PHYSICAL REVIEW D VOLUME 39, NUMBER 12 15 JUNE 1989

Search for four-dimensional string models

David Sénéchal
Newman Laboratory for Nuclear Studies, Cornell University, Ithaca, New York 14853-5001
(Received 22 February 1989)

Within the framework of the spin-structure construction of fermionic string models, we describe a computerized method for generating four-dimensional heterotic models and analyzing their massless spectrum. Thus, starting with the spin structure, we easily identify the low-energy gauge group and all the massless representations, including any U(1) charges. We attempt a systematic and complete generation of all four-dimensional heterotic string models based on world-sheet periodic and antiperiodic complex fermions and compare this with the generation of models from the same category obtained by randomly generating spin structures. We thus explicitly construct over 900 four-dimensional models having $N = 4$ supersymmetry, and over 32 000 models having $N = 1$ supersymmetry. Some phenomenologically interesting examples are presented.

I. INTRODUCTION

In recent years superstring theory[1] has drawn a considerable amount of interest and is a serious candidate for a unified, fundamental theory of nature including gravity and the other known interactions. The heterotic theory of closed strings[2] is especially interesting. Originally, it was only formulated in ten space-time dimensions, and it was known that only a few classical ground states existed. By a classical ground state we mean a ground state of the string field theory in the classical approximation to its effective action. Presumably, the (not-yet-formulated) correct string field theory has a unique ground state, and it was assumed that, in that state, six of the ten space-time dimensions would be compactified. However, the notion of compactification is not essential, because all that is required from a classical ground state of the string is that it still display the symmetry of the original first-quantized string action, i.e., conformal invariance. In general then, a classical ground state of the heterotic string in four space-time dimensions will be described by left-moving string coordinates associated with a $c = 22$ two-dimensional conformal field theory, and by right-moving string coordinates and transverse fermions associated with a $c = 9$ two-dimensional superconformal field theory.

Recently, conformal field theories involving parafermions have been investigated,[3,4] but it remains simpler to consider conformal field theories built out of bosons ($c = 1$) and real fermions ($c = \frac{1}{2}$). In this case many classes of solutions (i.e., classical ground states) have been found, some of them overlapping. A class of solutions was found by imposing conditions on the momentum lattice of the world-sheet bosons,[5–7] or equivalently the charge lattice of world-sheet complex fermions. Another large class of solutions was found in the related *orbifold* construction.[8–10] Alternatively, imposing constraints on the boundary conditions, i.e., the spin structure, of world-sheet fermions as they go around noncontractible loops led to the discovery of still another large class of solutions.[11,12] More recently, the existence of additional solutions was shown by investigating right-moving supercurrent built from bosonic fields.[13,14] Needless to say, the number of classical ground states of the heterotic string is very large, and all of these are equally worthy of study, short of a comprehensive string field theory. However, not all classes of string ground states can be analyzed with the same ease; therefore, if we wish to conduct a search for an attractive string model on which to base more detailed analysis, the facility with which we can examine these ground states is a major factor.

For this reason the spin-structure construction of Kawai, Lewellen, and Tye is very useful, as a computerized search may be set up relatively easily. The hope is to generate and classify all classical string ground states reached by using free periodic or antiperiodic complex world-sheet fermions, entertaining the idea that these form a representative fraction of all string ground states. It is interesting to have an estimate for the fraction of that restricted set of ground states which is phenomenologically close to our low-energy universe, and to find a few viable models as well as a few generic properties of that class of classical ground states. We have devised a computerized method for rapidly generating consistent spin structures of the world-sheet fermions, and for obtaining the massless spectrum of the so-specified string ground states. Our method relies mainly on the random generation of spin structures, but is shown to be fairly complete by comparison with a more systematic generation. Ground states having $N = 4$ supersymmetry are relatively few, and over 900 of them have been explicitly constructed using this method. There are considerably more ground states with $N = 1$ supersymmetry at the Planck scale. Over 30 000 of them have been explicitly constructed. Conditions have been imposed on the massless spectrum corresponding to a few phenomenological low-energy scenarios, and string models (as we loosely refer to classical ground states of the string) obeying these conditions have been obtained. The scenarios envisaged include a direct embedding of the $SU(3)\times SU(2)\times U(1)$ standard model in the string massless spectrum, as well as embeddings of larger gauge theories, such as the $SU(4)\times SU(2)_L\times SU(2)_R$ theory of Pati and Salam and the SU(5) grand unified theory, with flipped quark and lepton assignment, or without. Of course not only the massless spectrum but the whole spectrum is given by the spin-structure construction. Moreover, the interactions between different string states are determined, and the perturbation expansion is shown to be consistently defined. We hope to describe a similar search for nonsupersymmetric models in a later publication, as well as a deeper analysis of particular examples obtained through this work.

This paper is divided as follows. In Sec. II we present a brief review of the spin-structure construction from a practical point of view; the goal is to describe a set of rules constraining the spin structure of the world-sheet fermions so as to make the theory modular invariant. In Sec. III we describe the general pattern of the massless spectrum coming from this construction. In Sec. IV we describe the general features of the computing algorithms attempting this model generation and analysis. Section V deals with conditions imposed on the massless fermion spectrum of phenomenologically interesting models. Finally, Sec. VI gives some examples.

II. A REVIEW OF THE SPIN-STRUCTURE METHOD

If we assume the *compactified* degrees of freedom of the string to be free fermions, then we must specify the boundary conditions of these fermions around noncontractible loops on the world sheet (the spin structure). Furthermore, if we want the theory to be modular invariant, i.e., invariant under global diffeomorphisms not connected to the identity, we must include several different spin structures in the theory, since modular transformations change spin structures into each other. A set of easily implemented rules has been devised in Ref. 11 to ensure that all the spin structures necessary to make the one-loop partition function modular invariant are included, and that a proper Gliozzi-Scherk-Olive (GSO) projection[15] is realized, thereby ensuring higher-loop modular invariance. In this section we summarize these results as applied to four-dimensional heterotic string models.

In the light-cone gauge we take the degrees of freedom of the four-dimensional heterotic string to be 44 left-moving and 20 right-moving Majorana fermions, in addition to the bosonic string coordinates. These numbers are fixed by the requirement that the conformal anomaly of the two-dimensional world-sheet field theory vanish, or alternatively that the one-loop partition function be modular invariant. It is believed that if the latter is modular invariant and can be interpreted as a trace over fermionic and bosonic states (the requirement for a proper GSO projection), then the multiloop amplitudes are modular invariant as well; therefore we may concentrate on the one-loop partition function. If we then parametrize the world-sheet torus by $\sigma_1+\sigma_2\tau$, τ being its modular parameter, the fermions must be periodic or antiperiodic as $\sigma_i\mapsto\sigma_i+1$. Moreover, world-sheet supersymmetry requires that the two right-moving transverse fermions (i.e., the world-sheet fermions carrying space-time indices) have the same periodicity (or spin structure), and that the remaining 18 right-moving fermions may be arranged in *triplets*, such that the product of all three members has the same periodicity as the transverse fermions; this ensures that their interaction term with the world-sheet gravitino is periodic, i.e., well defined. The left-moving fermions are not part of any world-sheet supersymmetry and therefore their 44 spin structures are independent. Moreover, one is allowed to perform SO(44) rotations among them as $\sigma_1\mapsto\sigma_1+1$ and $\sigma_2\mapsto\sigma_2+1$. Since the two rotations must commute, we may simultaneously diagonalize them, in terms of complex fermions and real fermions $\psi^l(\sigma_1,\sigma_2)$, resulting in the conditions

$$\psi^l(\sigma_1+1,\sigma_2)=e^{-2\pi iW^l}\psi^l\ ,$$
$$\psi^l(\sigma_1,\sigma_2+1)=e^{-2\pi iW'^l}\psi^l\ . \tag{2.1}$$

In general, combinations of rotations and sign changes cannot be diagonalized by complex fermions or real fermions alone. However, we will concentrate on the case where only complex fermions are needed; this also assumes that the right-moving fermion triplets can be paired to form complex fermion triplets. We then end up with 10 right-moving and 22 left-moving free complex fermions satisfying the above periodicity conditions, where the W^l's and the W'^l's form (10+22)-dimensional spin-structure vectors $\mathbf{W}$ and $\mathbf{W}'$. Given the spin structure we can calculate the fermionic partition function $Z^{\mathbf{W}}_{\mathbf{W}'}(\tau)$ and its change under the modular transformations $\tau\mapsto\tau+1$ and $\tau\mapsto 1/\tau$. Only by taking a linear combination

$$Z_{\rm fer}(\tau)=\sum_{\mathbf{W},\mathbf{W}'}C^{\mathbf{W}}_{\mathbf{W}'}Z^{\mathbf{W}}_{\mathbf{W}'}(\tau)$$

of partition functions from different spin structures can we form a modular-invariant function. In addition, if we require that, keeping $\mathbf{W}$ constant, the sum over included $\mathbf{W}'$'s give a factor of 0, 1, or -1 times $e^{2\pi i\tau(H^L_W-H^R_W)}$, that is, if we require a proper GSO projection, we arrive at the following conditions for the spin structures and the states entering the model.

First, the spin-structure vectors are generated by a *basis* $\mathbf{W}_0,\mathbf{W}_1,\ldots,\mathbf{W}_{n_W}$ as follows. If m_i is the smallest integer such that $m_i\mathbf{W}_i$ has only integer components, then the set of spin structures is given by all possible combinations $\alpha W\doteq\sum\alpha_i\mathbf{W}_i$ with $\alpha_i=0,1,\ldots,m_i-1$.

Here $\stackrel{1}{=}$ means equality modulo 1. The vector

$$\mathbf{W}_0 \equiv ((\tfrac{1}{2})^{10} \mid (\tfrac{1}{2})^{22})$$

must always be part of the basis. To each combination αW corresponds a sector of states with right and left fermionic Hamiltonians $H^R_{\alpha W}, H^L_{\alpha W}$ with vacuum energies

$$E^R_{\alpha W} = \tfrac{1}{2}\sum_{l=1}^{10}[(\alpha W^l)^2 - \alpha W^l + \tfrac{1}{6}] ,$$
$$E^L_{\alpha W} = \tfrac{1}{2}\sum_{l=11}^{32}[(\alpha W^l)^2 - \alpha W^l + \tfrac{1}{6}] . \tag{2.2}$$

If we add to this the contribution from the string coordinates, we have the following expression for the energy levels of the left-moving part of the string:

$$m_L^2 = E^L_{\alpha W} + \sum_{l=11}^{32}\left(\sum_{q=1}^{\infty}[(q-\alpha W^l)\overline{n}^l_q + (q+\alpha W^l-1)n^l_q]\right) + \sum_{i=1}^{D-2}\sum_{q=1}^{\infty} q M_q^{iL} - \tfrac{1}{24}(D-2) , \tag{2.3}$$

where n^l_q and $\overline{n}^l_q$ are the lth complex fermion's occupation numbers in the qth Fourier mode; they each may take the value 0 or 1, and one defines the lth fermion number

$$N^l_{\alpha W} = \sum_q (n^l_q - \overline{n}^l_q) .$$

M_q^{iL} is the occupation number for the ith string coordinate in the qth Fourier mode, $D=4$ is the dimension of space-time, and i runs from 1 to $D-2$ since we work in the light-cone gauge. $\frac{1}{24}(D-2)$ is the contribution to the vacuum energies from the string coordinates. The corresponding expression for the right-moving part is similar: l is summed from 1 to 10 instead. Occupation numbers must be such that $m_L^2 = m_R^2$; physically this means that the states are invariant under shifts $\sigma_1 \to \sigma_1$+const of the first world-sheet coordinate. $m^2_{R,L}$ is then the mass of the string state in units of the Planck mass.

Second, the spin-structure basis vectors $\mathbf{W}_i$ must satisfy the following constraints:

$$\mathbf{W}_i \cdot \mathbf{W}_j \stackrel{1}{=} k_{ij} + k_{ji} , \tag{2.4}$$

$$m_j k_{ij} \stackrel{1}{=} 0 , \tag{2.5}$$

$$k_{ii} + k_{i0} + s_i - \tfrac{1}{2}\mathbf{W}_i \cdot \mathbf{W}_i \stackrel{1}{=} 0 , \tag{2.6}$$

where $s_i = \mathbf{W}_i^1$, the space-time part of the spin structure, and where (k_{ij}) is some auxiliary matrix related to projection choices. We use the inner product

$$\mathbf{V}\cdot\mathbf{W} \stackrel{\text{def}}{=} \sum_{j=11}^{32} V^j W^j - \sum_{j=1}^{10} V^j W^j . \tag{2.7}$$

The *triplet constraint* on the right-moving spin structures imposed by world-sheet supersymmetry may be written as

$$s_i \stackrel{1}{=} W_i^2 + W_i^3 + W_i^4 \stackrel{1}{=} W_i^5 + W_i^6 + W_i^7 \stackrel{1}{=} W_i^8 + W_i^9 + W_i^{10} . \tag{2.8}$$

Note that if, for simplicity's sake, we restrict the complex fermions to be periodic or antiperiodic only, i.e., $W_i^l = 0$ or $\frac{1}{2}$, and $m_i = 2$, then the above constraints amount to the following: The product $\mathbf{W}_i \cdot \mathbf{W}_i$ must be integral, $\mathbf{W}_i \cdot \mathbf{W}_j$ must be a multiple of $\frac{1}{2}$, the elements of the k matrix are either 0 or $\frac{1}{2}$ and only the ones below the diagonal are independent.

Third, the allowed states in the sector αW, i.e., those which are not eliminated by the GSO projection, must have fermion number $\mathbf{N}_{\alpha W}$ such that

$$\mathbf{W}_i \cdot \mathbf{N}_{\alpha W} \stackrel{1}{=} s_i + \sum_j k_{ij}\alpha_j + k_{0i} - \mathbf{W}_i \cdot \alpha W . \tag{2.9}$$

A state is a space-time fermion or boson depending on whether αs $(=\alpha W^1)$ is 0 or $\frac{1}{2}$. All of this remains valid if we have Majorana fermions along with complex ones; in such cases the above formulas have to be slightly modified. A weight of $\frac{1}{2}$ must be introduced in the Lorentzian inner product in front of the spin structures of real fermions, and the antifermion occupation numbers $\overline{n}^l_q$ must be absent. The advantage of complex fermions is that we can then define a U(1) fermionic charge $Q^l = \mathbf{N}^l_{\alpha W} + \alpha W^l - \frac{1}{2}$ for each, thereby obtaining a charge vector facilitating the identification of the gauge group and matter representations. The allowed states then correspond to fermionic charge vectors $\mathbf{Q} = \mathbf{N} - \mathbf{W}_0 + \alpha W$ satisfying

$$\mathbf{Q}_{\alpha W} \cdot \mathbf{W}_i \stackrel{1}{=} s_i + k_{i0} + \sum_j k_{ij}\alpha_j . \tag{2.10}$$

Given the above rules, it is then straightforward to find the spectrum of the model. Many examples have been given in Ref. 6; the next section will describe the content of a generic massless spectrum.

III. GENERAL REMARKS ABOUT THE MASSLESS SPECTRUM

Only sectors with vacuum energies ≤ 0 will contribute to the massless spectrum. Moreover, one sees from Eq. (2.3) that only the first Fourier modes $(q=1)$ contribute as well. If the lth fermion is antiperiodic $(\alpha W^l = \frac{1}{2})$, then n_1^l or $\bar{n}_1^l$ may be 1, thus raising the energy by $\frac{1}{2}$ unit each and giving a charge $Q^l_{\alpha W} = 1, -1$ (or 0 if they are both equal to 1). If the fermion is periodic, i.e., if it has a zero mode, then only n_1^l may be 1, in which case it does not raise the energy and gives a charge $Q^l_{\alpha W} = \frac{1}{2}$ (or $-\frac{1}{2}$ if it is not excited at all). Bosonic oscillators of course do not affect the charge vector $\mathbf{Q}$, but the first Fourier mode will raise the energy by one unit. It is straightforward to show, using Eq. (2.3) and the definition of $\mathbf{Q}_{\alpha W}$, that any massless state has a charge vector such that its left-moving part $\mathbf{Q}^L$ has norm 2 or 0. In general we use the superscripts R, L to indicate the right- or left-moving part of a quantity.

Only in the $\mathbf{W}_0$ sector, which has the lowest possible vacuum energies, i.e., $(E^R \mid E^L) = (-\frac{1}{2} \mid 1)$, can we excite a bosonic oscillator to obtain a massless state. Indeed, by exciting the left-moving bosonic oscillators and the right-moving transverse fermion, we create a tensor state comprising the graviton, the antisymmetric tensor state and the dilaton. If instead we excite the other right-moving fermions, we create *right* vector bosons belonging to the graviton supermultiplet ($N = 4$). Whatever the spin-structure basis is, however, the graviton will always be part of the spectrum: its charge vector is

$$\mathbf{Q}_{\text{grav}} = (\pm 1, 0^9 \mid 0^{22}) ,$$

which satisfies Eq. (2.10) for all i.

The massless vector states created by exciting left-moving fermions and the right-moving antiperiodic transverse fermion are the gauge bosons of the low-energy effective theory. The corresponding charge vectors $\mathbf{Q}$ are the roots of the gauge group's Lie algebra. For those sectors which contain gauge bosons, αs is $\frac{1}{2}$ and the vacuum energies are $(-\frac{1}{2} \mid -1)$, $(-\frac{1}{2} \mid -\frac{1}{2})$, or $(-\frac{1}{2} \mid 0)$. In the first case, we must excite two left-moving fermions, resulting in a left charge vector of the form $\mathbf{Q}^L = (0,\ldots,\pm 1,\ldots,\pm 1,\ldots,0)$, or we may take $n_1^l = \overline{n}_1^l = 1$, giving a vanishing $\mathbf{Q}^L$. This only happens in the $\mathbf{W}_0$ sector, in which case Eq. (2.10) reduces to $\mathbf{Q}^L \cdot \mathbf{W}_i \stackrel{1}{=} 0$. In the second case, there are four left-moving zero modes which do not raise the energy if occupied; therefore we must excite one left-moving antiperiodic fermion and some zero modes, resulting in a charge vector with four $\pm\frac{1}{2}$'s and one ± 1. In the third case, there are eight left-moving zero modes, and no antiperiodic fermion must be excited if the state is to be massless; the charge vector then contains eight $\pm\frac{1}{2}$'s.

There is a systematic way to determine the Lie algebra from the set of roots: we simply have to pick a set of simple roots and calculate the Cartan matrix. A simple algorithm has been devised to accomplish this task on a computer, and will be described in the next section. The algebras arising from this construction can only be simply laced, restricting the simple factors to be SO($2n$), SU($n+1$), E_6, E_7, or E_8. If there are U(1) factors in the gauge group, only part of the 22-dimensional root space is spanned by the simple roots. Its orthogonal complement contains the U(1) components of the charge vectors, and it is always possible to pick an integral, orthogonal basis $\{\mathbf{e}_a\}$ in this subspace; we will refer to it as the U(1) basis. This basis must be orthogonal in order for the corresponding U(1) gauge bosons not to mix. Many bases are possible, all of them related by orthogonal transformations and normalization changes.

There are possibly tachyons in the $\mathbf{W}_0$ sector. They are obtained by exciting one left-moving fermion, and fall into the fundamental representation of SO(44) if no other constraint suppresses them. In order to eliminate them all, we must introduce in the spin-structure basis a vector $\mathbf{W}_1$ such that $\mathbf{W}_1^L = \mathbf{W}_0^L$ and $s_1 = 0$. This leaves us with two possibilities, of which we select the one introducing space-time supersymmetry, i.e.,

$$\mathbf{W}_1 = (0,(0,\tfrac{1}{2},\tfrac{1}{2})^3 \mid (\tfrac{1}{2})^{22}). \tag{3.1}$$

Henceforth we assume that the above vector is always part of the spin-structure basis, as it ensures that no tachyon will be present in any other sector as well.

Sectors containing massless fermions must have four right-moving zero modes, since αs must equal 0 and E^R must be a multiple of $\frac{1}{2}$. The vacuum energies are then $(0 \mid -1)$, $(0 \mid -\frac{1}{2})$, or $(0 \mid 0)$. The structure of the left charge vectors is the same as for the gauge bosons; they form weight vectors of representations of the gauge group. At the massless level, these representations are sufficiently small that all nonzero dominant weights, i.e., weights with non-negative inner product with all the simple roots, are highest weights. This makes the identification of the irreducible representations straightforward in terms of Dynkin labels (see Sec. IV). If there are U(1) factors in the gauge group, we project the charge vectors onto the U(1) basis vectors, thus obtaining the various U(1) charges for that state.

Because of the extra right-moving zero modes, a degeneracy of weights is possible and as many as four copies of the same representation can coexist. This reflects the presence of space-time supersymmetry in the model. The number of supersymmetric charges (which we denote by N) is equal to the number of gravitino states in the spectrum. Gravitinos are obtained by exciting the left-moving bosonic coordinate and the right-moving transverse fermion in a fermionic sector; extra right-moving zero modes may also be occupied. This occurs in sectors with vacuum energies $(0 \mid -1)$, i.e., mostly in the $\mathbf{W}_1$ sector. For a fixed helicity, there are eight possible occupation configurations for the three extra zero modes, at least half of which will be suppressed by the constraint

$$\mathbf{Q} \cdot \mathbf{W}_i \stackrel{1}{=} s_i + k_{i1} . \tag{3.2}$$

We set k_{i0} equal to zero, a choice that can always be made.[16] The charge vectors of the positive-helicity gravitinos are

$$\begin{aligned}
\mathbf{Q}_1^R &= (\tfrac{1}{2},(\tfrac{1}{2}00),(\tfrac{1}{2}00),(-\tfrac{1}{2}00)) ,\\
\mathbf{Q}_2^R &= (\tfrac{1}{2},(\tfrac{1}{2}00),(-\tfrac{1}{2}00),(\tfrac{1}{2}00)) ,\\
\mathbf{Q}_3^R &= (\tfrac{1}{2},(-\tfrac{1}{2}00),(\tfrac{1}{2}00),(\tfrac{1}{2}00)) ,\\
\mathbf{Q}_4^R &= (\tfrac{1}{2},(-\tfrac{1}{2}00),(-\tfrac{1}{2}00),(-\tfrac{1}{2}00)) .
\end{aligned} \tag{3.3}$$

It is easy to see what the $\mathbf{W}_i^R$'s should be in order to preserve $N = 4$ supersymmetry, or to reduce it to $N = 2$, 1, or 0. First, we always have the freedom to add $\mathbf{W}_0$ and $\mathbf{W}_1$ to the other basis vectors to ensure that $s_i = 0$ and that $\mathbf{W}_i^L$ and $\mathbf{W}_i^R$ $(i > 1)$ have integer norm separately. Then the only four *right movers* preserving $N = 4$ supersymmetry are

$$\begin{aligned} R_0 &= (0,(000),(000),(000))\,,\\ R_1 &= (0,(0\tfrac{1}{2}\tfrac{1}{2}),(0\tfrac{1}{2}\tfrac{1}{2}),(000))\,,\\ R_2 &= (0,(000),(0\tfrac{1}{2}\tfrac{1}{2}),(0\tfrac{1}{2}\tfrac{1}{2}))\,,\\ R_3 &= (0,(0\tfrac{1}{2}\tfrac{1}{2}),(000),(0\tfrac{1}{2}\tfrac{1}{2}))\,. \end{aligned} \tag{3.4}$$

If two of the gravitinos are to be eliminated, say $\mathbf{Q}_2$ and $\mathbf{Q}_3$, then we should introduce some of the following right-movers:

$$\begin{aligned} R_4 &= (0,(\tfrac{1}{2}0\tfrac{1}{2}),(\tfrac{1}{2}0\tfrac{1}{2}),(000))\,,\\ R_5 &= (0,(\tfrac{1}{2}0\tfrac{1}{2}),(\tfrac{1}{2}\tfrac{1}{2}0),(000))\,,\\ R_6 &= (0,(\tfrac{1}{2}\tfrac{1}{2}0),(\tfrac{1}{2}0\tfrac{1}{2}),(000))\,,\\ R_7 &= (0,(\tfrac{1}{2}\tfrac{1}{2}0),(\tfrac{1}{2}\tfrac{1}{2}0),(000))\,. \end{aligned} \tag{3.5}$$

For each of these set $k_{i1} = \frac{1}{2}$. We can also use R_{8-11}, having the same structure as the above but with the last two triplets interchanged, and R_{12-15}, with the first and last interchanged. If we set $k_{i1} = \frac{1}{2}$ for each of the basis vectors having $\mathbf{W}_i^R = R_{4-15}$, then only the fourth gravitino will survive the constraints, and $N = 1$ supersymmetry will be assured. Of course other prescriptions are possible if we choose to keep a different gravitino state. There are also many prescriptions to suppress any space-time supersymmetry.

In general, a large number of massless scalars will appear in the spectrum. They occur in sectors having the same vacuum energies as the sectors containing massless fermions, with the difference that $\alpha s = \frac{1}{2}$. The mere fact that one gravitino state exists implies that for each massless fermion representation there is a corresponding massless scalar representation with the same quantum numbers (except for the adjoint representation, which corresponds to the vector bosons). In fact the two chiralities of the massless fermion correspond to the scalar and its complex conjugate, as only complex scalars are present.

Finally, let us say a word about U(1) anomalies. Even though our string models are consistent, there may appear anomalies in the triangle diagrams at one loop, with one or three U(1) gauge bosons as external legs. This apparent anomaly is canceled according to the Green and Schwarz mechanism, through a coupling between the U(1) gauge fields A_μ^i and a pseudoscalar field ϕ related to the antisymmetric tensor field $B_{\mu\nu}$, arising in the one-loop effective action.[17–19] This coupling has the form

$$(\partial_\mu\phi + e_1 A_\mu^1 + e_2 A_\mu^2 + \cdots)^2 \tag{3.6}$$

and so only one linear combination of the Abelian gauge fields couples to ϕ, and therefore only one linear combination of U(1) charges is anomalous. We should then be able to find a U(1) basis such that

$$\sum_i Q_i^a = \delta_{a,1} A\,, \tag{3.7}$$

$$\sum_i Q_i^a Q_i^b Q_i^c = 0 \;\text{ if }\; a,b,c \neq 1\,. \tag{3.8}$$

The anomalous U(1) gauge boson will acquire a mass of the order of the Planck scale, as we can see from the coupling above.

IV. SOME COMPUTING ALGORITHMS

This section is concerned with the practical work of generating and analyzing string models in the spin structure formalism with periodic or antiperiodic complex fermions. Readers not interested in the details of the algorithms may skip the following and proceed to the next section.

The reasons for our restriction to complex world-sheet fermions are practical ones: the existence of fermionic charge vectors makes the identification of representations easy, and the restriction to periodic or antiperiodic fermions makes a systematic generation of models in that class a (barely) tractable task. Part of the software written in the course of this research is devoted to the analysis of the massless spectrum determined by a given spin-structure basis, and other parts attempt to generate as many such bases as possible.

First let us describe a few algorithms which identify the gauge group and the fermion representations for a given spin structure and k matrix. Let us assume that we have found all the root vectors $\mathbf{Q}^L$ of the gauge bosons by sweeping over all sectors possibly containing them and trying the candidate roots against the constraint

$$\mathbf{Q}_{\alpha W}^L \cdot \mathbf{W}_i \overset{1}{=} \sum_j k_{ij}\alpha_j\,.$$

The reader is referred to any good text on semisimple Lie algebras[20] for missing details. In the space of root and weight vectors, which we refer to as *root space*, even if it also contains U(1) charges, we must define an ordering. A dictionarylike order is the simplest possibility: a vector $\mathbf{Q}$ is defined to be positive if its first nonzero component is < 0 and vice versa; $\mathbf{Q}_1 > \mathbf{Q}_2$ then means that $\mathbf{Q}_1 - \mathbf{Q}_2 > 0$. In fact, we need only to have at our disposal the set of positive roots. Next we find a set of simple roots $\{\alpha_i\}$, i.e., positive roots which are not the sum of two other positive roots; a property of simple roots is that they have nonpositive inner products with each other. To find such a set, we use the following recursive procedure: The smallest positive root has to be simple; all positive roots having positive inner product with it cannot be simple and are discarded; the smallest of the remaining set of positive roots has to be simple too and all remaining positive roots having positive inner product with it are discarded, and so on. Finally, we end up with a number $r \leq 22$ of simple roots. The other $22 - r$ dimensions are attributed to U(1) factors. We then compute the Cartan matrix

$$A_{ij} = \alpha_i \cdot \alpha_j \ .$$

The corresponding Dynkin diagram is made of a dot for each i and of links between the dots for which $A_{ij} \neq 0$; in general it is made of several disconnected pieces and some internal representation of it is built in terms of vertices and leg lengths. From this the computer can identify the gauge group and put the simple roots in a conventional order, in view of calculating the Dynkin labels of the massless representations.

We find the weights in the same way as the roots, i.e., by sweeping over all relevant sectors and trying out all positive helicity candidates (in the case of fermions). Note that *CPT* invariance is reflected in the fact that if a charge vector $\mathbf{Q}$ is part of the spectrum, so is $-\mathbf{Q}$; hence fixing the space-time helicity $(\mathbf{Q}^1)$ removes some redundancy. Note, however, that whereas roots come in pairs of opposite sign, weights in general do not, and therefore negative as well as positive weights must be considered. Each allowed weight having a non-negative inner product with each root is called *dominant.* At the massless level all dominant weights are also the highest (i.e., largest) weight of some irreducible representation. The Dynkin labels of a highest weight are its inner products with the simple roots, and they specify the irreducible representation uniquely. The most common irreducible representations and their Dynkin labels are listed in Appendix A. We still have to assign U(1) charges to the fermions; this is done by projecting the highest weights onto an integer-component, orthogonal basis on the $(22-r)$-dimensional orthogonal complement of the space spanned by the simple roots. To obtain such a basis, we start from the canonical Cartesian basis for a 22-dimensional space and apply a Gram-Schmidt procedure giving priority to a known orthogonal basis for the simple roots; this U(1) assignment is of course arbitrary. Given the list of massless fermion representations (of a fixed handedness) and their U(1) charges, we can check which ones are chiral by constructing the complex-conjugate representation of each and looking down the list for a match: nonchiral representations will occur in pairs, each being the complex conjugate of the other with the same handedness.

We of course check the U(1) anomalies by summing

$$\sum_i \mathbf{Q}_i^a \equiv A^a \tag{4.1}$$

over massless fermions for each charge. The correct anomalous linear combination of charges is then obtained by projecting the weights onto the vector

$$\mathbf{e}_{\text{anom}} = \sum_a \frac{A^a \mathbf{e}_a}{\mathbf{e}_a \cdot \mathbf{e}_a} \tag{4.2}$$

or a multiple thereof. The remaining charges have to be redefined, that is a new orthogonal U(1) basis has to be found for the orthogonal complement of $\mathbf{e}_{\text{anom}}$. A similar procedure must be done when a particular linear combination of the remaining charges is being considered, as a possible hypercharge candidate.

So far in this section we did not assume the complex world-sheet fermions to be periodic and antiperiodic only. In fact, computer codes obtaining the massless representations and the gauge group have been written for the case where the fermions are periodic and antiperiodic as well as for the case where they acquire arbitrary rational phases after going around noncontractible loops of the world-sheet torus. However, in the remainder of this section, we will restrict ourselves to the periodic-antiperiodic case, for the sake of simplicity.

Let us now discuss two ways of generating large numbers of periodic and antiperiodic spin structures. The ambitious one is to devise a large, recursive algorithm that sweeps over all spin-structure bases, thereby exhausting systematically the class of classical ground states of the string we are considering. A more practical way is to sample the space of string models at random, quickly generating sets of consistent spin structures and hoping that some saturation occurs. A saturation would indicate that the space has been scanned *almost* completely. In both cases, we only want to generate the left part of the spin structure, and attach thereon the known right movers compatible with the desired space-time supersymmetry (see Sec. III). It is also very convenient to work solely with integers, and we therefore multiply spin structures, the k matrix, and charges by 2 and work in modulo-2 arithmetic.

The random generation is very simple. The bits of a random integer are used to form a spin-structure basis vector (provided the number of 1's is a multiple of 4), so that a whole vector is obtained from a single random-number call. It is preferable however for the random-number generator not to be uniform. Uniformity tends to give too much importance to numbers having about as many 1's and 0's in their binary expression, compared to numbers having very few 1's or very few 0's; in fact the distribution of the number of 1's is binomial, whereas we would like it to be more constant. This is easy to fix, however, and a distribution where the number of 1's is relatively uniform may be obtained. The right movers are then added, randomly if choice permits, so as to make the overlaps $\mathbf{W}_i \cdot \mathbf{W}_j$ even. The linear independence of the vectors is directly checked by computing all the sectors αW and making sure that none of them vanishes if $\alpha \neq 0$. The k_{ij}'s below the diagonal are also randomly set, except for the k_{i1}'s (which are fixed by supersymmetry) and the k_{i0}'s (which are always set equal to zero). This method is not recursive, allowing fairly large bases to be generated in little time.

A systematic generation of all periodic-antiperiodic spin structures is considerably more arduous, the main difficulty being the enormous redundancy of bases for a given set of spin structures. The basic problem, as far as the left movers are concerned, is to generate all possible vector spaces of 22-component vectors made of 0's and 1's, of even norm, obeying modulo-2 arithmetic. For such a vector space of dimension n, there are 2^n elements, which we refer to as *sectors*, and many possible bases, especially if permutations of the 22 components are included. Since we are going to generate bases directly, such a redundancy of bases is a problem. One way to avoid this redundancy is to define, for each space, a *standard basis* which can be recovered unambiguously from any other basis. The general idea is to define an ordering on the vector space, and to pick the largest possible linearly independent sectors of norm divisible by 4 to be the basis vectors. Because we want the ordering to be unaffected by permutations of the 22 components, a dictionarylike order is not suitable; instead we define a *semiordering* based on overlaps between sectors. The sectors are first ordered with respect to their norm, and we thus end up with several groups, each containing sectors of the same norm; then each group is further subdivided according to the sum of overlaps with members of the highest group, and then with members of the next highest group, etc. At the end of this iteration, the only ambiguities left in the ordering are irrelevant. A basis is then picked among sectors of norm divisible by 4, starting with the longest and going down, eliminating linearly dependent sectors along the way. This eliminates the redundancy due to the arbitrariness in the choice of the spin-structure basis, given a space of sectors. Finally, to eliminate the arbitrariness due to permutations of the 22 fermions, we order them in a dictionarylike way based on the already fixed order of the basis vectors; this means that the 1's are stacked to the left as much as possible, from top to bottom, as in the following example:

$$\begin{aligned} \mathbf{W}_2^L &= (1111111111111111000000)\ , \\ \mathbf{W}_3^L &= (1111111100000000111100)\ , \\ \mathbf{W}_4^L &= (1111000011110000110011)\ , \\ \mathbf{W}_5^L &= (1100110011101110101000)\ . \end{aligned} \tag{4.3}$$

Recall that we scaled all $\frac{1}{2}$'s to 1's for convenience.

Once we know how to obtain a standard basis from a space of sectors, we can generate them recursively in an intelligent order. First we start with the one-dimensional spaces, of which there are only five since we require the basis vectors to have a norm divisible by 4. Then we go to two-dimensional spaces by adding all possible vectors of norm 20, 16, 12, ... to the one-dimensional bases, starting with the longest. The two-dimensional bases thus obtained are then standardized, and compared with the ones already obtained to see if they occurred before (in which case they are not stored). We then go to three-dimensional spaces, etc. We refer to the norms (in decreasing order) of the members of the standard basis as the *composition* of that basis. Spaces of a given dimension are thus generated in order of decreasing composition, facilitating the bookkeeping. The reason behind storing all the standard bases is that it saves a tremendous amount of computing to check if a space of sectors has been obtained previously before going on to further, time-consuming analysis.

A given basis of left movers may be symmetric under interchange of one or more vectors. This has to be taken into account when tacking on the possible right movers in all ways consistent with the requirement that $\mathbf{W}_i \cdot \mathbf{W}_j$ be even. Adding the right movers can also lead to other redundancies, arising from permutation symmetry of the *triplets* and of the last two fermions of each triplet. To remove this redundancy in the case of $N=1$ supersymmetry, we artificially pick R_4 and R_8 (or R_{10}, depending on the overlaps of the left movers) to be attached respectively to $\mathbf{W}_2$ and $\mathbf{W}_3$.

At this point the redundancy due to independent permutations of the right- and left-moving components has been completely eliminated. However, more redundancy arises from rotational symmetry of the *charge lattice*;[16] some of this can be eliminated by fixing k_{i0} to be 0, and more by demanding that none of the sectors have both norm 4 and vanishing right movers. However, some redundancy will remain that cannot be eliminated by easy means, making a systematic and complete generation of models slower and slower as the number of basis vectors increases.

Once the spin structure has been generated, there is still the problem of sweeping over the possible k matrices. The number of independent k_{ij}'s is $\frac{1}{2}(n-2)(n-3)$ where n is the number of spin-structure basis vectors (including $\mathbf{W}_0$); the number of possible k matrices is then $2^{(n-2)(n-3)/2}$. The task of generating them becomes rapidly intractable as n increases. Methods to bypass the redundancy lying in the k matrix have not so far been explored.

So far, the systematic approach has been carried out to $n=5$, that is, spin structures involving 32 sectors or less. If, *for the sake of counting*, we distinguish $N=1$ models by their gauge group, the number of massless representations, the number of chiral fermion representations and the number of chiral fermions, then the number of distinct models we have constructed for $n \leq 5$ is 1532. To increase this systematic sweep over to $n=6$ would require a substantial increase in computing time, although it is by no means impossible. To illustrate this point, let us mention that the number of bases we generated for the left-moving part of the spin structure is 1376 for $n \leq 5$, compared with 42 927 for $n=6$ alone; adding the right moving parts and sweeping over the k matrices is also considerably lengthier for $n=6$ than it is for $n=5$.

The random generation of spin structures is more efficient and fruitful. We are not limited to small numbers of basis vectors since the approach is not recursive, and sets of nine basis vectors are trivially generated. To check if the random sweep is complete, we may compare its results with those obtained with the more complicated systematic sweep. The results for $n \leq 4$ are encouraging: out of the 1532 models (as specified above), 1489 were also found by a random sweep, that is, 97% of them. Had the sweep run longer, that proportion would have further increased. Another way of checking the completeness of a random sweep is to look for saturation in the number of new models generated compared with the rate of spin-structure generation. This may be seen for the case of $N = 4$ supersymmetry, for which the number of models is much less. This question is discussed in Appendix C. As for $N = 1$ models, their number is much larger and we expect a saturation to appear much later. As of this writing, the statistics are as follows: for $N = 4$ supersymmetry, 918 distinct models were constructed, after having generated over two million complete spin-structure bases at random. In that case the saturation in the production is such that we may estimate the actual number of $N = 4$ models to be about one thousand. For $N = 1$ supersymmetry, 32 756 distinct models were constructed, after having generated over 60 000 spin structures; in that case almost no saturation was visible, and we may expect the actual figure to be an order of magnitude larger, if not more. To illustrate this, let us point out that among the 32 756 $N = 1$ models constructed, 3249 included SU(3)×SU(2)×U(1) as a factor in the gauge group; if we run the program and keep only the models having an SU(3)×SU(2)×U(1) factor, which is faster, we can construct many more of these models, namely over 30 000, as of this writing.

In practice, to avoid the use of large quantities of computer memory, we choose an embedding scenario for the standard model, write up a computerized set of conditions on the spectrum for the model to be "interesting," and run a random generation for a few days. Only the models passing the programmed conditions will be kept to be checked further, by hand. On a typical run looking for SU(5) models, the computer generated 1 259 165 (not necessarily distinct) models, of which 10 917 explicitly contained SU(5) as a gauge group factor; of these, only 15 passed the computerized spectrum test, of which one example will be discussed below. Looking for an explicit appearance of SU(3)×SU(2)×U(1), the computer generated 1 667 707 (not necessarily distinct) models and found 223 486 with the right group factors, of which twenty or so passed the computerized spectrum test. Of these, only seven survived a quick look at the possibilities for a hypercharge.

V. EMBEDDING THE STANDARD MODEL

Presumably only a few of the ground states thus found can reasonably accommodate the particle spectrum of the standard SU(3)×SU(2)×U(1) gauge theory. Let us examine a few scenarios for the inclusion of the elementary particles we know today in our string models. For the moment we ignore the question of supersymmetry breaking for those models that have $N = 1$ supersymmetry (SUSY) at the Planck scale.

TABLE I. Standard-model massless fermion representation.

Dimension	Dynkin label	Description
$(\mathbf{2},\mathbf{3})(-\frac{1}{3})$	$(1)(01)\ (-\frac{1}{3})$	Quark doublet q_L^c
$(\mathbf{2},\mathbf{1})(+1)$	$(1)(00)\ (+1)$	Lepton doublet l_L^c
$(\mathbf{1},\overline{\mathbf{3}})(\frac{4}{3})$	$(0)(10)\ (\frac{4}{3})$	$\frac{2}{3}$ charged q_R
$(\mathbf{1},\overline{\mathbf{3}})(-\frac{2}{3})$	$(0)(10)\ (-\frac{2}{3})$	$-\frac{1}{3}$ charged q_R
$(\mathbf{1},\mathbf{1})(-2)$	$(0)(00)\ (-2)$	Lepton l_R

Our first scenario requires that SU(3), SU(2), and U(1) appear explicitly as factors of the gauge group. The desired massless fermion representations of a given chirality are shown in Table I.

The first set of labels gives the dimensions and the hypercharges of the SU(2)×SU(3)^c representations, and the second set gives the Dynkin labels. The fermions of opposite chirality live in the complex-conjugate representations, and no representation of the same chirality is the complex conjugate of one of the above, i.e., they are *chiral* representations. Family replication may occur by direct repetition of the above set or through some horizontal symmetry, a remnant of a local symmetry for which the gauge bosons have become very heavy. For instance, the above representations might also be part of the **4** of an SU(4) broken at a very high-energy scale. The number of such families should be at least three and no more than six; as it turns out, four is by far the most frequent number encountered in this range. Among all the U(1) factors in the model's gauge group, there must be a nonanomalous linear combination which fits the standard hypercharge assignment. Any other representation having SU(3)^c×SU(2)×U(1) quantum numbers should be massive enough not to interfere with low-energy phenomena; for instance, this may be achieved if that representation is not chiral, or is part of a heavy composite bound by a confining force. Finally, a complex Higgs isodoublet with hypercharge 1 must be present. Of course every model has its particularities, and only by going through an example can we give more details.

We may also consider scenarios where the SU(3) ×SU(2)×U(1) group factors do not appear explicitly at the Planck scale, but where a grand-unified-theory (GUT) group appears instead, breaking at some intermediate scale via a Higgs mechanism. For instance, we might have a Pati-Salam-type gauge group, such as $SU(2)_R \times SU(2)_L \times SU(4)^c$ (Ref. 21). The desired massless fermion representations are then (0)(1)(001) [or $(\mathbf{1},\mathbf{2},\mathbf{4})$] for quarks and leptons and (1)(0)(100) [or $(\mathbf{2},\mathbf{1},\overline{\mathbf{4}})$] for antiquarks and antileptons. Lepton number then plays the role of a fourth color. The correct hypercharge arises when $SU(2)_R \times SU(4)^c$ is broken to $SU(3)^c \times U(1)$. This first stage of spontaneous symmetry breaking occurs through a Higgs scalar belonging to $(\mathbf{2},\mathbf{1},\overline{\mathbf{4}})$ (Ref. 22) and its complex conjugate. The branching rules are as follows:

$$(\mathbf{1},\mathbf{2},\mathbf{4}) \rightarrow (\mathbf{2},\mathbf{3})(-\tfrac{1}{3}) + (\mathbf{2},\mathbf{1})(1)\ ,$$
$$(\mathbf{2},\mathbf{1},\overline{\mathbf{4}}) \rightarrow (\mathbf{1},\overline{\mathbf{3}})(\tfrac{1}{3}) + (\mathbf{1},\overline{\mathbf{3}})(\tfrac{4}{3}) + (\mathbf{1},\mathbf{1})(0) + (\mathbf{1},\mathbf{1})(-2)\ .$$

The second stage of symmetry breaking occurs through the usual Higgs mechanism:

$$(\mathbf{2},\mathbf{2},\mathbf{1}) \rightarrow (\mathbf{2},\mathbf{1})(1) + (\mathbf{2},\mathbf{1})(-1)\ .$$

The problem of fast proton decay may be remedied by introducing an additional scalar multiplet in the $(\mathbf{1},\mathbf{1},\mathbf{6})$ (Ref. 23). This extra representation appears naturally in the breaking of the **10** of SO(10) and is very common. All chiral representations that are not singlets under $SU(2)_L$ or SU(4) should fall in this pattern. Again, the family structure may arise from distinct representations or from further horizontal symmetry.

In another example, an SU(5) simple factor appears. The three representations wanted are the **5**, the $\overline{\mathbf{10}}$, and a singlet. In the minimal SU(5) of Georgi and Glashow,[24] the singlet is a right-handed neutrino. The breaking occurs through an adjoint of Higgs field **24** as follows:

$$\mathrm{SU}(5) \rightarrow \mathrm{SU}(2) \times \mathrm{SU}(3)^c \times \mathrm{U}(1)_Y\ ,$$

$$\mathbf{5} \rightarrow (\mathbf{2},\mathbf{1})(1) + (\mathbf{1},\mathbf{3})(-\tfrac{2}{3})\ ,$$

$$\overline{\mathbf{10}} \rightarrow (\mathbf{1},\mathbf{1})(-\tfrac{2}{3}) + (\mathbf{2},\overline{\mathbf{3}})(-\tfrac{1}{3}) + (\mathbf{1},\mathbf{3})(\tfrac{4}{3})\ .$$

The second stage of symmetry breaking occurs through a Higgs mechanism in the **5** of SU(5). This scenario is not likely to occur in our string models with $N = 1$ supersymmetry, for the following reason: in those models, only one adjoint representation of massless fermions is present, and its superpartner is the representation of gauge bosons; hence there is no adjoint of scalars. The only way an adjoint of the Higgs field could occur is by compositeness; for instance, the Higgs field could be a composite of a $(\mathbf{5},\mathbf{10})$ of SU(5)×SO(10) and its complex conjugate. Assuming the SO(10) is confined at a higher scale than the SU(5) GUT scale, the composite would reduce to a **24** + **1** of SU(5), singlet of SO(10). However, such large representations seem to be rare. There is another way of putting quarks and leptons in SU(5) multiplets, though, in the so-called flipped SU(5)×U(1) (Refs. 25–27). In this model the quarks and leptons live in the three representations $\mathbf{10}(\tfrac{1}{2})$, $\overline{\mathbf{5}}(-\tfrac{3}{2})$, and $\mathbf{1}(\tfrac{5}{2})$. The singlet is no longer a right-handed neutrino, but a right-handed positron; also, the u and d quarks are interchanged in the SU(5) multiplets. The breaking to the standard model occurs through a Higgs field in the $\mathbf{10}(\tfrac{1}{2})$ (and complex conjugate), and the usual Higgs field lies in the $\mathbf{5}(-1)$ (and complex conjugate).

Finally, another scenario could rely on the SO(10) GUT (Refs. 28–30), in which all above groups are contained. The only necessary representation is the $\overline{\mathbf{16}}$, or (00001). Under the breaking SO(10) →SU(5)×U(1), we have the branching rule

$$\overline{\mathbf{16}} \rightarrow \mathbf{1}(5) + \mathbf{5}(-3) + \overline{\mathbf{10}}(1)\ .$$

This happens with a Higgs field lying in the spinor **16** of SO(10) and the adjoint **45**. The same difficulties mentioned above for the minimal SU(5) case still exist here. For the minimal SU(5), one does not need the adjoint to break SO(10) →SU(5), but rather to produce the adjoint of Higgs field that breaks SU(5). SO(10) can also break to SU(4)×SU(2)×SU(2), through a symmetric tensor **54** of Higgs field, or a **210**, two representations that do not occur at the massless level of our string models. It seems therefore that if SO(10) is going to play a role, it will already be broken at the Planck scale.

VI. EXAMPLES

Before presenting specific examples of models obtained through our program, a few general, qualitative statements can be made based on a cursory look at hundreds of models. As far as gauge group factors are concerned, almost every simple factor of rank less than 22 occurs; exceptional groups are not exceptional in this respect. Some groups like SO(44) do not occur, as they are incompatible with $N = 1$ supersymmetry in this construction. Others such as SU(n) ($n \geq 5$) have a complicated embedding of their root lattice in an integer basis, and need a lot of elbow room to fit; hence they are very rare or not seen at all. In fact SU(n) has been seen up to $n = 12$. Observations have confirmed the fact that only the simplest representations are present at the massless level. For SU(n) these are the fundamental and its conjugate, the antisymmetric tensors and their conjugates, and the adjoint; for SO($2n$) they are the vector, the adjoint, and the spinor and its conjugate. Anomalous U(1) charges are very common; in fact, upon verification on a set of 2337 random models with U(1) charges, it was found that 1163 of these models had an anomalous U(1), which is remarkably close to 50%.

Let us now present three examples of models that were selected according to criteria described in Sec. V. A deeper analysis of each is needed to see if they are viable models or not. In particular, we do not pretend that all the couplings needed are present. We hope to present such careful analyses in a later publication. For the moment only the spectrum will be described, accompanied by a few hopeful comments. The reader is referred to Sec. II for explanations concerning the spin-structure construction.

In our first example, we consider a *flipped* SU(5) model, with the spin structure

$$\mathbf{W}_0 = (1111111111 \mid 1111111111111111111111)\ ,$$

$$\mathbf{W}_1 = (0011011011 \mid 1111111111111111111111)\ ,$$

$$\mathbf{W}_2 = (0101101000 \mid 1111111111110000000000)\ ,$$

$$\mathbf{W}_3 = (0110000101 \mid 1110000000001000000000)\ ,$$

$$\mathbf{W}_4 = (0101000110 \mid 11011000000000000000000) ,$$

$$\mathbf{W}_5 = (0101110000 \mid 10010100000000111110000) ,$$

$$\mathbf{W}_6 = (0101000110 \mid 11011111100001110010000) .$$

We have scaled all the $\frac{1}{2}$'s to 1's for convenience. All the k_{ij}'s below the diagonal, except for k_{i0}, k_{52}, and k_{64}, are set equal to 1. The gauge group is

$$SU(2)^2 \times SU(4) \times SU(5) \times SU(8) \times U(1)^6 .$$

The first U(1) charge is anomalous; the second (call it $4Y_f$) is the one associated with the SU(5) in the *flipped* model. The representations for the massless spectrum are listed in Appendix B. There are 84 of them (not counting the adjoints), of which 38 are chiral. We have four generations of quarks and leptons (see representations Nos. 9–12, 29–32, and 44–47). Many other chiral representations exist, of which only nonsinglets under SU(8) have nonzero Y_f. Since SU(8) is likely to confine at high energies (the β function has been checked to be negative if we include all states in the spectrum), we expect these representations to beome quite massive. The Higgs candidates live in representations Nos. 33 and 50. To show that there are symmetry-breaking patterns in which the other representations do not affect low-energy physics too much is clearly a hard problem, given the large number of parameters and possible minima in the superpotentials we may construct.

In our second example we consider a direct embedding of SU(3)×SU(2)×U(1) with some horizontal symmetry. The spin structure is

$$\mathbf{W}_2 = (0101101000 \mid 11110000000000000000000) ,$$

$$\mathbf{W}_3 = (0110000101 \mid 10001111111111110000000) ,$$

$$\mathbf{W}_4 = (0110101000 \mid 01001111000000001110000) ,$$

$$\mathbf{W}_5 = (0110000101 \mid 01000000110000001000000) ,$$

$$\mathbf{W}_6 = (0101000101 \mid 11111111111111100111111) ,$$

$$\mathbf{W}_7 = (0101101000 \mid 10101100000100011001000) ,$$

in addition to $\mathbf{W}_0$ and $\mathbf{W}_1$. All the k_{ij}'s below the diagonal are 1 except for $k_{i0}, k_{42}, k_{52}, k_{54}, k_{63}, k_{65}$, and k_{74}. The gauge group is

$$SU(2)^5 \times SU(3) \times SU(6)^2 \times U(1)^5 .$$

Again, Appendix B lists the massless representations. There are 44 representations, 32 of which are chiral. The first U(1) is anomalous and the second one is identified with the hypercharge (up to a factor of 6). Four generations of quarks and leptons are contained in representations Nos. 2–5, 8, 11, 13–16, 19, 20, 24, and 25. We are forced to hope that the gauge bosons corresponding to the third and fourth SU(2)'s are very heavy, as these two factors provide a horizontal symmetry for family replication (at least for the quark doublets, the conjugate d quarks, and the positron). The only chiral representations not falling in the standard pattern and having nonzero hypercharge are nonsinglets under the two SU(6) factors, which were checked to have negative β functions and are therefore expected to confine. The Higgs candidate lives in representation No. 12.

In our third and last example we consider the SU(3)×SU(2)×U(1) model with spin structure

$$\mathbf{W}_2 = (0101101000 \mid 11111111111111111111100) ,$$

$$\mathbf{W}_3 = (0110000101 \mid 11111111111100000000010) ,$$

$$\mathbf{W}_4 = (0101110000 \mid 11111000000111111100001) ,$$

$$\mathbf{W}_5 = (0000110110 \mid 10000111100110000000001) ,$$

$$\mathbf{W}_6 = (0110110000 \mid 01110111010111111011011) ,$$

$$\mathbf{W}_7 = (0110000110 \mid 11000100100101100000010) ,$$

in addition to $\mathbf{W}_0$ and $\mathbf{W}_1$. All the k_{ij}'s below the diagonal are 1 except for $k_{i0}, k_{32}, k_{53}, k_{62}, k_{63}, k_{65}, k_{72}, k_{74}$, and k_{75}. The gauge group is

$$SU(2)^6 \times SU(3)^2 \times U(1)^{12} .$$

There are 150 representations, and 104 are chiral. This is a large number and we can see how a detailed analysis could be lengthy. The spin structure has 256 sectors. There is one anomalous U(1), and another charge may be identified with the hypercharge. There are four families of quarks and leptons, without horizontal symmetry. The spectrum consists roughly of the following: 16 chiral singlets, of which 4 have $Y = -2$ and 12 have $Y = 0$ (in the conventional normalization); 56 chiral doublets in various SU(2)'s, with hypercharge $Y = \pm 1$; the $SU(2)_W$ doublets of course have $Y = 1$; there are eight chiral $SU(3)^c$ triplets with $Y = \frac{4}{3}, -\frac{2}{3}$ and eight chiral triplets in $SU(3)'$ with $Y = 0$; there are four sextets of $SU(3)^c \times SU(2)_W$ with $Y = -\frac{1}{3}$ and four sextets of $SU(3)' \times SU(2)'$ with

TABLE II. Examples of models constructed embedding the standard model.

$SU(4)^3 \times SU(3)^c \times SU(2)^5 \times U(1)^6$	[horizontal SU(2) symmetry]
$SU(4)^2 \times SU(3)^c \times SU(2)^7 \times U(1)^7$	[horizontal SU(2) symmetry]
$SU(3)^c \times SU(2)^{11} \times U(1)^9$	[horizontal SU(2) symmetry]
$SU(3) \times SU(3)^c \times SU(2)^8 \times U(1)^{10}$	[no anomalous U(1)]
$SU(3)^3 \times SU(2)^5 \times U(1)^{11}$	

$Y = 0$; there are eight chiral quartets in SU(2)×SU(2), singlets under $SU(2)_W$ and with $Y = 0$. Finally, there are some nonchiral representations, including two possibilities for the Higgs field. The representations for this model are too numerous to be listed in the Appendix.

Among all the models constructed that contained explicitly SU(3)×SU(2)×U(1), only a few survived a quick glance at their spectrum. In addition to the two models mentioned above, we have found the models, all having four families of quarks and leptons, shown in Table II.

We also constructed many models whose gauge group contains an SU(2)×SU(2)×SU(4) factor, and imposed conditions on their spectra as described in Sec. V. As of this writing, out of over 240 000 randomly generated models containing the Pati-Salam gauge group as a factor, only one has the required massless representations. The family replication in this model occurs by some SU(2)×SU(2) horizontal symmetry, with no apparent hope for these two factors to break spontaneously. We are still trying to find an interesting model in this category. We hope to provide in a future publication a more detailed analysis of some of the models we found.

ACKNOWLEDGMENTS

We thank our colleagues at Cornell for valuable advice; in particular we thank Shyamoli Chaudhuri, Keith Dienes, Hikaru Kawai, David C. Lewellen, Joel Schwartz, and Henry Tye for numerous helpful discussions. Throughout this research, we used Sun 350 work stations (with four megabytes of real memory). This work was supported in part by the National Science Foundation.

APPENDIX A

We present a short list of "basic" representations with the corresponding Dynkin labels.[31] They are useful in identifying the representations listed in Appendix B. See Table III.

APPENDIX B

In Tables IV and V we list the massless fermion representations for the models described in Sec. VI. The models have $N = 1$ supersymmetry, so these tables give the massless scalars as well. The first column is a label; a "c" after the label indicates that the representation is chiral, and an "r" indicates that it is real. If the representation is not chiral, it is implied that its complex conjugate is also part of the spectrum (even though it is not listed here in order to save space). We do not include in this list the adjoint representations, i.e., the gauge bosons and the gauginos, since they are obviously part of the spectrum. The second column gives the dimension; the third lists the Dynkin labels of the representations with respect to the simple factors of the gauge group, in the same order as they are listed in the gauge group. The fourth column gives the U(1) charges. The representations are given in the order of increasing dimension. Table IV lists the representations pertaining to

$$SU(2)^2 \times SU(4) \times SU(5) \times SU(8) \times U(1)^6$$

while Table V lists those pertaining to

$$SU(2)^5 \times SU(3) \times SU(6)^2 \times U(1)^5 .$$

APPENDIX C

In this appendix we briefly discuss the problems involved when trying to guess the total number of models in a set by looking at the saturation of a random generation of models from that set. Let us consider the following elementary example. Suppose that we have a set of N objects, and that we select them at random without removing them from the set. The probability p_n that after n picks we reach for an object that we have not chosen before is

$$p_n = \left(1 - \frac{1}{N}\right)^n$$

provided every object is equally likely to be picked each

TABLE III. Dynkin labels for common representations.

Algebra	Dynkin labels	Dimension	Representation
SU($n+1$)	(100...0)	$n+1$	Fundamental
	(0...001)	$n+1$	Fundamental conj.
	(10...01)	$n(n+2)$	Adjoint
	(010...010)	$\frac{1}{2}n(n+1)$	
SO($2n$)	(100...0)	$2n$	Vector
	(0100...0)	$n(2n-1)$	Adjoint
	(0...001)	2^{n-1}	Spinor
	(0...010)	2^{n-1}	Spinor conj.
E_6	(100000)	27	Fundamental
	(000010)	27	Fundamental conj.
	(000001)	78	Adjoint
E_7	(0000010)	56	
	(1000000)	133	Adjoint
E_8	(00000010)	248	Adjoint

TABLE IV. Massless representations of $SU(2)^2 \times SU(4) \times SU(5) \times SU(8) \times U(1)^6$.

Label		Dimension	Dynkin labels	U(1) charges					
1	c	1	0,0,000,0000,0000000	1	0	−24	−2	−2	−1
2	c	1	0,0,000,0000,0000000	1	0	−24	−2	−2	−1
3	c	1	0,0,000,0000,0000000	1	0	−24	−2	−2	1
4	c	1	0,0,000,0000,0000000	1	0	−24	−2	−2	1
5	c	1	0,0,000,0000,0000000	1	0	−24	−2	2	−1
6	c	1	0,0,000,0000,0000000	1	0	−24	−2	2	−1
7	c	1	0,0,000,0000,0000000	1	0	−24	−2	2	1
8	c	1	0,0,000,0000,0000000	1	0	−24	−2	2	1
9	c	1	0,0,000,0000,0000000	1	10	6	6	0	−1
10	c	1	0,0,000,0000,0000000	1	10	6	6	0	−1
11	c	1	0,0,000,0000,0000000	1	10	6	6	0	1
12	c	1	0,0,000,0000,0000000	1	10	6	6	0	1
13		1	0,0,000,0000,0000000	0	0	0	0	2	−2
14		1	0,0,000,0000,0000000	0	0	0	0	2	2
15		1	0,0,000,0000,0000000	0	5	15	4	−1	−2
16		1	0,0,000,0000,0000000	0	5	15	4	−1	2
17		1	0,0,000,0000,0000000	0	5	15	4	1	−2
18		1	0,0,000,0000,0000000	0	5	15	4	1	2
19		1	0,0,000,0000,0000000	2	5	−33	0	−1	0
20		1	0,0,000,0000,0000000	2	5	−33	0	1	0
21		1	0,0,000,0000,0000000	2	10	−18	4	0	0
22	c	4	1,1,000,0000,0000000	−1	0	−16	6	0	−1
23	c	4	1,1,000,0000,0000000	−1	0	−16	6	0	−1
24	c	4	1,1,000,0000,0000000	−1	0	−16	6	0	1
25	c	4	1,1,000,0000,0000000	−1	0	−16	6	0	1
26		4	1,1,000,0000,0000000	2	0	−8	−8	0	0
27		4	1,1,000,0000,0000000	2	5	7	−4	−1	0
28		4	1,1,000,0000,0000000	2	5	7	−4	1	0
29	c	5	0,0,000,0001,0000000	1	−6	6	6	0	−1
30	c	5	0,0,000,0001,0000000	1	−6	6	6	0	−1
31	c	5	0,0,000,0001,0000000	1	−6	6	6	0	1
32	c	5	0,0,000,0001,0000000	1	−6	6	6	0	1
33		5	0,0,000,0001,0000000	−4	4	−4	−4	0	0
34		5	0,0,000,0001,0000000	2	−6	−18	4	0	0
35		5	0,0,000,0001,0000000	2	−1	−3	8	−1	0
36		5	0,0,000,0001,0000000	2	−1	−3	8	1	0
37	c	6	0,0,010,0000,0000000	1	0	16	−6	0	−1
38	c	6	0,0,010,0000,0000000	1	0	16	−6	0	−1
39	c	6	0,0,010,0000,0000000	1	0	16	−6	0	1
40	c	6	0,0,010,0000,0000000	1	0	16	−6	0	1
41		6	0,0,010,0000,0000000	2	0	−8	−8	0	0
42		6	0,0,010,0000,0000000	2	5	7	−4	−1	0
43		6	0,0,010,0000,0000000	2	5	7	−4	1	0
44	c	10	0,0,000,0100,0000000	3	2	−2	−2	0	−1
45	c	10	0,0,000,0100,0000000	3	2	−2	−2	0	−1
46	c	10	0,0,000,0100,0000000	3	2	−2	−2	0	1
47	c	10	0,0,000,0100,0000000	3	2	−2	−2	0	1
48		10	0,0,000,0100,0000000	2	−3	7	−4	−1	0
49		10	0,0,000,0100,0000000	2	−3	7	−4	1	0
50		10	0,0,000,0100,0000000	2	2	22	0	0	0
51	c	16	0,1,000,0000,0000001	0	5	−5	−5	0	0
52	c	16	0,1,000,0000,1000000	2	−5	−3	−3	0	0
53	c	16	0,1,000,0000,1000000	2	0	12	1	−1	0
54	c	16	0,1,000,0000,1000000	2	0	12	1	1	0
55	c	16	1,0,000,0000,0000001	0	5	−5	−5	0	0
56	c	16	1,0,000,0000,1000000	2	−5	−3	−3	0	0
57	c	16	1,0,000,0000,1000000	2	0	12	1	−1	0
58	c	16	1,0,000,0000,1000000	2	0	12	1	1	0
59	r	24	1,1,010,0000,0000000	0	0	0	0	0	0
60	c	32	0,0,001,0000,0000001	2	0	−8	3	0	0
61	c	32	0,0,100,0000,0000001	2	0	−8	3	0	0
62	r	70	0,0,000,0000,0001000	0	0	0	0	0	0

TABLE V. Massless representations of $SU(2)^5 \times SU(3) \times SU(6)^2 \times U(1)^5$.

Label		Dimension	Dynkin labels	U(1) Charges				
1		1	0,0,0,0,0,00,00000,00000	4	−12	0	−12	0
2	c	2	0,0,0,0,1,00,00000,00000	−1	6	−3	−9	−1
3	c	2	0,0,0,0,1,00,00000,00000	−1	6	−3	−9	1
4	c	2	0,0,0,0,1,00,00000,00000	−1	6	3	−9	−1
5	c	2	0,0,0,0,1,00,00000,00000	−1	6	3	−9	1
6	c	2	0,0,0,1,0,00,00000,00000	−1	0	0	15	−2
7	c	2	0,0,0,1,0,00,00000,00000	−1	0	0	15	2
8	c	2	0,0,0,1,0,00,00000,00000	3	−12	0	3	0
9	c	2	0,0,1,0,0,00,00000,00000	−1	0	0	15	−2
10	c	2	0,0,1,0,0,00,00000,00000	−1	0	0	15	2
11	c	2	0,0,1,0,0,00,00000,00000	3	−12	0	3	0
12		2	0,0,0,0,1,00,00000,00000	6	6	0	6	0
13	c	3	0,0,0,0,0,10,00000,00000	−1	−4	−3	−9	−1
14	c	3	0,0,0,0,0,10,00000,00000	−1	−4	−3	−9	1
15	c	3	0,0,0,0,0,10,00000,00000	−1	−4	3	−9	−1
16	c	3	0,0,0,0,0,10,00000,00000	−1	−4	3	−9	1
17		3	0,0,0,0,0,10,00000,00000	4	8	0	−12	0
18		4	1,1,0,0,0,00,00000,00000	0	0	0	0	2
19	c	6	0,0,0,1,0,10,00000,00000	3	8	0	3	0
20	c	6	0,0,1,0,0,10,00000,00000	3	8	0	3	0
21		6	0,0,0,0,1,01,00000,00000	2	−2	0	18	0
22	c	8	1,1,0,1,0,00,00000,00000	1	0	0	−15	0
23	c	8	1,1,1,0,0,00,00000,00000	1	0	0	−15	0
24	c	12	0,0,0,1,1,01,00000,00000	3	−2	0	3	0
25	c	12	0,0,1,0,1,01,00000,00000	3	−2	0	3	0
26	c	12	0,1,0,0,0,00,00000,00001	1	6	−2	1	0
27	c	12	0,1,0,0,0,00,00001,00000	−1	−6	−2	−1	0
28	c	12	1,0,0,0,0,00,00000,10000	−1	−6	2	−1	0
29	c	12	1,0,0,0,0,00,10000,00000	1	6	2	1	0
30	c	12	0,1,0,0,0,00,10000,00000	2	0	2	10	0
31	c	12	0,1,0,0,0,00,10000,00000	2	0	2	10	0
32	c	12	1,0,0,0,0,00,00000,00001	2	0	−2	10	0
33	c	12	1,0,0,0,0,00,00000,00001	2	0	−2	10	0
34	c	15	0,0,0,0,0,00,00000,01000	3	0	1	−5	−1
35	c	15	0,0,0,0,0,00,00000,01000	3	0	1	−5	1
36	c	15	0,0,0,0,0,00,00010,00000	3	0	−1	−5	−1
37	c	15	0,0,0,0,0,00,00010,00000	3	0	−1	−5	1
38		36	0,0,0,0,0,00,00001,00001	0	0	2	0	0

time. It follows that the average number of objects that have been chosen at least once after n picks is

$$f_n = N\left[1-\left(1-\frac{1}{N}\right)^n\right] \simeq N - Ne^{-n/N} .$$

We see that the saturation as $n \gg N$ is purely exponential, like a charging capacitor in fact. This idea is too simple to be applied to our random generation of models, since not all of them are equally likely to be generated. Instead, the number of models having a probability between x and $x+dx$ to be picked is some function $h(x)dx$, such that

$$\int_0^1 h(x)dx = N , \quad \int_0^1 xh(x)dx = 1$$

and the average number of distinct models obtained after n random generations is

$$f_n = N - \int_0^1 h(x)e^{-nx}dx .$$

If we substitute $N\delta(x-1/N)$ for $h(x)$, we recover the equal-likelihood case. Without the knowledge of $h(x)$, it seems hard to fit the known results for n and f_n to an expected saturation curve in order to obtain a value for N. Moreover, it is easily seen that the saturation (i.e., the way f_n approaches N) no longer has to be exponential. For instance, if we assume the distribution

$$h(x) = \frac{N^2}{2} \quad \left(x < \frac{2}{N}\right) ,$$

$$h(x) = 0 \quad \left(x > \frac{2}{N}\right) ,$$

we then have

$$N - f_n = \frac{N^2}{2n}(1 - e^{-2n/N}) \ .$$

We see in this case that the saturation occurs according to a power law, i.e., much more slowly than for the uniform-likelihood case, even if the exponential component is faster. It would be interesting to know what $h(x)$ really looks like, knowing the values of f_n. Such an inversion appears difficult.

[1] M.B. Green and J.H. Schwarz, Nucl. Phys. **B255**, 92 (1985); J.H. Schwarz, Phys. Rep. **89**, 223 (1982); M.B. Green, J.H. Schwarz, and E. Witten, *Superstring Theory* (Cambridge University Press, Cambridge, England, 1987).

[2] D.J. Gross, J.A. Harvey, E. Martinec, and R. Rohm, Nucl. Phys. **B256**, 253 (1985); **B267**, 75 (1986).

[3] D. Gepner, Phys. Lett. B **199**, 380 (1987).

[4] J. Distler and B. Greene, Nucl. Phys. **B309**, 295 (1988).

[5] K.S. Narain, Phys. Lett. **169B**, 41 (1986); K.S. Narain, M.H. Sarmadi, and E. Witten, Nucl. Phys. **B279**, 369 (1987).

[6] P. Ginsparg, Phys. Rev. D **35**, 648 (1987).

[7] N. Lerche, D. Lüst, and A.N. Schellekens, Nucl. Phys. **B287**, 477 (1987).

[8] L. Dixon, J. Harvey, C. Vafa, and E. Witten, Nucl. Phys. **B261**, 651 (1985); **B274**, 285 (1986).

[9] K.S. Narain, M.H. Sarmadi, and C. Vafa, Nucl. Phys. **B288**, 551 (1987).

[10] L. Dixon and J. Harvey, Nucl. Phys. **B274**, 93 (1986).

[11] H. Kawai, D.C. Lewellen, and S.-H.H. Tye, Phys. Rev. Lett. **57**, 1832 (1986); **58**, 429(E) (1987); Nucl. Phys. **B288**, 1 (1987).

[12] I. Antoniadis, C. Bachas, and C. Kounnas, Nucl. Phys. **B289**, 87 (1987); I. Antoniadis and C. Bachas, *ibid.* **B298**, 586 (1987).

[13] S. Chaudhuri, H. Kawai, and S-H.H. Tye, Nucl. Phys. B (to be published).

[14] A.N. Schellekens and N.P. Warner, Nucl. Phys. **B308**, 397 (1988).

[15] F. Gliozzi, J. Scherk, and D. Olive, Nucl. Phys. **B122**, 253 (1977).

[16] D.C. Lewellen, Ph.D. thesis, Cornell University, 1987.

[17] M. Dine, N. Seiberg, and E. Witten, Nucl. Phys. **B289**, 589 (1987).

[18] J.J. Atick and A. Sen, Nucl. Phys. **B292**, 109 (1987).

[19] W. Lerche, B.E.W. Nilsson, and A.N. Schellekens, Nucl. Phys. **B289**, 609 (1987).

[20] N. Jacobson, *Lie Algebras* (Dover, New York, 1962); J.E. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, Berlin, 1972).

[21] J.C. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974).

[22] J.C. Pati, S. Rajpoot, and A. Salam, Phys. Rev. D **17**, 131 (1980).

[23] I. Antoniadis and G.K. Leontaris, Phys. Lett. B **216**, 333 (1989).

[24] H. Georgi and S.L. Glashow, Phys. Rev. Lett. **32**, 438 (1974).

[25] S.M. Barr, Phys. Lett. **112B**, 219 (1982).

[26] J.-P. Derendinger, J.E. Kim, and D.V. Nanopoulos, Phys. Lett. **139B**, 170 (1984).

[27] I. Antoniadis, J. Ellis, J.S. Hagelin, and D.V. Nanopoulos, Phys. Lett. B **194**, 23 (1987).

[28] H. Georgi, in *Particles and Fields—1974 (APS/DPF Williamsburgh)*, proceedings of the 1974 Meeting of the APS Division of Particles and Fields, edited by C.E. Carlson (AIP Conf. Proc. No. 23) (AIP, New York, 1975).

[29] H. Fritzsch and P. Minkowski, Ann. Phys. (N.Y.) **93**, 193 (1975).

[30] S.Rajpoot, Phys. Rev. D **22**, 2244 (1980).

[31] See, for instance, R. Slansky, Phys. Rep. **79**, 1 (1981).