PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Gauge Symmetry in String Theory

P. Goddard

Phil. Trans. R. Soc. Lond. A 1989 **329**, 329-342
doi: 10.1098/rsta.1989.0080

Phil. Trans. R. Soc. Lond. A **329**, 329–342 (1989)
Printed in Great Britain

Gauge symmetry in string theory

By P. Goddard

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, U.K.

An account is given of the way symmetry is incorporated into string theory by using the Frenkel–Kac–Segal mechanism, taken from the representation theory of affine Kac–Moody algebras. The intrinsically quantum mechanical nature of this mechanism is emphasized, and the present stage of development of string theory is compared with the 'old quantum theory'. The corresponding method of incorporating gauge symmetry into superstring theories is discussed and arguments that appear to prevent the construction of realistic theories of this type are reviewed in outline.

1. Introduction

The renaissance of string theory, which started in 1984 with the discovery of the anomaly cancellation by Green & Schwarz (1984), quickly led to the proposal of the heterotic string theory (Gross *et al.* 1985*a*, *b*, *c*). This contained, in particular, two important conceptual developments. Each of these had been adumbrated in previous work but were now realized in a concrete, one might almost say blatant, fashion. Firstly, there was the realization that the left- and right-moving waves on a closed string can be treated to a large degree independently (though they are linked through considerations of modular invariance). This idea was there to some extent in the formulation of the Gliozzi–Olive–Scherk (GSO) projection (Gliozzi *et al.* 1976, 1977), but in the heterotic string it reaches the extremity of having the left-moving modes vibrating in 26 dimensions in some sense, whereas the right-moving modes vibrate in 10, but have more degrees of freedom. (Because the extra dimensions for the left-moving modes are, in a sense we shall explain, intrinsically quantum mechanical, this perhaps somewhat disturbing picture need not be taken too literally.)

Secondly, and at least equally importantly, the heterotic string incorporated gauge symmetry in a new way, by using the Frenkel–Kac–Segal (FKS) mechanism (Frenkel & Kac 1980; Segal 1981). Originally, internal symmetry was incorporated into string theory by using the Chan–Paton procedure (Paton & Chan 1969), which could be applied only to open string theories. It could be pictured as attaching to the ends of the string objects that had no inertia but merely carried the 'colour' quantum numbers of quarks (in the context of building the gauge symmetry of the strong interactions).

Although this procedure enabled Neveu & Scherk to show that non-abelian gauge symmetry could be incorporated inside string theory (Neveu & Scherk 1972), it was not wholly satisfactory for a number of reasons. It meant that closed strings were intrinsically neutral. Perhaps more fundamentally, it seems contrary to the general spirit of string theory to have the charges concentrated at two (or more) points of the string, rather than spread out along its length. The extra consistency claimed for string theories, as opposed to theories of point particles, is usually ascribed to the fact that strings are not all localized at points, and so it would seem somewhat at variance to the general philosophy of string theory to concentrate the charges in this way.

The FKS mechanism, which applies to closed string theories, describes the internal symmetry generators (and so the charges) as the integrals of local densities along the string. Thus the charges are, in a sense, smeared out. The work of Frenkel & Kac (1980), and of Segal (1981), was directed at constructing representations of affine Kac–Moody algebras using the vertex operators that had occurred in the development of dual models, before these models were reinterpreted as string theory. (In fact, such constructions had already occurred in special cases in the physics literature (Halpern 1975; Banks *et al.* 1976).) An attempt to elucidate the FKS mechanism in the context of the formalism commonly used in string theory, and to develop it further was made by Goddard & Olive (1984), who pointed out that this approach suggested that $\mathrm{Spin}(32)/\mathbb{Z}_2$ and $E_8 \times E_8$ might be particularly interesting possibilities for gauge groups. Of course, it was precisely for these gauge groups that Green & Schwarz subsequently found the anomaly cancellation, and for which the heterotic string theories could be constructed.

An important feature of the FKS mechanism, which has not always been stressed as much as it might in the literature, is that it is intrinsically quantum mechanical. We shall discuss this in some detail in §2 (see also Goddard 1987). This is to be contrasted with the Kaluza–Klein mechanism for internal symmetries, which incorporates the symmetry at the classical level. In this mechanism, space-time is enlarged from $\mathbb{R}^{3,1}$ to $\mathbb{R}^{3,1} \times M$, where M is a compact manifold, and the internal symmetries are the isometries of M. In the FKS mechanism, we consider closed strings moving on a space of the form $\mathbb{R}^{3,1} \times T$, a torus of dimension r, say. Classically, this gives rise to an internal symmetry of $U(1)^r$, through a Kaluza–Klein type mechanism, but this symmetry is enhanced in the quantum theory to a non-abelian symmetry group G for a torus T of suitable dimensions, which are intrinsically quantum mechanical. The rank of G is the dimension r of the torus.

In the past four years there has been much frenetic activity in string theory, which has brought the subject to the attention of the wider scientific community, and even beyond, attracting much bemusement and a little derision. All this may be encouraging but it might well be argued that there has been much less conceptual progress in this period of intense activity since 1984 than in the first period of development, from 1968 to 1974 or 1976, when the subject was being pursued by a smaller band of physicists, whose efforts were less welcomed by the community at large. At that time there were major conceptual developments every year (e.g. Veneziano formula, n-point functions, higher-order contributions, Virasoro–Shapiro model, Virasoro conditions, fermionic models, supersymmetry, conformal field theory, no-ghost theorem, string picture, projection operators and the calculation of fermion scattering, connection with non-abelian gauge theory and gravity, GSO projection). In the intervening period, up to 1984, there were also a few major advances in our understanding of string theory, most notably Polyakov's (1981 *a*, *b*) path integral formulation (which made the connection with the methods of field theory much more immediate, thus making string theory more accessible and acceptable to a wider audience, and also facilitated the treatment of curved backgrounds, an essential step in substantiating the claims that string theory provides a successful quantum theory of gravity) and the proving of the space-time supersymmetry of the 10-dimensional theory, then renamed superstring theory (Green & Schwarz 1981).

Among the areas in which there has been conceptual development in the past four years, two related ones might be singled out (leaving aside string field theory, particularly Witten's (1986)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

beautiful approach, the ultimate status of which is unclear so far), namely modular invariance and conformal field theory. In the realization that ideas associated with the modular transformation properties of string partition functions could be useful in the classification of possible consistent string theories dates back to the work of Nahm (1976, 1977), which might have provided a greater stimulus had it not been published at a time when interest in string theory was rapidly diminishing. One of the recent advances in understanding has been to appreciate how crucial these properties are for the consistency of closed string theories, anomaly cancellation for example, and how they parallel considerations recently introduced in the study of the critical behaviour of two-dimensional statistical systems by Cardy (1986).

The mathematical language that unites two-dimensional critical phenomena with string theory is that of conformal field theory. In the context of string theory, the conformal field theory is the theory of the string degrees regarded as at theory defined over the two-dimensional world-sheet of the string. Elucidating the relation between the properties of the string in space-time and the properties of the conformal field theory defined by considering the dynamics of its degrees of freedom over its world-sheet, viewed as a two-dimensional 'world' itself, has been another area in which our understanding has increased recently. For example, a correspondence has been found between $N = 2$ superconformal invariance on the world-sheet and space-time supersymmetry (D. Friedan, A. Kent, S. Shenker & E. Witten, unpublished work 1987; Banks *et al.* 1988). In fact a programme has been outlined in which the conformal structure of string theory has put forward as the essential concept, with space and time being in some sense derived phenomenological concepts (Friedan & Shenker 1986) (though, a little paradoxically, this sophisticated approach might be regarded from a certain point of view as a step backwards, because realizing that the string world-sheet was to be thought of as embedded in space-time (Y. Nambu, unpublished work 1970; Goto 1971) was a very significant advance in the development of string theory). In any case it is widely expected that further advances in string theory will come from a progress in our understanding of conformal field theory.

In our review of some aspects of gauge symmetry in string theory, the ideas relating to conformal field theory and, although we shall not be able to discuss it in detail, modular invariance, will play an important role. Under the heading of conformal field theory may be subsumed considerations relating to the infinite-dimensional Lie algebras that enable us to understand much of the symmetry structure of string theories. (For a review of infinite-dimensional algebras in relation to quantum physics see Goddard & Olive (1986).) In §2 I shall discuss the FKS mechanism for the introduction of gauge symmetry into string theory, emphasizing its quantum-mechanical nature, and draw attention to some potentially paradoxical aspects of it. In §3 I discuss an analogous mechanism for gauge symmetry in superstring theories. This procedure enables me to obtain the observed gauge group, but not the correct representations for the quarks and leptons. I sketch the present obstacles to obtaining a realistic physical this way.

2. The Frenkel–Kac–Segal mechanism

The simplest geometric approach to string theory, and in many ways the one with the most intuitive physical picture, is that of Y. Nambu (unpublished work, 1970) and Goto (1971). This based on the action principle, directly generalizing the prescription that point particles

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

move along geodesics, that classically the string moves in such a way as to extremize the area that it sweeps out in d-dimensional space-time (the area of its 'world-sheet'). Thus one has to find the extremum of the action

$$\mathscr{A} = \frac{T_0}{c}\int[(\dot{x}x')^2 - \dot{x}^2x'^2]^{\frac{1}{2}}\,\mathrm{d}\sigma\,\mathrm{d}\tau, \tag{1}$$

where the world-sheet of the string is the two-dimensional surface $x^\mu(\sigma,\tau)$, $0 \leqslant \sigma \leqslant 2\pi$, $-\infty < \tau < \infty$, and the constant T_0 has been introduced to ensure that $\mathscr{A}$ has the correct dimensions of action; T_0 can be interpreted as the 'rest tension' in the string (Goddard *et al.* 1973). Here we shall only consider closed strings and so we have the condition $x(\sigma,\tau) = x(\sigma+2\pi,\tau)$. The velocity of light has been denoted by c as usual and we use the notation

$$\dot{x} = \partial x/\partial\tau, \quad x' = \partial x/\partial\sigma. \tag{2}$$

If we quantize the simplest motions of the string, which are those in which it rotates, doubled up on itself into a straight line, in a plane, we see that the lowest quantum mechanical modes of the string have a characteristic length,

$$l = \surd(c\hbar/\pi T_0) \tag{3}$$

(which up to a constant could have been deduced on dimensional grounds). If the theory is accommodate gravity on a realistic scale, this length has to be comparable in magnitude with the Planck length

$$l_\mathrm{P} = \surd(G\hbar/c^3) \approx 10^{-33}\ \mathrm{cm}, \tag{4}$$

and then all the massive states of the string will have masses on the scale of the Planck mass $m_\mathrm{P} = \surd(\hbar c/G)$.

This leads to a potential paradox, for which there may be only partial explanations so far. The FKS mechanism, currently thought to be responsible for providing the observed symmetries of nature within the context of a string theory, requires that the string be moving in some of the 'internal' dimensions, i.e. those in excess of the familiar four, on a torus with a particular shape, as we shall discuss in greater detail in this section. The dimensions of this torus are fixed at values which are comparable with the Planck scale (4). Although the theory of a free string moving on such a toroidal background would be consistent in itself, we would not be content with this of course and we would wish to consider a theory of interacting strings. It is an inevitable consequence of the uncertainty principle in a quantum theory containing gravity that these fluctuations should be able to disturb the background. The fact that the fluctuations of the string have energies on a scale $m_\mathrm{P}c^2$ means that there will be fluctuations in the background geometry on a scale l_P, the scale of the torus itself. Given such large fluctuations in the background geometry responsible for it, one would expect the FKS symmetry to disappear, and we cannot consider a torus with dimension large on the Planck scale, because this would not possess the FKS symmetry.

One way round this potential difficulty is to suppose that the interaction between strings to be very weak, so that the shape providing the symmetry is disturbed very little. This would be a valid approach in string perturbation theory, sufficing perhaps to give a mathematical definition of string theory, but realistically we do not expect strings to be weakly coupled on the Planck scale in a realistic theory. Another way round this difficulty is to note that the observed symmetries in the heterotic string theory, for example, are supposed to come from the extra 16 dimensions available to the left-moving modes but not to the right-moving ones. This

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

provides us with a rigid structure, incapable of deformation, like the affine Kac–Moody algebra resulting from it, and so unresponsive to the fluctuations we have been discussing. This is a potentially mathematically consistent answer to the 'paradox', but it might be thought not to be completely convincing physically, even if these extra 16 dimensions are not to be taken too seriously. It sounds rather like introducing a rigid wall into a physical problem, and this is procedure which, though superficially consistent, can often lead to inconsistencies on closer examination.

If we are not content with this explanation, we might regard the 'paradox' as rather reminiscent of that associated with the stability of the Bohr orbits of electrons in the old quantum theory. Bohr's original quantization conditions might be seen as analogous to rigidly fixing the dimensions of the torus in terms of multiples of l, the string length. Just as the Bohr orbits could be destroyed, within the pseudo-classical framework of their formulation, by allowing the electron to radiate (thus removing the explanation of the observed discrete spectral lines, which was their *raison d'être*) so the torus can be destroyed by allowing the strings to interact with their background. If there is a real difficulty here, it may be that its resolution will require a conceptual revolution, associated with understanding physical phenomena on length scales of the order of l_P, just as quantum theory required a revolution associated with the description of phenomena with action comparable with $\hbar$. Because string theory has introduced a third, and *presumably* the last, dimension-bearing constant, l (after c and $\hbar$), it would be surprising, and perhaps disappointing, if it did not require a complete revision of our concepts. These conceptual revolutions (relativity and quantum mechanics) have been successively more drastic and have taken us further from the world of common experience. They have been less immediately accessible. This should be even more so in this third case. A major aspect of such a revolution may well be the abolition of the space-time continuum, and there have been various suggestions along these lines in the context of string field theory (Witten 1986) and conformal field theory (Friedan & Shenker 1986), but it may be that something very much more radical, making a departure from quantum theory as well, is required. A more recent extremely interesting suggestion for the nature of a theory on length scales less than l_P, both radical and plausible, is provided by Witten (1988).

As a preliminary to describing the FKS mechanism, we shall now review in outline the quantization of the string (Goddard *et al.* 1973). Choosing orthonormal coordinates σ, τ such that

$$\dot{x}^2 + x'^2 = 0, \quad x'\dot{x} = 0, \tag{5}$$

the equation of motion of the string, which follows from extremizing the action (1), is

$$\ddot{x} = x''. \tag{6}$$

The momentum conjugate to $x^\mu(\sigma,\tau)$ is

$$\pi^\mu(\sigma,\tau) = (T_0/c)\,\dot{x}^\mu(\sigma,\tau), \tag{7}$$

so that the canonical commutation relations are

$$[x^\mu(\sigma,\tau), \pi^\nu(\sigma',\tau)] = i\hbar\delta(\sigma-\sigma')\,\eta^{\mu\nu}, \tag{8}$$

where $\eta^{\mu\nu}$ denotes the space-time metric (taken to be $\eta^{00} = -1$; $\eta^{jj} = 1$ for each j, $1 \leqslant j \leqslant d-1$; $\eta^{\mu\nu} = 0$, $\mu \neq \nu$).

Consider first a closed string moving in $\mathbb{R}^{25,1}$. (For the bosonic string theory that we are

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

considering, the space has to be 26-dimensional for consistency of the theory.) Then, the general solution to (6) can be written

$$x(\sigma,\tau)/l = \tfrac{1}{2}X_{\mathrm{L}}(\mathrm{e}^{\mathrm{i}(\tau+\sigma)}) + \tfrac{1}{2}X_{\mathrm{R}}(\mathrm{e}^{\mathrm{i}(\tau-\sigma)}), \tag{9}$$

where

$$X_{\mathrm{L}}(z) = q - \mathrm{i}p_{\mathrm{L}} \ln z + \mathrm{i} \sum_{n\neq 0} \frac{\alpha_n^{\mathrm{L}}}{n} z^{-n} \tag{10a}$$

and

$$X_{\mathrm{R}}(z) = q - \mathrm{i}p_{\mathrm{R}} \ln z + \mathrm{i} \sum_{n\neq 0} \frac{\alpha_n^{\mathrm{R}}}{n} z^{-n}, \tag{10b}$$

and the factor of l^{-1} has been included to leave q, p and α dimensionless.

The periodicity boundary condition $x(\sigma,\tau) = x(\sigma+2\pi)$, applied to all the components of x, then implies

$$p_{\mathrm{L}} = p_{\mathrm{R}} = p, \tag{11}$$

say. The canonical commutation relations (8) are then equivalent to

$$[\alpha_m^{\mathrm{R}\mu}, \alpha_n^{\mathrm{R}\nu}] = [\alpha_m^{\mathrm{L}\mu}, \alpha_n^{\mathrm{L}\mu}] = m\delta_{m,-n}\eta^{\mu\nu}, \quad [\alpha_m^{\mathrm{L}\mu}, \alpha_n^{\mathrm{R}\mu}] = 0, \tag{12}$$

$$[q^{\mu}, p^{\nu}] = \tfrac{1}{2}\mathrm{i}\eta^{\mu\nu}. \tag{13}$$

This means that p^{μ} has the representation

$$p_{\mu} = \tfrac{1}{2}\mathrm{i}\,\partial/\partial q^{\mu}. \tag{14}$$

We consider the case in which the whole space is not $\mathbb{R}^{25,1}$ but some of the directions are compact. Consider the simplest case in which one of the directions, x^1, is compactified into a circle of radius $R = la$ (so that a is a dimensionless parameter). Because the string might wind round the circle some integral number of times, m, say, the periodicity condition for the x^1 component is replaced by

$$x^1(\sigma,\tau) = x^1(\sigma+2\pi,\tau) + 2\pi mR. \tag{15}$$

Note that the integer m, the winding number is a *classical* concept. The condition (15) implies that, rather than (11), $p_{\mathrm{L}}^1, p_{\mathrm{R}}^1$ satisfy

$$p_{\mathrm{L}}^1 - p_{\mathrm{R}}^1 = 2ma, \tag{16}$$

and the corresponding quantum conditions are

$$[q^1, p_{\mathrm{L}}^1] = [q^1, p_{\mathrm{R}}^1] = \tfrac{1}{2}\mathrm{i}, \tag{17}$$

which have the representation

$$p_{\mathrm{L}}^1 = \tfrac{1}{2}\mathrm{i}\,\partial/\partial q^1 + ma, \quad p_{\mathrm{R}}^1 = \tfrac{1}{2}\mathrm{i}\,\partial/\partial q^1 - ma. \tag{18}$$

Now the wave function of the string has to be single valued on the circle and so its dependence on q^1 must be a sum of terms of the form

$$\mathrm{e}^{\mathrm{i}nq^1/a} \tag{19}$$

for some integer n. This is the familiar quantization of momentum resulting from the periodicity on a circle. The quantum number n, in contrast to the winding number m, is indeed a *quantum* concept.

If we consider the two parts of the momentum, left-moving and right-moving, as part of a single vector we can write

$$(p_L^1; p_R^1) = (n/2a + ma; n/2a - ma). \tag{20}$$

Because m and n, for apparently very different reasons, are integers, we see that the vector (20) lies on a lattice in $\mathbb{R}^2$. If we introduce an indefinite inner product on this two-dimensional real space, so regarding it as $\mathbb{R}^{1,1}$

$$(p_L^1; p_R^1)^2 \equiv (p_L^1)^2 - (p_R^1)^2 = 2mn, \tag{21}$$

we see that the inner product of any two points of this momentum lattice is an even integer. A lattice with this property is an *even* lorentzian lattice. In fact this lattice is also *self-dual*, that is it is the same as its dual or reciprocal lattice, the lattice consisting of all points having integral inner products with all points of the lattice defined by (20). Such even self-dual lattices can only exist in spaces $\mathbb{R}^{M,N}$ for which $M-N$ is a multiple of 8. Their particular significance in relation to string theory is one of the realizations of the past few years (Frenkel 1985; Goddard & Olive 1984; Narain 1986; for a recent review see Lerche *et al.* 1989).

We note that this lattice is invariant under the replacement

$$a \to 1/2a. \tag{22}$$

This operation interchanges the classical winding number m with the quantum number n. Such a symmetry is very analogous to the symmetry (Goddard *et al.* 1977) between the quantum-mechanical electric charges and the classical magnetic charges that exist in spontaneously broken gauge theories that possess classical solutions, of the 't Hooft–Polyakov type, bearing magnetic charge. The classical and quantum nature of m and n is made clear by writing down the corresponding component of the total left-moving momentum of the string

$$\mathscr{P}_L^1 = 2\pi T_0 l p_L^1/c = mRT_0/c + n\hbar/2R, \tag{23}$$

the first term on the right-hand side being classical and the second quantum. In terms of the original radius R, the replacement (22) is

$$R \to c\hbar/\pi T_0 R, \tag{24}$$

making clear the quantum-mechanical nature of the transformation.

The FKS symmetry occurs at the fixed point of this transformation, namely at $a = \frac{1}{\sqrt{2}}$. At this point the symmetry of the quantum theory of the closed string increases from U(1) to SU(2) (or, more precisely, from $\mathrm{U}(1) \times \mathrm{U}(1)$ to $\mathrm{SU}(2) \times \mathrm{SU}(2)$). Thus the symmetry is indeed intrinsically quantum mechanical.

A signal of the increased gauge symmetry appearing at $a = \frac{1}{\sqrt{2}}$ is the appearance of extra massless states. Suppose we denote states of the string with momentum but no excitations by

$$|k; k_L^1; k_R^1\rangle, \tag{25}$$

where k denotes the momentum, both left- and right-moving, in the non-compact directions and k_L^1 and k_R^1 denote the components in the compact direction. Then they are always massless states of the string

$$\alpha_{-1}^{L1} \alpha_{-1}^{R\mu} |k; 0; 0\rangle, \tag{26}$$

where $\mu \neq 1$ and $k^2 = 0$. (We are using 'mass' in the usual sense of being associated with the non-compact dimensions of the space-time in which the string is moving.) These states correspond to a massless vector particle, the gauge particle associated with the U(1) symmetry. (There is a similar gauge particle, obtained by interchanging left and right, associated with the other U(1) symmetry.) At $a = \frac{1}{\sqrt{2}}$, extra massless states appear:

$$\alpha_{-1}^{R\mu}|k; \pm\sqrt{2}; 0\rangle, \tag{27}$$

where again $\mu \neq 1$ and $k^2 = 0$. These are the extra gauge bosons corresponding to the enhancement of the U(1) symmetry to SU(2) through the FKS mechanism.

The states (26) and (27) satisfy the constraints required of physical states of the closed string in this covariant formalism

$$L_n^L|\psi\rangle = 0, \quad L_n^R|\psi\rangle = 0, \quad n > 0, \tag{28a}$$

$$(L_0^L + L_0^R)|\psi\rangle = 2|\psi\rangle, \tag{28b}$$

$$L_0^L|\psi\rangle = L_0^R|\psi\rangle, \tag{29}$$

where L_n^L is defined in terms of α^L and $\alpha_0^L = p^L$, and L_n^R is defined in terms of α^R and $\alpha_0^R = p^R$, by equations of the form:

$$L(z) \equiv \sum_{n=-\infty}^{\infty} L_n z^{-n-2} = \frac{1}{2}{:}\left(\mathrm{i}\frac{\mathrm{d}X}{\mathrm{d}z}\right)^2{:}, \tag{30}$$

$$L_n = \frac{1}{2}\sum_{m=-\infty}^{\infty} {:}\alpha_m \alpha_{n-m}{:}. \tag{31}$$

The colons denote normal ordering with respect to the operators α_m. The L_n defined by these equations satisfy the Virasoro algebra

$$[L_m, L_n] = (m-n) L_{m+n} + \tfrac{1}{12}cm(m^2-1)\,\delta_{m,-n}, \tag{32}$$

with the *central charge* $c = 26$, the dimension of space-time required for consistency.

The considerations we have discussed here generalize from compactification on a circle to compactification on some more general torus. We can construct such a torus by identifying points in space related by displacements lying on some lattice Λ, whose points span the dimensions that are being compactified. This means that we view space-time as a quotient $\mathbb{R}^{25,1}/\Lambda$. In the simple example of compactification on a circle, we have $\Lambda = 2\pi R\mathbb{Z}$. Generalizing $R = la$, it is convenient to write $\Lambda = 2\pi l\Gamma$. Then, instead of (16) we have

$$p_L - p_R = 2\gamma \quad \text{for some} \quad \gamma \in \Gamma. \tag{33}$$

For the string's wave function to be single valued, it has to be a sum of terms whose dependence on q is of the form $e^{i\beta q}$, where β has integral inner product with all the points of the lattice Γ. This means that the projection β' of β onto the space spanned by Γ has to lie on the lattice Γ^* dual to Γ. It then follows that the corresponding projections p'_L, p'_R, of p_L and p_R, respectively, can be put together to form the vector

$$(p'_L; p'_R) = (\tfrac{1}{2}\beta' + \gamma, \tfrac{1}{2}\beta' - \gamma), \quad \gamma \in \Gamma, \quad \beta' \in \Gamma^*. \tag{34}$$

This again forms an even self-dual lattice, in $\mathbb{R}^{M,M}$, where $M = \dim\Gamma$ (Englert & Neveu 1985; Narain 1986).

As we saw in the case of compactification on the circle, for particular quantum-mechanical values of the dimensions of such a torus, the extra massless particles appear, indicating enhancement of symmetry. The generators of such symmetries can be constructed from the vertex operators describing the emission of these massless vector particles. For definiteness, consider massless particles whose momentum in the compact dimensions is entirely associated with the left-moving modes on the string and whose momentum in the uncompactified dimensions is associated with the right-moving modes. The vertex describing the emission of such a particle will consist of the product of a right-moving part and a left-moving part. The right-moving part is of the form,

$$\mathrm{i}\,\mathrm{d}X^{\mu}_{\mathrm{R}}/\mathrm{d}z\!:\!\mathrm{e}^{\mathrm{i}kX_{\mathrm{R}}}\!:. \tag{35}$$

If we denote the left-moving part of the vertex by $T^a(z)$, under quite general considerations the moments of $T^a(z)$ will satisfy an affine Kac–Moody algebra, $\hat{g}$,

$$[T^a_m, T^b_n] = \mathrm{i}f^{abc}T^c_{m+n} + km\delta_{m,-n}\delta^{ab}, \tag{36}$$

where

$$T^a(z) = \sum_{n=-\infty}^{\infty} T^a_n z^{-n-1}. \tag{37}$$

The moments will also satisfy

$$[L_m, T^a_n] = -nT^a_{m+n}. \tag{38}$$

The zero moments T^a_0 satisfy the Lie algebra of a compact group, with Lie algebra g,

$$[T^a_0, T^b_0] = \mathrm{i}f^{abc}T^c_0, \tag{39}$$

which is the gauge group provided by the FKS mechanism. The quantity k determines the *level* of the affine Kac–Moody algebra. (For further explanation of the FKS mechanism see, for example, Green *et al.* (1987) or Goddard (1987); for a review of Kac–Moody algebras in relation to their applications in physics see Goddard & Olive (1986).)

3. Type II strings

As we have remarked, to get consistency in the bosonic string theory discussed in the last section we need $c = 26$ in (32). To get a theory in four-dimensional space-time, we need to compactify 22 of the dimensions. Even compactifying these, on a suitable torus for example, we would still be left with a tachyonic state (a state of negative squared mass) as a bar to any sort of realistic physical interpretation. At present there are basically two sorts of ways of avoiding tachyons. Firstly, there are the superstring theories, in which fermionic degrees of freedom are added to the bosonic. This reduces the 'critical dimension', in which the theory achieves consistency, from 26 to 10, and the value of c in (32) is now 15, receiving a contribution of one from each bosonic degree of freedom and a half from each fermionic one. We restrict attention here to the closed superstring theories (type II theories), considering the introduction of symmetry by a fermionic analogue of the FKS procedure. Secondly, there are the heterotic string theories, which, as we mentioned in §1, are a sort of hybrid between the superstring and the bosonic string. The gauge symmetry here comes from compactification of the extra bosonic left-moving degrees of freedom, using the FKS mechanism, and so producing, before any symmetry breaking, a rank sixteen gauge group, $E_8 \times E_8$ or $\mathrm{Spin}(32)/\mathbb{Z}_2$.

The large extra number of dimensions in the heterotic string theory results in a much larger gauge symmetry than has so far been observed. After compactification of the 16 extra bosonic degrees of freedom, one is left with a 10-dimensional theory. It is still necessary to compactify a further six to obtain a four-dimensional world. This further compactification has been typically used to break the large symmetry resulting from the FKS mechanism to something more compatible with observation. However, if it were done in a symmetrical way, and it might be that in such a further compactification one is driven to the sort of fixed of a transformation like (22), the symmetry would be further enhanced rather than reduced.

We shall consider here attempts that have taken place in the last two years to incorporate gauge symmetry into type II superstring theories (Bluhm *et al.* 1987, 1988; Kawai *et al.* 1987; Antoniadis *et al.* 1987). These attempts have succeeded in showing that, contrary to previous expectations, it is possible to incorporate non-abelian symmetries into such theories, including ones large enough to contain the gauge group, SU(3) × SU(2) × U(1), of the 'standard model'. At the moment there seems unfortunately to be a firm obstacle to obtaining a realistic spectrum of quarks and leptons (Dixon *et al.* 1987).

In the superstring, in addition to the bosonic degrees of freedom moving each way, described by $X^{\mu}(z)$, there are fermionic degrees of freedom described by fermion fields, $\psi^{\mu}(z)$, $1 \leqslant \mu \leqslant d$, where d is the dimension of space-time. Now the expression for the Virasoro generators has to include both bosonic and fermionic degrees of freedom,

$$L(z) = \frac{1}{2}{:}\left(\mathrm{i}\frac{\mathrm{d}X}{\mathrm{d}z}\right)^2{:} + \frac{1}{2}{:}\frac{\mathrm{d}\psi}{\mathrm{d}z}\psi{:}, \tag{40}$$

and the algebra defining the physical state conditions is extended to include 'super-Virasoro' or 'superconformal' generators,

$$G(z) \equiv \sum_r G_r z^{-r-\frac{3}{2}} = \psi(z)\,\mathrm{i}\frac{\mathrm{d}X}{\mathrm{d}z}. \tag{41}$$

The final algebra is now

$$[L_m, L_n] = (m-n)\,L_{m+n} + \tfrac{1}{12}cm(m^2-1)\,\delta_{m,-n}, \tag{42a}$$

$$[L_m, G_r] = (\tfrac{1}{2}m - r)\,G_{m+r}, \tag{42b}$$

$$\{G_r, G_s\} = 2L_{r+s} + \tfrac{1}{3}c(r^2-\tfrac{1}{4})\,\delta_{r,-s}. \tag{42c}$$

Here, m, n are integers and r, s can be either half-odd-integers (Neveu–Schwarz case) or integers (Ramond case), the former corresponding to excitations being space-time bosons and the latter to their being space-time fermions. In (42), $c = 15$ for consistency, corresponding to a space-time dimension of 10, with a contribution of 1 from each boson and $\frac{1}{2}$ from each fermion. The fermion field has an expansion of the form

$$\psi(z) = \sum_r \psi_r z^{-r-\frac{1}{2}}. \tag{43}$$

The physical state conditions take the form

$$L_n|\psi\rangle = 0, \quad G_r|\psi\rangle = 0, \quad n, r > 0, \tag{44a}$$

$$(L_0 - \tfrac{1}{2})\,|\psi\rangle = 0. \tag{44b}$$

In the case of the type II superstring, there is an algebra and set of conditions like this for both the left- and the right-moving modes. (For the heterotic string, the left-moving modes have only bosonic degrees of freedom, carrying the internal symmetry, $c = d = 26$, whereas the right-moving modes have $d = 10$, $c = 15$.) To get a four-dimensional theory, one must compactify six of the 10 dimensions, regarding them as internal.

Initially, it was thought that type II strings could only have U(1) gauge symmetry. Then it was realized that for suitable correlations of left- and right-moving modes, at the critical radius, one could get $SU(2)^6$ symmetry (Bluhm & Dolan 1986), which was non-abelian but not very interesting phenomenologically. Finally, it was seen how this could be generalized to the groups $SU(2) \times SU(4)$ and $SO(5) \times SU(3)$, or subgroups thereof. To describe this procedure for incorporating gauge symmetry into the superstring, it is best to replace the six internal boson degrees of freedom by 12 internal fermions, making 18 internal fermions in all. Then the part of (40) corresponding to the six internal dimensions can be rewritten

$$L_{\text{int}}(z) = \tfrac{1}{2}{:}\psi^a \, \mathrm{d}\psi^a/\mathrm{d}z{:}, \tag{45}$$

where the sum is over $a = 1$ to 16. The method of introducing gauge symmetry is to change the internal part of (40) to

$$G_{\text{int}}(z) = -\tfrac{1}{6}\mathrm{i} f_{abc} \psi^a(z)\, \psi^b(z)\, \psi^c(z), \tag{46}$$

where f_{abc} are the structure constants of an 18-dimensional semisimple Lie algebra, g, which gives the possibilities $SU(2)^6$, $SU(2) \times SU(4)$ and $SO(5) \times SU(3)$ (Bluhm *et al.* 1987). It can be shown (Goddard & Olive 1985) that (45) and (46) satisfy a super-Virasoro algebra with a central charge $c_{\text{int}} = 9$, and the theory can be arranged so that it has a corresponding gauge symmetry whose generators T_0^a are moments of

$$T^a(z) = -\tfrac{1}{2}\mathrm{i} f_{abc} \psi^b(z)\, \psi^c(z). \tag{47}$$

This symmetry can easily be broken from g to $g' \subset g$, where g/g' is a symmetric space. It is in fact possible to obtain a type II theory with a gauge symmetry group which is any subgroup of $SU(2)^6$, $SU(4) \times SU(2)$, $SO(5) \times SU(3)$, $SU(3) \times SU(2)^2$, $SU(3)^2$ or G_2 (Dixon *et al.* 1987).

In many ways these models appear very attractive. It is possible to obtain representation contents which are tantalizingly close to those observed in nature (Bluhm *et al.* 1988). The states one focuses on are those that are massless. According to the currently accepted interpretation of string theory, it is only these states that should be directly physically observable in practice, because all other states will have masses of the order of m_P, far above accessible energies. The massless states will have to acquire their masses as the result of some symmetry breaking, often ascribed to some conjectured non-perturbative effect. It seems not too difficult to construct models that seem fairly natural, and in which the observed states and few others occur, but as the massless level and the first excited level rather than all at the massless level. It is difficult to know what significance to attach to these 'near misses'. In fact there is the rather general argument of Dixon *et al.* (1987) to the effect that it is impossible to construct a type II model with a gauge group containing that of the standard model, $SU(3) \times SU(2) \times U(1)$, with massless states transforming non-trivially under both $SU(3)$ and $SU(2)$, as we need for certain of the quark states in nature.

As a preliminary to sketching this argument, note that if we have an affine algebra as in (36) and (38), we can place a lower bound on the value of the central charge of the Virasoro algebra. To do this we make use of the Sugawara construction of a Virasoro algebra out of an affine Kac–Moody algebra, defining

$$\mathscr{L}_n^g = \frac{1}{2k + Q^g} \sum_m {}^{\times}_{\times} T^a_m \, T^a_{n-m} {}^{\times}_{\times}, \tag{48}$$

where the normal ordering operation denoted by the crosses is defined by

$$\left.\begin{aligned} {}^{\times}_{\times} T^a_m \, T^a_n {}^{\times}_{\times} &= T^a_m \, T^a_n, \quad \text{if } n \geqslant 0, \\ &= T^a_n \, T^a_m, \quad \text{if } n \leqslant 0, \end{aligned}\right\} \tag{49}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

the quantity Q^g is the quadratic Casimir operator of g in the adjoint representation,

$$f^{abc}f^{abd} = Q^g\delta^{cd}. \tag{50}$$

The $\mathscr{L}^g$ satisfy the Virasoro algebra with central charge

$$c^g = 2k\dim g/(2k+Q^g) \tag{51a}$$

$$= x\dim g/(x+h^g). \tag{51b}$$

The number $x = 2k/\Psi^2$, where Ψ is a long root of the Lie algebra g, is called the *level* of the representation of $\hat{g}$; $h^g = Q^g/\Psi^2$ is called the dual Coxeter number of g. Both x and h^g have to be integers. If we consider

$$K_n = L_n - \mathscr{L}^g_n, \tag{52}$$

we obtain a Virasoro algebra with a value of the central charge $c^K = c - c^g$. Because the representation of the Virasoro algebra K is easily seen to be unitary, at least if we restrict attention to L_{int}, and with the spectrum of K_0 bounded below, it follows that $c^K \geqslant 0$ and so

$$c_{\text{int}} \geqslant c^g. \tag{53}$$

This can be viewed as an example of the coset construction (Goddard *et al.* 1985, 1986).

In the case of type II theories, the gauge particles could come from either both the Neveu–Schwarz sectors of both the left- and right-moving modes, or the Ramond sector of both modes. (The particles that come from the Neveu–Schwarz sector of one and the Ramond sector of the other are space-time fermions.) Assuming that they come from the Neveu–Schwarz sector of both, let us consider the gauge symmetry resulting from the construction (46) applied to the left-moving modes. There are in superstring theories two parts to the vertex describing such a particle, which in this case consists of a bosonic part, $T^a(z)$, and a fermionic part, $\chi^a(z)$. One can show on quite general grounds that the moments of these satisfy

$$[T^a_m, T^b_n] = \mathrm{i}f^{abc}T^c_{m+n} + km\delta_{m,-n}\delta^{ab}, \tag{54a}$$

$$[T^a_m, \chi^b_r] = \mathrm{i}f^{abc}\chi^c_{m+r}, \tag{54b}$$

$$\{\chi^a_r, \chi^b_s\} = \delta_{r,s}\delta^{a,b}, \tag{54c}$$

a super-Kac–Moody algebra. Given such a construction, we can perform a subtraction (rather as in the coset construction), and set

$$\tilde{T}^a(z) = T^a(z) - \tfrac{1}{2}\mathrm{i}f^{abc}\chi^b(z)\,\chi^c(z). \tag{55}$$

Then $\tilde{T}^a(z)$ will provide a representation of the affine Kac–Moody algebra (54*a*) which commutes with $\chi^b(z)$. We can then construct two commuting Virasoro algebras, the Sugawara construction applied to $\tilde{T}^a(z)$, $\tilde{\mathscr{L}}^g(z)$, and the Virasoro algebra, $L^\chi(z)$, associated with the free fermion fields $\chi^a(z)$. Again

$$K(z) = L_{\text{int}}(z) - \tilde{\mathscr{L}}^g(z) - L^\chi(z) \tag{56}$$

is a Virasoro algebra, commuting with both $T^a(z)$ and $\chi^b(z)$, and with central charge $c^K = c_{\text{int}} - \tilde{c}^g - \frac{1}{2}\dim g$, where $\tilde{c}$ is the central charge for $\tilde{\mathscr{L}}^g$ and $\frac{1}{2}\dim g$ is that for L^χ. Thus it follows that

$$c_{\text{int}} \geqslant \tilde{c} + \tfrac{1}{2}\dim g. \tag{57}$$

We are interested in the case $g = su(3) \oplus su(2) \oplus u(1)$, for which $\dim g = 12$, and

$$\tilde{c}^{g} = \tilde{c}^{su(3)} + \tilde{c}^{su(2)} + \tilde{c}^{u(1)}, \tag{58}$$

so that

$$c_{\text{int}} \geqslant 10, \tag{59}$$

whereas we should have $c_{\text{int}} = 9$ for a type II theory, as there are six internal dimensions. This is the basic contradiction, or obstacle to getting a realistic type II string theory at present.

There are some further points to be made. If the non-abelian gauge symmetry is associated with the left-moving modes, it can be shown that the fermions for which the left-moving modes are of Ramond type cannot be massless (and so correspond to particles that might in practice be observable). Hence, the non-abelian gauge symmetry cannot be divided between the left- and right-moving modes. Moreover, if there is a right-moving U(1) gauge symmetry, the massless fermions can be shown to be neutral under it, so in practice it is of little relevance. The only remaining obvious possibility is that some or all of the gauge particles are of Ramond–Ramond type. Dixon *et al.* also present argument to exclude this, though it is perhaps in this region that the best hope for a chink in the argument lies.

All this said, it is intriguing that just one unit of c, just a '10 % effect', bars us from producing a more nearly realistic type II string theory. The progress of string theory, and our understanding of its structure and interpretation, has been punctuated by 'no-go theorems' of this sort. Usually, the way forward has been to the side (which could perhaps point to heterotic string theory), or to take more seriously what string theory itself is insisting on as its own interpretation. Given that the philosophy of string theory is that, once all the consistency conditions are known, the correct physical theory should be uniquely determined, some reason has in any case to be found why these type II theories are inconsistent, not merely phenomenologically unacceptable. In any case, the delicacy of the discrepancy might suggest that the mechanisms described in this section deserve further consideration.

References

Antoniadis, I., Bachas, C. & Kounnas, C. 1987 *Nucl. Phys.* B **289**, 87.
Antoniadis, I. & Bachas, C. 1988 *Nucl. Phys.* B **298**, 586.
Banks, T., Dixon, L. J., Friedan, D. & Martinec, E. 1988 *Nucl. Phys.* B **299**, 613.
Banks, T., Horn, D. & Neuberger, H. 1976 *Nucl. Phys.* B **108**, 119.
Bluhm, R. & Dolan, L. 1986 *Phys. Lett.* B **169**, 347.
Bluhm, R., Dolan, L. & Goddard, P. 1987 *Nucl. Phys.* B **289**, 364.
Bluhm, R., Dolan, L. & Goddard, P. 1988 *Nucl. Phys.* B **309**, 330.
Cardy, J. L. 1986 *Nucl. Phys.* B **270** [FS18], 445.
Dixon, L., Kaplunovsky, V. & Vafa, C. 1987 *Nucl. Phys.* B **294**, 43.
Englert, F. & Neveu, A. 1985 *Phys. Lett.* B **163**, 349.
Frenkel, I. B. 1985 *Lectures in applied mathematics*, vol. 21, p. 325. American Mathematical Society.
Frenkel, I. B. & Kac, V. G. 1980 *Invent. Math.* **62**, 23.
Friedan, D. & Shenker, S. 1986 *Phys. Lett.* B **175**, 287.
Gliozzi, F., Scherk, J. & Olive, D. 1976 *Phys. Lett.* B **65**, 282.
Gliozzi, F., Scherk, J. & Olive, D. 1977 *Nucl. Phys.* B **122**, 253.
Goddard, P. 1987 In *Superstrings '87* (ed. L. Alvarez-Gaumé *et al.*), pp. 157–189. Singapore: World Scientific.
Goddard, P., Goldstone, J., Rebbi, C. & Thorn, C. B. 1973 *Nucl. Phys.* B **56**, 109.
Goddard, P., Kent, A. & Olive, D. 1985 *Phys. Lett.* B **152**, 88.
Goddard, P., Kent, A. & Olive, D. 1986 *Communs math. Phys.* **103**, 105.
Goddard, P., Nuyts, J. & Olive, D. 1977 *Nucl. Phys.* B **125**, 1.
Goddard, P. & Olive, D. 1984 In *Vertex operators in mathematics and physics*, *Proc. Conf.*, November 10–17, 1983 (ed. J. Lepowsky, S. Mandelstam & I. M. Singer), pp. 51–96. New York: Springer-Verlag.
Goddard, P. & Olive, D. 1985 *Nucl. Phys.* B **257** [FS14], 226.

Goddard, P. & Olive, D. 1986 *Int. J. mod. Phys.* A **1**, 303.
Goto, T. 1971 *Prog. theor. Phys.* **46**, 1560.
Green, M. B. & Schwarz, J. H. 1981 *Nucl. Phys.* B **181**, 502.
Green, M. B. & Schwarz, J. H. 1984 *Phys. Lett.* B **149**, 117.
Green, M. B., Schwarz, J. H. & Witten, E. 1987 *Superstring theory*, vol. 1 (*Introduction*). Cambridge University Press.
Gross, D. J., Harvey, J. A., Martinec, E. & Rohm, R. 1985*a* *Phys. Rev. Lett.* **54**, 502.
Gross, D. J., Harvey, J. A., Martinec, E. & Rohm, R. 1985*b* *Nucl. Phys.* B **256**, 253.
Gross, D. J., Harvey, J. A., Martinec, E. & Rohm, R. 1985*c* *Nucl. Phys.* B **267**, 75.
Halpern, M. 1975 *Phys. Rev.* D **4**, 1684.
Kawai, H., Lewellen, D. C. & Tye, S.-H. H. 1987 *Phys. Lett.* B **191**, 63.
Lerche, W., Schellekens, A. N. & Warner, N. P. 1989 *Phys. Rep.* (In the press.)
Nahm, W. 1976 *Nucl. Phys.* B **114**, 174.
Nahm, W. 1977 *Nucl. Phys.* B **120**, 125.
Narain, K. S. 1986 *Phys. Lett.* B **169**, 41.
Neveu, A. & Scherk, J. 1972 *Nucl. Phys.* B **36**, 35.
Paton, J. E. & Chan, H. M. 1969 *Nucl. Phys.* B **10**, 516.
Polyakov, A. M. 1981*a* *Phys. Lett.* B **103**, 207.
Polyakov, A. M. 1981*b* *Phys. Lett.* B **103**, 211.
Segal, G. 1981 *Communs math. Phys.* **80**, 301.
Witten, E. 1986 *Nucl. Phys.* B **268**, 23.
Witten, E. 1988 *Communs math. Phys.* **117**, 353.