Eq. (2.50) in view of (2.68).

A standard reference on conformal anomalies is Coleman and Jackiw (1971), and a discussion of global and local conformal symmetry is given by Polchinski (1988a). The original calculation of the above Weyl anomaly is due to Polyakov (1981a) and has been clarified by Di Vecchia *et al.* (1982a, 1982b), Friedan (1982), Fujikawa (1982), Alvarez (1983), and Ambjørn *et al.* (1986); the articles of Di Vecchia *et al.*, Alvarez, and Ambjørn *et al.* also treat the case with boundaries. The first careful account of the crucial zero-mode factors is that of Alvarez (1983). Different aspects of Weyl invariance in string theory were treated by Fujikawa (1987) and Tanii (1987).

G. Amplitudes as integrals over moduli space for $h \geq 2$

We finally come to a detailed discussion of cancellation of Weyl anomalies. We shall deal with the case $h \geq 2$ first and present the cases of the torus and the sphere in the next section. The only modification will involve the presence of conformal Killing vectors.

1. The quantum measure and conformal invariance

To carry out the Dg integral we parametrize the space of metrics by $g = \exp(\delta v) e^{2\sigma}\hat g$, with $\hat g$ in a slice $\hat S$ transversal to the orbits of Weyl(M) and of Diff$_0$(M). Such a slice may be taken, for example, within $\mathcal{M}_{\text{const}}$, which guarantees right away that it is transverse to Weyl(M), by the uniqueness arguments of Sec. II.D. Here $\exp(\delta v)$ denotes integrated elements of Diff$_0$(M). [Recall that, for a vector field δv, the action $\exp(\delta v)$ on a metric is to replace its value at a given point on M by its value at the point on the integral curve of δv, a unit of time away.]

The change of variables $g \to (\sigma, v, \hat g)$ requires a Jacobian which can be evaluated from the decomposition (2.24).

It will be helpful to keep the following picture in mind. A Riemannian manifold, parametrized by a set of coordinates $x_1, \ldots, x_n$, is endowed with the standard volume element $\sqrt{g}\, d^n x$, which may be viewed as the volume (with respect to this metric) of the n coordinate vectors. In our case, we have the coordinates σ and v, which are functions on M, together with real coordinates m_j, $j = 1, \ldots, 6h-6$ for $\hat S$. Each element of $\hat S$ is a metric $\hat g(m)$, and tangents to $\hat S$ are symmetric two-tensors f_j defined by

$$\delta g(m) = \sum_{j=1}^{3h-3} \delta m_j \hat f_j .$$

Thus the coordinate vectors along $\hat S$ are $\delta\sigma\hat g$, $\hat P_1 \delta v$, and $\hat f_j$. Since we are interested in computing the Jacobian at an arbitrary point in $\mathcal{M}$, we shall apply the diffeomorphism $\exp(\delta v)$ and the Weyl rescaling $e^{2\sigma}$ under which $\hat f_j$ scales as $f_j = e^{2\sigma}\hat f_j$. The measure is

$$Dg_{mn} = \operatorname{Vol}_g(g\delta\sigma, P_1(\delta v), f_j) D\sigma\, Dv\, dm , \tag{2.75}$$

and P_1, whose definition [Eq. (2.23)] requires a metric, is always chosen with respect to g. Using the orthogonal decomposition (2.24) of δg_{mn}, we see that the first two entries are orthogonal, and that the last one may be restricted to the orthogonal projection of f_j onto $\operatorname{Ker}P_1^\dagger$ (see Fig. 6). When we use the orthogonality, the volume then decomposes into a product, and we obtain

$$Dg_{mn} = \operatorname{Vol}_g(g\delta\sigma)\operatorname{Vol}_g(P_1\delta v) \times \operatorname{Vol}_g(f_j \perp \text{proj. } \operatorname{Ker}P_1^\dagger) D\sigma\, Dv\, dm .$$

Ultralocality implies that the first factor is proportional to a factor of the type (2.14) and may be ignored. Moreover,

$$\operatorname{Vol}_g(P_1\delta v) = (\det P_1^\dagger P_1)^{1/2} .$$

Finally, recalling that $\operatorname{Ker}P_1^\dagger$ was spanned by basis vectors ϕ_j,

$$\operatorname{Vol}_g(f_j \perp \text{proj. } \operatorname{Ker}P_1^\dagger) = \frac{\det\langle f_j \mid \phi_k \rangle_g}{\det\langle \phi_j \mid \phi_k \rangle_g^{1/2}} .$$

Putting all together, we have

$$Dg_{mn} = \frac{\det\langle f_j \mid \phi_k \rangle_g}{\det\langle \phi_j \mid \phi_k \rangle_g^{1/2}} (\det P_1^\dagger P_1)^{1/2} D\sigma\, Dv\, dm . \tag{2.76}$$

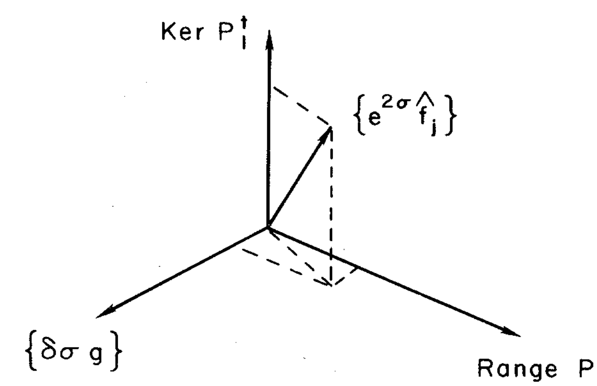

FIG. 6. Orthogonal decomposition of $f_j = e^{2\sigma}\hat f_j$.

The Weyl dependence of this measure is readily exhibited if we notice that

$$\langle f_j \mid \phi_k \rangle_g = \langle \hat{f}_j \mid \phi_k \rangle_{\hat{g}} = \langle \mu_j \mid \phi_k \rangle ,$$

where

$$\mu^z_{j\bar z} = \hat{g}^{z\bar z} \hat{f}_{j\bar z\bar z}$$

are the Beltrami differentials corresponding to moduli deformations along the slice $\hat{S}$. Using the Weyl rescaling formulas of (2.72), we get

$$Dg_{mn} = \left[\frac{\det \hat{P}_1^\dagger \hat{P}_1}{\det \langle \phi_j \mid \phi_k \rangle_{\hat{g}}} \right]^{1/2} \det \langle \mu_j \mid \phi_k \rangle \times e^{-26 S_L(\sigma)} D\sigma\, Dv^m \prod_{j=1}^{6h-6} dm_j \; . \qquad (2.77)$$

This is the desired result for $h \geq 2$, and we shall see that only a slight extension of it is necessary for $h=0,1$; the extension leaves the Weyl independence unaltered. We now discuss the choice of the normalization factor $\mathcal{N}$ in Eqs. (2.11) and (2.12). There are three possibilities.

(a) The metric $G_{\mu\nu}$ in $I(x,g)$ is flat Euclidean, in which case the x^μ integration may be carried out, and we find for genus h

$$\langle V_{i_1}(k_1) \cdots V_{i_n}(k_n) \rangle_h = \int_{\mathcal{T}_h} \prod_{j=1}^{6h-6} dm_j \frac{\det \langle \mu_j \mid \phi_k \rangle}{\det \langle \phi_j \mid \phi_k \rangle_{\hat{g}}^{1/2}} \int \frac{D\sigma\, Dv}{\mathcal{N}} e^{-\lambda\chi} (\det \hat{P}_1^\dagger \hat{P}_1)^{1/2} \left[\frac{8\pi^2}{\int d^2\xi \sqrt{g}} \det{}' \hat{\Delta} \right]^{-d/2} \times e^{-(26-d) S_L(\sigma)} \langle\langle V_{i_1}(k_1) \cdots V_{i_n}(k_n) \rangle\rangle \; . \qquad (2.78)$$

For $d=26$, we have Weyl invariance, and it is natural to set that $\mathcal{N} = \mathrm{Vol}(\mathrm{Diff}(M)) \times \mathrm{Vol}(\mathrm{Weyl}(M))$. Note that the inserted vertex operators are constructed in such a way that possible Weyl anomalies are required to cancel (see Sec. VIII). Upon integrating out the reparametrization vector fields, we produce a factor of $\mathrm{Vol}(\mathrm{Diff}_0(M))$, which partially cancels the analogous factor in $\mathcal{N}$ and, in view of Eq. (2.33), reduces the integral from $\mathcal{T}_h$ to $\mathcal{M}_h$:

$$\langle V_{i_1}(k_1) \cdots V_{i_n}(k_n) \rangle_h = e^{-\lambda\chi} \int_{\mathcal{M}_h} \prod_{j=1}^{6h-6} dm_j \frac{\det \langle \mu_j \mid \phi_k \rangle}{\det \langle \phi_j \mid \phi_k \rangle_{\hat{g}}^{1/2}} (\det \hat{P}_1^\dagger \hat{P}_1)^{1/2} \times \left[\frac{8\pi^2}{\int_M d^2\xi \sqrt{\hat{g}}} \det{}' \Delta_{\hat{g}} \right]^{-13} \langle\langle \hat{V}_{i_1}(k_1) \cdots \hat{V}_{i_n}(k_n) \rangle\rangle \; . \qquad (2.79)$$

Here $\langle\langle \;\; \rangle\rangle$ denotes the expectation value where only the x integral has been performed.

If we choose $\hat{S}$ to be a $6h-6$ dimensional slice within $\mathcal{M}_{\rm const}$ and transversal to the orbits of $\mathrm{Diff}_0(M)$, the Weil-Peterson measure $d(\mathrm{WP})$ is related to the measure $\prod dm_j$ on the slice $\hat{S}$ by

$$d(\mathrm{WP}) = \frac{\det \langle \mu_j \mid \phi_k \rangle}{\det \langle \phi_j \mid \phi_k \rangle_{\hat{g}}^{1/2}} \prod_{j=1}^{6h-6} dm_j \; , \qquad (2.80)$$

and we conclude

$$\langle V_{i_1}(k_1) \cdots V_{i_n}(k_n) \rangle_h = e^{-\lambda\chi} \int_{\mathcal{M}_h} d(\mathrm{WP}) (\det \hat{P}_1^\dagger \hat{P}_1)^{1/2} \left[\frac{8\pi^2}{\int_M d^2\xi \sqrt{\hat{g}}} \det{}' \Delta_{\hat{g}} \right]^{-13} \langle\langle \hat{V}_{i_1}(k_1) \cdots \hat{V}_{i_n}(k_n) \rangle\rangle \; . \qquad (2.81)$$

Observe that, unlike in gauge theories, the normalization factor $\mathcal{N}$ depends on the metric and has been placed inside the functional integrals. The justification for the above procedure may be given by appealing to the principle of ultralocality. Define a norm on vector fields δv^m analogous to Eqs. (2.13) and (2.21),

$$\|\delta v^m\|^2 = \int_M d^2\xi \sqrt{g}\, g_{mn} \delta v^m \delta v^n \; , \qquad (2.82)$$

and consider the integral

$$\int Dv^m e^{-\lambda \|\delta v^m\|^2} \; . \qquad (2.83)$$

The principle of ultralocality implies that this integral must be given by an expression of the form

$$\exp \left[-\mu_2^2(\lambda) \int_M d^2\xi \sqrt{g} \right] \qquad (2.84)$$

for some function $\mu_2^2(\lambda)$. As $\lambda \to 0$, Eq. (2.83) tends to $\mathrm{Vol}(\mathrm{Diff}_0(M))$, while (2.84) simply leads to a renormalization of the area term in Eq. (2.10) analogous to the one discussed in Eq. (2.14). Thus $\mathrm{Vol}(\mathrm{Diff}_0(M))$ is in effect irrelevant. The same argument applies to $\mathrm{Vol}(\mathrm{Weyl}(M))$. As for the "volume" of the mapping class group (=cardinality of MCG_h), it does not depend on g_{mn} but only on the topology. The only nontrivial volume element in $\mathcal{N}$ is that of the mapping class group, and it can thus be pulled out of under the integration, just as in the case of gauge theories, with the difference that the group is now discrete. The net effect, as mentioned above, is to reduce the integral over Teichmüller space to one over a fundamental domain for the mapping class group, which is the same as moduli space.

Critical dimensions of string theory in a flat Euclidean or Minkowski background have been obtained in a variety of ways over the years. In the light-cone gauge, it arises by insisting on Lorentz invariance, as pointed out by Brink and Nielsen (1973), Goddard *et al.* (1973), and Mandelstam (1974a, 1974b, 1974c). In the covariant approach, it appears by insisting on decoupling from the scattering amplitudes of negative norm states, as analyzed, for example, by Brower and Thorn (1971). Notice that, in our case, Weyl invariance will eliminate vertex operators producing unphysical states as well. In Sec. II.I we shall discuss how the critical dimension arises in a treatment with ghosts.

The above formulas for the Polyakov measure were obtained by D'Hoker and Phong (1986a) and independently by Moore and Nelson (1986). They provide a starting point for calculations of covariant multiloop amplitudes in the bosonic string.

In the days of dual models, multiloop amplitudes were considered by Kaku and Yu (1970), Lovelace (1970), Alessandrini (1971), Alessandrini and Amati (1971), and Kaku and Scherk (1971). These constructions were based on the assumption that the on-shell scattering vertices could actually be used as off-shell internal vertices, and unphysical states were generally not projected out. These shortcomings have been overcome more recently by Montonen (1974) and Neveu and West (1987a, 1987b) and through the introduction of ghost fields by Di Vecchia, Frau, *et al.* (1987a, 1987b) and Petersen and Sidenius (1987). It seems, however, that a precise definition of the integration region for moduli space is not available in these approaches. In the light-cone gauge, Mandelstam (1986a, 1986b) has also obtained explicit formulas for multiloop amplitudes. Finally, open-string amplitudes in the Polyakov string have been dealt with, for example, by Boulware and Newman (1986) and Burgess and Morris (1987a, 1987b).

From the above point of view, it is natural to divide by the volume of the mapping class group. This choice cannot really be justified within the context of the Polyakov ansatz; a further physical principle is required. This principle is the unitarity of the scattering amplitudes. To investigate unitarity, one may compare the Polyakov results with those of the manifestly unitary interacting string picture of Mandelstam (1973a, 1974a, 1986a, 1986b). Such a study was carried out by D'Hoker and Giddings (1987), and it was found that dividing out exactly once by the volume of the mapping class group leads to unitarity of all scattering amplitudes. For the bosonic string, unitarity is of course formal because of the presence of the tachyon. We shall come back to this question in Sec. V.G.

(b) $I(x^\mu,g)$ does not lead to a Weyl-invariant theory, for example, when d is different from 26. The $D\sigma$ integral is then not redundant, and the value $\mathcal{N}=\mathrm{Vol}(\mathrm{Diff}(M))$ is presumably the correct choice if a unitary theory exists at all. Since all determinants are real, there are no global gravitational anomalies, and we may factor out the mapping class group by restricting integrals from Teichmüller to moduli space:

$$\langle V_{i_1}(k_1)\cdots V_{i_n}(k_n)\rangle=\sum_{h=0}^{\infty}\int_{\mathcal{M}_h}d(\mathrm{WP})\int D\sigma\, Dx^\mu(\det\hat{P}_1^\dagger\hat{P}_1)^{1/2}V_{i_1}(k_1)\cdots V_{i_n}(k_n)e^{-26S_L(\sigma)}e^{-I(x,\hat{g})}\ . \tag{2.85}$$

No satisfactory quantization of the Liouville model has been achieved to date for closed strings, despite attempts by D'Hoker, Freedman, and Jackiw (1983) and Braaten *et al.* (1984). In the case of open strings alone, it seems possible to obtain a tachyon-free string theory in special dimensions 1, 7, 13, 19, and 25 with discrete mass spectrum, as discussed by Gervais and Neveu (1982, 1983, 1984, 1986), Marnelius (1983), and Bilal and Gervais (1987a, 1987b). An interesting proposal has been made very recently by Polyakov (1987a) based on SL(2,**R**) current algebra.

(c) For general background metric $G_{\mu\nu}(x)$ (and possibly antisymmetric tensor and dilaton fields), insisting upon Weyl invariance for the string tree-level theories leads to equations for the background fields and for the dimension of space-time. These equations were studied by Friedan (1980), Lovelace (1984, 1986), Callan *et al.* (1985), Fradkin and Tseytlin (1985a, 1985b), Callan, Klebanov, and Perry (1986), and Fridling and Jevicki (1986). The critical dimension $d=26$ emerges then as the condition of vanishing of the first coefficient of the dilation β function. It is possible that the dimension of space-time is dynamical, as suggested by Polyakov (1986). The solutions to these equations provide possible spaces for consistent string propagation. Higher string loop effects will in general again spoil the Weyl invariance, even if these background equations are satisfied. Fischler and Susskind (1986a, 1986b) have, however, argued that such effects should be understood as loop corrections to the string background equations of motion. Explicit examples of how this might happen have been presented for the open-string case by Callan *et al.* (1987, 1988).

2. Scalar Green's function and amplitudes

The Green's function is defined by $G(z,w)=\langle x(z)x(w)\rangle$, and in locally conformal coordinates z, with metric $ds^2=\rho\, dz\, d\bar{z}$, satisfies

$$\int d^2z\sqrt{g}\,G(z,w)=0\ ,$$
$$\partial_z\partial_{\bar{z}}G(z,w)=-2\pi\delta(z-w)+\frac{2\pi g_{z\bar{z}}}{\int d^2z\sqrt{g}}\ , \tag{2.86}$$
$$\partial_z\partial_{\bar{w}}G(z,w)=2\pi\delta(z-w)-\pi\sum_{I,J}\omega_I(z)(\mathrm{Im}\Omega)^{-1}_{IJ}\bar{\omega}_J(w)\ .$$

The "period matrix" Ω_{IJ} is the matrix of periods of the

Abelian integrals associated with ω_I. It will be defined in detail in Sec. VI.D. It is invariant under $\mathrm{Diff}_0(M)$ and $\mathrm{Weyl}(M)$, and characterizes the conformal structure of the surface. In Eq. (2.86), the additional terms besides the δ functions result from projections on the spaces of zero modes of ∇_0^z and ∇_1^z, respectively. They break Weyl invariance and, in fact, under scalings $g=e^{2\sigma}\hat{g}$ the two-point functions will transform as

$$G(z,w)=\hat{G}(z,w)-\frac{1}{\int d^2z\sqrt{g}}\int d^2v\sqrt{g}\,[\hat{G}(z,v)+\hat{G}(v,w)]+\frac{1}{\left[\int d^2z\sqrt{g}\right]^2}\int\int d^2v\,d^2y\sqrt{g(v)}\sqrt{g(y)}\hat{G}(v,y)\ . \tag{2.87}$$

For coincident points, we may again regularize $G(z,w)$ by a heat kernel with small-time cutoff procedure. The regularized Green's function $G_R(z,z)$ at coincident points will satisfy a similar scaling law:

$$G_R(z,z)=\hat{G}_R(z,z)+2\sigma(z)-\frac{2}{\int d^2z\sqrt{g}}\int d^2v\sqrt{g}\,\hat{G}(z,v)+\frac{1}{\left[\int d^2z\sqrt{g}\right]^2}\int\int d^2v\,d^2y\sqrt{g(v)}\sqrt{g(y)}\hat{G}(v,y)\ , \tag{2.88}$$

with the key additional term $\sigma(z)$ on the right-hand side.

The field x itself does not have a definite conformal dimension, but derivatives of x as well as $\exp(ikx)$ do, and our task is to determine the (Weyl-invariant) forms of their correlation functions, taking into account proper renormalization for composites such as $\exp(ikx)$. For example,

$$\langle \partial_z x\,\partial_w x\rangle=\partial_z\partial_w G(z,w)$$

is Weyl invariant in view of Eq. (2.87), and $\partial_{\bar z}x$ has conformal dimension (1,0). More subtle are correlation functions of the operator

$$V_k(z)=\rho^{k^2/2}e^{ikx(z)}\ . \tag{2.89}$$

Since the exponential should be viewed as normal ordered, we can replace the Green's function at points (z,z) and (w,w) by their regularizations, and we arrive at

$$\begin{aligned}&\langle V_k(z)V_{k'}(w)\rangle=\delta(k+k')F(z,w)^{-k^2}\ ,\\&F(z,w)=[\rho(z)\rho(w)]^{-1/2}\\&\qquad\times\exp[-G(z,w)+\tfrac{1}{2}G_R(z,z)+\tfrac{1}{2}G_R(w,w)]\ .\end{aligned} \tag{2.90}$$

The crucial feature of $F(z,w)$ is that it is Weyl invariant, as can be deduced from Eqs. (2.87) and (2.88) and behaves like $|z-w|^2$ for z near w. Thus the conformal dimension of the vertex operator $V_k(z)$ is well defined and equal to $(k^2/2,k^2/2)$.

Although the scalar Green's function $G(z,w)$ depends in a more complicated way on the metric, the function $F(z,w)$ can actually be written explicitly in terms of the prime form. In fact, Eq. (2.90) shows that $F(z,w)$ is a single-valued real symmetric expression, transforming in each variable as a $(-\frac{1}{2},-\frac{1}{2})$ differential. Furthermore, it satisfies the equation

$$\partial_z\partial_{\bar z}\ln F(z,w)=2\pi\delta(z-w)-\pi\sum_{I,J}\omega_I(z)(\mathrm{Im}\Omega)^{-1}_{IJ}\bar{\omega}_J(z)\ .$$

All these properties characterize $F(z,w)$ as

$$F(z,w)=\exp\left[-2\pi\,\mathrm{Im}\int_w^z\omega(\mathrm{Im}\Omega)^{-1}\mathrm{Im}\int_w^z\omega\right]\times|E(z,w)|^2\ . \tag{2.91}$$

Here, Ω is again the period matrix, and $E(z,w)$ is a holomorphic $(-\frac{1}{2},0)$ form in z and w with a single zero at $z=w$, called the prime form. We shall define it in detail in Sec. VI and give a representation for it in terms of ϑ functions. At present, we need only the above-mentioned properties.

The above Green's function allows us to evaluate the scattering amplitudes very explicitly. Consider first the insertion of an exponential factor, universal to all vertex operators, leaving its position on the surface free. Thus we deal with multiple insertions of $V_k(z)$ of Eq. (2.89),

$$\langle\langle V_{k_1}(z_1)\cdots V_{k_n}(z_n)\rangle\rangle=(2\pi)^{26}\delta(k)\exp\left[-\tfrac{1}{2}\sum_{i\neq j=1}^{n}k_i\cdot k_j G(z_i,z_j)+\sum_{i=1}^{n}\tfrac{1}{2}k_i^2[\ln\rho(z_i)-G_R(z_i,z_i)]\right]\ . \tag{2.92}$$

Using Eq. (2.90) one may recast this solely in terms of F,

$$\langle\langle V_{k_1}(z_1)\cdots V_{k_n}(z_n)\rangle\rangle=(2\pi)^{26}\delta(k)\prod_{i<j}F(z_i,z_j)^{k_i\cdot k_j}\ . \tag{2.93}$$

In the special case of the tachyon, we have

$$V_k=\int d^2z\,V_k(z),\quad V_k(z)=\sqrt{g}\,(z)e^{ik\cdot x(z)}\ ,$$

with $k_\mu k^\mu=2$, and the above formula may be applied directly.

For massless particles, one should rather start from the generating function for amplitudes with one derivative on x,

$$V_k^*(z,\zeta)=\exp[ik\cdot x(z)+\zeta_\mu\partial_z x^\mu(z)+\bar{\zeta}_\mu\partial_{\bar z}x^\mu(z)]\ , \tag{2.94}$$

so that amplitudes may be gotten from the above by retaining only terms linear in ζ and $\bar{\zeta}$. Since $k^2=0$, no ρ-dependent prefactor occurs. Correlation functions of V_k^*

may be worked out as easily as those of V_k. Let us just notice here that such correlation functions are already implicit in Eq. (2.93). Indeed, it suffices to replace ζ_μ by the difference between two momenta, taking their insertion points infinitesimally far apart, so Eq. (2.93) may be viewed as a generating function for all amplitudes. Its Weyl invariance guarantees the Weyl invariance of the original vertex operators, as will be explained in Sec. VIII.

To conclude this section, we discuss the role of internal loop momenta. In accord with the radial quantization procedure, the momentum operator measuring the momentum flowing through a contour C is given by

$$P_C^\mu = \oint_C \frac{dx}{2\pi}\partial_z x^\mu(z) . \tag{2.95}$$

The internal (or loop) momenta may be defined as the momenta flowing through the A_I-homology cycles (see Fig. 10 below)

$$P_I^\mu = \oint_{A_I} \frac{dz}{2\pi}\partial_z x^\mu(z) , \tag{2.96}$$

and amplitudes at fixed internal momenta (and fixed genus h) may be introduced by inserting δ functions in the functional integral:

$$\langle\langle V_1 \cdots V_n \rangle\rangle(p_I^\mu) = \left[\frac{8\pi^2 \det{}'\Delta_g}{\int_M d^2\xi\sqrt{g}}\right]^{13} \int Dx \prod_{I,\mu}\delta\left[p_I^\mu - \oint_{A_I}\frac{dz}{2\pi}\partial_z x^\mu\right] V_1 \cdots V_n e^{-I_0(x)} . \tag{2.97}$$

These amplitudes produce Weyl(M)- and Diff$_0$(M)-invariant amplitudes after the Faddeev-Popov ghost determinant has been taken into account, and as long as $V_1 \cdots V_n$ are physical vertex operators. However, we do not have invariance under the full mapping class group because a choice of homology basis has been made. The full amplitude is of course obtained after integrating over p,

$$\langle\langle V_1 \cdots V_n \rangle\rangle = \int dp_I^\mu \langle\langle V_1 \cdots V_n \rangle\rangle(p_I^\mu) . \tag{2.98}$$

For the special case of exponential insertions, Eq. (2.97) is easily evaluated, and we get

$$\langle\langle V_1 \cdots V_n \rangle\rangle(p_I^\mu) = (\det \mathrm{Im}\Omega)^{13}\left|\exp\left[i\pi p_I^\mu \Omega_{IJ} p_J^\mu + 2\pi i p_I^\mu \sum_i k_i^\mu \int_P^{z_i}\omega_I\right]\right|^2 \prod_{i<j} |E(z_i,z_j)|^{2k_i\cdot k_j} . \tag{2.99}$$

Later on we shall more fully explore the meaning of the above formulas in function of the holomorphic structure of moduli space.

The important observation that regularization produces a factor of $\sigma(z)$ leading to well-defined conformal dimensions is due to Polyakov (1981a). It is the starting point for determining the mass spectrum of the string by requiring Weyl invariance, an issue that will be discussed at length in Sec. VIII. The above careful discussion of scaling laws for two-point functions taking into account zero modes and global issues is due to Verlinde and Verlinde (1987a). They also point out that Eq. (2.90) can be inverted, producing a formula of type (2.87) for the Green's function, with $\hat{G}(z,w)$ on the right-hand side replaced by $\ln F(z,w)$. The basic ideas and some examples of radial quantization are in Fubini, Hanson, and Jackiw (1973). Internal momenta are of course familiar from the dual-model theories, but in the above form they were rediscovered by Verlinde and Verlinde (1987b).

H. Amplitudes for tree and one-loop level

As was explained in Sec. II.C, the tree and one-loop cases do not follow the pattern exhibited for $h \geq 2$. The main complication is that there now exist conformal Killing vectors ψ_α belonging to KerP_1. Thus the operators P_1 and $P_1^\dagger P_1$ should be acting only on the reparametrizations in $(\mathrm{Ker}P_1)^\perp$. It is convenient to treat the cases $h=0$ and 1 separately.

1. Tree-level amplitudes

For $h=0$, we have six real conformal Killing vectors and no moduli parameters. The measure (2.76) must be modified to

$$Dg_{mm} = (\det{}'P_1^\dagger P_1)^{1/2} D'v^m D\sigma , \tag{2.100}$$

where the prime on $D'v^m$ denotes the fact that it is restricted to $(\mathrm{Ker}P_1)^\perp$. Under a Weyl transformation we obtain from Eq. (2.70) that

$$\left[\frac{\det{}'P_1^\dagger P_1}{\det\langle\psi_\alpha|\psi_\beta\rangle_g}\right]^{1/2} = \left[\frac{\det{}'\hat{P}_1^\dagger \hat{P}_1}{\det\langle\hat{\psi}_\alpha|\hat{\psi}_\beta\rangle_{\hat{g}}}\right]^{1/2} e^{-26S_L(\sigma)} \tag{2.101}$$

and hence

$$Dg_{mn} = (\det{}'\hat{P}_1^\dagger\hat{P}_1)^{1/2}\left[\frac{\det\langle\psi_\alpha|\psi_\beta\rangle_g}{\det\langle\hat{\psi}_\alpha|\hat{\psi}_\beta\rangle_{\hat{g}}}\right]^{1/2} \times e^{-26S_L(\sigma)} D\sigma\, D'v^m . \tag{2.102}$$

Note that the above ratio of finite-dimensional determinants is precisely the ratio of volumes[4] of $\mathrm{Ker}P_1$ and $\mathrm{Ker}\hat{P}_1$, so that we have for the analog of Eq. (2.77)

$$Dg_{mn}=\frac{1}{\mathrm{Vol}(\mathrm{Ker}\hat{P}_1)}(\det{}'\hat{P}_1^\dagger\hat{P}_1)^{1/2}e^{-26S_L(\sigma)}D\sigma\, Dv^m\ , \tag{2.103}$$

where all reparametrizations v^m are now included in Dv^m. Assuming that we work in the critical dimension $d=26$ and adopt the normalization factor $\mathcal{N}=\mathrm{Vol}(\mathrm{Diff}(M))\times\mathrm{Vol}(\mathrm{Weyl}(M))$ as in Eq. (2.78), we have a general formula for tree-level scattering amplitudes:

$$\langle V_{i_1}(k_1^\mu)\cdots V_{i_n}(k_n^\mu)\rangle=ce^{-2\lambda}\langle\langle V_{i_1}(k_1^\mu)\cdots V_{i_n}(k_n^\mu)\rangle\rangle \times\frac{1}{\mathrm{Vol}(\mathrm{Ker}\hat{P}_1)}\ . \tag{2.104}$$

Here the symbol $\langle\langle\ \rangle\rangle$ denotes again the fact that the functional integral over x alone has been performed. The determinants of $\Delta_{\hat{g}}$ and $\hat{P}_1^\dagger\hat{P}_1$ are constants, since there are no moduli parameters left, and their contribution has been denoted by c. This constant has been computed by Weisberger (1987a).

Of course, $\mathrm{Vol}(\mathrm{Ker}\hat{P}_1)$ is infinite for the sphere, so the above expression can be nonzero only if the vacuum expectation value of the vertex operators involves a similar infinite factor. This can indeed happen because the vertex operators are integrals over the sphere of local functions:

$$V_j(k_j^\mu)=\int_M d^2z_j\sqrt{g}\,U_j[\varepsilon_j,Dx]e^{ik_j^\mu x_\mu}\ , \tag{2.105}$$

where U_j is a polynomial in derivatives of x and depends linearly on the polarization tensor ε_j (see Sec. VIII for details). Since V_j is reparametrization invariant, it is in particular invariant under the group $\mathrm{PSL}(2,C)=\mathrm{SL}(2,C)/\{\pm 1\}$ of conformal Killing transformations, acting on the coordinates z_j of the compactified plane (i.e., the sphere) by Möbius transformations

$$z_j\mapsto\frac{az_j+b}{cz_j+d}\quad\text{with } ad-bc=1\ . \tag{2.106}$$

This group acts freely on all the z_j's, and we may fix three arbitrary distinct points with the help of a unique Möbius transformation in $\mathrm{PSL}(2,C)$ and factor out the $\mathrm{PSL}(2,C)$-invariant volume. The latter is constructed by recalling that, under the Möbius transformation of Eq. (2.106), we have

$$dz_j\mapsto\frac{dz_j}{(cz_j+d)^2}\ ,\qquad z_i-z_j\mapsto\frac{z_i-z_j}{(cz_i+d)(cz_j+d)}\ , \tag{2.107}$$

so that the volume element on $\mathrm{Ker}\hat{P}_1$ is a constant times

$$d\mu=\frac{d^2z_id^2z_jd^2z_k}{|z_i-z_j|^2|z_j-z_k|^2|z_k-z_i|^2}\ , \tag{2.108}$$

where i,j,k denote any three distinct points among $1,2,\ldots,n$. The fact that three such points must be fixed, and the appearance of the difference factors in Eq. (2.108), are familiar from dual-model calculations.

To see how the volume of the conformal Killing group is factored out, it is instructive to compute the scattering amplitude for tachyonic particles only. In this case we have $U_j[\varepsilon_j,Dx]=\hat{\varepsilon}$ in Eq. (2.105) with $\hat{\varepsilon}$ constant and $k_\mu k^\mu=2$. We need the Green's function $G(z,z')$ for scalars on the sphere. Due to the zero mode of Δ, we can invert Δ only on the space of functions orthogonal to constants

$$\Delta_z G(z,z')=4\pi\delta^2(z,z')-\frac{4\pi}{\int_M d^2z\sqrt{g}}\ . \tag{2.109}$$

With the standard metric on the sphere

$$g_{mn}=\frac{2\delta_{mn}}{(1+|z|^2)^2} \tag{2.110}$$

it is easily checked that

$$G(z,z')=-\ln\frac{|z-z'|^2}{(1+|z|^2)(1+|z'|^2)}\ . \tag{2.111}$$

Evaluating the contractions over x fields in Eq. (2.104), we get

$$\langle\langle V(k_1)\cdots V(k_n)\rangle\rangle=\hat{\varepsilon}^n(2\pi)^{26}\delta(k)\prod_{j=1}^n\int\frac{2\,d^2z_j}{(1+|z_j|^2)^2}\exp\left[-\tfrac{1}{2}\sum_{i,j=1}^n k_i\cdot k_jG(z_i,z_j)\right]\ , \tag{2.112}$$

where $k=k_1+\cdots+k_n$ is the total momentum. From $k_i^2=2$ and $k=0$, it follows that the $(1+|z|^2)$ factors cancel out and we have

$$\langle\langle V(k_1)\cdots V(k_n)\rangle\rangle=\hat{\varepsilon}^n(2\pi)^{26}\delta(k)\prod_{j=1}^n\int d^2z_j\prod_{i<j}|z_i-z_j|^{2k_i\cdot k_j}\exp\left[\sum_{i=1}^n G(z_i,z_i)\right]\ . \tag{2.113}$$

The singularity that arises from considering the Green's function at coincident points should be regularized in a $\mathrm{PSL}(2,C)$-invariant way. This requires setting $G(z_i,z_i)$ to a constant independent of z_i. This constant arises once for

[4]More accurately it should be understood as the volume of the corresponding Lie group.

each vertex operator and thus effectively modifies $\hat{\varepsilon}$ to $\varepsilon=\hat{\varepsilon}e^{G(z_i,z_i)}$. The remaining expression in Eq. (2.113) is PSL(2,C) invariant and thus divergent due to the infinite volume of Ker$\hat{P}_1$. We now fix three points z_{n-2},z_{n-1},z_n and isolate the invariant measure associated with them, as given in Eq. (2.108):

$$\langle\langle V(k_1)\cdots V(k_n)\rangle\rangle=\varepsilon^n(2\pi)^{26}\delta(k)\int d\mu\left[|z_{n-2}-z_{n-1}|^2|z_{n-1}-z_n|^2|z_n-z_{n-2}|^2 \times\prod_{j=1}^{n-3}\int d^2z_j\prod_{1\le i<j\le n}|z_i-z_j|^{2k_i\cdot k_j}\right]. \tag{2.114}$$

It is not hard to see that the object in large parentheses is PSL(2,C) invariant all by itself upon transformation of all z_j with $j=1,2,\ldots,n$, so that the first integral yields Vol(Ker$\hat{P}_1$). It is customary to fix $z_{n-2}=0$, $z_{n-1}=1$, $z_n=\infty$, and one then obtains with the help of Eq. (2.104)

$$\langle V(k_1)\cdots V(k_n)\rangle=ce^{-2\lambda}\varepsilon^n(2\pi)^{26}\delta(k)\prod_{j=1}^{n-3}\int d^2z_j\prod_{1\le i<j\le n-1}|z_i-z_j|^{2k_i\cdot k_j}. \tag{2.115}$$

Introducing the generalized Mandelstam variables[5]

$$s_{ij}=-(k_i+k_j)^2$$

we see that Eq. (2.115) is absolutely convergent only if each s_{ij} is below the tachyon threshold,

$$\mathrm{Re}(s_{ij})<-2\ ,$$

and if

$$1<\sum_{j=1,j\neq i}^{n-1}\left[1+\frac{s_{ij}}{4}\right].$$

The function elsewhere is defined by analytic continuation, which will introduce imaginary components to the amplitude, so that it can obey the correct factorization properties. Note that, for real s_{ij}, the amplitude (2.115) cannot be correct for all ranges of s_{ij}, otherwise the full amplitude would be real, which is inconsistent with unitarity.

These closed-string amplitudes are analogous to those obtained by Koba and Nielsen (1969) for the open string. For the three-point function we have

$$\langle V(k_1)V(k_2)V(k_3)\rangle=ce^{-2\lambda}\varepsilon^3(2\pi)^{26}\delta(k)\ , \tag{2.116}$$

whereas for the four-point function we find the manifestly dual amplitude of Shapiro (1970) and Virasoro (1969),

$$\langle V(k_1)\cdots V(k_4)=ce^{-2\lambda}\varepsilon^4(2\pi)^{26}\delta(k)\int d^2z_1|z_1|^{2k_1\cdot k_2}|1-z_1|^{2k_1\cdot k_3} =ce^{-2\lambda}\varepsilon^4\pi(2\pi)^{26}\delta(k)\frac{\Gamma(-1-s/2)\Gamma(-1-t/2)\Gamma(-1-u/2)}{\Gamma(2+s/2)\Gamma(2+t/2)\Gamma(2+u/2)}\ , \tag{2.117}$$

where the Mandelstam variables for the four-particle amplitude are as usual denoted by $s=-(k_1+k_2)^2$, $t=-(k_2+k_3)^2$, $u=-(k_1+k_3)^2$. Identification of the amplitude with a combination of Γ functions has also freed us from the necessity for separate analytic continuation. The Γ function exhibits all the required factorization properties. Factorization in the s channel at the tachyon pole $s\sim-2$ imposes an additional relation on the constants λ and ε (recall that c was in principle calculated),

$$(ce^{-2\lambda}\varepsilon^3)^2=8\pi^2ce^{-2\lambda}\varepsilon^4$$

so that (2.118)

$$\varepsilon^2=\frac{8\pi^2}{c}e^{2\lambda}\ ,$$

and the normalization ε of the vertex operator is completely determined by the unique coupling constant λ, as pointed out by Weinberg (1985).

Our analysis has tacitly assumed that at least three vertex operators were inserted. When no vertex operator is inserted one should replace $(2\pi)^{26}\delta(k)$ by the volume Ω of space-time; we have $\langle\langle 1\rangle\rangle=1$, and by virtue of Eq. (2.104) the full amplitude vanishes. Physically, this indicates that the space-time cosmological constant vanishes at tree level. When one vertex is inserted, only one point on the sphere can be fixed, and after fixing that point, one should no longer factor out Vol(Ker$\hat{P}_1$), but rather the volume of the isotropy subgroup leaving that point invariant. If one chooses the fixed point at infinity, then this subgroup is generated by translations, rotations, and dilations in the plane and still has infinite volume under the PSL(2,C)-invariant measure on this group, which can be parametrized by two points z_1,z_2 in the plane. The invariant volume element is then $d^2z_1d^2z_2/|z_1$

[5]Our convention for s has a $-$ sign, because k_i is really Euclidean. Upon analytic continuation to Minkowski space-time, s is the usual Mandelstam variable.

$-z_2|^4$, whose integral indeed diverges. Thus the one-point function vanishes. Physically, this indicates that flat space-time is a tree-level solution to the string equations of motion. Finally, if two vertex operators are inserted, one should fix two points, and it remains to divide out by the volume of the isotropy subgroup leaving two points invariant. When one fixes these points at zero and infinity, the group is that of rotations and dilations and again has infinite volume, so that the two-point function also vanishes to tree level. Physically this means that tree-level mass and wave-function corrections are absent.

Finally, as is well known from dual-model theory, the amplitude (2.115) completely factorizes, a procedure that may be used to compute amplitudes for particles other than tachyons.

Tree-level amplitudes in the Polyakov formulation were studied by Nepomechie (1982) and Aoymama, Dhar, and Namazie (1986), and a prescription for linking their calculation to that of open strings (especially convenient for amplitudes of particles with spin) was suggested by Kawai, Lewellen, and Tye (1986).

2. One-loop-level amplitudes

For $h=1$, we have two real moduli τ_1,τ_2 and two real conformal Killing vectors, and the measure (2.77) is thus modified to

$$Dg_{mn}=(\det{}'P_1^\dagger P_1)^{1/2}\frac{\det\langle\mu_j\,|\,\phi_k\rangle}{\det\langle\phi_j\,|\,\phi_k\rangle^{1/2}}D\sigma\, D'v^m d^2\tau\ , \tag{2.119}$$

where the prime on $D'v^m$ denotes the omission of the two conformal Killing vectors. Under Weyl rescaling, any metric g can be mapped into a flat metric $\hat g$ with unit area, and in view of Eq. (2.70) the measure (2.119) becomes

$$Dg_{mn}=\left[\frac{\det{}'\hat P_1^\dagger \hat P_1}{\det\langle\phi_j\,|\,\phi_k\rangle_{\hat g}}\right]^{1/2}\det\langle\mu_j\,|\,\phi_k\rangle \times\left[\frac{\det\langle\psi_\alpha\,|\,\psi_\beta\rangle_g}{\det\langle\hat\psi_\alpha\,|\,\hat\psi_\beta\rangle_{\hat g}}\right]^{1/2} e^{-26S_L(\sigma)}D\sigma\, D'v^m d^2\tau\ . \tag{2.120}$$

Now Eq. (2.80) may be used to rewrite the τ integral as the Weil-Petersson measure, and the conformal Killing determinants of Eq. (2.120) may be handled as for the case of the sphere. Thus we obtain

$$Dg_{mn}=\frac{1}{\mathrm{Vol}(\mathrm{Ker}\hat P_1)}(\det{}'\hat P_1^\dagger \hat P_1)^{1/2}e^{-26S_L(\sigma)} \times D\sigma\, Dv^m d(\mathrm{WP})\ . \tag{2.121}$$

The standard representation of the torus is by a parallelogram in the complex plane, with sides 1 and $\tau=\tau_1+i\tau_2$ and $\mathrm{Im}(\tau)>0$, periodic boundary conditions, and the Euclidean metric. This slice is actually not of unit area (instead $\int d^2\xi\sqrt{g}=2\tau_2$), but because of Weyl invariance this choice is equivalent. The space of all tori obtained this way spans Teichmüller space and is parametrized by τ in the complex upper half-plane $\mathbf{H}=\{\tau=\tau_1+i\tau_2;\tau_2>0\}$. In Sec. IV.A, we shall describe an explicit construction of the Weil-Petersson measure for a slice of unit area yielding

$$d(\mathrm{WP})=2\frac{d^2\tau}{\tau_2^2}\ . \tag{2.122}$$

The torus obtained in this fashion is equivalent under Diff(M) to any torus with modular parameter $\tau'=(a\tau+b)/(c\tau+d)$, $ad-bc=1$, and a,b,c,d integers. These transformations form a group PSL(2,Z). However the group of "large" diffeomorphisms, i.e., the mapping class group (or modular group for the torus), in addition includes the transformation that flips the sign of both sides of the parallelogram, corresponding to the element $-I$ of SL(2,Z). Thus the full modular group should be taken to be SL(2,Z)=MCG$_1$. Moduli space is obtained from Teichmüller space by identification under the mapping class group. For simplicity, we still identify it with the fundamental domain $\mathcal{M}_1$ of PSL(2,R),

$$\mathcal{M}_1=\{\tau=\tau_1+i\tau_2 \text{ with } \tau_2>0, -\tfrac{1}{2}\le\tau_1\le\tfrac{1}{2}, |\tau|\ge 1\}\ , \tag{2.123}$$

represented in Fig. 7 on the condition of including a fac-

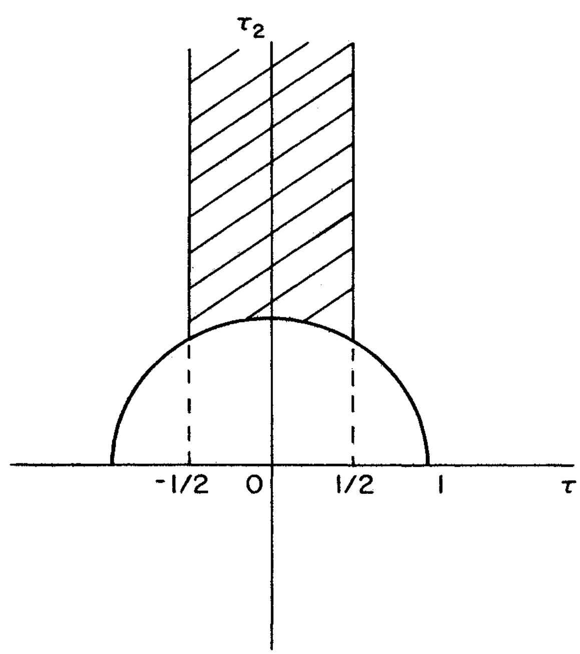

FIG. 7. A fundamental domain for PSL(2,Z), representing the torus.

tor of $\frac{1}{2}$ whenever we replace an integral over moduli space by an integral over $\mathcal{M}_1$. The Weil-Petersson measure $d(\mathrm{WP})$ is clearly invariant under MCG_1, so it may be projected down onto moduli space $\mathcal{M}_1$. More on this subject will be said in Sec. IV.B.

We are now in a position to write down the expression for one-loop scattering amplitudes:

$$\langle V_{i_1}(k_1)\cdots V_{i_n}(k_n)\rangle=\int_{\mathcal{M}_1}\frac{d^2\tau}{\tau_2^2}(\det{}'\hat{P}_1^\dagger\hat{P}_1)^{1/2}\left[\frac{4\pi^2}{\tau_2}\det{}'\Delta_{\hat{g}}\right]^{-13}\langle\langle V_{i_1}(k_1)\cdots V_{i_n}(k_n)\rangle\rangle\frac{1}{\mathrm{Vol}(\mathrm{Ker}\hat{P}_1)}\ . \tag{2.124}$$

The determinants are evaluated in Sec. V.A, and one finds

$$(\det{}'\hat{P}_1^\dagger\hat{P}_1)^{1/2}=\det{}'2\Delta_{\hat{g}}=\tfrac{1}{2}\det{}'\Delta_{\hat{g}} \tag{2.125}$$

and

$$\det{}'\Delta_{\hat{g}}=\tau_2^2\,|\eta(\tau)|^4\ , \tag{2.126}$$

where the Dedekind η function is defined in Appendix E [Eq. (E9)]. The volume of $\mathrm{Ker}(\hat{P}_1)$ is easily obtained, since conformal Killing vectors on the torus are again constants, and one finds $\mathrm{Vol}(\mathrm{Ker}\hat{P}_1)=2\tau_2$, so that

$$\frac{d^2\tau}{\tau_2^2}(\det{}'\hat{P}_1^\dagger\hat{P}_1)^{1/2}\left[\frac{4\pi^2}{\tau_2}\det{}'\Delta_{\hat{g}}\right]^{-13}\frac{1}{\mathrm{Vol}(\mathrm{Ker}\hat{P}_1)}=\frac{d^2\tau}{8\pi^2\tau_2^2}\frac{1}{(4\pi^2\tau_2)^{12}}\,|\eta(\tau)|^{-48}\ . \tag{2.127}$$

With the help of the transformation law of $\eta(\tau)$ given in Eq. (2.45), it is easy to check that Eq. (2.127) is invariant under $SL(2,Z)$ as expected, since the calculation was manifestly reparametrization invariant throughout. The one-loop cosmological constant follows immediately,

$$\Lambda_{h=1}=\int_{\mathcal{M}_h}\frac{d^2\tau}{8\pi^2(\tau_2)^2}\frac{1}{(4\pi^2\tau_2)^{12}}\,|\eta(\tau)|^{-48}\ ,$$

and it is divergent due to the presence of the tachyon.

To obtain the scattering amplitudes, one needs the Green's function $G(z,z';\tau)$ satisfying Eq. (2.109), but now for the torus. It may be obtained by the method of images or, equivalently, from the translation properties of the standard ϑ_1 function for the torus, defined in Appendix C,

$$\vartheta_1(z+1,\tau)=-\vartheta_1(z,\tau)\ ,\qquad \vartheta_1(z+\tau,\tau)=e^{-i\pi\tau-2\pi iz}\vartheta_1(z,\tau)\ . \tag{2.128}$$

The only zero of $\vartheta_1(z,\tau)$ as a function of z is at $z=0$, and hence

$$-\ln\left|\frac{\vartheta_1(z-z',\tau)}{\vartheta_1'(0,\tau)}\right|^2 \tag{2.129}$$

satisfies the Laplace equation everywhere, except at $z=z'$, where it has the correct short-distance singularity. It can be made single valued on the torus by the addition of a function quadratic in $z-z'$:

$$G(z,z';\tau)=-\ln\left|\frac{\vartheta_1(z-z',\tau)}{\vartheta_1'(0,\tau)}\right|^2-\frac{\pi}{2\tau_2}(z-\bar{z}-z'+\bar{z}')^2\ . \tag{2.130}$$

Thus it is the unique candidate for the Green's function on the torus, and indeed satisfies Eq. (2.109). Using a translation-invariant cutoff when $z'\to z$, one finds that $G(z,z;\tau)$ is independent of z and τ. It may be used to rescale the coupling constant $\hat{\varepsilon}$ to ε, exactly as in the case of the tree-level amplitudes.

As an example, we present the explicit expression for the one-loop amplitudes for the scattering of tachyonic states only. The measure is given by Eq. (2.124), and the vertex operator part yields

$$\begin{aligned}\langle\langle V(k_1)\cdots V(k_n)\rangle\rangle&=\varepsilon^n(2\pi)^{26}\delta(k)\prod_{j=1}^n\int d^2z_j\exp\left[-\sum_{i<j}k_i\cdot k_j G(z_i,z_j;\tau)\right]\\&=(2\pi)^{26}\delta(k)\varepsilon^n\prod_{j=1}^n\int d^2z_j\prod_{i<j}F(z_i,z_j)^{k_i\cdot k_j}\ ,\end{aligned} \tag{2.131}$$

where we use the function F defined in Sec. II.G:

$$F(z_i,z_j)=\exp\left[\frac{\pi(z_i-\bar{z}_i-z_j+\bar{z}_j)^2}{2\tau_2}\right]\left|\frac{\vartheta_1(z_i-z_j,\tau)}{\vartheta_1'(0,\tau)}\right|^2\ . \tag{2.132}$$

Putting all together, we obtain

$$\langle V(k_1)\cdots V(k_n)\rangle=\delta(k)\varepsilon^n\int_{\mathcal{M}_1}\frac{d^2\tau}{2\tau_2^2}\frac{1}{(\tau_2)^{12}}\,|\eta(\tau)|^{-48}\prod_{j=1}^n\int d^2z_j\prod_{i<j}F(z_i,z_j)^{k_i\cdot k_j}\ . \tag{2.133}$$

These amplitudes are divergent for all values of momenta, due to the presence of the factor $|\eta(\tau)|^{-48}$, which contributes a factor exponential in τ_2 as $\tau_2\to\infty$. Ultimately, this is connected with the presence of the tachyon. Other manifestations of the instability of Minkowski space-time are the fact that the one-loop dilaton tadpole [computed, say, from Eq. (2.133) by factorization] is nonvanishing, indicating that the Minkowski space-time is not a solution to the string equations of motion to this order. Fischler and Susskind (1986a, 1986b) suggested that de Sitter space-time, on the other hand, does solve the equations of motion to this order.

In the dual model of open strings, one-loop amplitudes were considered by Gross, Neveu, Scherk, and Schwarz (1970). The closed bosonic string amplitudes to one-loop order were computed by Shapiro (1972), who also correctly identified the fundamental domain for moduli space (up to the above-mentioned factor of 2). In the new era, they were reevaluated first by Polchinski (1986) and subsequently by D'Hoker and Phong (1986a, 1986b) and by Panda (1987). Investigations for open strings are found in Cohen *et al.* (1987), Varughese and Weisberger (1987), and Weisberger (1987a, 1987b), where it is argued that the cylinder graph with boundary conditions may be used as an off-shell propagator for closed strings.

I. Formulation with ghosts

In the presence of an infinite-dimensional symmetry, factoring out the symmetry group and enforcing the correct measure in loop amplitudes can also be accomplished by introducing Faddeev-Popov ghosts. Our previous discussion shows that for the first-quantized bosonic string this is not strictly necessary, and any scattering amplitude can be computed without appealing to ghosts. However, a ghost formulation will put at our disposal the powerful tools of conformal field theory and exhibit the key Becchi-Rouet-Stora-Tyutin (BRST) invariance. These are also crucial ingredients in the construction of a full-fledged second-quantized string field theory, as was realized by Siegel (1985), Banks and Peskin (1986), Siegel and Zwiebach (1986, 1987), Witten (1986a, 1986b), and Neveu, Nicolai, and West (1986). Furthermore, in fermionic string theories, ghosts will be indispensable, as they will couple to fermion emission vertices.

Following the standard Faddeev-Popov procedure, we replace the gauge parameter δv^z for reparametrization invariance by an anticommuting ghost field c^z. Introducing its conjugate antighost field b_{zz}, we can now write down the reparametrization ghost action:

$$I_{\rm gh}(b,c)=\frac{1}{2\pi}\int d^2z\sqrt{g}\,b_{zz}\nabla^z c^z+\text{c.c.}\ , \tag{2.134}$$

which is Weyl and reparametrization invariant. The Weyl symmetry again will be anomalous, since the natural metrics on the ghost field space,

$$\begin{aligned}\|c\|^2&=\int d^2z\sqrt{g}\,g_{z\bar z}c^z c^{\bar z}\ ,\\ \|b\|^2&=\int d^2z\sqrt{g}\,(g^{z\bar z})^2 b_{zz}b_{\bar z\bar z}\ ,\end{aligned} \tag{2.135}$$

are not Weyl invariant. We also have an important global symmetry generated by

$$\begin{aligned}&c^z\to e^{-i\theta_z}c^z,\quad b_{zz}\to e^{+i\theta_z}b_{zz}\ ,\\ &c^{\bar z}\to e^{+i\theta_{\bar z}}c^{\bar z},\quad b_{\bar z\bar z}\to e^{-i\theta_{\bar z}}b_{\bar z\bar z}\ .\end{aligned} \tag{2.136}$$

Even though $c^{\bar z}$ is formally the complex conjugate of c^z, their analogs in Minkowski conventions would be independent. The metrics on the ghost space, however, are only invariant under a U(1) subgroup of Eq. (2.136), generated by $\theta_z=\theta_{\bar z}$. Thus the ghost number current

$$j_z=-b_{zz}c^z \tag{2.137}$$

can be expected to be anomalous in the full quantum theory. Indeed, in Appendix B, a heat-kernel regularization shows that

$$\nabla^z j_z=-\tfrac{3}{2}R\ . \tag{2.138}$$

The integrated version of this anomaly agrees with the index theorem of Eq. (2.50), which asserts that

$$8(c\text{ zero modes})-\#(b\text{ zero modes})=\tfrac{3}{2}\chi(M)\ . \tag{2.139}$$

Now recall that the gauge-fixing operators P_1 and $P_1^\dagger$ decompose as $P_1=\nabla_z^1\oplus\nabla_{-1}^z$, $P_1^\dagger=-(\nabla_2^z\oplus\nabla_z^{-2})$, so that the Faddeev-Popov determinant $(\det P_1^\dagger P_1)^{1/2}$ naively can be represented as

$$\int D(b\bar b c\bar c)e^{-I_{\rm gh}(b,c)}\ . \tag{2.140}$$

However, in the presence of zero modes this functional integral would vanish. For convenience, let us restrict our discussion to the case of genus $h\geq 2$, the case of the torus requiring straightforward modifications. In this case ∇_{-1}^z has no zero mode, while its adjoint ∇_2^z admits $3h-3$ zero modes, namely the holomorphic quadratic differentials (cf. Sec. II.E). To absorb these zero modes, we need $3h-3$ insertions. Thus the key nonvanishing functional integral of interest is

$$\int D(b\bar b c\bar c)\prod_{i=1}^{3h-3}b(z_i)\bar b(z_i)e^{-I_{\rm gh}(b,c)}\ , \tag{2.141a}$$

which can be evaluated to be

$$(\det{}'P_1^\dagger P_1)^{1/2}\frac{|\det\phi_k(z_j)|^2}{\det\langle\phi_j|\phi_k\rangle}\ , \tag{2.141b}$$

where ϕ_j are any basis of $3h-3$ holomorphic quadratic differentials. If we substitute this in the expression for the Polyakov string measure (2.79), we obtain

$$\begin{aligned}Z_B=\int dm_1\cdots dm_{6h-6}&\left|\frac{\det\langle\mu_j|\phi_k\rangle}{\det\phi_k(z_j)}\right|^2\\ &\times\int D(b\bar b c\bar c)Dx'\\ &\times\prod_{i=1}^{3h-3}b(z_i)\bar b(z_i)e^{-[I_{\rm gh}(b,c)+I_m(x)]}\ .\end{aligned} \tag{2.142}$$

In this expression it should be understood that for each value of the $6h-6$ moduli parameters $m_1,\ldots,m_{6h-6}$ characterizes a background metric with respect to which all the functional integrals and finite-dimensional determinants are evaluated. The "matter" action is just the Polyakov action [Eq. (2.5)] with the worldsheet metric as background:

$$I_m(x)=\frac{1}{4\pi}\int d^2z\,\partial_z x^\mu \partial_{\bar z} x^\mu \ . \tag{2.143}$$

Finally Dx' denotes omission of the constant zero mode, and $z_1,\ldots,z_{3h-3}$ are arbitrary points on the surface M.

Other useful formulations of the string amplitudes are also readily derivable from Eq. (2.141). In particular,

$$\left|\int Db\,Dc\prod_{i=1}^{3h-3}\langle \mu_i\,|\,b\rangle e^{-I_{\rm gh}(b,c)}\right|^2 =\frac{(\det' P_1^\dagger P_1)^{1/2}}{\det\langle \phi_j\,|\,\phi_k\rangle^{1/2}}\det\langle \mu_j\,|\,\phi_k\rangle \ , \tag{2.144}$$

where we have used real quadratic differentials on the right-hand side, to conform with Eq. (2.79). Thus the bosonic string amplitude can also be written as

$$Z_B=\int dm_1\cdots d\bar m_{3h-3}W\overline{W} \tag{2.145}$$

with

$$W=\int Db\,Dc\,Dx'\prod_{i=1}^{3h-3}\langle \mu_i\,|\,b\rangle e^{-I_{\rm gh}(b,c)-I_m(x)} \ . \tag{2.146}$$

We have used the standard notation for the pairing between the b field and the Beltrami differentials

$$\langle \mu\,|\,b\rangle=\int d^2z\,\mu^z_{\bar z}b_{zz} \ . \tag{2.147}$$

Here the integral in x is assumed to have been split as well into holomorphic and antiholomorphic parts, and we have kept in W the holomorphic one. How this can be done exactly requires careful treatment and will be taken up in Sec. VII.

The above offers a remarkably simple procedure for guessing the right measure: simply insert the right number of b's to absorb the zero modes, and pair off with the tangents to the slice for moduli. We shall see that this procedure generalizes to the superstring as well.

The above formulas simplify to some extent if we choose to represent moduli space by slices with Beltrami differentials $\mu^z_{i,\bar z}$ admitting jump discontinuities. As observed earlier in Sec. II.E, a smooth Beltrami differential cannot be represented as $\nabla_{\bar z}v^z$ for a smooth vector field if it deforms the conformal structure. However, deformations can be achieved with discontinuous vector fields and Beltrami differentials. Thus let our slice satisfy

$$\mu^z_{i,\bar z}=\nabla_{\bar z}v_i^z \ , \tag{2.148}$$

where v_i^z are quasiconformal vector fields, i.e., v_i^z are smooth vector fields with a unit jump δv^z across a closed contour C_i. Contours generating a basis of quasiconformal deformations can be chosen in a variety of ways; for which we refer, for example, to the discussion of Mandelstam diagrams (Sec. IV.G) and Fenchel-Nielsen coordinates (Sec. IV.E) of deformations. If we substitute them into fermionic functional integrals of the form (2.144), we note that the $3h-3$ insertions $b(z_j)$ can effectively be viewed as holomorphic, because all $3h-3$ factors are required to produce zero modes to compensate for the ghost number anomaly. Thus the insertion becomes

$$\int d^2z\sqrt{g}\,g^{z\bar z}\mu^z_{i\bar z}b_{zz}=\int d^2z\sqrt{g}\,\nabla^z v_i^z b_{zz} =\oint_{C_i}dz\,b_{zz}\delta v_i^z=\oint_{C_i}dz\,b_{zz} \tag{2.149}$$

and Eq. (2.146) reduces to

$$W=\int Db\,Dc\,Dx'\left[\prod_{i=1}^{3h-3}\oint_{C_i}dz\,b_{zz}\right]e^{-[I_{\rm gh}(b,c)+I_m(x)]} \ . \tag{2.150}$$

If the worldsheet is viewed as a surface with punctures, one may use the insertion of the operator identity

$$\oint_{C_z}dw\,b_{ww}(w)c^z(z)=1 \ , \tag{2.151}$$

where C_z is a small contour surrounding the point with coordinate z. If we choose the points z to coincide with the punctures (i.e., vertex operator insertions), then the ghost formulation of Eq. (2.79) reads

$$\langle V_{i_1}(k_1)\cdots V_{i_n}(k_n)\rangle=\sum_{h=0}^\infty e^{-\lambda\chi}\int_{\mathcal{M}_h}[dm]\int Dx^\mu\int Dc\,D\bar c\,Db\,D\bar b\,e^{-I_m(x)-I_{\rm gh}(b,c)} \times[\overline{V}_{i_1}(k_1)\cdots \overline{V}_{i_n}(k_n)]\prod_{j=1}^{3h-3+n}\oint_{C_j}dz\,b_{zz}\oint_{C_j}d\bar z\,b_{\bar z\bar z} \ . \tag{2.152}$$

Here the modified vertex operators are given by

$$\overline{V}(k)=\int d^2z\,c^z c^{\bar z}U(\varepsilon,x)e^{ikx} \ , \tag{2.153}$$

where the vertex operator without ghosts reads

$$V(k)=\int d^2\xi\sqrt{g}\,U(\varepsilon,g,x)e^{ikx} \ . \tag{2.154}$$

The crucial role of ghosts in formulating string theories emerged first out of Polyakov's original work (1981a). The b,c system was explored further in Friedan (1984) and string partition functions and amplitudes expressed in terms of ghost insertions in Friedan, Martinec and Shenker (1986). Equation (2.152) in terms of insertions of contour integrals was proposed by Martinec (1986) and Giddings and Martinec (1986), who derived it from a slightly different formalism of extended path integrals instead of ghost insertions that we adopted here. An alternative ghost action including the square of the ghost current was considered in Freedman and Warner (1986a, 1986b) and Freedman *et al.* (1987). Relations with the harmonic gauge are discussed in Freedman, Latorre, and Pilch (1988).

J. Conformal field theory

In the previous section we have presented string amplitudes in terms of correlation functions of the matter fields x^μ and the reparametrization ghosts b_{zz} and c^z. These are basic examples of conformal field theories, i.e., theories invariant under conformal transformations. In two dimensions conformal invariance is an especially powerful constraint, and we give now a brief discussion of the properties of these conformal fields.

The discussion will actually be clearer from a more general point of view, so we consider a theory of chiral fermions $b(dz)^n, c(dz)^{1-n}$ with action

$$I(b,c)=\frac{1}{2\pi}\int d^2z\sqrt{g}\, b\nabla^z_{1-n}c \ . \tag{2.155}$$

Classically the theory is invariant under Weyl transformations, so the stress tensor

$$T_{mn}=-\frac{4\pi}{\sqrt{g}}\delta I/\delta g^{mn}$$

is traceless and given by[6]

$$T_{zz}=-nb\partial_z c+(1-n)(\partial_z b)c \ . \tag{2.156}$$

In particular, the equations of motion imply that T_{zz} is holomorphic. Quantum mechanically, the Weyl symmetry is anomalous and will prevent the stress tensor from being simultaneously covariant and holomorphic. This can readily be seen from Ward identities for reparametrization invariance. Indeed, if we insist on reparametrization invariance, the correlation function

$$Z_F(z_1,\ldots,w_N)=\int D(bc)e^{-I(b,c)}\prod_{i=1}^M b(z_i)\prod_{i=1}^N c(w_i) \tag{2.157}$$

will transform by

$$\delta Z_F=\sum_{i=1}^M (n\nabla_{z_i}v^{z_i}+v^{z_i}\nabla_{z_i})Z_F + \sum_{i=1}^N[(1-n)\nabla_{w_i}v^{w_i}+v^{w_i}\nabla_{w_i}]Z_F \tag{2.158}$$

under a reparametrization $z\to z+v^z$. On the other hand, we can also write

$$\begin{aligned}\delta Z_F&=\frac{1}{4\pi}\int d^2z\sqrt{g}\,\delta g^{zz}\left\langle T_{zz}\prod_{i=1}^M b(z_i)\prod_{i=1}^N c(w_i)\right\rangle + \int d^2z\sqrt{g}\,\delta g^{z\bar z}\left[\frac{1}{\sqrt{g}}\frac{\delta Z_F}{\delta g^{z\bar z}}\right]\\ &=\frac{1}{2\pi}\int d^2z\sqrt{g}\,v^z\nabla^z\left\langle T_{zz}\prod_{i=1}^M b(z_i)\prod_{i=1}^N c(w_i)\right\rangle + \int d^2z\sqrt{g}\,v^z\nabla^{\bar z}\left[\frac{1}{\sqrt{g}}\frac{\delta Z_F}{\delta g^{z\bar z}}\right] .\end{aligned} \tag{2.159}$$

The variation of Z with respect to the trace of g is determined by the conformal anomaly. As in the case of the reparametrization ghosts, there will be a fermion number violation, measured by the index of ∇^z_{1-n}. If, say, $n\geq 2$, ∇^z_{1-n} will have no zero modes, while $(\nabla^z_{1-n})^\dagger$ will have $\Upsilon=(h-1)(2n-1)$ zero modes. The only nonvanishing correlation function (2.157) must satisfy $M=N+\Upsilon$ and can then be expressed as

$$Z_F=\frac{\det'\nabla^z_n}{\det\langle\phi_a|\phi_b\rangle^{1/2}}\sum_{i_1,\ldots,i_I}(-1)^{\sum i_j}\det\phi_k(z_{i_j}) \times \det_{l\neq i_1,\ldots,i_I} G(z_l,w_j) \ , \tag{2.160}$$

where $\phi_1,\ldots,\phi_I$ are zero modes and $G(z,w)$ is a propagator for ∇^z_n. The zero modes and propagators are unchanged under Weyl scalings, so

$$\frac{1}{\sqrt{g}}\frac{\delta}{\delta g^{z\bar z}}\ln Z_F=-\frac{1}{2}\frac{\delta}{\delta\sigma}\ln\left[\frac{\det'\nabla^z_n}{\det\langle\phi_a|\phi_b\rangle^{1/2}}\right] =\frac{c_n}{24\pi}Rg_{z\bar z} \ , \tag{2.161a}$$

where the central charge $-2c_n$ is given by

$$c_n=6n^2-6n+1 \tag{2.161b}$$

in view of Eq. (2.69). Equating the two expressions (2.158) and (2.159) for the variation of Z_F under reparametrizations gives

$$\frac{1}{2\pi}\nabla^z\left\langle T_{zz}\prod_{i=1}^M b(z_i)\prod_{i=1}^N c(w_i)\right\rangle+\frac{c_n}{24\pi}\nabla_z RZ_F=\sum_{i=1}^M[n\nabla_{z_i}\delta(z,z_i)+\delta(z,z_i)\nabla_{z_i}]Z_F + \sum_{i=1}^N[(1-n)\nabla_{w_i}\delta(z,w_i)+\delta(z,w_i)\nabla_{w_i}]Z_F \ .$$

[6]In this expression, the Christoffel symbols have canceled out.

This is an equation for $\langle T_{zz}\prod_1^M b(z_i)\prod_1^N c(w_i)\rangle$ which shows in particular that it is not meromorphic in z, with the obstruction arising precisely from the conformal anomaly. The equation can be solved using any propagator orthogonal to the zero modes $\phi_1,\ldots,\phi_{3h-3}$ of ∇_2^z. Such a propagator will satisfy

$$\begin{aligned}&\nabla^z G^\xi_{zz}=2\pi\delta(\xi,z)\ ,\\&\nabla^\xi G^\xi_{zz}=-2\pi\delta(\xi,z)+2\pi\sum_{a=1}^{3h-3} g^{\xi\bar\xi}\mu^\xi_{a,\bar\xi}\phi_{a,zz}\ ,\end{aligned}\tag{2.162}$$

where the $\mu^z_{a,\bar z}$ are the dual basis of Beltrami differentials

$$\int d^2z\sqrt{g}\,g^{z\bar z}\mu^z_{a,\bar z}\phi_{b,zz}=\delta_{ab}\ .$$

The first Ward identity for reparametrization invariance takes the form

$$\begin{aligned}&\Big\langle T_{zz}\prod_1^M b(z_i)\prod_1^N c(w_i)\Big\rangle-\sum_{a=1}^{3h-3}\phi_{a,zz}\int d^2\xi\sqrt{g}\,g^{\xi\bar\xi}\mu^\xi_{a,\bar\xi}\Big\langle T_{\xi\xi}\prod_1^M b(z_i)\prod_1^N c(w_i)\Big\rangle\\&\qquad=-\frac{c_n}{24\pi}Z_F\int d^2\xi\sqrt{g}\,G^\xi_{zz}\partial_\xi R+\sum_{i=1}^M(n\nabla_{z_i}G^{z_i}_{zz}+G^{z_i}_{zz}\nabla_{z_i})Z_F+\sum_{i=1}^N[(1-n)\nabla_{w_i}G^{w_i}_{zz}+G^{w_i}_{zz}\nabla_{w_i}]Z_F\ .\end{aligned}\tag{2.163}$$

A second Ward identity is obtained by differentiating the first with respect to $g^{-1/2}\delta/\delta g^{ww}$. The result is

$$\begin{aligned}&\Big\langle T_{zz}T_{ww}\prod_1^M b(z_i)\prod_1^N c(w_i)\Big\rangle-\sum_{a=1}^{3h-3}\phi_{a,zz}\int d^2\xi\sqrt{g}\,g^{\xi\bar\xi}\mu^\xi_{a,\bar\xi}\Big\langle T_{\xi\xi}T_{ww}\prod_1^N b(z_i)\prod_1^N c(w_i)\Big\rangle\\&\qquad=-\frac{c_n}{6}Z_F\nabla^3_w G^w_{zz}+\Big[\frac{c_n}{24\pi}\int d^2\xi\sqrt{g}\,G^\xi_{zz}\partial_\xi R+G^w_{zz}\nabla_w+2(\nabla_w G^w_{zz})-\sum_1^M(n\nabla_{z_i}G^{z_i}_{zz}+G^{z_i}_{zz}\nabla_{z_i})\\&\qquad\quad-\sum_1^N[(1-n)\nabla_{w_i}G^{w_i}_{zz}+G^{w_i}_{zz}\nabla_{w_i}]\Big]\Big\langle T_{ww}\prod_1^M b(z_i)\prod_1^N c(w_i)\Big\rangle\ .\end{aligned}\tag{2.164}$$

We can now read off operator product expansions by looking at short-distance $z-z_1$, $z-w_1$, $z-w$ singularities. Since the Green's function $G^{z_1}_{zz}$ is equal to $1/(z-z_1)$ up to smooth terms, the first and second Ward identities lead to

$$\begin{aligned}&T_{zz}b(\xi)\sim\left[\frac{n}{(z-\xi)^2}+\frac{1}{z-\xi}\partial_\xi\right]b(\xi)\ ,\\&T_{zz}c(\xi)\sim\left[\frac{1-n}{(z-\xi)^2}+\frac{1}{z-\xi}\partial_\xi\right]c(\xi)\ ,\\&T_{zz}T_{ww}\sim\left[\frac{2}{(z-w)^2}+\frac{1}{z-w}\partial_w\right]T_{ww}\\&\qquad-\tfrac{1}{6}c_n(\partial_w-\Gamma^w_{ww})\partial_w(\partial_w+\Gamma^w_{ww})\frac{1}{z-w}\ .\end{aligned}\tag{2.165}$$

Introducing the local counterterm $\partial_w\Gamma^w_{ww}-\frac{1}{2}(\Gamma^w_{ww})^2$ and the chiral stress tensor

$$T^{\rm chi}_{zz}=T_{ww}-\tfrac{1}{6}c_n[\partial_w\Gamma^w_{ww}-\tfrac{1}{2}(\Gamma^w_{ww})^2]\ ,$$

we can rewrite Eq. (2.165) as

$$\begin{aligned}&T^{\rm chi}_{zz}b(\xi)\sim\left[\frac{n}{(z-\xi)^2}+\frac{1}{z-\xi}\partial_\xi\right]b(\xi)\ ,\\&T^{\rm chi}_{zz}c(\xi)\sim\left[\frac{1-n}{(z-\xi)^2}+\frac{1}{z-\xi}\partial_\xi\right]c(\xi)\ ,\\&T^{\rm chi}_{zz}T^{\rm chi}_{ww}\sim\frac{-c_n}{(z-w)^4}+\left[\frac{2}{(z-w)^2}+\frac{1}{z-w}\partial_w\right]T^{\rm chi}_{ww}\ ,\end{aligned}\tag{2.166}$$

where this time $T^{\rm chi}_{zz}$ is holomorphic. Note that as a result of the third operator product expansion (OPE) above, $T^{\rm chi}_{zz}$ does not transform as a rank-two tensor. Rather, under a holomorphic reparametrization, $T^{\rm chi}_{zz}$ will transform with a Schwarzian derivative:

$$\begin{aligned}&T^{\rm chi}_{zz}=T^{\rm chi}_{ww}\left[\frac{dw}{dz}\right]^2-\frac{c_n}{6}S(w,z)\ ,\\&S(w,z)=\left[\frac{d^3w}{dz^3}\right]\Big/(dw/dz)-\frac{3}{2}\left[\frac{d^2w}{dz^2}\Big/\frac{dw}{dz}\right]^2\ .\end{aligned}\tag{2.167}$$

This is yet another way of expressing the fact that the conformal anomaly prevents simultaneous holomorphicity and covariance.

In the above discussion we chose to maintain manifest covariance. If we had chosen instead to maintain manifest holomorphicity, we could have defined the chiral stress tensor by the following normal ordering procedure:

$$T^{\rm chi}_{zz}=\lim_{w\to z}\left[-nb(w)\partial c(z)+(1-n)\partial b(w)c(z)+\frac{1}{(w-z)^2}\right]\ .\tag{2.168}$$

A routine calculation will again lead to the transformation law (2.167). Henceforth by stress tensor we shall actually designate the chiral one, which we denote simply by T_{zz}. Similarly all composite operators requiring regularization will be normal ordered as in Eq. (2.168), by splitting points and subtracting the singular part of the OPE.

The stress tensor can be viewed as generator of local conformal transformations

$$\delta_\varepsilon b(\xi)=\oint_{C_\xi}\frac{dz}{2\pi i}\varepsilon(z)T_{zz}b(\xi)\ . \tag{2.169}$$

As such it will give rise to a Virasoro algebra with central charge exactly the coefficient of the conformal anomaly. In fact, if we introduce the Virasoro generators

$$L_m=\oint_{C_0}\frac{dz}{2\pi i}z^{m+1}T_{zz}\ , \tag{2.170}$$

the matrix elements will be given by

$$\begin{aligned}\langle [L_m,T_{zz}]\rangle&=\oint_{C_{0,z}}\frac{dw}{2\pi i}w^{m+1}\langle T_{ww}T_{zz}\rangle\\&\quad-\oint_{C_0}\frac{dw}{2\pi i}w^{m+1}\langle T_{ww}T_{zz}\rangle\\&=\oint_{C_z}\frac{dw}{2\pi i}w^{m+1}\langle T_{ww}T_{zz}\rangle\ .\end{aligned} \tag{2.171}$$

Here $C_{0,z}$ is a curve enclosing both 0 and z (see Fig. 8). Substituting in Eq. (2.171) the OPE of (2.166) yields

$$[L_m,L_p]=(m-p)L_{m+p}+\frac{-c_n}{6}(m^3-m)\delta_{m+p,0}\ . \tag{2.172}$$

Equations (2.166), (2.169), and (2.172) are the local equations characterizing a conformal field theory. In the case at hand the b,c fields are primary fields of conformal weights n and $(1-n)$, respectively. They live on a Riemann surface M (more precisely are sections of the line bundles K^n and K^{1-n} where K is the canonical bundle of M), and the global version of the operator product expansions is given by the Ward identities (2.163) and (2.164). The negative sign in front of c_n is due to the quantization of b and c as fermions; it would be absent if b and c were quantized as bosons. This will be the case of the superghosts of Secs. III and VIII.

The theory of b,c fields is actually completely characterized by its current algebra. As pointed out before [Eq. (2.134)] for the reparametrization ghosts, the theory admits a symmetry $b\to e^{i\theta}b, c\to e^{-i\theta}c$. The (chiral) fermion number current $j_z=-bc$ is anomalous and satisfies

$$\nabla_{\bar z}j_z=-\tfrac{1}{2}(2n-1)\sqrt{g}R\ . \tag{2.173}$$

This can be seen by heat-kernel regularization exactly as in Eq. (2.138), which corresponds to $n=2$. Integrating this relation gives back the violation of fermionic number $\Upsilon=(2n-1)(h-1)$ determined earlier through index theorems [cf. Eq. (2.139)]. From the short-distance expansion[7]

$$b(z)c(w)\sim\frac{1}{z-w} \tag{2.174}$$

[7]In such relations it should always be assumed that Υ insertions have been made to absorb zero modes and ensure a meromorphic propagator.

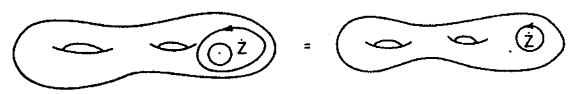

FIG. 8. Deformation of contour integrals.

it is easy to derive the OPE's:

$$\begin{aligned}j_zb(w)&\sim\frac{-1}{z-w}b(w)\ ,\\ j_zc(w)&\sim\frac{1}{z-w}c(w)\ ,\\ j_zj_w&\sim\frac{1}{(z-w)^2}\ .\end{aligned} \tag{2.175}$$

Finally the stress tensor T_{zz} can be recovered from the number current,

$$T_{zz}=\tfrac{1}{2}j_z^2+\tfrac{1}{2}Q\partial_z j_z\ , \tag{2.176}$$

and will satisfy the OPE

$$T_{zz}j_w\sim\frac{-Q}{(z-w)^3}+\frac{1}{(z-w)^2}j_z\ , \tag{2.177}$$

where $Q=(2n-1)$.

We turn now to bosons. The free scalar fields x^μ offer the simplest example of a bosonic conformal field theory, with the action that of the bosonic Polyakov string, the stress tensor equal to

$$T_{zz}=-\tfrac{1}{2}(\partial_z x^\mu)^2\ , \tag{2.178}$$

and its central charge given by $c^x=d$. Strictly speaking, we have $\langle x^\mu(z)x^\nu(w)\rangle\sim-\delta^{\mu\nu}\ln|z-w|^2$, so that x^μ does not have a conformal dimension. This is no hindrance, however, since fields built out of $\exp(ik^\mu x_\mu)$ and $\partial_z x^\mu, \partial_{\bar z}x^\mu$ do have well-defined dimensions, as we saw in Sec. II.G. Chiral scalar fields with gravitational anomalies will be defined in Sec. VII.B.

More sophisticated theories arise when the bosonic fields φ are multiple valued and possibly coupled to a background charge Q. Indeed, one of the fundamental features of two-dimensional quantum field theory is the Bose-Fermi correspondence, and the bosonization of the chiral fermions b and c of ranks n and $1-n$ discussed above will lead precisely to such theories. Recall that for free bosons

$$\begin{aligned}\langle :e^{i\varphi(z)}::e^{-i\varphi(w)}:\rangle&=e^{\langle\varphi(z)\varphi(w)\rangle}\\&=|z-w|^{-2}\\&=\langle b(z)c(w)\bar b(z)\bar c(w)\rangle\ ,\end{aligned} \tag{2.179}$$

so we expect $\exp[i\varphi(z)]$ to correspond to a fermion bilinear $b\bar b$. With this we can exhibit a bosonic action that will reproduce the current algebra of the b,c system:

$$I(\varphi)=\frac{1}{4\pi}\int d^2z(\partial_z\varphi\partial_{\bar z}\varphi-iQ\sqrt{g}R\varphi)\ . \tag{2.180}$$

Here $Q=2n-1$. The necessity of including the curvature term can be seen from several points of view. For example, the symmetry $b \to e^{i\theta}b$, $c \to e^{-i\theta}c$ should correspond to the shift $\varphi \to \varphi + \theta$, in which case the curvature term will produce the correct anomaly $\Upsilon = (2n-1)(h-1)$. The local version of this statement is that the current $j_z = -i\partial_z\varphi$ will satisfy the same equation as the fermion number current,

$$\partial_{\bar{z}}(-i\partial_z\varphi) = \frac{-Q}{2}\sqrt{g}R \ . \tag{2.181}$$

Moreover, the additional term $-iQ\partial_z^2\varphi/2$ that arises then in the stress tensor

$$T_{zz} = -\tfrac{1}{2}(\partial_z\varphi)^2 - \frac{i}{2}Q\partial_z^2\varphi \tag{2.182}$$

is precisely the modification needed to ensure that T_{zz} obey the Virasoro algebra with the correct central charge $1-3Q^2 = c_n$. Similarly we can check that the operator product expansions

$$\begin{aligned} T_{zz}j_w &\sim \frac{-Q}{(z-w)^3} + \frac{1}{(z-w)^2}j_z \ , \\ j_z j_w &\sim \frac{1}{(z-w)^2} \ , \end{aligned} \tag{2.183}$$

also reproduce Eqs. (2.175) and (2.177).

We can now confirm that vertex operators correspond to fermion bilinears by considering the OPE

$$\begin{aligned} T_{zz}e^{iq\varphi(w)} &\sim \frac{q(q+Q)/2}{(z-w)^2}e^{iq\varphi(w)} + \frac{1}{z-w}\partial_w e^{iq\varphi(w)} \ , \\ j_z e^{iq\varphi(w)} &\sim \frac{-q}{z-w}e^{iq\varphi(w)} \end{aligned} \tag{2.184}$$

The background charge has shifted the conformal weight of $\exp(iq\varphi)$ from $q^2/2$ to $q(q+Q)/2$. Thus q should be taken to be 1 and -1 for $b\bar{b}$ and $c\bar{c}$, in agreement with Eq. (2.179).

We have up to this point discussed only the formal and local aspects of the bosonic theories, unlike the fermionic theories for which we started from global formulas before examining singularities to obtain the local ones. Global issues do play an important role here, however. In fact, the operators $\exp(\pm i\varphi)$ suggest that φ is an angle. The correct statement on a worldsheet with nontrivial topology is that $d\varphi$ is a closed but in general not exact form, so that φ is a multiple-valued function. This requires the action $I(\varphi)$ to be suitably modified so as to be well defined and the path integral $D\varphi$ over all configurations to be correctly interpreted, as well. As usual with solitons, $d\varphi$ can be characterized up to an exact form by its winding numbers, so $D\varphi$ decomposes into a sum over soliton sectors indexed by winding numbers. The required machinery to do this as well as compute correlations and establish bosonization will be developed in Sec. VII. We postpone until then a careful study of global issues.

The program of using conformal invariance to classify critical points of statistical systems was pioneered by Polyakov (1969). The importance of conformal (primary) fields in string theory was recognized early on by Gervais and Sakita (1971a) and Andrić and Gervais (1972). The foundations of modern conformal field theory were laid out by Belavin, Polyakov, and Zamolodchikov (1984). Unitary conformal field theories with $c<1$ were classified by Friedan, Qiu, and Shenker (1984) and constructed explicitly by Goddard, Kent, and Olive (1986). Unitary $c=1$ models were studied by Dijkgraaf, Verlinde, and Verlinde (1988). Operator product expansions for the ghost system of the bosonic string appear in Friedan (1984). Our treatment here is an adaption to the higher-loop case of his arguments. Global versions of Ward identities are also derived by Eguchi and Ooguri (1987) and Sonoda (1987a). The stress tensor as a projective connection is studied in Alvarez and Windey (1987) and Dugan and Sonoda (1987). Bosonization of higher-spin free fermions b,c is due to Marnelius (1983) and Friedan, Martinec, and Shenker (1986). The corresponding bosonic system coupled to a background charge had appeared earlier in the work of Dotsenko and Fateev (1984). The importance of modular invariance in conformal field theory was stressed by Cardy (1986), Gepner and Witten (1986), Itzykson and Zuber (1986), Capelli, Itzykson, and Zuber (1987) and Gepner (1987a, 1987b).

K. Becchi-Rouet-Stora-Tyutin (BRST) invariance

We have seen in Sec. II.I that the Polyakov model for the bosonic string can be represented as a sum over moduli parameters of the full theory including ghosts with action

$$I_{\text{tot}} = \frac{1}{2\pi}\int d^2z(\tfrac{1}{2}\partial_z x^\mu \partial_{\bar{z}} x^\mu + b_{zz}\nabla_{\bar{z}}c^z + b_{\bar{z}\bar{z}}\nabla_z c^{\bar{z}}) \ . \tag{2.185}$$

In this formulation, the cancellation of the conformal anomaly in the critical dimension $d=26$ corresponds to the cancellation of the central charge in the total stress tensor,

$$T_{zz}^{\text{tot}} = T_{zz}^x + T_{zz}^{\text{gh}} \ . \tag{2.186}$$

From Eqs. (2.156), (2.168), and (2.178), the stress tensors for the ghost and matter parts are given by the following ordering prescription:

$$\begin{aligned} T_{zz}^x &= \lim_{w\to z} -\frac{1}{2}\left[\partial_z x^\mu \partial_w x^\mu + \frac{d}{(z-w)^2}\right] \ , \\ T_{zz}^{\text{gh}} &= \lim_{w\to z}\left[-2b(w)\partial_z c(z) - \partial_w b(w)c(z) + \frac{1}{(w-z)^2}\right] \ . \end{aligned}$$

The resulting central charges are $c^{\text{gh}} = -c_2 = -13$ and $c^x = dc_0/2 = d/2$, so that T_{zz}^{tot} is now a globally defined holomorphic rank-2 tensor. This is the property allowing decoupling of physical states by Virasoro gauge conditions.

As usual, the total action incorporating Faddeev-Popov ghosts exhibits a new symmetry, known as BRST symmetry:

$$\delta x^\mu = \lambda c^z \partial_z x^\mu, \quad \delta c^z = -\lambda c^z \nabla_z c^z ,$$
$$\delta b_{zz} = -\lambda[-\tfrac{1}{2}\partial_z x^\mu \partial_z x^\mu + c^z \nabla_z b_{zz} + 2(\nabla_z c^z) b_{zz}] . \tag{2.187}$$

The parameter λ is infinitesimal and Grassmann valued. The fact that these transformations indeed generate a symmetry is read off from the transformations of the matter and ghost parts separately,

$$\delta \partial_z x^\mu \partial^z x^\mu = \lambda(\nabla^z c^z)\partial_z x^\mu \partial_z x^\mu + \nabla_z(\lambda c^z \partial_z x^\mu \partial^z x^\mu) ,$$

$$\delta b_{zz} \nabla^z c^z = -\lambda \partial_z x^\mu \partial_z x^\mu \nabla^z c^z - \nabla_z(\lambda c^z b_{zz} \nabla^z c^z) .$$

The BRST current is

$$j_z^{\rm BRST} = c^z T_{zz}^x + \tfrac{1}{2} c^z T_{zz}^{\rm gh} + \tfrac{3}{2}(\nabla_z)^2 c^z . \tag{2.188}$$

The ordering prescription is again taken to be

$$j_z^{\rm BRST} = c^z T_{zz}^x + \tfrac{1}{2}\lim_{w\to z}\left[c^w T_{zz}^{\rm gh} + \frac{1}{(z-w)^2} c^w - \frac{2}{z-w}\partial_w c^w \right] + \tfrac{3}{2}(\nabla_z)^2 c^z .$$

In view of the transformation laws (2.167) for the stress tensors, the BRST current will transform as a genuine holomorphic rank-1 tensor. The corresponding BRST charge is given by the contour integral

$$Q_{\rm BRST} = \oint \frac{dz}{2\pi i} j_z^{\rm BRST} , \tag{2.189}$$

where the contour surrounds insertions. From operator product expansions, we can deduce that

$$[Q_{\rm BRST}, x^\mu] = c^z \partial_z x^\mu ,$$
$$\{Q_{\rm BRST}, c^z\} = c^z \nabla_z c^z , \tag{2.190}$$
$$\{Q_{\rm BRST}, b_{zz}\} = T_{zz}^{\rm tot} .$$

As already noted in Eq. (2.174) it is essential in deriving Eqs. (2.190) to have meromorphic propagators. This is automatic in string theory, since the string measure incorporates exactly the right number of insertions to absorb the effects of the ghost zero modes.

At the classical level $Q_{\rm BRST}$ is just a Grassmann quantity that squares to 0. At the quantum level the operator statement $Q_{\rm BRST}^2 = 0$ holds only in the critical dimension $d=26$. In this case the Virasoro gauge conditions for physical states translate into

$$Q_{\rm BRST}\,|\,{\rm phys}\rangle = 0 , \tag{2.191}$$

and states of the form

$$Q_{\rm BRST}\,|\,{\rm anything}\rangle \tag{2.192}$$

are spurious and decouple from physical processes. In other words, physical states should rather be viewed as BRST cohomology classes, i.e., elements of the coset space ${\rm Ker}Q_{\rm BRST}/{\rm Image}Q_{\rm BRST}$.

In the Polyakov path-integral formulation of strings, the decoupling of spurious states (2.192) translates into the fact that amplitudes with an insertion of the BRST current around N arbitrary vertex insertions can be written as total derivatives on moduli space. More precisely, let $V_i(w_i)$ be vertices of conformal dimensions (λ_i, λ_i), let $C_{w_1,\ldots,w_N}$ be a contour surrounding $w_1,\ldots,w_N$, and parametrize moduli space by coordinates m_j,

$$\delta g_{\bar z \bar z} = \sum_{j=1}^{3h-3} \delta m_j g_{z\bar z} \mu_{j\bar z}^z , \tag{2.193}$$

where $\mu_{j\bar z}^z$ are $3h-3$ Beltrami differentials. Then

$$\prod_{l=1}^N d^2 w_l \sqrt{g}\,(g^{w_l \bar w_l})^{\lambda_l} \left\langle \prod_{j=1}^{3h-3} \langle \mu_j \,|\, b_{zz}\rangle \oint_{C_{w_1,\ldots,w_N}} j_w^{\rm BRST} dw \prod_{l=1}^N V(w_l) \right\rangle$$
$$= \sum_{k=1}^{3h-3} \frac{\partial}{\partial m_k}\left[(-)^k \prod_{l=1}^N d^2 w_l \sqrt{g}\,(g^{w_l\bar w_l})^{\lambda_l} \left\langle \prod_{j\neq k} \langle \mu_j \,|\, b_{zz}\rangle \prod_{l=1}^N V(w_l) \right\rangle \right] . \tag{2.194}$$

This relation between BRST invariance and total derivatives on moduli space can be easily seen from the OPE's (2.190) and a deformation-of-contours argument when the Beltrami differentials $\mu_{j\bar z}^z$ are generated by quasiconformal vector fields as in Sec. II.I. In this setup the insertions w_l remain separated from the supports of $\mu_{j\bar z}^z$. We can deform the contour $C_{w_1,\ldots,w_N}$ away from the insertions and pull it off the worldsheet, leaving in the process only the residues at the $3h-3$ insertions of b_{zz}. By the third equation in (2.190) the residue is exactly an insertion of the stress tensor T_{zz}. Such a term $\langle \mu_k \,|\, T_{zz}\rangle$ accounts for the piece of the right-hand side of Eq. (2.194) that arises when $\partial/\partial m_k$ lands on the action. In general the vertices V_l will depend on the moduli parameters. Their variations with respect to the trace of the metric cancel the variations of the volume forms in Eq. (2.194). Finally their variations with respect to δg^{zz} proper are Dirac functions supported only at w_l. They will vanish when paired with Beltrami differentials μ_k arising from quasiconformal deformations, since these are supported along disjoint contours.

This argument requires modifications if the supports of the Beltrami differentials cover the whole surface. The reason is that insertions of b_{zz} resulting from $\langle \mu \,|\, b_{zz}\rangle$, insertions of $V(w_l)$ as well as points on the contour $C_{w_1,\ldots,w_N}$, may come arbitrarily close together, invalidating operator product expansions such as (2.165) and (2.190), where only two points come close. Ignoring this effect would cause us to miss the variations with respect to moduli of the vertex operators.

A thorough justification of deformations-of-contours arguments along these lines seems quite involved at this point. Instead, we shall present an alternative argument, which does not rely on analyticity, and generalizes easily to superstrings. In this formulation, the basic object is the generating functional including ghosts and sources

$$Z[x^*,b^*,c^*]=\int D(xbc)e^{-I_{\text{tot}}(x,b,c)+I_s} , \qquad (2.195)$$

where the source Lagrangian is given by

$$I_s=\int_M d^2z\sqrt{g}\,(x^{*\mu}x^\mu+b^{*zz}b_{zz}+c_z^*c^z) . \qquad (2.196)$$

The BRST charge Q, the partition function, and the correlation functions can all be expressed in terms of combinations of the operators,

$$\hat{x}=\frac{1}{\sqrt{g}}\frac{\delta}{\delta x^*},\quad \hat{b}_{zz}=\frac{1}{\sqrt{g}}\frac{\delta}{\delta b^{*zz}} ,$$
$$\hat{c}^{\,z}=\frac{1}{\sqrt{g}}\frac{\delta}{\delta c_z^*},\quad \hat{g}_{zz}=\frac{1}{\sqrt{g}}\frac{\delta}{\delta g^{zz}} , \qquad (2.197)$$

acting on the generating functional. For example, the partition function is given by

$$Z_h=\int_{\mathcal{M}_h}\prod_{j=1}^{3h-3}d^2m_j\langle\mu_j\,|\,\hat{b}\,\rangle Z[x^*,b^*,c^*]\Bigg|_{*=0} \qquad (2.198)$$

and the expectation value of a general operator $V(\hat{x},\hat{b},\hat{c},g)$ is

$$\langle V\rangle=\int\prod_{j=1}^{3h-3}d^2m_j\langle\mu_j\,|\,\hat{b}\,\rangle V(\hat{x},\hat{b},\hat{c},g)Z[x^*,b^*,c^*] . \qquad (2.199)$$

The BRST transform of an operator V is then defined by

$$\delta V=[\lambda Q_{\text{BRST}},V(\hat{x},\hat{b},\hat{c},g)] , \qquad (2.200)$$

with the BRST operator given by

$$Q_{\text{BRST}}=\int d^2z\sqrt{g}\,(x^*\hat{c}^{\,z}\partial_z\hat{x}-2b^{*zz}\hat{g}_{zz}-c^*\hat{c}^{\,z}\nabla_z\hat{c}) .$$

The OPE of Eq. (2.194) is replaced by

$$[\lambda Q_{\text{BRST}},\hat{b}_{zz}]=\lambda\hat{g}_{zz} , \qquad (2.201)$$

and pulling through of the BRST contour can be justified by the following commutation rules for the BRST charge:

$$\{\langle\mu_k\,|\,\hat{b}\,\rangle,Q_{\text{BRST}}\}=-2\frac{\partial}{\partial m_k}+2\int d^2z_k\int d^2z\sqrt{g}\,b^{*zz}\frac{\delta\mu_{k\bar{z}}^{z}}{\delta g^{zz}}\hat{b}_{z_kz_k} ,$$
$$\left[\prod_{j\neq k}\langle\mu_j\,|\,\hat{b}\,\rangle,\{\langle\mu_k\,|\,\hat{b}\,\rangle,Q_{\text{BRST}}\}\right]=0 . \qquad (2.202)$$

The decoupling of spurious states in this formalism can now be established. First observe that Ward identities for BRST invariance can be stated as

$$Q_{\text{BRST}}Z[x^*,b^*,c^*]=0 \qquad (2.203)$$

for all values of the sources, and in the critical dimension. To show this, one must use the Ward identities for reparametrization invariance, as well as for Weyl invariance. The Weyl Ward identities are anomalous in general, but in the critical dimension a cancellation of the matter and ghost contributions reduces them to the naive Ward identities, which are the ones needed to prove Eq. (2.203). Furthermore, in an expectation value

$$\langle\delta_{\text{BRST}}V\rangle=\int\prod_j d^2m_j\langle\mu_j\,|\,\hat{b}\,\rangle[\lambda Q_{\text{BRST}},V]\times Z[x^*,b^*,c^*]\Big|_{*=0} \qquad (2.204)$$

we may replace the commutator $[\lambda Q_{\text{BRST}},V]$ by $\lambda Q_{\text{BRST}}V$ and permute Q_{BRST} through all b insertions to obtain a total derivative on moduli. This establishes Eq. (2.194)

Strictly speaking, BRST invariance is at this point purely formal, since in principle it could be broken by contributions from the boundary of moduli space. A geometric discussion of the boundary of moduli space is provided in Sec. IV.H. For the bosonic string the amplitudes diverge, and a proper discussion of BRST invariance will require some renormalization (e.g., Fischler and Susskind, 1986a, 1986b; Seiberg, 1987; Sen, 1987). For superstrings where amplitudes are expected to be finite, whether the boundary of moduli space does contribute is a major issue, here as well as in questions of supersymmetry breaking.

The original BRST invariance of gauge-fixed Yang-Mills theories was introduced by Becchi, Rouet, and Stora (1976) and Tyutin (1975). That the BRST operator of string theory is nilpotent exactly in the critical dimension is due to Kato and Ogawa (1983), who also gave the interpretation of physical states as BRST cohomology classes. BRST invariance of multiloop amplitudes and deformation of contour arguments were stressed by Friedan, Martinec, and Shenker (1986). Arguments along these lines based on special meromorphic propagators are given in Sonoda (1987b). The setup in the functional language with external sources which we presented here to establish BRST invariance for the bosonic string is due to Mansfield (1987). The corresponding Ward identities are also given in Cohen, Gomez, and Mansfield (1986).

The requirement that spurious states decouple is what led originally to the discovery of the critical dimension, and the fact that this decoupling can be carried out consistently was one of the great successes of dual-model theories. It was established by Brower and Thorn (1971), Del Giudice, Di Vecchia, and Fubini (1972), and Goddard and Thorn (1972). The BRST formulation can of course be used to recapture many properties of the dual models in the operator formalism. A BRST proof of the no-ghost theorem is given by Freeman and Olive (1986), Frenkel, Garland, and Zuckerman, (1986), Spiegelglas (1987), and Thorn (1987). A proof based on the Kac (1983) determinant was given by Thorn (1984).

L. Formulation on surfaces with punctures

The main formula (2.79) for the scattering amplitudes of n particles was derived in the vertex operator formalism. In the introduction to this section, we saw that one can also formulate string perturbation theory on surfaces with punctures and wave functions. We now compare the two formulations.

At the h-loop level, the worldsheet for the scattering of n particles is a surface[8] M^* with h handles and n punctures $\xi_1,\dots,\xi_n$. Quantization can be carried out as for Eq. (2.12),

$$\langle V_1(k_1)\cdots V_n(k_n)\rangle^* = \int \frac{Dg_{mn}}{\mathcal{N}^*}\int Dx^\mu W_1(\xi_1)\cdots \times W_n(\xi_n)e^{-I(x)} . \tag{2.205}$$

This time the normalization $\mathcal{N}^*$ should be taken to be

$$\mathcal{N}^* = \mathrm{Vol}(\mathrm{Diff}(M^*))\times \mathrm{Vol}(\mathrm{Weyl}(M^*)) \tag{2.206}$$

and $W_i(\xi_i)$ are wave functions evaluated at the punctures. In the critical dimension $d=26$, the amplitude (2.205) reduces to an integral over the moduli space $\mathcal{M}_{h,n}$ of Riemann surfaces of genus h and with n punctures:

$$\langle V_1(k_1)\cdots V_n(k_n)\rangle^* = \int_{\mathcal{M}_{h,n}}[dm^*]\frac{\det\langle\mu_\alpha\mid\phi_\beta\rangle}{\det\langle\phi_\alpha\mid\phi_\beta\rangle^{1/2}}(\det{}^*P_1^\dagger P_1)^{1/2}\left[\frac{8\pi^2}{\int_M d^2\xi\sqrt{g}}\det{}'\Delta_g\right]^{-13}\langle\langle W_1(\xi_1)\cdots W_n(\xi_n)\rangle\rangle . \tag{2.207}$$

In appearance Eq. (2.207) is very similar to Eq. (2.79) of the vertex operator formalism. However, the definitions of the parameters m, Beltrami differentials μ_α, and quadratic differentials ϕ_β and $\det{}^*P_1^\dagger P_1$ in Eq. (2.207) have to be adapted to the fact that the worldsheet M^* is now viewed as having punctures. First a moduli parameter m^* for $\mathcal{M}_{h,n}$ will consist of a moduli parameter for $\mathcal{M}_h=\mathcal{M}_{h,0}$ and of n points on the surface. Thus m^* should correspond to $m_1,\dots,m_{6h-6}$ and $\xi_1,\dots,\xi_n$ of n points on the surface and the complex dimension of $\mathcal{M}_{h,n}$ is

$$\dim\mathcal{M}_{h,n}=3h-3+n .$$

(A more precise geometric description of $\mathcal{M}_{h,n}$ as a fiber bundle over moduli space can be given by Teichmüller universal curve constructions, which are treated in Sec. IV, but we shall not need it here.) The number of Beltrami differentials μ_α is correspondingly increased to $3h-3+n$. We can choose the slice representing $\mathcal{M}_{h,n}$ so that the first $\{\mu_j\}_{j=1,\dots,3h-3}$ Beltrami differentials arise from a slice representing $\mathcal{M}_h$, while the remaining $\{\mu_p\}_{p=3h-3+1,\dots,3h-3+n}$ are generated by vector fields v_p which move the punctures by a unit displacement. Similarly the ϕ_β's are now holomorphic quadratic differentials on the surface M^*. They can be divided into earlier differentials $\{\phi_j\}_{j=1,\dots,3h-3}$, which are holomorphic on the whole surface M, and n meromorphic differentials $\{\phi_p\}_{p=3h-3+1,\dots,3h-3+n}$ with each ϕ_p having a simple pole[9] at ξ_p. (Such ϕ_p's exist in view of the Riemann-Roch theorem, which we shall discuss later in Sec. VII.C.) Finally, $\det{}^*P_1^\dagger P_1$ is the determinant of the operator $P_1^\dagger P_1$ restricted to the subspace of vector fields that vanish at the punctures. The reason is that the "small" diffeomorphisms of the punctured surface m^* are the small diffeomorphisms of the full surface M that leave the punctures fixed, and those are generated only by vector fields in the above restricted subspace.

We can take the vector fields v_p to be smooth and supported in a neighborhood of small size δ around each puncture. Since $\mu_p=\nabla_{\bar z}v_p^z$, it follows readily that $\langle\mu_p\mid\phi_j\rangle=0$. If the meromorphic differentials ϕ_p are chosen so that $\langle\phi_j\mid\phi_p\rangle=0$, we shall have

$$\frac{\det\langle\mu_\alpha\mid\phi_\beta\rangle}{\det\langle\phi_\alpha\mid\phi_\beta\rangle^{1/2}}=\frac{\det\langle\mu_j\mid\phi_k\rangle}{\det\langle\phi_j\mid\phi_k\rangle^{1/2}}\times\frac{\det\langle\mu_p\mid\phi_q\rangle}{\det\langle\phi_p\mid\phi_q\rangle^{1/2}} . \tag{2.208}$$

On the other hand, the Faddeev-Popov determinants are related by

$$\det{}^*P_1^\dagger P_1=\det P_1^\dagger P_1\frac{\det\langle\phi_p\mid\phi_q\rangle\det\langle v_p\mid v_q\rangle}{[\det\langle\phi_p\mid\mu_q\rangle]^2} . \tag{2.209}$$

In view of the support of v_p, it is easy to see that

$$\det\langle v_p\mid v_q\rangle^{1/2}=\delta^{2n}\prod_{i=1}^n g(\xi_p)+O(\delta^{2n+1}) . \tag{2.210}$$

Combining Eq. (2.208) with (2.209) and (2.210) and absorbing the factor $\delta^2 g(\xi_p)$ into a redefinition of the wave function, we arrive at

[8] Objects considered on the punctured surface will be denoted with an asterisk.

[9] Meromorphic differentials with a simple pole are precisely dual to the reparametrization vector fields with a simple zero at the puncture, so one need not consider differentials with poles of higher order.

$$\langle V_1(k_1)\cdots V_n(k_n)\rangle^* = \int_{\mathcal{M}_{h,0}}[dm]\frac{\det\langle \mu_j \mid \phi_k\rangle}{\det\langle \phi_j \mid \phi_k\rangle^{1/2}}(\det P_1^\dagger P_1)^{1/2}\left[\frac{8\pi^2}{\int d^2\xi\sqrt{g}}\det{}'\Delta_g\right]^{-13}$$
$$\times\int d^2\xi_1\sqrt{g(\xi_1)}\cdots\int d^2\xi_n\sqrt{g(\xi_n)}\langle\langle W(\xi_1)\cdots W(\xi_n)\rangle\rangle\ . \tag{2.211}$$

It remains to express the wave functions W in terms of the vertex operators for on-shell particle emission V. One starts by considering a surface where the puncture is replaced with a boundary of finite size. The amplitude computed by the insertion of a vertex operator is the same as the one computed on the surface with a boundary component and a wave function gotten by doing the path-integral operator over a disc D (of radius δ), including the vertex operator, and fitting into the boundary, as indicated in Fig. 9. Thus the wave function equals the integral over the disc with the vertex operator inserted. This is easily computed, and we obtain

$$W[x(\sigma)] = \int^{x(\sigma)} Dx^\mu P(\varepsilon, Dx)e^{ik_\mu x^\mu(\xi)}e^{-I[x]}\ , \tag{2.212}$$

where the vertex operator was of the form

$$V(\varepsilon,k,x) = P(\varepsilon,\partial x)(\xi)e^{ik_\mu x^\mu(\xi)}\ . \tag{2.213}$$

Splitting $x^\mu(z,\bar z)$ into a harmonic piece $\tilde x^\mu(z,\bar z)$ with boundary values $x^\mu(\sigma)$ and a fluctuation $y^\mu(z,\bar z)$, we find

$$W[x(\sigma)] = \int Dy^\mu P[\varepsilon,\partial\tilde x^\mu + \partial y^\mu]e^{ik_\mu\tilde x^\mu + ik_\mu y^\mu}\exp\left[-\frac{1}{8\pi}\int d^2\xi\,\partial_m y^\mu\partial^m y_\mu\right]\exp\left[-\frac{1}{8\pi}\oint dn^m\tilde x^\mu\partial_m\tilde x_\mu\right]\ . \tag{2.214}$$

Now since vertex operators are constructed so that they are normal ordered, we should not contract two legs on the same vertex. Hence if we let $\delta\to 0$, the Gaussian factor in Eq. (2.214) tends to 1, and we recover the desired relation between vertex operators and wave functions at punctures:

$$W[x(\sigma)] = P[\varepsilon,\partial\tilde x^\mu](0)e^{ik_\mu\tilde x^\mu(0)}$$
$$\times\exp\left[-\frac{1}{8\pi}\oint dn^m\tilde x^\mu\partial_m\tilde x^\mu\right]\ . \tag{2.215}$$

The above arguments are due to D'Hoker and Giddings (1987).

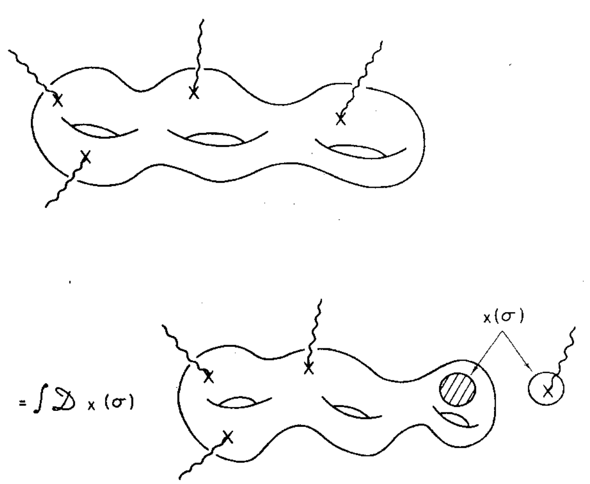

FIG. 9. Relation between amplitudes computed by inserting vertex operators and by giving wave functionals on boundary components.

III. CLOSED ORIENTED FERMIONIC STRINGS

Soon after the discovery of bosonic strings, it was realized that worldsheet spinors (fermions) ψ^μ carrying a space-time vector index μ could also be incorporated in the theory. As the number of negative norm states is now doubled compared to the bosonic string, an additional local symmetry is required to decouple these states. Local supersymmetry, discovered in this context by Gervais and Sakita (1971b , 1971c), is the appropriate invariance to do this, and so from a geometric point of view the starting point for the fermionic string is two-dimensional supergravity as developed by Zumino (1974), Brink, Di Vecchia, and Howe (1976), and by Deser and Zumino (1976b). For a general reference to supersymmetry and supergravity, we refer the reader to Ferrara and Fayet (1977), van Nieuwenhuizen (1981), Wess and Bagger (1983), Gates *et al.* (1984), Ferrara (1987), and West (1987).

Whereas the original model of Ramond (1971) contains space-time fermions, the model of Neveu and Schwarz (1971) also incorporates space-time bosons. In both models, one has worldsheet local supersymmetry, and space-time Lorentz invariance is manifest. These theories are consistent only in ten space-time dimensions. Once the Lorentz-covariant form is known, one may construct the associated light-cone gauge formulation. The light-cone Ramond-Neveu-Schwarz (RNS) formulation was used in a major development by Gliozzi, Scherk, and Olive (1976), who suggested that the even G parity sector of the Neveu-Schwarz theory together with a chiral truncation of the Ramond theory—the so-called GSO projection—yields a space-time supersymmetric spectrum. It took several years before Green and Schwarz (1981) proved the presence of a genuine supersymmetry by constructing the supercharge using the fermion emission vertex of the dual model. This fermion vertex had been introduced by

Thorn (1971) and by Mandelstam (1973b) in the light-cone formulation. Once the presence of supersymmetry was established, Green and Schwarz (1982) discovered a light-cone reformulation different from the Ramond-Neveu-Schwarz theory. Here, only physical bosonic space-time vectors and fermionic space-time spinors are present, no GSO projection is needed, and space-time supersymmetry is manifest. It is known as the Green-Schwarz formulation. Superstrings are classified into three groups: type I, type II, and heterotic.

Type-I superstring theories contain both open and closed unoriented strings. The open-string sector can support non-Abelian gauge fields when one attaches non-Abelian charges to the ends of the string. Mathematically, such charges are incorporated through the Chan-Paton rule, but factorization and duality limit the gauge groups to be orthogonal or symplectic. Ultimately it was discovered by Green and Schwarz (1984) that only O(32) can yield an anomaly-free theory, and thus the type-I superstring is unique. In a Minkowski space-time low-energy limit, it reduces to an $N=1$ supergravity plus Yang-Mills theory.

Type-II superstrings contain only closed oriented strings, and the only freedom left is the relative parity of the two gravitinos, producing the nonchiral type-IIA and the chiral type-IIB theories. In a Minkowski space-time low-energy limit, these theories reduce to $N=2$ supergravity without Yang-Mills multiplet.

Heterotic strings contain closed oriented strings only and are obtained as a hybrid (hence the nomenclature) between the type-II superstring and the closed oriented bosonic string. This hybrid is possible because on closed oriented worldsheets left- and right-movers are independent degrees of freedom, except for their collective momentum, to the point that one-half of one string theory can be replaced by that of another string theory. The 16 extra dimensions of the bosonic string component are compactified and yield a $\mathrm{Spin}(32)/Z_2$ or $E_8\times E_8$ gauge groups only. In a Minkowski space-time low-energy limit, it reduces to $N=1$ supergravity plus Yang-Mills theory.

In this section we shall derive the basic formulas for loop amplitudes for any of the above closed fermionic strings. We shall always be interested in theories with manifest space-time Lorentz invariance, and hence work with the covariant RNS and Polyakov formulations. As a drawback, space-time supersymmetry will not be manifest. From the worldsheet point of view, type-II theories are formulated as $N=1$ two-dimensional supergravity with "matter" multiplets x^μ and ψ^μ, and a supergravity multiplet consisting of a zweibein e_m^a and a two-dimensional spin-$\frac{3}{2}$ gravitino field χ_m. The worldsheet matter action for these fields reads

$$I_m=\frac{1}{4\pi}\int_M d^2\xi\sqrt{g}\,[\tfrac{1}{2}g^{mn}\partial_m x^\mu\partial_n x_\mu +\psi^\mu\gamma^m\partial_m\psi_\mu-\psi^\mu\gamma^a\gamma^m\chi_a\partial_m x_\mu -\tfrac{1}{4}\psi^\mu\gamma^a\gamma^b\chi_a(\chi_b\psi_\mu)]+\lambda\chi(M)\ . \tag{3.1}$$

Heterotic strings, on the other hand, correspond to $N=\frac{1}{2}$ supergravity with the same position and supergravity multiplets, except that ψ^μ and χ_m are of definite chirality:

$$\gamma^z\psi^\mu=0,\quad \gamma^{\bar z}\chi_m=0\ . \tag{3.2}$$

In addition there are internal degrees of freedom, which we represent by a fermionic variable ψ^a also of definite chirality:

$$\gamma^{\bar z}\psi^a=0,\quad a=1,2,\ldots,P\ .$$

For heterotic strings, the corresponding worldsheet action is

$$I_m=\frac{1}{4\pi}\int_M d^2\xi\sqrt{g}\,(\tfrac{1}{2}g^{mn}\partial_m x^\mu\partial_n x_\mu+\psi^\mu\gamma^{\bar z}e_{\bar z}{}^m\partial_m\psi_\mu -\psi_+^\mu\chi_{\bar z}^+e_z{}^m\partial_m x_\mu+\psi^a\gamma^z e_z{}^m\partial_m\psi^a) +\lambda\chi(M)\ . \tag{3.3}$$

Note that the internal degrees of freedom described here by ψ^a may alternatively be introduced as so-called left-moving bosonic fields x^a. This is the approach originally taken by Gross *et al.* (1985a).

Spinors on a worldsheet of nontrivial topology must be appropriately defined. Indeed their phase shifts under parallel transport around closed loops are half of those of vectors and hence are ambiguous. We shall see that for closed oriented strings there are exactly 2^{2h} consistent choices of phase shifts for a worldsheet of genus h. Each choice is called a spin structure.

A crucial issue for fermionic strings is the assignment of spin structures. In the functional quantization formalism, the GSO projection for the type-II string is enforced by separating spinors of left chirality from spinors of right chirality, assigning each group independent spin structures ν and $\bar\nu$, and summing over ν and $\bar\nu$. This is the natural prescription to avoid global anomalies, since no spin structure is preferred, and the mapping class group will interchange them. That the spin structures within each group must be the same is a requirement of space-time Lorentz invariance. For the heterotic string, $\mathrm{Spin}(32)/Z_2$ symmetry forces the spin structures of all 32 ψ^a's to be identical, whereas $O(16)\times O(16)$ possibly extended to $E_8\times E_8$ requires the spin structures to be the same within each group of 16 ψ^a's, although the spin structures for the two groups need not be equal (see Witten, 1985b; D'Hoker and Phong, 1986d; and Seiberg and Witten, 1986).

In practice this principle of splitting left- from right-movers requires more specific prescriptions. In fact, actions are formulated with a Minkowski signature on g_{mn}, and we analytically continue to Euclidean signature. In the Minkowski metric ψ^μ,ψ^a,χ_m are Majorana-Weyl spinors. In the Euclidean metric, however, there are no Majorana-Weyl spinors, and the two chiral components of a Majorana spinor are complex conjugates of one another and must carry the same spin structure. To get around this difficulty, we start from a real spinor (sum of a Weyl spinor and its complex conjugate) and have to separate only upon quantization the contributions of the complex-conjugate factors. Each factor may then be thought of as the contribution of one Majorana-Weyl fermion. A major difficulty in this task is caused by the contributions of the bosonic fields x^μ and the terms $\chi\bar{\chi}\psi_+\psi_-$, which must be separated as well. We shall see in Sec. III.K below that this separation can only be enforced by introducing internal loop momenta p_I^μ, and contributions of left and right spinors after assignment of independent spin structures must be matched at the same value of p_I^μ. The precise prescriptions are given in Eqs. (3.196)–(3.201). In Secs. VII.F and VII.G we shall discuss their relations with the holomorphic structure of string amplitudes on supermoduli space.

When dealing with supersymmetric theories in general and with two-dimensional supergravity in particular, one may either use the component field formalism, in terms of the fields defined above, or group different component fields that transform into one another under supersymmetry transformations into the same multiplet or superfield. The superfield formalism is more appropriate for cancellation of local anomalies and enforcing the correct quantum measure. The natural setting for superstrings is $N=1$ supergeometry, and the analog of Riemann surfaces and moduli space will be super Riemann surfaces and supermoduli space. The structure of supermoduli space and its relation to moduli space are of great importance, and we shall explore them in this section as well as in Sec. VII. The superfield approach will be taken as a starting point in Secs. III.B, III.D–III.J, and III.L, and the component field formalism will be related to it in Secs. III.C, III.K, and III.M–III.P.

There are also string theories with larger worldsheet supersymmetry classified by Ademollo *et al.* (1976a). There is the $N=2$ superstring constructed by Ademollo *et al.* (1976b) for which a locally supersymmetric formulation was given by Brink and Schwarz (1977), which is critical in two (complex) space-time dimensions; it was recently explored by Cohn (1987) and D'Adda and Lizzi (1987). There is an $N=4$ theory constructed by Ademollo *et al.* (1976c) whose covariant formulation is due to Pernici and van Niewenhuizen (1986) and that is critical in -2 (quaternionic) dimensions. In the covariant formulation these string theories involve also a nondynamical gauge field on the worldsheet. String theories with gauge fields on the worldsheet have also been considered by Tomboulis (1987) and Porrati and Tomboulis (1988).

A compendium of standard conventions and reference formulas, including the Dirac matrices, is given in Appendix A.

A. Spinors on a Riemann surface

Before constructing the amplitudes for fermionic strings, it is useful to recall some standard terminology of the theory of Riemann surfaces needed for a proper definition of fermions on the surface. The *first homology group* of a compact surface M without boundaries and with h handles is given by

$$H^1(M)=\mathbf{Z}^{2h} . \tag{3.4}$$

A canonical basis for this group is provided by closed curves A_I and B_I, $I=1,2,\ldots,h$, with canonical intersection matrix

$$\#(A_I,A_J)=0, \quad \#(A_I,B_J)=\delta_{IJ}, \quad \#(B_I,B_J)=0 . \tag{3.5}$$

Recall that the intersection form is antisymmetric. An example of such an assignment of A and B curves is given in Fig. 10. The choice of canonical basis is clearly not unique. If (A_I,B_I) is a canonical basis, then so is (A_I',B_I') with

$$B_I'=B_{IJ}A_J+A_{IJ}B_J, \quad A_I'=D_{IJ}A_J+C_{IJ}B_J , \tag{3.6}$$

where the $(2h\times 2h)$-dimensional matrix

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

belongs to the symplectic group with integer coefficients $\mathrm{Sp}(2h,\mathbf{Z})$. This group is the so-called *Siegel modular* group or simply modular group. One may think of it as being generated by 2π twists about A and B cycles. Such 2π twists about a closed curve are usually called *Dehn twists.*

Actually, the modular group is a subgroup of the mapping class group encountered earlier. To generate the mapping class group by Dehn twists, one needs twists about A and B cycles, but also about curves "linking consecutive handles" D_n as indicated in Fig. 10.

The quotient of the mapping class group by the modular group is the so-called *Torelli group,* which no longer acts on the homology basis.

Using the canonical homology decomposition in A and B cycles, we may cut the surface apart and represent it as a simply connected region of the plane—the fundamental region—on which sides are pairwise identified. As indicated in Fig. 11, it is convenient to perform this cutting process loop by loop, so that the boundary consists of unions of segments $A_IB_IA_I^{-1}B_I^{-1}$. Conversely, having such a fundamental region, one may reassemble the surface loop by loop, as shown in four stages in Fig. 12.

We now come to spin structures. In the Introduction

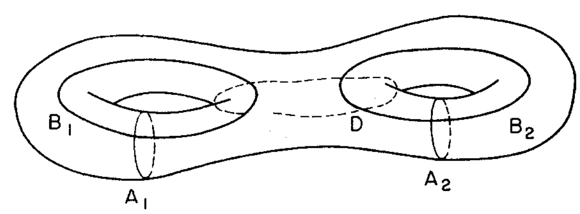

FIG. 10. A genus-2 surface with its canonical homology basis, generated by closed curves A_I and B_I. The Dehn twist D has also been indicated.

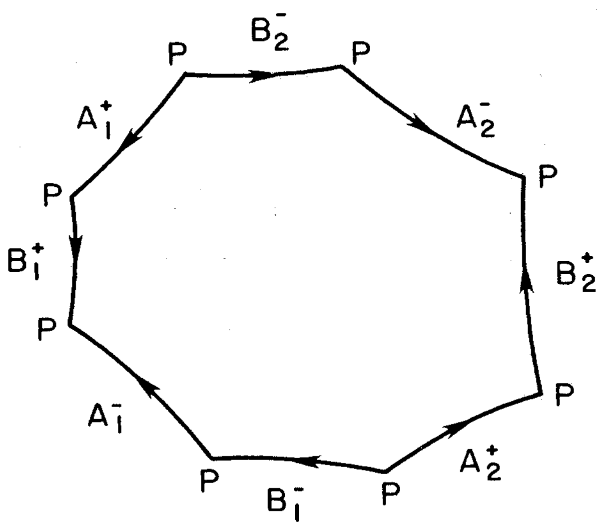

FIG. 11. A genus-2 surface cut along canonical homology cycle. All cycles pass through a common point P.

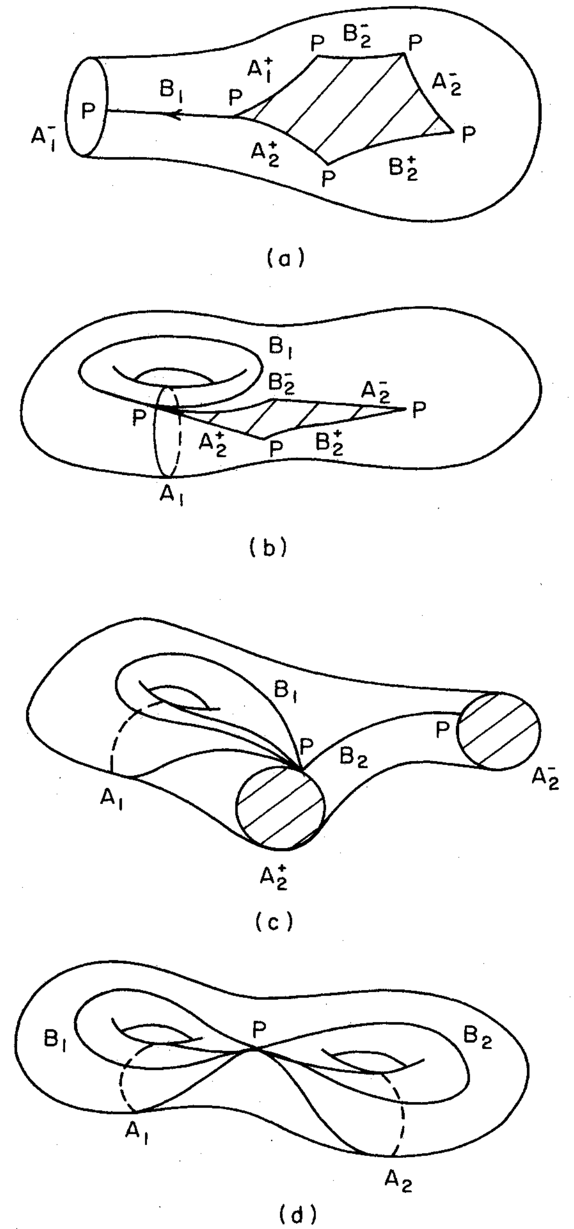

FIG. 12. Reconstruction of the genus-2 surface from the cut representation: (a) gluing B_1^+ and B_1^-; (b) gluing A_1^+ and A_1^-; (c) gluing B_2^+ and B_2^-; (d) gluing A_2^+ and A_2^-.

we already mentioned that the phase shift of a spinor after parallel transport along a closed curve should be half that of a vector and is ambiguous. Thus spinors exist only on manifolds for which a consistent choice of phases along all closed curves can be made. In general there is a topological obstruction to doing this, which is the second Stiefel-Whitney class. For oriented surfaces, however, this class vanishes and spin structures can be visualized as follows. If we fix a reference spin structure ν, the phase shifts around each of the homology cycles A_I and B_I of any other spin structure will differ from those of ν by 0 or π. Thus there are altogether 2^{2h} different spin structures [see Atiyah (1971) and Dabrowski and Percacci (1986), and the discussion in Sec. VI.F].

Each spin structure defines a distinct class of spinors, which does not interact with the others, and a corresponding Dirac operator. This implies immediately a natural classification of spin structures into even and odd ones, corresponding to the parity of the number of zero modes of the Dirac operator. It will be seen in Secs. V.C and VI.F that spin structures can be more conveniently expressed in terms of multipliers or theta characteristics, and that generically the number of Dirac zero modes is always 0 to 1. A diffeomorphism of the worldsheet M may transform a spin structure ν into a different one ν'. Since the parity of Dirac zero modes is invariant, Diff(M) will preserve the parity of the spin structure. It is an important fact that within each parity they can actually all be permuted under the mapping class group. (If we represent spin structures by theta characteristics, this will follow at once from the transformation law of theta functions; see Sec. VI.E and Appendix E.) This property will fix the relative phases of Dirac determinants within each group, and the relative phases between the two parities themselves will ultimately be determined from factorization requirements.

B. $N=1$ supergravity, supercomplex structures, and super Riemann surfaces

Locally, $N=1$ superspace is parametrized by two real $\xi^m=(\xi^1,\xi^2)$ or one complex coordinate $\xi=(1/\sqrt{2})(\xi^1+i\xi^2)$ and two real odd coordinates $\theta^\mu=(\theta^1,\theta^2)$, or one complex odd coordinate $\theta=(1/\sqrt{2})(\theta^1+i\theta^2)$ and its complex conjugate $\bar\theta$. These coordinates are collected into one supercoordinate $z^M=(\xi,\bar\xi;\theta,\bar\theta)$, where the index M is a coordinate or Einstein index. Correspondingly, we have the partial derivatives $\partial_M=(\partial/\partial\xi,\partial/\partial\bar\xi;\partial/\partial\theta,\partial/\partial\bar\theta)$. We shall also use a local U(1) frame with indices $A=(z,\bar{z};+,-)$, where z and $\bar{z}$ refer to the *vector* representation of the U(1) frame group and + and − refer to the *spinor* representation. The corresponding lower-case latin and greek letters correspond to the even and odd parts of these coordinates, respectively.

The $N=1$ supergravity multiplet consists of the superzweibein $E_M{}^A$ and the U(1) superconnection Ω_M, from which a U(1)-covariant superderivative $\mathcal{D}_M$ may be constructed. When this derivative acts on U(1) tensors V of weight n, it is given by[10]

$$\mathcal{D}^n_M V=\partial_M V+in\Omega_M V . \tag{3.7}$$

In particular, on one-forms and vector fields we have

$$\mathcal{D}_M V_A=\partial_M V_A+\Omega_M E_A{}^B V_B ,$$
$$\mathcal{D}_M V^A=\partial_M V^A-(-)^{mb}V^B E_B{}^A\Omega_M ,$$

where

$$E_a{}^b=\varepsilon_a{}^b,\quad E_a{}^\beta=E_\alpha{}^b=0,\quad E_\alpha{}^\beta=\tfrac{1}{2}(\gamma_5)_\alpha{}^\beta .$$

In differential form notation we have

$$\mathcal{D}^n=dz^M\mathcal{D}^n_M=d+in\Omega \text{ with } \Omega=dz^M\Omega_M ,$$

and d stands for the ordinary differential $d=dz^M\partial_M$. We shall mostly be using the covariant derivatives with U(1) indices $\mathcal{D}^n_A=E_A{}^M\mathcal{D}^n_M$, because they are manifestly super-reparametrization invariant. It will also prove useful to employ the real operators

$$\mathcal{P}_n=\mathcal{D}^n_+\oplus\mathcal{D}^{-n}_- \tag{3.8}$$

acting on the direct sum of superfields of U(1) weights n and $-n$, analogous to the operators P_n of Sec. II. We also introduce the Laplacians

$$\Box^{(+)}_n=\mathcal{D}^{n+1/2}_-\mathcal{D}^n_+,\quad \Box^{(-)}_n=\mathcal{D}^{n-1/2}_+\mathcal{D}^n_- , \tag{3.9}$$

so that (as we shall see later)

$$\mathcal{P}^\dagger_n\mathcal{P}_n=\Box^{(+)}_n\oplus\Box^{(-)}_{-n} .$$

The Laplacian on scalar superfields will be denoted by $\Box_0=-\Box^{(+)}_0=\Box^{(-)}_0$.

Torsion $T_{BC}{}^A$ and curvature R_{AB} tensors are defined by

$$[\mathcal{D}_A,\mathcal{D}_B]=T_{AB}{}^C\mathcal{D}_C+inR_{AB} , \tag{3.10}$$

where [,] is understood to be a commutator except when both A and B are spinor indices in which case it is an anticommutator. The supergeometry may be specified by imposing the standard torsion constraints

$$T_{ab}{}^c=T_{\alpha\beta}{}^\gamma=0,\quad T_{\alpha\beta}{}^c=2(\gamma^c)_{\alpha\beta} . \tag{3.11a}$$

[10] $U(1)$ weights are normalized so that $n=1\ (-1)$ for a lower z ($\bar{z}$) index; in flat space, $U(1)$ indices then agree with the conventions used in Sec. II.

Equivalently—and more usefully—we may replace Eq. (3.11a) by the constraint that the curvature $R_{\alpha\beta}$ be proportional to $(\gamma_5)_{\alpha\beta}$, instead of $T_{ab}{}^c=0$. Thus Eq. (3.11a) is equivalent to

$$R_{\alpha\beta}=-i(\gamma_5)_{\alpha\beta}R_{+-},\quad T_{\alpha\beta}{}^\gamma=0,\quad T_{\alpha\beta}{}^c=2(\gamma^c)_{\alpha\beta} . \tag{3.11b}$$

Another way of looking at the latter constraints is that they entirely specify the commutation relations between $\mathcal{D}_\alpha$ and $\mathcal{D}_\beta$, so that with the help of Eq. (3.10) all components of torsion and curvature can be computed in terms of the single scalar superfield R_{+-}. Thus, Eq. (3.11) implies the torsion formulas

$$\begin{aligned} &T_{\beta c}{}^a=0 ,\\ &T_{b\gamma}{}^\alpha=-\frac{i}{2}(\gamma_b)^\alpha_\gamma R_{+-} ,\\ &T_{bc}{}^\alpha=-\frac{i}{2}\varepsilon_{bc}(\gamma_5)^{\alpha\beta}\mathcal{D}_\beta R_{+-} ,\end{aligned} \tag{3.12}$$

and the curvature formulas

$$\begin{aligned} &R_{b\gamma}=i(\gamma_5\gamma_b)_\gamma{}^\delta\mathcal{D}_\delta R_{+-} ,\\ &R_{ab}=-\frac{i}{2}\varepsilon_{ab}\mathcal{D}^\alpha\mathcal{D}_\alpha R_{+-}-\tfrac{1}{2}\varepsilon_{ab}(R_{+-})^2 .\end{aligned} \tag{3.13}$$

All other components vanish. As a consequence of the torsion constrains, one may express the components of the superconnection in terms of the superzweibein:

$$\Omega_+=2iE_+{}^M(\partial_M E_+{}^N)E_N{}^+,\quad \Omega_z=E_+{}^M\partial_M\Omega_+ . \tag{3.14}$$

1. Symmetries

The supergeometry is invariant under transformations that preserve the torsion constraints. We list them below, together with their infinitesimal versions expressed in terms of the infinitesimal changes in the superzweibein $H_A{}^B$,

$$H_A{}^B=E_A{}^M\delta E_M{}^B . \tag{3.15}$$

Thus the symmetries of the supergeometry are as follows.

(i) Local U(1) transformations, forming a group sU(1). These are generated by a real superfield L acting by

$$\begin{aligned} &E^\pm_M=e^{\pm(i/2)L}\hat{E}^\pm_M,\quad \mathcal{D}^n_+=e^{-i(n+1/2)L}\hat{\mathcal{D}}^n_+e^{inL} ,\\ &E_M{}^z=e^{iL}\hat{E}_M{}^z,\quad \mathcal{D}^n_-=e^{-i(n-1/2)L}\hat{\mathcal{D}}^n_-e^{inL} ,\\ &E_M{}^{\bar{z}}=e^{-iL}\hat{E}_M{}^{\bar{z}},\quad \Omega_M=\hat{\Omega}_M+\partial_M L ,\end{aligned} \tag{3.16}$$

and infinitesimal transformations given by

$$\delta E_M{}^A=-E_M{}^B E_B{}^A\delta L,\quad H_A{}^B=-\delta L E_A{}^B . \tag{3.17}$$

(ii) Super-reparametrizations, forming a group sDiff(M). The infinitesimal ones are generated by super vector fields δV^M and are given by