Nuclear Physics B256 (1985) 253-284

HETEROTIC STRING THEORY*
(I). The free heterotic string

David J. GROSS, Jeffrey A. HARVEY, Emil MARTINEC and Ryan ROHM

Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544, USA

Received 26 February 1985

A new theory of closed orientable superstrings is constructed as a chiral combination of the closed $D = 26$ bosonic and $D = 10$ fermionic strings. The resulting ten-dimensional theory is $N = 1$ supersymmetric, Lorentz invariant, and free of tachyons. An alternate formulation where sixteen of the bosonic coordinates are replaced by fermionic coordinates is presented and used to construct a manifestly Lorentz covariant and reparametrization invariant action for the theory. The theory is consistent only for gauge group $\mathrm{Spin}(32)/Z_2$ or $E_8 \times E_8$.

1. Introduction

Contemporary attempts to construct unified field theories have largely focused on the search for new and hidden symmetries of nature. The most promising of these are based on the beautiful supersymmetric extension of Poincaré invariance and on the revival of the ideas of Kaluza and Klein of hidden, compact spatial dimensions. These ideas have the potential of providing a framework for a truly unified theory of gravity and matter, which can provide an explanation for the known low-energy gauge theory of matter and predict its full particle content. Within the framework of ordinary field theory, however, these attempts have reached an impasse. First, they do not provide a cure for the nonrenormalizability of quantum gravity. Therefore new physics must appear at the Planck length. Even if one ignores this problem and focuses on the low-energy structure of such theories, it appears to be impossible to construct realistic theories without a great loss of predictive power. In order to generate the observed spectrum of chiral fermions it appears necessary to retreat from the most ambitious Kaluza-Klein program, which would uniquely determine the low-energy gauge group as isometries of some compact space, and introduce gauge fields by hand [1]. Finally no realistic and compelling model has emerged.

String theories offer a way out of this impasse. They are based on an enormous increase of the number of degrees of freedom, with a corresponding enlargement of the symmetry group of nature. It has been clear for some time that the ten-dimensional superstring provides a consistent, finite theory which has as its low-energy manifestation ten-dimensional supergravity [2]. However, in order to generate

* Supported in part by NSF grant PHY-80-19754, and DOE contract DE-AC02-76ER03072.

the observed chiral fermions one must start with a chiral string theory which has potential anomalies. Until recently the only known anomaly-free chiral superstring theory was that of orientable, type II, $N=2$ supersymmetric closed strings [3]. This ten-dimensional theory does not contain Yang-Mills interactions; it is unlikely that they, together with the known fermions, could emerge upon compactification. The recent discovery that nonorientable, type I, open and closed superstrings with $N=1$ supersymmetry are finite and anomaly free if the gauge group is SO(32) [4] has increased the phenomenological prospects of unified string theories.

The anomaly cancellation mechanism of Green and Schwarz can be understood in terms of the low-energy field theory that emerges from the superstring, which is a slightly modified version of $D=10$ supergravity. The dangerous Lorentz and Yang-Mills chiral anomalies cancel if and only if the gauge group of the theory is SO(32) or $E_8 \times E_8$. The ordinary superstring theory cannot incorporate $E_8 \times E_8$. However the apparent correspondence between the low-energy limit of anomaly-free superstring theories and anomaly-free supergravity theories suggests the existence of a new kind of string theory whose low-energy limit would contain an $E_8 \times E_8$ gauge group. This is the motivation that led to the construction of the heterotic string. It is of more than academic interest to construct such a theory, since the phenomenological prospects for an $E_8 \times E_8$ theory appear to be much brighter.

In the standard formulation of string theory gauge groups are introduced by attaching (Chan-Paton) labels to the edges of open strings [5]. This procedure can only yield, in a consistent theory, the gauge groups SO(N) and Sp($2N$); it cannot produce $E_8 \times E_8$ [6]. Alternatively one might hope to obtain a gauge group from the compactification of a closed string theory, à la Kaluza-Klein. Such a possibility was suggested to us by the construction due to Frenkel and Kac of representations of Kac-Moody algebras in terms of string vertex operators [7]. The heterotic string theory is a theory of closed strings, in which the compactification of sixteen internal dimensions produces the gauge group SO(32) or $E_8 \times E_8$*.

Another motivation behind the construction of the heterotic string was the search for new consistent string theories. Previously known string theories are the bosonic theory (the Veneziano model) in 26 dimensions and the fermionic, superstring theory in ten dimensions. The new string theory is constructed as a chiral hybrid of these. The possibility of doing this for closed strings becomes evident once one realizes that the physical degrees of freedom of closed strings are two-dimensional massless fields - the transverse coordinates (and their supersymmetric partners) of the world sheet of a propagating string. These fields can be separated into right and left movers, even in the presence of string interactions.

For the consistency of the string theory it is necessary and sufficient that each sector be separately consistent. Thus the heterotic string is constructed as a combina-

* The observation that rank 16 is equal to the difference of the 26 dimensions of the bosonic string theory and the 10 of the superstring was made by Freund [8].

tion of the right-moving coordinates of the ten-dimensional superstring and the left-moving coordinates of the 26-dimensional bosonic string. This yields a theory which is Lorentz invariant in $D=10$ and which has appealing features not found in either theory separately*.

In this paper we shall construct in some detail the free heterotic string. In sect. 2 we present an operator construction using light-cone gauge quantization. Much of this is simply putting together the bosonic and fermionic string components. The new feature is the treatment of the 16 internal, left-moving, bosonic coordinates. We show that these must, for consistency, be compactified on a particular torus. This compactification leads to a set of 496 massless vector bosons. Some of these (16) can be thought of as ordinary Kaluza–Klein gauge bosons corresponding to the 16 isometries of the torus; the others (480) are massless solitons. Together they fill out the adjoint representation of $G = E_8 \times E_8$ or $G = \mathrm{Spin}(32)/Z_2$. We show that the theory is free of tachyons and exhibit the massless states and the first massive excitations. We demonstrate that the theory is Lorentz invariant, $N=1$ supersymmetric and that all the states form unitary representations of the gauge group G.

In sect. 3 we present an alternative formulation of the heterotic string theory in terms of internal fermionic coordinates. This amounts to a "fermionization" of the internal coordinates, with minor complications due to the finite length of the closed string. This formulation has some advantages, particularly in setting up a manifestly covariant formalism. In sect. 4 we show that one can describe the heterotic string in a covariant fashion, with no anomalies. Sect. 5 consists of some concluding remarks.

In the following paper [9] we shall discuss the interacting heterotic string, construct the vertex operators for the emission of massless strings, calculate various tree and loop amplitudes, derive the low-energy field theoretic limit of the theory, and verify that it is finite.

2. Construction of the free heterotic string

In this section we shall present the operator formulation of the free heterotic string in light-cone gauge. Before proceeding let us review the philosophy behind the construction.

Noninteracting strings are constructed by first quantization of an action given by the invariant area of the world sheet swept out by the string, or by its supersymmetric generalization. The reparametrization invariance of the classical action permits one to choose a "gauge" in which the timelike parameter of the world sheet, τ, is identified with light-cone time, $X^+ = \sqrt{\frac{1}{2}}\,(X^0 + X^9)$. In this light-cone gauge the theory reduces to a two-dimensional free field theory of the physical degrees of

* Thus the name "heterotic".

freedom (the 24 transverse coordinates, X^i, of the bosonic string or the 8 transverse coordinates, X^i, and their fermionic partners, of the superstring) subject to constraints. This procedure is only valid, however, in the critical dimension of 26 for the bosonic string and 10 for the superstring. In other space-time dimensions the existence of anomalies implies that the conformal degree of freedom of the two-dimensional metric does not decouple [10]. If this degree of freedom is ignored then there is a breakdown of world-sheet reparametrization invariance [11].

In the critical dimension the physical degrees of freedom, being massless two-dimensional free fields, can be decomposed into right and left movers, i.e. functions of $\tau-\sigma$ and $\tau+\sigma$ respectively. If we consider only closed strings then the right- and left-moving modes never mix. The separation of these modes will persist even in the presence of interactions, as long as we consider orientable world surfaces. This is because interactions between closed strings are constructed, order by order in perturbation theory, by modifying the topology of the two-dimensional world sheet on which the strings propagate. In terms of the first quantized two-dimensional theory, however, no interaction is introduced; the right and left movers still propagate independently. Thus there is no obstacle to treating the right movers and the left movers differently, as long as each sector is separately consistent. This is the idea behind the construction of the heterotic string, which combines the right movers of the superstring with the left movers of the bosonic string. It is necessarily a theory of closed orientable strings, since open string boundary conditions would mix right and left movers, and since one can clearly define the orientation of such a string.

In this section we shall construct the heterotic string in light-cone gauge and show that it is indeed consistent, Lorentz invariant, supersymmetric and free of tachyons. The only subtlety arises in the quantization and compactification of the sixteen extra left-moving bosonic coordinates. An alternate version of the heterotic string (see sect. 3) can be given in which these coordinates are replaced by 32 fermionic coordinates. This formulation allows one to construct a manifestly covariant form of the action, as will be shown in sect. 4.

2.1. THE PHYSICAL DEGREES OF FREEDOM

The physical degrees of freedom of the right-moving sector of the fermionic superstring consists of 8 transverse coordinates $X^i(\tau-\sigma)$ $(i=1\cdots8)$ and 8 Majorana-Weyl fermionic coordinates $S^a(\tau-\sigma)$ [12]. The physical degrees of freedom of the left-moving sector of the bosonic string consist of 24 transverse coordinates, $X^i(\tau+\sigma)$ and $X^I(\tau+\sigma)$ $(i=1\cdots8,\ I=1\cdots16)$. Together they comprise the physical degrees of freedom of the heterotic string. The eight transverse right- and left-moving coordinates combine with the longitudinal light-cone coordinates $X^{\pm}(\tau,\sigma)$ to describe the position of the string embedded in ten-dimensional flat space-time.

The dynamics of the free heterotic string is governed by the light-cone gauge action ($\alpha = 1, 2$)

$$S = -\int d\tau \int_0^\pi d\sigma \frac{1}{4\pi\alpha'}[\partial_\alpha X^i \partial^\alpha X^i + \partial_\alpha X^I \partial^\alpha X^I + i\bar{S}\gamma^- (\partial_\tau + \partial_\sigma)S], \tag{2.1}$$

where S is a ten-dimensional Majorana-Weyl spinor, with components S^a, i.e.

$$\gamma^+ S = hS = 0,$$
$$h \equiv \tfrac{1}{2}(1 + \gamma_{11}). \tag{2.2}$$

(From here on, unless otherwise noted, we shall work in units in which the string tension $T_0 = 1/2\pi\alpha' = 1/\pi$ and follow the notation of ref. [12].) The coordinates X^i will contain both right movers, $X^i(\tau - \sigma)$, and left movers, $X^i(\tau + \sigma)$; however, we must impose a constraint restricting X^I to consist of left-moving modes alone. This can be achieved by demanding that

$$\Phi^I \equiv (\partial_\tau - \partial_\sigma)X^I = 0, \qquad I = 1 \cdots 16. \tag{2.3}$$

Alternatively, following [13], we can add to the action a term proportional to $\lambda^I[(\partial_\tau - \partial_\sigma)X^I]^2$, which has the same effect. Siegel has shown that the resulting action possesses a local gauge symmetry which allows the Lagrange multipliers λ^I to be gauged away, leaving one with the constraint of eq. (2.3).

The above action is invariant under a supersymmetry transformation that relates right-moving bosons and fermions, while the left-moving bosons remain unchanged:

$$\delta X^i = \frac{1}{\sqrt{p^+}} \bar{\varepsilon}\gamma^i S,$$

$$\delta S = \frac{i}{\sqrt{p^+}} \gamma_- \gamma_\mu (\partial_\tau - \partial_\sigma) X^\mu \varepsilon. \tag{2.4}$$

where ε is a right-moving Majorana-Weyl light-cone spinor. In terms of two-dimensional field theory we have $N = \frac{1}{2}$ supersymmetry.

Canonical quantization of the transverse and fermonic coordinates is straightforward. Since we are dealing with closed strings, the fields are periodic functions of $0 \leq \sigma \leq \pi$ and can be expanded as

$$X^i(\tau - \sigma) = \tfrac{1}{2}x^i + \tfrac{1}{2}p^i(\tau - \sigma) + \tfrac{1}{2}i \sum_{n \neq 0} \frac{\alpha_n^i}{n} e^{2in(\tau - \sigma)},$$

$$X^i(\tau + \sigma) = \tfrac{1}{2}x^i + \tfrac{1}{2}p^i(\tau + \sigma) + \tfrac{1}{2}i \sum_{n \neq 0} \frac{\tilde{\alpha}_n^i}{n} e^{-2in(\tau + \sigma)} \tag{2.5}$$

$$S^a(\tau - \sigma) = \sum_{n=-\infty}^{+\infty} S_n^a e^{-2in(\tau - \sigma)}, \tag{2.6}$$

where

$$[x^i, p^j] = i\delta^{ij}, \qquad [\alpha^i_n, \alpha^j_m] = [\tilde{\alpha}^i_n, \tilde{\alpha}^j_m] = n\delta_{n+m,0}\delta^{ij},$$
$$[\alpha^n_i, \tilde{\alpha}^m_j] = 0,$$
$$\{S^a_n, \bar{S}^b_m\} = (\gamma^+ h)^{ab}\delta_{n+m,0}. \tag{2.7}$$

In the case of the left-moving internal coordinates, X^I, we must introduce Dirac brackets to take into account the second-class constraints $\Phi^I = 0$. Thus the Poisson bracket of X^I with its canonical momentum, $P^I = (1/\pi)\partial_\tau X^I$, which is canonically given by

$$\{X^I(\sigma, \tau), P^I(\sigma', \tau)\} = \delta(\sigma - \sigma')\delta^{IJ} \tag{2.8}$$

is modified to yield the Dirac bracket

$$\{X^I(\sigma, \tau), P^J(\sigma', \tau)\}_{\rm DB} = \{X^I(\sigma, \tau), P^J(\sigma', \tau)\} - \int d\sigma'' \, d\sigma'''\{X^I(\sigma, \tau), \Phi^K(\sigma'', \tau)\}$$
$$\times C_{KK'}(\sigma'', \sigma''')\{\Phi^{K'}(\sigma''', \tau), P^J(\sigma', \tau)\}, \tag{2.9}$$

where $C_{KK'}(\sigma, \sigma')$ is the inverse of the Poisson bracket of $\Phi^K(\sigma)$ and $\Phi^K(\sigma')$. Since $C_{KK'}(\sigma, \sigma') = \delta^{KK'}(1/2\pi\partial_\sigma)\delta(\sigma - \sigma')$, one finds that

$$\{X^I(\sigma, \tau), P^J(\sigma', \tau)\}_{\rm DB} = \tfrac{1}{2}\delta(\sigma - \sigma')\delta^{IJ}. \tag{2.10}$$

Therefore X^I has the expansion

$$X^I(\tau + \sigma) = x^I + p^I(\tau + \sigma) + \tfrac{1}{2}i \sum \frac{\tilde{\alpha}^I_n}{n} e^{-2in(\tau+\sigma)}, \tag{2.11}$$

where upon quantization

$$[\tilde{\alpha}^I_n, \tilde{\alpha}^J_m] = n\delta_{n+m,0}\delta^{IJ}, \tag{2.12}$$

$$[x^I, p^J] = \tfrac{1}{2}i\delta^{IJ}. \tag{2.13}$$

The factor of $\frac{1}{2}$ in the commutator of x^I and p^I, the center-of-mass position and momenta in the internal space, arises due to the constraint that ensures that X^I is a function of $\tau + \sigma$. It means that $2p^I$, $I = 1 \cdots 16$ are the generators of translations in the internal space. The allowed values of p^I will be determined, as we shall see below, by the structure of the internal sixteen-dimensional space.

In light-cone gauge $X^+(\tau, \sigma) = x^+ + p^+\tau$ and X^- is determined by solving the constraints that result from choosing a conformally flat metric on the world sheet [11]. If we expand $X^-(\tau, \sigma)$ as

$$X^-(\tau, \sigma) = x^- + p^-\tau + \tfrac{1}{2}i \sum_{n\neq 0} \frac{1}{n}(\alpha^-_n e^{-2in(\tau-\sigma)} + \tilde{\alpha}^-_n e^{-2in(\tau+\sigma)}), \tag{2.14}$$

then α^-_n is given, as in the fermionic string, as ($n \neq 0$)

$$\alpha^-_n = \frac{1}{p^+} \sum_m \alpha^i_m \alpha^i_{n-m} + \frac{1}{2p^+} \sum_m (m - \tfrac{1}{2}n) \bar{S}_{n-m} \gamma^- S_m , \tag{2.15}$$

and $\tilde{\alpha}^-_n$ is constructed, as in the bosonic string, as ($n \neq 0$)

$$\tilde{\alpha}^-_n = \frac{1}{p^+} \sum_m (\tilde{\alpha}^i_m \tilde{\alpha}^i_{n-m} + \tilde{\alpha}^I_m \tilde{\alpha}^I_{n-m}) . \tag{2.16}$$

(Note, in these formulas $\alpha^i_0 = \tilde{\alpha}^i_0 = \frac{1}{2}p^i$, $\tilde{\alpha}^I_0 = p^I$.)

These same constraints determine p^-, the generator of τ-translations conjugate to X^+, and thereby the mass operator $m^2 = 2p^+p^- - (p^i)^2$ of the string

$$\tfrac{1}{4}(\text{mass})^2 = N + (\tilde{N} - 1) + \tfrac{1}{2} \sum_{I=1}^{16} (p^I)^2 , \tag{2.17}$$

where $N(\tilde{N})$ are the normal-ordered number operators for the right (left) movers

$$N = \sum_{n=1}^{\infty} (\alpha^i_{-n} \alpha^i_n + \tfrac{1}{2} n \bar{S}_{-n} \gamma^- S_n) ,$$

$$\tilde{N} = \sum_{n=1}^{\infty} (\tilde{\alpha}^i_{-n} \tilde{\alpha}^i_n + \tilde{\alpha}^I_{-n} \tilde{\alpha}^I_n) . \tag{2.18}$$

The subtraction of -1 in (2.17) is due to the normal ordering of $\tilde{N}$, which is unnecessary in the case of the right movers due to fermion-boson cancellations. This subtraction can also be seen to be necessary to ensure ten-dimensional Lorentz invariance. Finally the factor of $\frac{1}{2}(p^I)^2$ comes from the internal momentum and winding number (see below) of the left movers.

In addition there is a further constraint that requires

$$N = \tilde{N} - 1 + \tfrac{1}{2} \sum_{I=1}^{16} (p^I)^2 . \tag{2.19}$$

This constraint has a simple physical explanation. Since there is no distinguished point on a closed string, we are free to shift the origin of the σ-coordinate by an arbitrary amount Δ. This is achieved by the unitary operator

$$U(\Delta) \equiv e^{2i\Delta(N - \tilde{N} + 1 - \frac{1}{2}\sum_I (p^I)^2)} , \tag{2.20}$$

which (recall that $[x^I, p^J] = \frac{1}{2}i\delta^{IJ}$) satisfies $U(\Delta)F(\tau, \sigma)U^+(\Delta) = F(\tau, \sigma + \Delta)$, where F can be X^i, X^I or S^a. The operator $U(\Delta)$ must therefore equal the identity operator on the space of physical states, which must therefore satisfy eq. (2.19). Again the subtraction constant, -1, is due to the normal ordering of $\tilde{N}$ and is necessary for Lorentz invariance. In sect. 4 these constraints will be rederived starting from a covariant formulation of the heterotic string.

2.2. COMPACTIFICATION OF THE INTERNAL DIMENSIONS

The extra 16 left-moving coordinates of the heterotic string can be interpreted as parametrizing an "internal" compact space T. It is unlikely that T could be curved without producing an inconsistent string theory, since then the resulting two-dimensional field theory of X^I would be an interacting nonlinear σ-model with inevitable conformal anomalies. Therefore we consider only flat compact manifolds, in other words T is a sixteen-dimensional torus*.

Since closed strings contain gravity in their low-energy limit one expects that a compactified closed string theory will contain massless vectors associated with the isometries of the compact space. In the case of a flat 16-dimensional torus this would yield the gauge bosons of $[\mathrm{U}(1)]^{16}$ A remarkable feature of closed string theories is that for special choices of the compact space there will exist additional massless vector mesons. These are in fact massless solitons of the closed string theory. They combine with the Kaluza-Klein gauge bosons to fill out the adjoint representation of a simple Lie group whose rank equals the dimension of T. In the case of the heterotic string the structure of T is so severely limited that only two choices are consistent. These produce the gauge mesons of $\mathrm{Spin}(32)/\mathrm{Z}_2$ or $\mathrm{E}_8\times\mathrm{E}_8$!

To illustrate this phenomenon consider the simple example of the bosonic (26-dimensional) string, with one coordinate compactified to a circle (say $0\leq X^{24}\leq 2\pi R$). The coordinates X^i, $i=1\cdots 23$, all have the normal mode expansion of eq. (2.5). However the expansion of $X^{24}(\tau,\sigma)$ is modified. First the momentum, p^{24}, in the compact dimension is quantized in units of $1/R$. In addition string theories give rise to a new phenomenon which has no counterpart in field theories of point particles, namely the existence of solitons which are consequences of the multiply connected structure of configuration space [15]. The function $X^{24}(\tau,\sigma)$ maps the circle $0\leq\sigma\leq\pi$ onto the circle $0\leq X^{24}\leq 2\pi R$, and such maps fall into homotopy classes characterized by a winding number L that counts how many times X^{24} wraps around the circle. Thus

$$X^{24}(\tau,\sigma)=x^{24}+p^{24}\tau+2LR\sigma+\tfrac{1}{2}i\sum_{n\neq 0}\frac{1}{n}\left(\alpha_n^{24}\,\mathrm{e}^{-2in(\tau-\sigma)}+\tilde{\alpha}_n^{24}\,\mathrm{e}^{-2in(\tau+\sigma)}\right)$$

$$p^{24}=\frac{M}{R}\,;\quad M,\,L\text{ are integers}\,. \tag{2.21}$$

String configurations with nonvanishing, integer L are solitons. They must be included in the spectrum of interacting closed string theories, since a string with $L=0$ can split to form two strings with $L=+1$ and $L=-1$ (see fig. 1). Such solitons will always arise whenever a closed string propagates on a nonsimply connected

* For consistent compactification of supersymmetric coordinates there is strong evidence that it suffices for T to be Ricci flat [14]. The only known consistent possibility for compactification of purely bosonic coordinates is for T to be a flat torus.

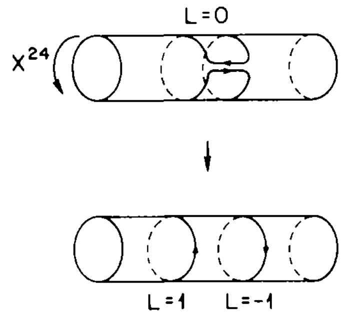

Fig. 1. A string with winding number $L=0$ splits into two strings of winding numbers $L=\pm 1$.

manifold. The excitations of the string with nonzero momentum (winding number) will have masses proportional to $1/R^2$ (R^2). The mass operator for the standard, 26-dimensional, closed string is given by

$$\tfrac{1}{4}(\text{mass})^2 = N+\tilde{N}-2+\frac{M^2}{4R^2}+L^2R^2\,, \tag{2.22}$$

where N and $\tilde{N}$ are the number operators for the left and right movers; the factor of $M^2/4R^2=\frac{1}{4}(p^{24})^2$ arises from the momentum in the compact dimension, and the term L^2R^2 is the energy required to wrap the string around the circle.

The physical states of the closed string satisfy a constraint which for an uncompactified string is simply $N=\tilde{N}$. This constraint implies that the unitary operator $U(\Delta)=\exp[2i(N-\tilde{N})\Delta]$, which shifts σ by an amount Δ, is equal to the identity operator on physical states, thus ensuring the absence of a distinguished point on the string. If the string contains a winding number term, $2LR\sigma$, one must modify $U(\Delta)$ so as to shift this as well. Thus

$$U(\Delta)=\exp[2i(N-\tilde{N}-p^{24}RL)\Delta] \tag{2.23}$$

and the physical states must satisfy

$$N=\tilde{N}+ML\,. \tag{2.24}$$

Let us now examine the massless spectrum of the effective 25-dimensional theory. It consists of the transverse components of the graviton, antisymmetric tensor and dilaton, $\tilde{\alpha}^i_{-1}\alpha^j_{-1}|0\rangle$, $1\leq i,j\leq 23$; the transverse components of two vector mesons, $\tilde{\alpha}^i_{-1}\alpha^{24}_{-1}|0\rangle$ and $\alpha^i_{-1}\tilde{\alpha}^{24}_{-1}|0\rangle$. These correspond to the gauge bosons expected from the usual Kaluza-Klein compactification on a circle with isometry $U(1)$. However for the particular choice of the radius of the circle

$$R=\sqrt{\alpha'}=\sqrt{\tfrac{1}{2}}\,, \tag{2.25}$$

the solitons with winding numbers $L=\pm 1$ and momenta $M=\pm 1$ (with $N-\tilde{N}=ML$) are also massless. Thus there are four more massless vectors $\alpha^i_{-1}|M=L=\pm 1\rangle$ and $\tilde{\alpha}^i_{-1}|M=-L=\pm 1\rangle$.

These six massless vectors form the adjoint representation of SU(2) × SU(2), with the diagonal generators having eigenvalues $\frac{1}{2}(M+L)$ and $\frac{1}{2}(M-L)$ respectively. This construction is a physical realization of the work of Frenkel and Kac on representations of Kac-Moody Lie algebras [7]. Their work suggests that this construction can be generalized to produce massless vector mesons in the adjoint representation of any simply laced Lie group G, i.e. a Lie group whose root vectors all have the same length. This restricts G to be SU(n), SO($2n$), E_6, E_7, E_8 or direct products of these. The method employs closed strings moving on a torus such that the lattice of allowed momenta coincides with the root lattice of the group. In other words the string moves on the maximal torus of the group G. However the quantization differs from the usual quantization of closed bosonic strings.

First consider compactification of the usual closed bosonic string on a D-dimensional torus, which may be thought of as R^D modulo a lattice Γ generated by D independent basis vectors e_i^I ($i=1\cdots D$). We shall choose the length of the e_i to be equal to $\sqrt{2}$ and identify the center-of-mass coordinate x^I with its translation by $2\pi R_i$ in the direction of e_i, so that

$$x^I \equiv x^I + \sqrt{2}\pi \sum_{i=1}^{D} n_i R_i e_i^I\,, \tag{2.26}$$

with n_i being integers. The mode expansion for the internal coordinates is

$$X^I(\sigma,\tau) = x^I + p^I\tau + 2L^I\sigma + \tfrac{1}{2}i \sum_{n\neq 0} \frac{1}{n}\left(\alpha_n^I\, e^{-2in(\tau-\sigma)} + \tilde{\alpha}_n^I\, e^{-2in(\tau-\sigma)}\right). \tag{2.27}$$

For a torus R^D/Γ the allowed center-of-mass momenta p^I lie on the dual lattice $\tilde{\Gamma}$, generated by e_i^{*I} ($i=1\cdots D$) defined by

$$\sum_{I=1}^{D} e_i^I e_j^{*I} = \delta_{ij}\,. \tag{2.28}$$

Since p^I generates translations of x^I we must have

$$p^I = \sqrt{2} \sum_{i=1}^{D} \frac{m_i}{R_i} e_i^{*I} \tag{2.29}$$

for integer m_i. The allowed winding numbers are

$$L^I = \sqrt{\tfrac{1}{2}} \sum_{i=1}^{D} R_i n_i e_i^I \tag{2.30}$$

for integer n_i. The constraint (2.24) is generalized to

$$N = \tilde{N} + \sum_I p^I L^I = \tilde{N} + \sum_i m_i n_i\,, \tag{2.31}$$

and the mass operator is now

$$\tfrac{1}{4}(\text{mass})^2 = N + \tilde{N} - 2 + \sum_I \left(\tfrac{1}{4}p^I p^I + L^I L^I\right)$$

$$= N + \tilde{N} - 2 + \sum_{i,j} \left(\frac{1}{2}\frac{m_i}{R_i} g_{ij} \frac{m_j}{R_j} + \tfrac{1}{2} n_i R_i g^*_{ij} n_j R_j\right), \tag{2.32}$$

where the metric $g_{ij} = \sum_{I=1}^{D} e_i^I e_j^I$ is the Cartan matrix for the Lie group with maximal torus R^D/Γ and $g^*_{ij} = \sum_{I=1}^{D} e_i^{*I} e_j^{*I} = (g_{ij})^{-1}$ is the dual metric. With our normalization g_{ij} is integer valued and $g_{ii} = 2$ for simply laced Lie groups.

As before we obtain $2D$ massless vector mesons $\tilde{\alpha}^i_{-1}\alpha^I_{-1}|0\rangle$ and $\alpha^i_{-1}\tilde{\alpha}^I_{-1}|0\rangle$ as expected since the isometry group of the torus is $\mathrm{U}(1)^D$. However the constraints make it clear that it is not possible to obtain the additional massless vectors with $N - \tilde{N} = \pm 1$ needed to complete the adjoint representation of a simple nonabelian group of rank $D > 1$ for any real R_i. The resolution of this difficulty provides a key to the construction of the heterotic string.

We return to the mode expansion (2.27) and write $X^I(\sigma, \tau) = X^I(\tau - \sigma) + \tilde{X}^I(\tau + \sigma)$ with

$$X^I(\tau - \sigma) = \tfrac{1}{2}x^I + (\tfrac{1}{2}p^I - L^I)(\tau - \sigma) + \tfrac{1}{2}i \sum_{n \neq 0} \frac{1}{n} \alpha_n^I \, \mathrm{e}^{-2in(\tau - \sigma)},$$

$$\tilde{X}^I(\tau + \sigma) = \tfrac{1}{2}x^I + (\tfrac{1}{2}p^I + L^I)(\tau + \sigma) + \tfrac{1}{2}i \sum_{n \neq 0} \frac{1}{n} \tilde{\alpha}_n^I \, \mathrm{e}^{-2in(\tau + \sigma)}. \tag{2.33}$$

Note that the left and right oscillators are treated independently but the center-of-mass coordinates and momenta are *not*. For ordinary space-time coordinates with $L = 0$ this is clearly necessary. However for an internal nonsimply-connected space we are free to regard X^I and $\tilde{X}^I$ as *completely independent* massless two-dimensional fields with expansions

$$X^I(\tau - \sigma) = x^I + p^I(\tau - \sigma) + \text{oscillators},$$

$$\tilde{X}^I(\tau + \sigma) = \tilde{x}^I + \tilde{p}^I(\tau + \sigma) + \text{oscillators}, \tag{2.34}$$

and $[x^I, \tilde{p}^J] = [\tilde{x}^I, p^J] = 0$, but $[x^I, p^J] = [\tilde{x}^I, \tilde{p}^J] = \tfrac{1}{2}i\delta^{IJ}$ as in (2.13). Now $2p^I(2\tilde{p}^I)$ generates translations of $x^I(\tilde{x}^I)$ so the allowed momenta are

$$p^I = \sqrt{\tfrac{1}{2}} \sum_{i=1}^{D} \frac{m_i}{R_i} e_i^{*I} \tag{2.35}$$

and similarly for $\tilde{p}^I$.

The left- and right-moving sectors are now related only by the constraint which follows from invariance under shifts in σ which becomes

$$N + \tfrac{1}{2}\sum_I (p^I)^2 = \tilde{N} + \tfrac{1}{2}\sum_I (\tilde{p}^I)^2. \tag{2.36}$$

The mass operator is now

$$\tfrac{1}{4}(\text{mass})^2 = N + \tilde{N} - 2 + \tfrac{1}{2}\sum_I \left(p^I p^I + \tilde{p}^I \tilde{p}^I\right). \tag{2.37}$$

Choosing the radius to be

$$R_i = R = \sqrt{\tfrac{1}{2}}\ (= \sqrt{\alpha'})\,, \tag{2.38}$$

we now find $2D$ U(1) gauge bosons as before but in addition there are massless vector meson states with $N - \tilde{N} = +1\ (-1)$ and $(\tilde{p}^I)^2 = 2((p^I)^2 = 2)$. These complete the adjoint representation of G×G, where G is a simply-laced Lie group with rank G $= D$.

The heterotic string is constructed by utilizing the left-moving part of the above construction with $D = 16$ to produce gauge bosons in the adjoint representation of $E_8 \times E_8$ or Spin(32)/Z_2. This choice is dictated by the following argument. Since X^I is to be a function of $\tau + \sigma$ only, the allowed momenta (2.35) must coincide with an allowed winding number (2.30). Thus the allowed states must have momenta which lie on the intersection of Γ and $\tilde{\Gamma}$. Furthermore, in order to satisfy the constraint, eq. (2.19), the square of the allowed momenta must be even integers. Finally, the lattice must be self-dual, i.e. Γ must equal $\tilde{\Gamma}$ (equivalently det $g_{ij} = 1$). This last requirement is necessary to have a consistent theory of *interacting* closed strings. We shall discuss this in detail in the following paper [9]. For the moment we simply note that the (euclidean) two-dimensional world sheet which describes a one-loop closed string amplitude is a torus. This torus is clearly symmetric under the interchange of σ and τ. To achieve this symmetry, necessary for global reparametrization invariance, the coefficients of τ and σ in X^I must span all the states of the *same* lattice. All of these requirements can be satisfied if and only if all the radii are determined as in (2.38) and the lattice is integral, even and self-dual. This means that the metric $g_{ij} = \sum_{I=1}^{16} e_i^I e_j^I$ is integer valued, g_{ii} is even and $\Gamma = \tilde{\Gamma}$. For this reason the obvious torus, given by R^{16}/Z^{16} is unacceptable. Z^{16} is self-dual if one chooses the basis vectors to be of length one - but then it is not even. The sublattice consisting only of even points is not self-dual.

Even, self-dual lattices are extremely rare. In fact, they only exist in $8n$ dimensions. In sixteen dimensions there are two such lattices, $\Gamma_8 \times \Gamma_8$ and Γ_{16}. The first is the direct product of two Γ_8, where Γ_8 is the root lattice of E_8, and g_{ij} is the Cartan matrix of E_8:

$$g_{ij}^{E_8} = \begin{vmatrix} 2 & -1 & & & & & & \\ -1 & 2 & -1 & & & & & \\ & -1 & 2 & -1 & & & & \\ & & -1 & 2 & -1 & & & \\ & & & -1 & 2 & -1 & & -1 \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 2 & 0 \\ & & & & -1 & & 0 & 2 \end{vmatrix}. \tag{2.39}$$

The construction of Γ_{16} is slightly more complicated. The basis vectors for the SO(32) root lattice can be written in terms of a set of orthonormal vectors $\{u_i\}$, $i=1\cdots 16$ as $e_i=u_i-u_{i+1}$, $i=1\cdots 15$, $e_{16}=u_{15}+u_{16}$. This root lattice is not self-dual since $\det g_{ij}=4$. However a self-dual lattice can be constructed from the root lattice by adding points that are multiples of *one* of the spinor weights of Spin(32). We can choose the spinor weight to be $s_1=\frac{1}{2}(u_1-u_2+u_3-u_4+\cdots+u_{15}-u_{16})$, $s_1^2=4$ and take the basis vectors for Γ_{16} to be s_1 and e_i, $i=2\cdots 16$. Since e_1 is an integer linear combination of the basis vectors, Γ_{16} contains all the points of the root lattice of SO(32) plus additional points that correspond to spinor weights of $\mathrm{Spin}(32)/\mathrm{Z}_2$. Note that this is not the same as SO(32). The center of Spin(32) is $\mathrm{Z}_2\times\mathrm{Z}_2$. Removing the diagonal combination of the two Z_2 factors eliminates all spinor representations, leaving SO(32). Removing one Z_2 eliminates one of the two spinor representations of Spin(32) and all representations in the same Z_2 conjugacy class as the spinor (including the vector).

On either of these tori x^I is given by eq. (2.26) where the allowed momenta are

$$p^I=\sum_{i=1}^{16} n_i e_i^I \qquad (n_i=\text{integer})\,, \tag{2.40}$$

and n_i also measures the winding number in the ith direction around the torus. The values of $(p_I)^2$ are all even integers, the minimal value is 2. Each of these sixteen-dimensional self-dual lattices has precisely 480 vectors of length squared 2. These will yield, in the heterotic string theory, 480 massless vector bosons. These 480 solitons are clearly labeled by the weights of the adjoint representations of $\mathrm{G}=\mathrm{E}_8\times\mathrm{E}_8$ or $\mathrm{Spin}(32)/\mathrm{Z}_2$. They combine with the 16 massless vectors with zero weights (the Kaluza-Klein gauge bosons) to fill out the adjoint representation of G. We shall see below that all the states of the heterotic string will form representations of G.

2.3. THE SPECTRUM

The physical states of the heterotic string are simply direct products of Fock space states $(|\rangle_{\mathrm{R}}\times|\rangle_{\mathrm{L}})$ of the right-moving fermionic string and the left-moving bosonic string, subject to the constraint (2.19). The right-moving ground state is annihilated by α_n^i, S_n^a $(n>0)$ and N. It forms an irreducible representation of the zero mode oscillators S_0^a and consists of 8 bosonic states $|i\rangle_{\mathrm{R}}$ and 8 fermionic states $|a\rangle_{\mathrm{R}}$, which satisfy

$$|a\rangle_{\mathrm{R}}=\tfrac{1}{8}i(\gamma_i S_0)^a|i\rangle_{\mathrm{R}}\,,\qquad |i\rangle_{\mathrm{R}}=\tfrac{1}{16}i(\bar{S}_0\{\gamma_i,\gamma_+\})^a|a\rangle_{\mathrm{R}}\,. \tag{2.41}$$

The left-handed ground state is annihilated by $\tilde{\alpha}_n^i$, $\tilde{\alpha}_n^I$ $(n>0)$, $\tilde{N}$ and p^I. By itself it is a tachyon, but it is removed from the physical spectrum by the constraint which requires that $\tilde{N}+\frac{1}{2}(p^I)^2=N+1>0$.

The spectrum of physical states contains only nonnegative masses due to the constraint. In fact

$$(\text{mass})^2 = 8N \geq 0 ,$$
$$N = \tilde{N} + \tfrac{1}{2}(p^I)^2 - 1 . \tag{2.42}$$

The ground state is therefore massless and consists of the direct product of $|i\rangle_R$ or $|a\rangle_R$ with $\tilde{\alpha}^j_{-1}|0\rangle_L$, $\tilde{\alpha}^I_{-1}|0\rangle_L$ or $|p^I\rangle_L$ (where $(p^I)^2=2$). All these states have $N=\tilde{N}+\frac{1}{2}(p^I)^2-1=0$. As expected they form irreducible 10-dimensional, massless, supermultiplets. The states $|i$ or $a\rangle_R \times \tilde{\alpha}^j_{-1}|0\rangle_L$ form the irreducible $N=1$, $D=10$ supergravity multiplet. Note that the antisymmetric tensor appears here in the antisymmetric product of two SO(8) vectors as compared to the type I closed string theory where it appears in the antisymmetric product of two SO(8) spinors. The states $|i\rangle_R \times \tilde{\alpha}^I_{-1}|0\rangle_L$ and $|a\rangle_R \times \tilde{\alpha}^I_{-1}|0\rangle_L$ together with the 16×480 states $|i$ or $a\rangle_R \times |p^I, (p^I)^2=2\rangle_L$ form an irreducible $N=1$, $D=10$ super-Yang-Mills multiplet of G (where $G=E_8\times E_8$ or Spin $(32)/Z_2$).

Although the spectrum of the massless states of the heterotic string (with $G=\text{Spin}(32)/Z_2$) coincides with that of the type I superstring with gauge group SO(32), the difference between these two theories is apparent in the structure of the massive levels. In the heterotic string we will obtain states with arbitrary p^I = weight vector of G, and thus states which are in arbitrarily large representations of G. The type I open string theory contains only the antisymmetric tensor (adjoint) and the symmetric tensor representations of SO(32).

We shall exhibit explicitly the structure of the first massive level with $N=1$, $(\tilde{N}, (p^I)^2) = (2,0)$, $(1,2)$ or $(0,4)$, and $(\text{mass})^2 = 8$. Already at this level there are 18 883 584 physical degrees of freedom! The left- and right-moving sectors separately can be assembled into SO(9) multiplets, since the Lorentz and supersymmetry generators act separately on each. Among the right movers we have 128 bosonic states $\alpha^i_{-1}|j\rangle_R$ and $S^a_{-1}|b\rangle_R$ which fill out the SO(9) representations $\square\!\square$ and $\begin{smallmatrix}\square\\ \square\\ \square\end{smallmatrix}$ of dimensions 44 and 84 respectively. In addition there are the fermion states, $\alpha^i_{-1}|a\rangle_R$ and $S^a_{-1}|i\rangle_R$ which form a single irreducible SO(9) representation of dimension 128.

The left-moving sector consists of the SO(8) scalars $|p^I, (p^I)^2=4\rangle_L$, $\tilde{\alpha}^I_{-1}|p^I, (p^I)^2=2\rangle_L$, $\tilde{\alpha}^I_{-1}\tilde{\alpha}^J_{-1}|0\rangle_L$ and $\tilde{\alpha}^I_{-2}|0\rangle_L$; the SO(8) vectors $\tilde{\alpha}^i_{-1}|p^I, (p^I)^2=2\rangle_L$, $\tilde{\alpha}^i_{-2}|0\rangle_L$, $\tilde{\alpha}^i_{-1}\tilde{\alpha}^I_{-1}|0\rangle_L$ and the SO(8) symmetric tensors $\tilde{\alpha}^i_{-1}\tilde{\alpha}^j_{-1}|0\rangle_L$. To count the number of these states we use the fact that the number of points of length squared $2m$ on the lattices Γ_{16} or $\Gamma_8\times\Gamma_8$ is given by $480\,\sigma_7(m)$, where $\sigma_7(m)$ is the sum of the seventh powers of the divisors of m [16]. Using this formula we find 69 752 SO(8) singlets which can be assembled into irreducible $E_8\times E_8$ representations as $(3875,1)+(1,3875)+(248,248)+(248,1)+(1,248)+(1,1)+(1,1)$. For $\text{Spin}(32)/Z_2$ the irreducible representations are $\begin{smallmatrix}\square\\ \square\\ \square\\ \square\end{smallmatrix}$ (dimension 35 960), $\begin{smallmatrix}\square\\ \square\end{smallmatrix}$ (496), $\square\!\square$ (527), the spinor of dimension (32 768) and one singlet (1). Note the appearance at the first massive level of the SO(32) spinor. This state cannot decay and might pose cosmological problems. Similarly the 497 SO(8) vectors transform as the adjoint representation of G plus a singlet while the SO(8) symmetric tensor is a singlet under G.

These states may be easily assembled into SO(9) multiplets, whereby the G singlets are in the SO(9) representation $\underset{\approx}{44}+\underset{\sim}{1}$ and the G adjoint representations form a SO(9) vector. The other G representations are SO(9) singlets. Finally the complete SO(9)$\times$G multiplets are obtained by taking the SO(9) tensor product of the representations $(\square\!\square+\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix})$ occurring in the right-moving sector with the representations $(\square\!\square+\square+\cdot)$ occurring in the left-moving sector. As we shall show below the physical states of the heterotic string form a representation of the super $D=10$ Poincaré group and of the algebra of $G=E_8\times E_8$ or Spin(32)/Z_2. This guarantees that a similar grouping of the states exists at each mass level.

The number of states of a string theory increases dramatically as a function of mass. In fact the degeneracy of the levels with mass M increases exponentially as $d(M)\sim\exp(\bar{\beta}M)$. This leads to a "limiting temperature" for the noninteracting string of $kT_L=1/\bar{\beta}$, since at this temperature the specific heat diverges [17]. Of course interactions might modify this conclusion. The value of $\bar{\beta}$ is easily determined by calculating the "partition" function $P(x)$ which, in the case of the standard bosonic closed string in 26 dimensions, is given by

$$P_B(x)=\left\{\prod_{n=1}^{\infty}(1-x^n)\right\}^{-24}=\sum_{N=0}^{\infty}P_B(N)x^N. \tag{2.43}$$

It is easy to see that $P_B(N)$ counts the number of string states with right-moving oscillator number N. Since for the closed string $N=\tilde{N}$ and $M^2=4(N+\tilde{N}-2)=8(N-1)$ we have that $d_B(M)=P_B^2(\frac{1}{8}M^2+1)$. Similarly the partition function of the fermionic superstring in 10 dimensions is given by

$$P_F(x)=\left[\frac{\prod_{n=1}^{\infty}(1+x^n)}{\prod_{n=1}^{\infty}(1-x^n)}\right]^8=\sum_{N=0}P_F(N)x^N. \tag{2.44}$$

Here however $M^2=4(N+\tilde{N})=8N$ and the ground state is 256-fold degenerate so that $d_F(M)=256P_F^2(\frac{1}{8}M^2)$. Since the heterotic string is a hybrid of the bosonic and fermionic strings and $M^2=4(N+\tilde{N}-1+\frac{1}{2}(p^I)^2)=8N$ we have

$$d_M(M)=16P_F(\tfrac{1}{8}M^2)P'_B(\tfrac{1}{8}M^2), \tag{2.45}$$

where $P'_B(x)$ is the partition function of the left movers

$$P'_B(x)=\frac{1}{x}\left\{\prod_{n=1}^{\infty}(1-x^n)\right\}^{-24}\left(1+\sum_{m=1,2,\ldots}480\sigma_7(m)x^m\right)=\sum_{\tilde{N}=-1}P'_B(\tilde{N})x^{\tilde{N}}. \tag{2.46}$$

The large-N behavior of the $P(N)$ can be determined from the behavior of $P(x)$ as $x \to 1$. Using the formulae

$$\prod_{n=1}^{\infty}(1\pm x^n) = \exp\sum_{n=1}^{\infty}\ln(1\pm x^n) \underset{x\to 1}{\simeq} \exp -\sum_{1}^{\infty}\frac{(\mp)^k}{k^2}\frac{1}{(1-x)},$$

we deduce that

$$P_{\mathrm{B}}(x) \sim P'_{\mathrm{B}}(x) \underset{x\to 1}{\simeq} \exp\frac{4\pi^2}{1-x}, \qquad P_{\mathrm{F}}(x) \sim \exp\frac{2\pi^2}{1-x}.$$

This translates into exponential growth of $P(N)$, which behaves as $P_{\mathrm{B}}(N) \sim P'_{\mathrm{B}}(N) \sim \exp(4\pi\sqrt{N})$; $P_{\mathrm{F}}(N) \sim \exp(2\sqrt{2}\pi\sqrt{N})$. In terms of the "limiting temperature" we then have

$$\bar{\beta}_{\mathrm{B}} = \sqrt{8}\pi = (4\pi\sqrt{\alpha'}),$$
$$\bar{\beta}_{\mathrm{F}} = 2\pi = (2\sqrt{2}\pi\sqrt{\alpha'}), \tag{2.47}$$

for the standard bosonic and fermionic strings. The limiting temperature of the heterotic string lies in between these

$$\bar{\beta}_{\mathrm{H}} = \tfrac{1}{2}(\bar{\beta}_{\mathrm{B}} + \bar{\beta}_{\mathrm{F}}) = (1+\sqrt{2})\pi = (2+\sqrt{2})\pi\sqrt{\alpha'}. \tag{2.48}$$

2.4. SYMMETRIES

It is straightforward to verify that the heterotic string is Lorentz invariant and $N=1$ supersymmetric in 10 dimensions. This is because the generators of these transformations act separately on the right and left movers. Since these are formulated in their critical dimensions of 10 and 26 respectively, the would-be anomalies cancel.

Thus the heterotic string contains one supersymmetry, generated by

$$Q^a = i\sqrt{p^+}(\gamma^+ S_0)^a + 2i\frac{1}{\sqrt{p^+}}\sum_n(\gamma_i S_{-n})^a \alpha_n^i, \tag{2.49}$$

which acts on the right-moving sector. The demonstration that

$$\{Q^a, Q^b\} = -2(h\gamma^\mu P_\mu)^{ab} \tag{2.50}$$

is identical to the demonstration of the right-moving supersymmetry of the type II closed superstring theory [12] (half of its $N=2$ supersymmetry). Similarly the generators of the 10-dimensional Lorentz group, $J^{\mu\nu}$, can be written as

$$J^{\mu\nu} = l^{\mu\nu} + J_{\mathrm{R}}^{\mu\nu} + J_{\mathrm{L}}^{\mu\nu}, \tag{2.51}$$

where

$$l^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu, \tag{2.52}$$

and $J_{\mathrm{R}}^{\mu\nu}(J_{\mathrm{L}}^{\mu\nu})$ are the Lorentz generators on the oscillator coordinates of the right-moving (left-moving) sector of the fermionic (bosonic) strings respectively. For example, $J_{\mathrm{L}}^{ij} = -i\sum_{n=1}^{\infty}(1/n)(\tilde{\alpha}_{-n}^{i}\tilde{\alpha}_{n}^{j} - \tilde{\alpha}_{-n}^{j}\alpha_{n}^{i})$, $J_{\mathrm{L}}^{+-} = J_{\mathrm{L}}^{-+} = 0$, $J_{\mathrm{L}}^{i-} = -i\sum_{n=1}^{\infty}(1/n)(\tilde{\alpha}_{-n}^{i}\tilde{\alpha}_{n}^{-} - \tilde{\alpha}_{-n}^{-}\tilde{\alpha}_{n}^{i})$. The demonstration that the $J^{\mu\nu}$ satisfy the 10-dimensional Lorentz algebra is identical to the demonstration of the Lorentz invariance of the 10-dimensional superstring [12] and the 26-dimensional bosonic string [11]. Note that these demonstrations do require the normal ordering subtractions of $\tilde{N}$ that appear in eqs. (2.17) and (2.19).

We shall now show that the states of the heterotic string form representations of the group $\mathrm{G} = \mathrm{E}_8 \times \mathrm{E}_8$ or $\mathrm{Spin}(32)/\mathrm{Z}_2$. We are guided by the work of Frenkel and Kac, and Segal who used string theory operators to construct representations of Kac-Moody algebras [7]. To this end we construct an operator $E(K^I)$ which acts on the left-moving states of the heterotic string; $E(K^I)$ represents the generators of G that translate states on the weight lattice by a root vector $K^I((K^I)^2 = 2)$. In order to understand the construction of Frenkel and Kac it is useful to consider a hamiltonian treatment of the left-moving sector of the heterotic string viewed as a two-dimensional field theory with solitons. The fields $X^I(\sigma)$ are maps from the circle S^1 parametrized by $0 \leq \sigma \leq \pi$ onto the 16-dimensional torus T. The configuration space can thus be thought of as the loop space $\Omega\mathrm{T}$ of maps $\mathrm{S}^1 \to \mathrm{T}$. This loop space has the structure of a group under pointwise multiplication of maps and falls into disconnected components labelled by the winding numbers of the torus, i.e. $\pi_0(\Omega\mathrm{T}) = \pi_1(\mathrm{T})$. In quantizing the theory we consider unitary operators $U(f_1)$ which when acting on the vacuum create states corresponding to the map $f_1 : \mathrm{S}^1 \to \mathrm{T}$, $|f_1\rangle = U(f_1)|0\rangle$. If we consider two maps f_1, f_2 and their composition $f_1 f_2$ then quantum mechanically we are only interested in rays in a Hilbert space so we only require

$$U(f_1)U(f_2) = \varepsilon(f_1, f_2)U(f_1 f_2)\,, \tag{2.53}$$

with $|\varepsilon(f_1, f_2)| = 1$. The phase factor ε is required by associativity to satisfy

$$\varepsilon(f_1, f_2)\varepsilon(f_1 f_2, f_3) = \varepsilon(f_1, f_2 f_3)\varepsilon(f_2, f_3)\,, \tag{2.54}$$

which is just the statement that ε is a two-cocycle. ε is nontrivial if it cannot be factored in the form $\mathrm{e}^{i\phi(f_1)}\mathrm{e}^{i\phi(f_2)}\mathrm{e}^{-i\phi(f_1 f_2)}$ and reabsorbed in the definition of U. In this case the operators U give a nontrivial projective representation of the group $\Omega\mathrm{T}$. These representations have been considered previously in the physics literature in connection with quantization of solitons in theories such as the sine-Gordon model [7] and the Skyrme model [18]. General conditions for the existence of nontrivial projective representations are given in a classic paper of Bargmann [19].

We are interested in representing the group of translations on the weight lattice. In order to reproduce the correct group commutation relations we must consider a nontrivial projective representation of this group. Points on the weight lattice correspond to allowed winding numbers (i.e. elements of $\pi_0(\Omega\mathrm{T})$) with addition of lattice vectors corresponding to group multiplication of elements in disconnected components of ΩT. For lattice vectors K^I, L^I, M^I the two-cocycle condition is thus

$$\varepsilon(K, L)\varepsilon(K+L, M) = \varepsilon(L, M)\varepsilon(K, L+M)\,. \tag{2.55}$$

For our purposes it is convenient to modify the translation operator directly and to consider the operators

$$E(K^I) = \int \frac{\mathrm{d}z}{2\pi i z} :\mathrm{e}^{2iK^I \cdot X^I(z)}: C(K^I)\,, \tag{2.56}$$

with $(K^I)^2 = 2$, $z = \mathrm{e}^{2i(\tau+\sigma)}$, and the contour runs around the unit circle. These operators contain the usual translation operator $\mathrm{e}^{2iK^I X^I}$ which translates the internal momenta (= winding numbers) by K^I. The additional term $C(K^I)$ can be viewed as an operator 1-cocycle which will be chosen so that the 480 $E(K^I)$ with $(K^I)^2 = 2$ plus the 16 operators p^I satisfy the Lie algebra of G (in the Chevalley basis).

Using the properties of harmonic oscillator coherent states it is straightforward to verify that

$$:\mathrm{e}^{2iK^I X^I(z)}::\mathrm{e}^{2iL^I X^I(w)}: = :\mathrm{e}^{2i(K\cdot X(z)+L\cdot X(w))}:(z-w)^{K\cdot L}(wz)^{-K\cdot L/2} \quad \text{for} \quad |w|<|z|\,. \tag{2.57}$$

Under the interchange $(K, z)\leftrightarrow(L, w)$ the RHS acquires a factor of $(-)^{K\cdot L}$. We therefore demand that $C(K)C(L) = (-1)^{K\cdot L}C(L)C(K)$ in order that the E have simple commutation relations. The analog of eq. (2.53) is just

$$C(K)C(L) = \varepsilon(K, L)C(K+L). \tag{2.58}$$

In addition to the two-cocycle condition (2.55) $\varepsilon(K, L)$ may be chosen to satisfy [7]

$$\begin{gathered} \varepsilon(K, L)\varepsilon(L, K) = (-1)^{K\cdot L} \\ \varepsilon(K, 0) = -\varepsilon(K, -K) = 1\,. \end{gathered} \tag{2.59}$$

Now, if we compute the commutator of $E(K)$ and $E(L)$ (K and L are root vectors), the integral over $(z-w)^{K\cdot L}$ will yield a nonzero result only if $K\cdot L = -1$ (and therefore $K+L$ is another root vector since $(K+L)^2 = 2+2-2 = 2$), or if $K\cdot L = -2$ which requires $K = -L$. This is a consequence of the fact that

$$\begin{aligned} :\mathrm{e}^{2iK\cdot X(z)}: = \exp\left[+\sum_{n=1}^{\infty} K\cdot\tilde{\alpha}_{-n}(z^n/n)\right]\exp\left[+2iK\cdot x + K\cdot p \ln z\right] \\ \times\exp\left[-\sum_{n=1}^{\infty} K\cdot\tilde{\alpha}_n(z^{-n}/n)\right] \end{aligned} \tag{2.60}$$

is a meromorphic function of z (recall that $K\cdot p$ is integer valued on the lattice of allowed momenta).

Therefore one easily shows that

$$[E(K), E(L)] = \begin{cases} \varepsilon(K, L)E(K+L)\,, & (K+L)^2 = 2\,, \\ K^I \cdot p^I\,, & K+L=0\,, \\ 0\,, & \text{otherwise}\,, \end{cases} \tag{2.61}$$
$$[p^I, E(K)] = K^I.$$

These are precisely the commutators of the generators of G, as long as the $\varepsilon(K, L)$ are chosen to equal the structure constants (± 1 in this basis) of the group G. Frenkel and Kac show that $C(K)$ and $\varepsilon(K, L)$ can be constructed to satisfy these properties. In fact $\varepsilon(K, L)$ is defined for arbitrary K and L on the lattice by extending the structure constants (defined for $K^2 = L^2 = (K+L)^2 = 2$) to the whole lattice using (2.59) and bilinearity.

Therefore we have an explicit representation of the generators of G on the Fock space of the left movers of the heterotic string. Since the right movers are G-singlets and the constraint, eq. (2.19), is G-invariant, it follows that the free heterotic string is G-invariant.

3. Fermionic representation

There is an alternate description of the internal degrees of freedom of the heterotic string (which are responsible for the Yang-Mills interactions) in terms of chiral, anticommuting, Lorentz-scalar coordinates. These fermionic coordinates are related to the bosonic coordinates discussed above by the well-studied bosonization of two-dimensional fermion fields, with minor complications due to the finite length of the string. For purposes of describing the spectrum and interactions of the string the two formulations are equivalent. They correspond to alternate constructions of a particular representation of the affine Lie algebra based on the gauge group of the theory, whose properties determine the structure of the string in the light-cone gauge*.

We shall now present the construction of the heterotic string in terms of internal fermionic coordinates, beginning with the model with Spin(32)/Z_2 gauge group. The construction of the $E_8 \times E_8$ model requires subtle modifications. The construction parallels that of the superstring in terms of the Neveu-Schwarz [21] and Ramond [22] models. Here we introduce a set of 32 real, anticommuting, left-moving coordinates $\psi^i(\tau+\sigma)$ which transform as the fundamental representation of SO(32). Regarded as two-dimensional fields on the string world-sheet the ψ^i are Majorana-Weyl fermions, however they transform as ten-dimensional Lorentz scalars. The light-cone gauge action of these fields is simply $(1/2\pi)\int d\tau \int_0^\pi d\sigma\, \psi^i\, (\partial/\partial\tau - \partial/\partial\sigma)\psi^i$, yielding the equation of motion $(\partial/\partial\tau - \partial/\partial\sigma)\psi^i = 0$.

* The use of fermionic variables to create representations of an internal symmetry group is not new; Bardacki and Halpern introduced this notion [20] as a possible way of incorporating hadronic symmetries in dual resonance models.

At this stage we have the option of choosing periodic or antiperiodic boundary conditions for ψ^i. As in the case of the "old superstring", we find it necessary to include in the Hilbert space of string states two sectors corresponding to the two possible boundary conditions. This is required to have a consistent, unitary description of string loop amplitudes. We have then one sector, analogous to the Ramond model, in which the fermions obey periodic boundary conditions in σ and thus have a Fourier expansion

$$\psi^i(\tau+\sigma) = \sum_{n=-\infty}^{+\infty} \psi^i_n \, \mathrm{e}^{-2in(\tau+\sigma)}, \qquad n \text{ integer}. \tag{3.1}$$

Upon quantization, the oscillators ψ^i_n obey

$$\{\psi^i_n, \psi^j_m\} = \delta^{ij}\delta_{n+m,0}\,. \tag{3.2}$$

The other (Neveu-Schwarz) sector contains fermions which obey antiperiodic boundary conditions in σ, so that the integer frequencies n in eqs. (3.1) and (3.2) are shifted by $\frac{1}{2}$. The choice of boundary condition determines the ground state in each sector. In the antiperiodic case the ground state is an SO(32) singlet state annihilated by all ψ^i_n $(n>0)$. In the periodic sector the zero frequency oscillators, ψ^i_0, form a Clifford algebra and therefore cannot annihilate the vacuum. Hence we must take the ground state to be the appropriate spinor representation of SO(32).

The hamiltonian describing the contribution of the left-moving bosonic (the eight transverse coordinates $\tilde{X}^i(\tau+\sigma)$) and fermionic modes to the mass operator of the string is

$$\begin{aligned} \tilde{N} &= \tilde{N}_{\mathrm{B}} + \tilde{N}_{\mathrm{F}}\,, \\ \tilde{N}_{\mathrm{F}} &= \sum_{n>0} n\psi^i_{-n}\psi^i_n - C_{\mathrm{A,P}}\,. \end{aligned} \tag{3.3}$$

$C_{\mathrm{A}}(C_{\mathrm{P}})$ is the normal ordering constant in the case of antiperiodic (periodic) fermions respectively. This subtraction, which includes the contribution from the spatial oscillations as well, is uniquely determined by the requirement of 10-dimensional Lorentz invariance. A straightforward calculation along the lines of ref. [12] provides the values

$$C_{\mathrm{A}} = 1\,, \qquad C_{\mathrm{P}} = -1\,. \tag{3.4}$$

We then obtain the spectrum pictured in fig. 2 (disregarding the spatial oscillations). The two sets of states so constructed each form irreducible Fock space representations of the respective sets of oscillators. However, as previously mentioned, we are actually interested in constructing representations of the affine Lie algebra based on SO(32), whose generators M^{ij}_n are bilinear in the fermions:

$$M^{ij}_n = \tfrac{1}{2}\sum_{m=-\infty}^{+\infty} \psi^i_{n-m}\psi^j_m \tag{3.5}$$

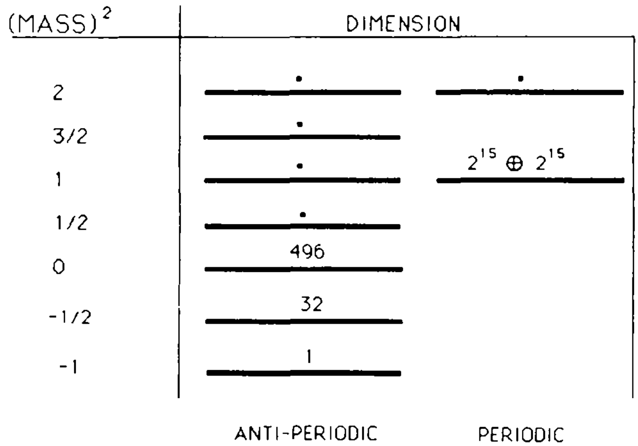

Fig. 2. Spectrum of states for the group Spin(32)/Z_2.

(n = integer or half-integer)

$$[T_1 \cdot M_m, T_2 \cdot M_n] = [T_1, T_2] \cdot M_{n+m} + \tfrac{1}{2} n \operatorname{Tr}(T_1 T_2) \delta_{m+n,0}, \tag{3.6}$$

where the T_{ij} are 32×32 antisymmetric matrices and $T \cdot M_n = T^{ij} M_n^{ij}$. Consequently the two fermion Fock space representations are each reducible, splitting into representations of the affine Lie algebra with even or odd "fermion number". (These correspond to the four conjugacy classes of SO(32) representations, whose characteristic members are singlet, spinor, spinor′ and vector, respectively.)

For the purpose of constructing the heterotic string it is necessary to exploit this reducibility by projecting onto the subspace of states which in the antiperiodic sector contain an even number of fermions (i.e. the integer mass levels), and in the periodic sector satisfy a Weyl projection with $\frac{1}{2}(1+\bar{\gamma})$ where $\bar{\gamma} = \psi_0^1 \psi_0^2 \cdots \psi_0^{32}$ on their SO(32) spinor index. The necessity of this truncation of the Hilbert space stems from the fact that the right-moving half of the theory contains only integer mass levels, whereas the unreduced spectrum of left movers has half-integer modes. Such states could never satisfy the constraint $N = \tilde{N}$, which is necessary to have a consistent spectrum, and so the half-integer levels must be removed. Since these correspond to the vector ($\underline{32}$) conjugacy class of SO(32) we must, for consistency, also discard half of the states in the periodic sector, since [spinor]×[spinor] = [vector].

The resulting theory does not contain any states of the vector or spinor types. Consequently the gauge group is Spin(32)/Z_2, and the oscillators (ψ_n^i) used to construct this group are not quite a representation of this group. This circumstance is familiar from the construction of superstring; although the unreduced Neveu-Schwarz and Ramond models were themselves consistent (at least at tree level), a truncation of this sort was necessary in order to find a supersymmetry generator which exchanged bosonic (Neveu-Schwarz) and fermionic (Ramond) sectors.

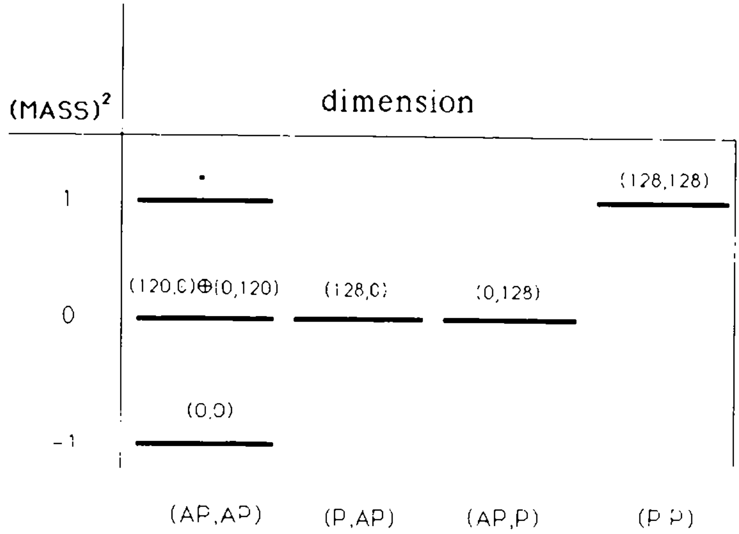

Fig. 3. Spectrum of states for the group $E_8 \times E_8$.

We now turn to the fermionic description of the $E_8 \times E_8$ heterotic string. Remarkably, it requires only a simple modification of the procedure used to obtain the $\mathrm{Spin}(32)/Z_2$ theory, although demonstrating that the correct model is obtained in this way is rather subtle. We proceed by defining two sets of 16 fermionic coordinates ψ^i and $\tilde{\psi}^i$ ($i = 1 \cdots 16$). These are quantized as before, with the boundary conditions and projections being imposed separately on each set of fermionic coordinates. We now have four sectors, depending on which boundary conditions are applied to ψ^i or $\tilde{\psi}^i$. We keep only those states with an even number of ψ^i excitations, independent of the boundary conditions applied to $\tilde{\psi}$ and of the $\tilde{\psi}$ internal quantum numbers. We then obtain states which are explicit representations of $SO(16) \times SO(16)$, which are restricted in the same way that we restricted the $SO(32)$ states.

The mass spectrum of the theory can be calculated if we include the correct normal-ordering constant. In the two sectors whose boundary conditions match we obtain the same result as before; in the two sectors where one set of fermions has antiperiodic boundary conditions and the other has periodic boundary conditions the normal-ordering constant is zero. Then the low-lying levels have the mass spectrum and $SO(16) \times SO(16)$ quantum numbers depicted in fig. 3 (again ignoring spatial excitations).

The massless states in the (AP, AP) sector are in the adjoint representation of $SO(16) \times SO(16)$, as is necessary in order for the heterotic string to contain Yang–Mills interactions. However, there are also massless states in the $\underline{128}$ spinor representation of each $SO(16)$. Since these also become gauge bosons (and gauginos) when we attach the right-moving modes, the theory must contain a gauge group larger than $SO(16) \times SO(16)$. The resolution of this problem lies in the fact that E_8

contains SO(16) as a maximal subgroup, with the adjoint of E_8 decomposing into SO(16) representations as $\underline{120}+\underline{128}$, and where the E_8 commutation relations are represented as $(i, j, \ldots = 1 \cdots 16)$

$$
\begin{aligned}
[M_{ij}, M_{kl}] &= (\delta_{ik} M_{jl} + \text{perms.})\,, \\
[M_{ij}, S^a] &= \tfrac{1}{4}(\Sigma^{ij})_{ab} S^b\,, \\
[S^a, S^b] &= (\Sigma^{ij})_{ab} M^{ij}\,,
\end{aligned}
\tag{3.7}
$$

where $\Sigma^{ij} = \frac{1}{2}i\,[\Gamma^i, \Gamma^j]$ and the Γ^i are gamma matrices for SO(16). We therefore conjecture that our construction yields a representation of the affine Lie algebra based on $E_8 \times E_8$, in which the SO(16) subgroup of E_8 is realized linearly on the operators M^{ij}; the remaining operators S^a are realized in a nonlinear fashion. These generators, however, must interchange the Fock spaces corresponding to the two kinds of boundary conditions in order to realize the algebra of E_8, and no simple expression involving the ψ^i could accomplish this. An analogous problem arises in the construction of the supersymmetry operator in the Neveu-Schwarz-Ramond realization of the superstring, which also is required to exchange the bosonic and fermionic Fock spaces. In that case, unfortunately, the supersymmetry operator did not have a simple expression in terms of oscillators of either sector. The superstring solution was to adopt a representation of the string coordinates for which the supersymmetry generators have a natural action. The analogous procedure in our case would be to abandon the fermionic coordinates altogether and use the algebra generated by $\{M_n^{ij}, S_n^a\}$ - the affine Lie algebra $\hat{E}_8$. However, the representations of this algebra are difficult to construct directly, so much so that mathematicians were led to use dual models as a means of constructing them [7, 23]. For practical purposes it is easier to calculate using the fermionic coordinates, wherein the SO(16) invariance is manifest and can be extended by gauge invariance to the full E_8 invariant expression.

The question remains whether we are sure that this construction actually produces a representation of the affine Lie algebra $(\hat{E}_8 \times \hat{E}_8)$. We are, but we will only present an outline of the argument here; for a more rigorous assurance we refer the reader to ref. [23] where it is shown that the representation of $\hat{E}_8$ which is of interest to us can be decomposed into $\widehat{\mathrm{SO}}(16)$ representations in precisely the combination found in our model, where the "basic" representation of $\hat{E}_8$ is the sum of the "basic" representation of $\widehat{\mathrm{SO}}(16)$ and the "spinor" representation of $\widehat{\mathrm{SO}}(16)$. This decomposition is based on an automorphism of the $\hat{E}_8$ Lie algebra, which in the notation of (3.7) leaves M_{ij} unchanged and takes S^a into $-S^a$. This automorphism can be extended trivially to the affine Lie algebra $\hat{E}_8$, and the commutation relations are unchanged. The representations of $\hat{E}_8$ can be decomposed into representations of the subalgebra of $\hat{E}_8$ which commutes with the automorphism, namely, the generators of $\widehat{\mathrm{SO}}(16)$; the states which are even under the automorphism are created by an even number of S_m^a, and those odd under the automorphism are created by an odd number of S^a_m. It is straightforward to show that the low-lying states of the $\hat{\mathrm{E}}_8$ representation, when decomposed this way, correspond to those of the $\widehat{\mathrm{SO}}(16)$ representations we constructed. There is also the circumstantial evidence that the number of states at each level of the $\hat{\mathrm{E}}_8$ "basic" representation match the number of states in the $\widehat{\mathrm{SO}}(16)$ construction, as seen by comparing the partition function $\mathrm{ch(G)} = \mathrm{Tr}\, q^{\tilde{N}_f}$ for the two cases (ch(G) is called the character function by the authors of ref. [7])

$$\mathrm{ch}\,(\hat{\mathrm{E}}_{8(\mathrm{basic})}) = \frac{\left(1 + 240\sum_{n=1}^{\infty} \sigma_3(n) q^n\right)}{\prod_{n=1}^{\infty} (1-q^m)^8}$$
$$= \mathrm{ch}\,(\widehat{\mathrm{SO}}(16)) + \mathrm{ch}\,(\widehat{\mathrm{SO}}(16)_{(1/2)})$$
$$= \frac{1}{2}\left(\prod_{n=1}^{\infty} (1-q^{n-1/2})^{16} + \prod_{n=1}^{\infty} (1+q^{n-1/2})^{16}\right) + 128 \prod_{n=1}^{\infty} (1+q^n)^{16}, \qquad (3.8)$$

where $\sigma_3(n)$ is the sum of the third powers of the divisors of n (a nontrivial identity, whose proof is based on the theory of modular forms [16, 24]). We conclude that our fermionic construction is indeed a representation of $\hat{\mathrm{E}}_8 \times \hat{\mathrm{E}}_8$; the generators of the subalgebra $\mathrm{SO}(16) \times \mathrm{SO}(16)$ are realized linearly in this construction, but the other generators unfortunately do not have a simple or useful realization in terms of the fermionic coordinates.

It is helpful to examine this result in relation to what we found in the bosonic realization. In that case, only the generators in the Cartan subalgebra were realized linearly, although the nonlinearly-realized operators also have a useful form. This is consistent with our expectations of the standard bosonization procedure, which gives linear realizations for the bosonic fields of only the flavor-diagonal fermionic currents. In this respect, both of the representations could be improved by turning to the third realization mentioned in sect. 4, which corresponds to a nonabelian bosonization of the fermionic theory; however this third realization is not practical for the kind of calculations currently contemplated.

From this point on the construction of the heterotic string proceeds as before, except that the internal bosonic operators must be replaced with the corresponding fermionic operators.

4. Covariant quantization

Thus far we have exhibited an apparently consistent light-cone quantization of the heterotic string. However, the light-cone formalism receives its justification as a geometric theory of surfaces by virtue of its being the gauge restriction of a reparametrization-invariant theory. Maintaining reparametrization invariance in the quantum theory is an essential requirement for the consistency of the theory; it is responsible for the critical dimensions 10 and 26. In this section we will describe both the path integral and operator formalisms for covariant quantization; the key point is that a proper BRS quantization shows that all potential anomalies cancel between the Faddeev-Popov ghosts and the string fields.

The path integral approach begins with a covariant action for the heterotic string. In principle, we would like an action which is manifestly group and supersymmetry invariant. The appropriate covariant framework for the $E_8 \times E_8$ and $\mathrm{Spin}(32)/Z_2$ currents is a nonlinear sigma model on the group manifold, with a Wess-Zumino term [25]. Of course the sigma model fields have to be constrained to be left moving. Similarly, a covariant action for the superstring exists [26] which may be thought of as a nonlinear sigma model on superspace (the super-translation group manifold) with a WZ term [27]. An arbitrary element of the super-translation group may be written as $h = \exp\{i[X \cdot P + \theta \cdot Q]\}$, where X_μ ($\mu = 0 \cdots 9$) and θ^a (a ten-dimensional Majorana-Weyl spinor) are the superspace coordinates and P and Q generate translations and supersymmetry transformations, respectively. In terms of the derivatives

$$h^{-1}(x)\partial_\alpha h(x) = (\partial_\alpha X^\mu - i\bar{\theta}\gamma^\mu \partial_\alpha \theta)P_\mu + \partial_\alpha \theta \cdot Q\,, \tag{4.1}$$

the covariant superstring action may be written

$$S_{\text{string}} = \int d^2\xi\, e\tfrac{1}{2}\,\mathrm{tr}\,[(h^{-1}e^\alpha_a \partial_\alpha h)(h^{-1}e^{a\beta}\partial_\beta h)] + \Gamma_h\,. \tag{4.2}$$

The e^α_a define a local tangent space frame in terms of the coordinates ξ^α which parametrize the world sheet ($e = \det e^a_\alpha$). We choose the arbitrary normalization of the generators $\mathrm{tr}\,(P_\mu P_\nu) = \eta_{\mu\nu}$; also $\mathrm{tr}\,(Q_a Q_b) = 0$. Γ_h is the Wess-Zumino action [25] which cannot be written as a local functional of the fields, but can be constructed in terms of the derivatives (4.1) as a three-dimensional integral when the world surface has spherical topology

$$\Gamma_h = \int_{M, \partial M = S^2} d^3\xi\, \varepsilon^{\alpha\beta\gamma}\, \mathrm{tr}\,[(h^{-1}\partial_\alpha h)(h^{-1}\partial_\beta h)(h^{-1}\partial_\gamma h)\,. \tag{4.3}$$

Γ_h is nonvanishing due to the relation $\mathrm{tr}\,[Q_a Q_b P^\mu] = (\gamma^\mu C^{-1})_{ab}$.

In the case of the heterotic string we want (4.2) to describe the dynamics of the "right" movers, where "right" is defined in terms of the local frame. We therefore add to the action the term

$$S_{\text{constraint}} = \int d^2\xi\, e\lambda^{++}(h^{-1}e^\alpha_+\partial_\alpha h)(h^{-1}e^\beta_+\partial_\beta h) \tag{4.4}$$

(here $\xi^\pm = \tau \pm \sigma$ are two-dimensional light-cone coordinates). This indeed enforces the constraint that the superstring coordinates are only right moving. Similarly, to describe the left-moving coordinates of the heterotic string, we introduce $g = \exp\{i[X \cdot P + \phi^a T^a]\}$ where T^a are the generators of $G = E_8 \times E_8$ or $\mathrm{Spin}(32)/Z_2$ and ϕ^a are the local coordinates on G. The dynamics of the left movers is then described by the same action (4.2), together with the constraint (4.4), but now with all +'s replaced by −'s. This ensures that g is only a function of left-moving coordinates (in the local frame).

However, the covariant superspace action is difficult to quantize due to the presence of certain second-class constraints [26], and the sigma-model approach is unwieldy for practical calculations. The heterotic string is no worse in this regard than the ordinary superstring, with the possible exception of the question of reconciling the local fermionic symmetry of (4.2) [26] with the new constraints (4.4).

Thus although it would be nice to have a manifestly reparametrization invariant, Lorentz invariant, group invariant, and supersymmetric formalism, such an approach is not yet practical. Two options are then available: first, one can give up manifest Lorentz and reparametrization invariance and construct the theory on the light cone as we have done in sects. 2 and 3. Here supersymmetry is manifest, but full Lorentz invariance is not and must be checked. The light-cone approach is the one most suitable for simple calculations, and we will employ it in the following paper on interactions [9]. The second alternative is to give up manifest supersymmetry while retaining Lorentz and reparametrization invariance; it is this approach that we will now discuss.

We have seen that the group theory may be represented in terms of left-moving, Lorentz scalar, bosons or fermions. Their action may be written in a manifestly covariant form as

$$S_{\mathrm{G,fermion}} = \int \mathrm{d}^2\xi\, e[\tfrac{1}{2} i\psi^I e_+^\alpha \partial_\alpha \psi^I]\,,$$

or

$$S_{\mathrm{G,boson}} = \int \mathrm{d}^2\xi\, e[\tfrac{1}{2}(e_a^\alpha \partial_\alpha X^I)^2 + \lambda^{--}(e_-^\alpha \partial_\alpha X^I)^2]\,. \tag{4.5}$$

Here $\psi^I (I = 1 \cdots 32)$ are the internal fermionic coordinates, $X^I (I = 1 \cdots 16)$ are the internal boson coordinates, and λ^{--} is a Lagrange multiplier which, as in (4.4), eliminates the right-moving bosons [13].

The ordinary superstring has a similar covariant formulation in terms of the Neveu-Schwarz-Ramond model. Here one describes the fermions by fields in the vector representation of O(9, 1). This is precisely analogous to the representation of the heterotic internal degrees of freedom given in sect. 3, where (say for G = Spin(32)/Z_2) one employs fermion fields ψ^I, in the vector representation of O(32). As we have explained, the states of the system will depend on the choice of boundary conditions. Periodic (antiperiodic) boundary conditions on the fermion fields will yield states which are O(9, 1) spinors (tensors). Space-time supersymmetry interchanges the two sectors and will not be manifest. In the case of the heterotic string the fermionic fields are further constrained to be right-moving, and thus the total space-time action is [28]

$$S_{\text{space-time}} = \int d^2\xi\, e[\tfrac{1}{2}(e_a^\alpha \partial_\alpha X^\mu)^2 - \tfrac{1}{2}i\psi^{\mu\rho^-} e_-^\alpha \partial_\alpha \psi^\mu + \tfrac{1}{2}i(\zeta_\alpha \rho^\beta \rho^\alpha \psi^\mu)\partial_\beta X^\mu]\,. \quad (4.6)$$

This action involves the space-time coordinates X^μ, O(9, 1) fermions ψ^μ, the zweibein e_a^α and its right-handed gravitino partner ζ_α (ρ^α are two-dimensional γ-matrices). In addition to local reparametrization invariance, this action has an "$N = \frac{1}{2}$" local supersymmetry

$$\begin{aligned} \delta e_-^\alpha &= i\varepsilon\rho_-\xi^\alpha\,, \\ \delta\zeta_\alpha &= -2\nabla_\alpha\varepsilon\,, \\ \delta X^\mu &= i\varepsilon\psi^\mu\,, \\ \delta\psi^\mu &= [\partial_\alpha X^\mu + \tfrac{1}{2}i(\zeta_\alpha\psi^\mu)]\rho^\alpha\varepsilon\,, \end{aligned} \quad (4.7)$$

with a right-handed spinor parameter that satisfies $\rho^-\varepsilon = \nabla_+\varepsilon = 0$.

Quantization is most easily performed in the superconformal gauge

$$e_a^\alpha = e^\phi \delta_a^\alpha\,, \qquad \zeta^\alpha = \rho^\alpha\eta\,. \quad (4.8)$$

With this choice the total heterotic string action, (4.5) + (4.6), describes a set of free fields (we choose the fermionic form of S_G for simplicity):

$$S_{\text{heterotic}} = \int d^2\xi[\tfrac{1}{2}(\partial_\alpha X^\mu)^2 + \tfrac{1}{2}i\psi^\mu\rho^-\partial_-\psi^\mu + \tfrac{1}{2}i\psi^I\rho^+\partial_+\psi^I] \quad (4.9)$$

since the conformal degrees of freedom ϕ and η decouple classically. This action still has a residual invariance under conformal reparametrizations (and their right-handed two-dimensional superpartners)

$$\xi^\pm \to f(\xi^\pm)\,, \quad (4.10)$$

which preserve the gauge (4.8). These transformations are generated by the components of the two-dimensional energy-momentum tensor and superconformal current

$$\begin{aligned} T_{++} &= (\partial_+ X^\mu)^2 + i\psi^\mu\rho_+\partial_+\psi^\mu\,, \\ T_{--} &= (\partial_- X^\mu)^2 + i\psi^I\rho_-\partial_- X^I\,, \\ S_+ &= i(\partial_+ X^\mu)\psi^\mu\,. \end{aligned} \quad (4.11)$$

The (super)conformal symmetries allow us to classically specify the further gauge choices

$$X^+ = x^+ + p^+\tau\,, \qquad \psi^+ = 0\,. \quad (4.12)$$

This procedure can be justified quantum mechanically in a hamiltonian framework [29]. The constraints $T_{++} = T_{--} = S_+ = 0$ enable us to solve for X^- and ψ^- in terms of the physical degrees of freedom X^i, ψ^i $(i=1\cdots8)$, ψ^I $(I=1\cdots32)$; this leaves only global τ and σ translations, generated by

$$H=\tfrac{1}{2}\int(\dot{X}^2+X'^2+i\psi^i\cdot\psi^{i\prime}+i\psi^I\psi^{I\prime})\,\mathrm{d}^2\xi\,,$$

$$P=\int(\dot{X}\cdot X'+i\psi^i\cdot\psi^{i\prime}-i\psi^I\psi^{I\prime})\,\mathrm{d}^2\xi\,, \tag{4.13}$$

as residual invariances. Finally we arrive at the light-cone gauge action

$$S=\int\mathrm{d}^2\xi[\tfrac{1}{2}(\partial_\alpha X^i)^2+\tfrac{1}{2}i\psi^i\rho\;\partial\;\psi^i+\tfrac{1}{2}i\psi^I\rho^+\partial_+\psi^I]\,. \tag{4.14}$$

Eqs. (4.13) and (4.14) are just the Neveu-Schwarz-Ramond version of the light-cone gauge heterotic string action (2.1) and constraints (2.17) and (2.19). Thus we see that one may derive the light-cone heterotic string from a manifestly reparametrization invariant action. Checking that Lorentz invariance is maintained upon quantization is the easiest way to ensure that no anomalies of reparametrization invariance have appeared.

At this point we are more interested in maintaining manifest Lorentz invariance as well as checking two-dimensional coordinate invariance explicitly. Hence we will not choose the light-cone gauge (4.12), but rather explore the quantum mechanical consequences of (4.8). In the path integral over the action, factoring out the gauge group volume via the choice (4.8) involves computing a jacobian measure factor from the fluctuations of the metric induced by infinitesimal reparametrizations

$$\delta g_{\alpha\beta}=\nabla_{(\alpha}\delta\xi_{\beta)}\,. \tag{4.15}$$

This jacobian, as well as the corresponding factor arising from fixing the $N=\frac{1}{2}$ supersymmetry, may be represented via standard manipulations [30] as a path integral over ghost fields. The ghost action is

$$S_{\text{ghost}}=\int\mathrm{d}^2\xi[b^{++}\nabla_+c_++b^{--}\nabla_-c_-+\beta\;\nabla_-\gamma]\,, \tag{4.16}$$

where b, c are the reparametrization ghosts and β, γ are the local supersymmetry ghosts.

Evaluating the path integral in the gauge (4.8) reduces to calculating a series of functional determinants of differential operators of various conformal spins. The conformal spin of a quantity is defined by the (possibly half-integral) phase $e^{in\theta}$ it acquires under a two-dimensional rotation by angle θ. The covariant derivative $\nabla_n=e^{-(n-1)\phi}\partial_+e^{n\phi}$ maps the tensor space of spin n into the tensor space of spin $(n+1)$. The effective action consists of the combination

$$\ln S_{\text{eff}}=\det{}^{-10/2}[\Delta_0=\nabla_0^+\nabla_0+\nabla_0\nabla_0^+]\det{}^{10}(\nabla_{1/2}^+)\det{}^{32}(\nabla_{-1/2})$$
$$\times\det(\nabla_{3/2}^+)\det(\nabla_1)\det(\nabla^+{}_1)\,. \tag{4.17}$$

Because of the chiral asymmetry of this expression there is an ambiguity in the phases of each of these determinants (except Δ_0). The last two are hermitian conjugates and therefore do not contribute to the phase anomaly. The remaining spin-$\frac{1}{2}$ and spin-$\frac{3}{2}$ anomalies have been computed by Alvarez-Gaumé and Witten [3] with the result that $22(=32-10)$ spin-$\frac{1}{2}$ determinants precisely cancel the anomaly of a spin-$\frac{3}{2}$ determinant*. Note that if we had worked with the bosonic representation of G, we could calculate the anomaly of a left-moving boson by regarding it as an antisymmetric 0-index tensor. One finds as expected that one such boson has the same gravitational anomaly as two Majorana-Weyl fermions. Thus we find that the heterotic string is anomaly free.

In addition to the potential phase anomaly, there is also a possibility of an anomaly in the modulus of the expression (4.17) since conformal invariance is an anomalous symmetry. The modulus may be regulated in any number of ways [31] and the anomaly again cancels in the critical dimension, 10 for the right movers and 26 for the left movers.

One may also study the anomaly cancellation in an operator approach [30]. There, analytic reparametrizations are generated by the stress-energy tensor T_{++}, anti-analytic ones by T_{--}; conformal invariance is expressed by the tracelessness condition $T_a{}^a=0$. The Fourier components

$$L_n = \int \mathrm{d}\xi\, \mathrm{e}^{in\xi} T_{++}(\xi)\,,$$

$$\tilde{L}_n = \int \mathrm{d}\xi\, \mathrm{e}^{in\xi} T_{--}(\xi)\,, \tag{4.18}$$

generate the algebra of reparametrizations (the Virasoro algebra)

$$[L_n, L_m] = (n-m)L_{n+m} + \tfrac{1}{12}c(n^3-n)\delta_{n,-m}\,,$$

$$[\tilde{L}_n, \tilde{L}_m] = (n-m)\tilde{L}_{n+m} + \tfrac{1}{12}\tilde{c}(n^3-n)\delta_{n,-m}\,, \tag{4.19}$$

which has, in general, a c-number anomaly given by the second term. When the stress tensor includes all the relevant string and ghost field contributions, one finds $c=\tilde{c}=0$. This is the operator equivalent of the anomaly cancellation found in the path integral approach.

It is this striking cancellation of anomalies that results in string theories being the only known examples of quantum field theories embodying quantum-mechanical reparametrization invariance; it is this geometric fact that accounts for their remarkable properties.

* We should note that the number 23 in sect. 12 of [3] has had the contribution of a single fermion subtracted.

5. Conclusions

In this paper we have constructed a new type of quantum string theory, the heterotic string, an amalgamation of the old fermionic and bosonic strings. We have shown that, at the noninteracting level, the heterotic string theory satisfies all the usual consistency requirements. Thus the formal reparametrization invariance, embodied in the geometrical formulation of the action, suffers no anomalies upon first quantization. Equivalently the light-cone gauge (unitary) formalism preserves the Lorentz invariance and ($N=1$) supersymmetry of the theory. The spectrum of the free heterotic string is ghost and tachyon free, with a massless ground state consisting of $D=10$, $N=1$, supergravity and super-Yang-Mills gauge bosons and their fermionic partners. In the following paper [9] we shall show that interactions can be consistently introduced and that the theory is one-loop finite. It is certainly true that the formulation of the heterotic string appears somewhat awkward and contrived. This is not a shortcoming of our theory; rather it is indicative of the present level of understanding of all quantum string theories, which leaves much to be desired. Many of the most remarkable features of these theories emerge without a full comprehension of their origin. Most mysterious are the general coordinate invariance and local gauge symmetries of string theories, whose appearance lacks a geometrical explanation. This suggests that there exists a more profound formulation of string theory, in which these features would be manifest. In such a formulation the heterotic theory might appear more natural. It certainly occupies a natural niche between the 26-dimensional closed bosonic string theory and the 10-dimensional, closed fermionic string theory.

The special feature of the heterotic string theory is the appearance of a (almost) uniquely determined local gauge symmetry group $G = E_8 \times E_8$ or $\mathrm{Spin}(32)/Z_2$. Heterotic strings are closed, so that there is no place to attach charges as one does in the case of open strings. Instead the gauge group arises via a Kaluza-Klein-type mechanism stemming from the compactification of 16 of the internal coordinates of the string. These are required, by consistency of the theory, to lie on a particular 16-dimensional torus. This compactification produces 16 gauge bosons associated with the isometries of the torus, but in addition it yields 480 massless solitons which wind around the torus. Together they fill out the adjoint representation of G, and all the physical states of the heterotic string form unitary representations of G. Nowhere do we require anomaly cancellation. Instead the gauge groups $E_8 \times E_8$ and $\mathrm{Spin}(32)/Z_2$ arise from other considerations. In fact we can now understand why $D=10$, $N=1$ supergravity is anomaly free only if one chooses these gauge groups. These anomaly-free field theories are simply the low-energy limit of the heterotic string (as we shall see explicitly in the following paper [9]).

Among the ill-understood features of quantum string theories is their high degree of uniqueness. The heterotic string provides us with two more entries (depending on whether $G = E_8 \times E_8$ or $\mathrm{Spin}(32)/Z_2$) to an extremely short list of consistent string theories. Are there any more? We do not know the complete answer to this question. However, the only new theories that we know how to construct, using the ideas that underlie the heterotic theory, are two-dimensional. Since the heterotic strings are formed by combining left and right movers of separately consistent closed string theories, the only additional theories of this type that we know are based on the $D=2$ fermionic string theory of ref. [32]. The right-moving sector of this theory could be combined with the left-moving sectors of the $D=10$ fermionic string or the $D=26$ bosonic string to yield new two-dimensional string theories. In the latter case we would again compactify the 24 internal dimensions of the string on a torus based on a 24-dimensional self-dual lattice, of which there are exactly 24 types (including the famous Leech lattice). It would be very amusing to construct these theories.

From a practical point of view the heterotic string, particularly in its $E_8 \times E_8$ manifestation, provides a very promising candidate for a unified theory of gravity and matter. Recent investigations have shown that this theory possesses features which may solve many long-standing problems in particle physics. Among these are: compactifications with zero cosmological constant of 6 of the 10 space-time dimensions to a four-dimensional theory with an unbroken $E_8 \times E_8'$ symmetry and some number of chiral fermion families of E_6 [33], a natural mechanism for breaking E_6 down to an acceptable low-energy gauge group [33], a mechanism for splitting weak Higgs doublets from color Higgs triplets [34], and a mechanism for supersymmetry breaking with zero cosmological constant in the "hidden" E_8' sector [35]. Although much work remains to be done there seem to be no insuperable obstacles to deriving all of known physics from the $E_8 \times E_8$ heterotic string.

References

[1] E. Witten, 1983 Shelter Island Conf. (MIT, 1984) to appear
[2] M.B. Green and J.H. Schwarz, Nucl. Phys. B181 (1981) 502; B198 (1982) 252, 441; Phys. Lett. 109B (1982) 444;
M.B. Green, J.H. Schwarz and L. Brink, Nucl. Phys. B198 (1982) 474
[3] L. Alvarez-Gaumé and E. Witten, Nucl. Phys. B234 (1983) 269
[4] M.B. Green and J.H. Schwarz, CALT-68-1182, CALT-68-1194
[5] J. Paton and H.M. Chan, Nucl. Phys. B10 (1969) 519
[6] J.H. Schwarz, Proc. Johns Hopkins Workshop, Florence (1982);
N. Marcus and A. Sagnotti, Phys. Lett. 119B (1982) 97
[7] I.B. Frenkel and V.G. Kac, Inv. Math. 62 (1980) 23;
G. Segal, Comm. Math. Phys. 80 (1982) 301;
I.B. Frenkel, Lectures in Applied Math, Vol. 21 (1985) AMS.
P. Goddard and D. Olive, Proc. Conf. on vertex operators in mathematics and physics, Publications of the Mathematical Sciences Research Institute no. 3 (Springer, Berlin, 1984).
[8] P.G.O. Freund, ITP preprint (1984)
[9] D.J. Gross, J.A. Harvey, E. Martinec and R. Rohm, Heterotic string theory. (II). The interacting heterotic string, Princeton preprint (1985).
[10] A.M. Polyakov, Phys. Lett. 103B (1981) 207

[11] P. Goddard, J. Goldstone, C. Rebbi and C. Thorn, Nucl. Phys. B56 (1973) 109
[12] J.H. Schwarz, Phys. Reports 89 (1982) 223;
M.B. Green, Surveys in high energy physics 3 (1983) 127;
L. Brink, Superstrings, CERN preprint-H4006 (1984)
[13] W. Siegel, Nucl. Phys. B238 (1984) 307
[14] L. Alvarez-Gaumé and D.Z. Freedman, Phys. Rev. D22 (1980) 846;
Comm. Math. Phys. 80 (1981) 282
[15] E. Cremmer and J. Scherk, Nucl. Phys. B103 (1976) 399
[16] J.P. Serre, A course in arithmetic (Springer, New York, 1973)
[17] K. Huan and S. Weinberg, Phys. Rev. Lett. 25 (1970) 895
[18] S.G. Rajeev, Phys. Rev. D29 (1984) 2944
[19] V. Bargmann, Ann. Math. 59 (1954) 1
[20] K. Bardakci and M.B. Halpern, Phys. Rev. D3 (1971) 2493
[21] A. Neveu and J.H. Schwarz, Nucl. Phys. B31 (1971) 86; Phys. Rev. D4 (1971) 1109
[22] P. Ramond, Phys. Rev. D3 (1971) 2415
[23] J. Leponsky and R.L. Wilson, Comm. Math. Phys. 62 (1978), 43;
V. Kac, D. Kazhdau, J. Lepowsky and R.L. Wilson, Adv. in Math. 42 (1981) 83;
I.B. Frenkel, J. Lepowsky and A. Meurman, An E_8 approach to F_1, Proc. 1982 Montreal Conf. on Finite group theory, ed. J. McKay, Springer Lecture Notes in Mathematics (1984); A moonshine module for the monster, Proc. Conf. on Vertex operators in mathematics and physics, Publications of the Mathematical Sciences Research Institute no. 3 (Springer, Berlin, 1984)
[24] A. Erdelyi et al., Higher transcendental functions, vol. 2 (McGraw-Hill, New York, 1953) ch. 13
[25] E. Witten, Comm. Math. Phys. 92 (1984) 451
[26] M. Green and J. Schwarz, Nucl. Phys. B243 (1984) 285
[27] M. Henneaux and L. Mezincescu, Texas preprint UTTG-26-84;
E. Martinec, unpublished
[28] L. Brink, P. Di Vecchia and P. Howe, Phys. Lett. 65B (1976) 471;
S. Deser and B. Zumino, Phys., Lett. 65B (1976) 369
[29] E. Martinec, Phys. Rev. D30 (1984) 1195
[30] D. Friedan, Proc. 1982 Les Houches Summer School, ed. J.B. Zuber and R. Stora (North-Holland, Amsterdam, 1984)
[31] A.M. Polyakov, Phys. Lett. 103B (1981) 207, 211;
P. Di Vecchia, B. Durhuus, P. Olesen and J.L. Peterson, Nucl. Phys. B198 (1982) 157; B207 (1982) 77;
O. Alvarez, Nucl. Phys. B216 (1983) 125;
E. Martinec, Phys. Rev. D28 (1983) 2604;
K. Fujikawa, Phys. Rev. D25 (1982) 2584
[32] M. Ademollo et al., Phys. Lett. 62B (1976) 105; Nucl. Phys. B111 (1976) 77
[33] P. Candelas, G.T. Horowitz, A. Strominger and E. Witten, Princeton preprint (1984), Nucl. Phys. B, to be published
[34] E. Witten, Princeton preprint (1985)
[35] M. Dine, R. Rohm, N. Seiberg and E. Witten, in preparation