What Are String Theories?

H. NICOLAI

Institut für theoretische Physik, Universität Karlsruhe, 7500 Karlsruhe 1, F.R.G.

ABSTRACT

Recent developments in string and superstring theory are reviewed at an introductory level.

KEYWORDS

Unification of fundamental interactions; strings; superstrings; conformal invariance.

1. INTRODUCTION

String theories have had a varied and curious history. In 1969, Veneziano wrote down a four-particle scattering amplitude that embodied many of the properties that physicists expected a future theory of hadronic interactions to possess (Veneziano, 1969). It did not take long until it was realized that the underlying dynamics, from which the Veneziano formula could be derived, was that of a relativistic string (i.e. an extended object) rather than of a pointlike particle (Nambu, 1970; Susskind, 1970). A short period of intense investigation ensued, and, at that time, several review articles were written¹. However, the further development came gradually to a halt with the advent of QCD as the (probably) correct theory of strong interactions (Fritzsch, Gell-Mann and Leutwyler, 1973). Moreover, it had become clear in 1972 that string theories could only be consistently quantized in 26 dimensions and predicted the existence of a tachyon (Brower, 1972; Goddard and Thorn, 1972). At the

---

¹ These are listed at the beginning of the reference list.

time, this result was perhaps even more puzzling than all the failed attempts to fit the predictions of string theory with hadronic phenomena. The existence of a "critical dimension" appeared as a rather unexpected phenomenon, and it was certainly not clear why a theory that at least shared some qualitative features with hadronic physics should only work in 26 dimensions. Although one could have hoped that some kind of string theory can be derived from QCD - after all, mesons can be thought of as pairs of quarks and anti-quarks bound together by a gluonic "string" - subsequent attempts in this direction have been largely unsuccessful. It appears now that a "stringy" description of hadronic physics will at best be an effective theory rather than a fundamental one.

When this point of view was adopted by the large majority of physicists, the subject went into a period of decline, and only a few hardy people continued to work on string theories. In 1974, Scherk and Schwarz suggested to radically alter the interpretation of string theories: after changing the fundamental energy scale of the theory from 200 MeV (i.e. hadronic physics) to 10^{19} GeV by almost 20 orders of magnitude, string theories should be viewed as providing a fundamental description of all physics rather than just hadronic physics (Scherk and Schwarz, 1974). Their rationale for this proposal was the unavoidable existence of a massless spin-2 particle in the closed string: such a particle was an embarrassment as long as one was dealing with strong interactions, but neatly fitted the properties of the graviton. However appealing as it was, this suggestion did not attract as much attention as it would have deserved; the tachyon was still there, and the critical dimension persisted to be $= 26$. Also, the model did not contain any fermions.

At about the same time, supersymmetry (Wess and Zumino, 1974) and supergravity (Ferrara, Freedman and van Nieuwenhuizen; Deser and Zumino, 1976) were formulated, and many theorists concentrated on these theories in the hope that they might provide a framework for the unification of all interactions. Especially the maximally extended $N=8$ supergravity theory in four dimensions (Cremmer and Julia, 1979; de Wit and Nicolai, 1981) seemed promising. Since supersymmetric theories tend to be less divergent than nonsupersymmetric ones (Wess and Zumino, 1974), one of the hopes was that the local $N=8$ supersymmetry of $N=8$ supergravity could cure the nonrenormalizable infinities of quantum gravity. This hope was thwarted when the existence of $N=8$ supersymmetric counterterms at higher orders was demonstrated². Moreover, all attempts to relate $N=8$ supergravity to known physics failed. On the string side, it was realized in 1976 that a spacetime supersymmetric string in ten dimensions could be manufactured out of the two sectors of the spinning string (Ramond, 1971; Neveu and

---

² To be sure, $N=8$ supergravity is suspected but has not been proven guilty of a divergence at this order (and beyond).

Schwarz, 1971) by a suitable truncation (Gliozzi, Scherk and Olive, 1977). This was the birth of superstring theory.

Superstring theories are superior to the old bosonic string in several respects. The troublesome tachyon is removed from the spectrum by the truncation introduced in (Gliozzi, Scherk and Olive, 1976) and the critical dimension is lowered to $D=10$. Moreover, the closed superstring theories contain $D=10$ supergravity (with or without matter), and one could therefore hope that the diseases of supergravity theory can be cured by "embedding" them in superstrings. Since superstrings contain infinitely many states, the problem of cancelling the divergences of quantum supergravity could now be reexamined in a completely new setting.

However, the formulation of Gliozzi, Olive and Scherk was not suitable to investigate these aspects; in particular, the explicit form of the supersymmetry operators relating bosonic and fermionic states was not known. For this reason, Green and Schwarz (1981) developed a "new formalism" in which the supersymmetry was explicit and which was more appropriate to study the properties of superstrings than the "old formalism". Soon after, they were able to show that certain superstring theories (the "type II theories") were one-loop finite (Green and Schwarz, 1982). Since, in contrast to point-field theories, string theories are probably finite to all orders if they are one-loop finite, these results fuelled hopes that superstring theories might provide the framework for a finite theory of quantum gravity.

A dramatic increase of interest in the subject took place in 1984. One of the major problems had been to find a unified theory that predicted the left-right asymmetry of present day physics. Such a theory must inevitably contain chiral fermions. On the other hand, any theory with chiral fermions is likely to be plagued by anomalies, that is quantum mechanical breakdowns of classical conservation laws. Only in some special cases do these anomalies cancel (for example, they cancel within a standard generation of quarks and leptons), and in higher dimensions, it becomes more and more difficult to achieve the requisite cancellations.

A first step had been taken in 1983 (Alvarez-Gaumé and Witten) where the cancellation of all anomalies was demonstrated for the (chiral) type II B theory, but this theory appeared to have no good phenomenological prospects. However, shortly thereafter, it was shown (Green and Schwarz, 1984) that, for the so-called type I theories, the requirement of anomaly cancellation in ten dimensions singles out two groups, namely $SO(32)$ and $E_8 \times E_8$ ³. Although a string theory with $SO(32)$ symmetry existed, no string theory with $E_8 \times E_8$ symmetry was known. However, soon after this discovery, a new type of string theory was discovered - the "heterotic string" which is a hybrid of the old $D=26$ bosonic

---

³ The cancellation of anomalies for $E_8 \times E_8$ was actually first pointed out by Thierry-Mieg (1984).

and the $D=10$ superstring (Gross et al., 1985). With this construction, it became possible to realize both $SO(32)$ and $E_8 \times E_8$. Subsequently, the compactification of the ten-dimensional supergravity coupled to $E_8 \times E_8$ matter was studied with the result that several generations of chiral fermions could be obtained (Candelas et al., 1985). It was the first time that a theory formulated with the ambitious aim of unifying all fundamental interactions led to low energy "predictions" which were not in immediate conflict with known physics. This fact and the hope that the theory will eventually yield unique predictions for low energy physics have sustained the enthusiasm of many theorists ever since. Many physicists regard the heterotic $E_8 \times E_8$ string as the most promising candidate for the ultimate unification of physics.

In the words of Gross et al. (1984) "Although much work remains to be done there seem to be no insuperable obstacles to deriving all of known physics from the $E_8 \times E_8$ heterotic string".

However, even if this optimistic assessment turns out to be erroneous, there are further reasons to believe in the relevance of string theories. These theories offer much better prospects for curing the ultraviolet divergences of quantum gravity than any known point field theory. The main reason for this is the "explosion of symmetry" that takes place in string theories and may be traced back to the special properties of the conformal group in two dimensions. This is the group of coordinate transformations that leaves distances and angles invariant up to factor (in more physical terms, conformal transformations leave the lightcone invariant). In two dimensions, all analytic functions $f(z)$ have this property because their derivatives $f'(z)$ do not depend on the direction (in contrast, the conformal transformations form the finitely generated Lie-group $SO(D,2)$ in D dimensions for $D>2$). Analytic mappings are generated by the operators

$$L_m = -z^{m+1} \frac{d}{dz} \quad (1.1)$$

which satisfy the commutation relations

$$[L_m, L_n] = (m-n)L_{m+n} \quad (1.2)$$

When reinterpreted in the framework of string theories, each of the operators L_{-m} ($m>1$) gives rise to an ordinary gauge invariance in the embedding space-time, and this "explains" why a string theory has "infinitely more" symmetry than ordinary point field theories (at this point, the reader must accept this assertion on faith; it is by no means obvious, how the two-dimensional symmetries generated by the operators L_{-m} are transmuted into higher dimensional symmetries. See, however, section 6). An important result of analytic function theory is the Riemann mapping theorem (see any standard textbook on complex function theory) which states that for any two connected regions G, G' of the complex plane there exists an analytic function $f(z)$ that maps one onto the other, i.e. $f(G)=G'$. Translated into string theory this means that the phys-

sics of string theories is independent of the shape of the "world-sheet" since this shape may be conformally deformed in an arbitrary manner - like an infinitely stretchable rubber surface, as it were. Thus, the physics only depends on the topology of the "world-sheet" and not on its metrical properties. To make this somewhat intuitive description mathematically precise requires a great deal of advanced mathematics.

There is also a more physical way to understand why string theories can help with the problem of quantum gravity, and this is by analogy with the theory of weak interactions. When theorists first tried to describe the decay of the neutron they did so by use of a Fermi-type Lagrangian where the interaction takes place at a point. It was later found that this theory is not renormalizable, i.e. gives rise to irremovable infinities at higher loop order. Nowadays, we know how to cure the problem: at sufficiently high energies (i.e. about 100 GeV), the pointlike vertex is dissolved and the weak force is mediated by a heavy boson, see Fig.1..

Fig.1. A famous example of how a vertex can be dissolved.

The four-point vertex is thus replaced by a three-point vertex at high energies and this is essentially what makes the new theory renormalizable and predictive beyond the tree approximation.

In string theories, it is conjectured that a similar effect takes place. In Einstein's theory of gravity, one obtains n-point vertices of arbitrarily high order when expanding the action $\sqrt{-g} R$. These give rise to even more severe infinities than the old Fermi theory of weak interactions: the number of infinities increases with each order in perturbation theory. Contrariwise, in string theories, these vertices are dissolved just as in the above example by the exchange of the massive string excitations (with masses quantized in units of 10^{19} GeV). In contrast to Fermi theory one has now infinitely many particles of arbitrarily high mass and "spin" to mediate these forces. In order to cope with this infinity, one needs a unifying principle, and this principle is provided by string theory.

However, there is now a much harder conceptual problem. Dissolving the gravitational vertices involves in some sense dissolving the very notion of space-time itself because the "composite" vertices of Einstein's theory are themselves derived from an action which is based on geometrical considerations of the structure of space-time. In string theories, this structure is replaced by something more fundamental but it is unknown what the new principle could be. We are still accustomed to thinking of strings as moving in a flat spacetime background, but it is clear that this picture can only serve as a "crutch" towards obtaining a better understanding. In particular, we must eventually abandon the conventional notion of spacetime which should be outcome rather than input in a complete formulation of string theories.

Here, we will explain some of the basic features of string theories to the non-expert reader. More emphasis will be placed on points of principle, and the choice of presentation will reflect the author's bias as to what might survive of string theories even if the attempts to relate the currently most popular model to known physics should fail. In any case, it is hoped that this article will convey some of the excitement these models have created among theorists. The organization of this article is as follows. In section 2, we treat the classical string in analogy with the relativistic point particle. The quantization of (bosonic) string theories, the emergence of the critical dimension and the spectrum will be discussed in section 3. Section 4 is intended as an elementary introduction to superstring theory, while section 5 is devoted to a more pictorial than mathematical description of strings in interaction. The last section is meant to be an "appetizer" for those readers who want to continue with their study of strings: it contains a short discussion of some of the topics that are currently under investigation by those on the forefront of research. Finally, we have included a list of some relevant references without aiming at completeness (which would be impossible anyhow); the interested reader is invited to have a look at some recent issues of the relevant journals for an "entrée" into the most recent literature.

2. THE CLASSICAL THEORY

The classical theory of the relativistic string can be developed in almost complete analogy with the classical theory of a relativistic point particle moving through space-time. The world line of such a particle is given by a function $x^\mu = x^\mu(\tau)$, see Fig.2. below.

Fig.2. The world-line of a relativistic point particle

We will leave the dimension of the space-time, in which the motion takes place, arbitrary and also assume the embedding space time R^D to be flat. To describe the dynamics of this particle and to determine its trajectory we need an action, and the simplest choice is simply the "length" ⁴ of the trajectory. Thus, we put

$$S = \text{"length"} = -m \int_a^b \sqrt{-\dot{x}^\mu \dot{x}_\mu} d\tau \quad (2.1)$$

where the parameter m has the dimension $\text{mass} = [\text{cm}]^{-1}$ and is needed to render the action dimensionless. A most important feature of (2.1) is its invariance under reparametrizations $\tau \rightarrow \tau'(\tau)$: the physics should not depend on how the trajectory is parametrized. The positivity of $-\dot{x}^2$ in the integral is equivalent to the requirement that the particle should not move faster than at the speed of light. With (2.1), we can calculate the canonical momenta p^μ which are associated with the coordinates $x^\mu(\tau)$; we get

$$p^\mu = \frac{\partial}{\partial \dot{x}_\mu} \left[-m \sqrt{-\dot{x}^2} \right] = m(-\dot{x}^2)^{-1/2} \dot{x}^\mu \quad (2.2)$$

These momenta are not independent but rather obey the constraint

$$p^2 + m^2 = 0 \quad (2.3)$$

as one can straightforwardly verify from (2.2). This is the well-known "dispersion relation" of a relativistic point particle. The fact that the canonical momenta are constrained complicates the Hamiltonian formalism somewhat; it is a reflection of the invariance of (2.1) under reparametrizations⁵. In the quantum theory, (2.3) becomes a constraint on the physical states and after the replacement $p^\mu \rightarrow i\partial/\partial x^\mu$, it is nothing but the Klein Gordon equation. This method of deriving the Klein Gordon equation from the classical theory of a relativistic point-particle may seem unusual, but it is possible to derive and develop all of quantum field theory on this basis.

⁴ This is not the Euclidean length but rather a Minkowskian length in a space of signature $(-+...+)$.

⁵ The Hamiltonian formalism with constraints has been developed by Dirac (1950).

Of course, we are more used to a formulation in terms of second quantization, i.e. involving quantum fields, but the merits of the above method have been recognized only recently in connection with string theories where the analog of the second quantized formulation is still being developed.

The classical physics of strings can be described in an analogous fashion. The basic object is now a string, that is an extended object, rather than a point. During its motion, it sweeps out a "world-sheet" rather than a world-line. To parametrize this world sheet, we need an extra parameter σ which conventionally is taken to lie in the interval $[0, \pi]$. Hence the motion is completely described by the functions $x^\mu = x^\mu(\sigma, \tau)$. Strings come in two varieties, namely "open" and "closed", which are distinguished through their boundary conditions.

Fig.3. Open and closed strings.

Again the dimension of the embedding space R^D is left arbitrary for the moment; it will, however, be restricted in the quantized theory in contrast to the case of the point particle. To describe the dynamics of the string, we again make the simplest choice, namely (Nambu, 1970)

$S = \text{"Surface Area"}$

$$= - \frac{1}{2\pi\alpha'} \int \sqrt{(\dot{x}x')^2 - \dot{x}^2 x'^2} \, d\sigma d\tau \quad (2.4)$$

where $\dot{x}^\mu = \frac{\partial x^\mu}{\partial \tau}$, $x'^\mu = \frac{\partial x^\mu}{\partial \sigma}$ and the parameter α' has dimension $[\text{mass}]^{-2} = [\text{cm}]^{+2}$; the quantity $\frac{1}{2\pi\alpha'}$, (of dimension $[\text{cm}]^{-2}$) is called the string tension. The expression (2.4) is invariant under reparametrizations $\sigma \rightarrow \sigma'(\sigma, \tau)$, $\tau \rightarrow \tau'(\sigma, \tau)$; the positivity of the integrand means that the strings moves no faster than at the speed of light at any of its points. The reparametrization invariance of (2.4) plays an even more important role than in the case of a point particle. In particular, it allows us to choose an "orthonormal" gauge where

$$\dot{x}^2 + x'^2 = 0, \quad \dot{x} \cdot x' = 0 \quad (2.5)$$

(2.5) simply means the following: if we cover the world sheet by a mesh of lines, the lines of constant σ and τ will intersect at right angles everywhere, see Fig.4..

Fig.4. The orthonormal gauge.

At this point, one can already see why two-dimensional surfaces (and therefore strings) are so special as opposed to higher dimensional objects (corresponding to membranes, etc.). The gauge choice (2.5) does not fix the gauge completely, but there is a huge residual invariance which preserves (2.5)! To find it, we use

$$\frac{\partial x^\mu}{\partial \sigma} = \frac{\partial x^\mu}{\partial \sigma'} \frac{\partial \sigma'}{\partial \sigma} + \frac{\partial x^\mu}{\partial \tau'} \frac{\partial \tau'}{\partial \sigma} \quad (2.6)$$

$$\frac{\partial x^\mu}{\partial \tau} = \frac{\partial x^\mu}{\partial \sigma'} \frac{\partial \sigma'}{\partial \tau} + \frac{\partial x^\mu}{\partial \tau'} \frac{\partial \tau'}{\partial \tau}$$

and require that (2.5) be valid also in terms of the new coordinates (σ', τ') . After a little calculation, one finds that this implies

$$\frac{\partial \sigma'}{\partial \tau} = \frac{\partial \tau'}{\partial \sigma}, \quad \frac{\partial \sigma'}{\partial \sigma} = \frac{\partial \tau'}{\partial \tau} \quad (2.7)$$

Introducing complex variables $z \equiv \sigma + i\tau$, $z' \equiv \sigma'(\sigma, \tau) + i\tau'(\sigma, \tau) = f(z)$, we see that (2.7) is equivalent to $(\bar{z} \equiv \sigma - i\tau)$

$$\frac{\partial f}{\partial \bar{z}} = \frac{1}{2} \left[\frac{\partial}{\partial \sigma} + i \frac{\partial}{\partial \tau} \right] f(\sigma, \tau) = 0 \quad (2.8)$$

Therefore, (2.7) is equivalent to the well-known Cauchy Riemann equations telling us that $f(z)$ is an analytic function. Now, the set of analytic functions constitutes a huge class of transformations of the complex plane, or, more generally, of a two-dimensional surface (possibly with handles and holes). This class is much larger than the corresponding set of conformal transformations in dimensions higher than two simply because there is no analog of analytic function theory in higher dimensions. There, the conformal transformations are given by a finite set of functions. In more technical terms, for $D > 2$, the conformal transformations (which preserve angles and distances) form the Lie-group $SO(D, 2)$ which is generated by finitely many transformations whereas, for $D = 2$, the group of analytic transformations has infinitely many generators, see (1.1).

Application of the canonical Hamiltonian formalism to the Lagrangian that one extracts from (2.4) leads to the same difficulties as for the point particle: owing to the reparametrization invariance of (2.4) the canonical momenta are constrained (there are now infinitely many of them, one corresponding to each $x^\mu(\sigma)$). Rather than discuss this in detail (see however Scherk, 1975) we now proceed directly to the equations of motion that follow from (2.4). To be able to drop surface terms, we impose the following boundary conditions

$$\begin{aligned} x'(0) &= x'(\pi) = 0 && \text{(open string)} \\ x(0) &= x(\pi). && \text{(closed string)} \end{aligned} \quad (2.9)$$

Varying (2.4) directly leads to some rather complicated equations which can be considerably simplified by use of (2.5). In this gauge, the equation of motion of the string is nothing but the free wave equation in two dimensions, namely

$$\ddot{x}^\mu - x''^\mu = 0. \quad (2.10)$$

This is now easy to solve; we get ⁶

$$x^\mu(\sigma, \tau) = q^\mu + p^\mu \tau + i \sum_{n \neq 0} \frac{1}{n} \alpha_n^\mu \cos n\sigma e^{-in\tau} \quad (2.11)$$

for the open string and

$$x^\mu(\sigma, \tau) = q^\mu + p^\mu \tau + \frac{i}{2} \sum_{n \neq 0} \frac{1}{n} \left[\alpha_n^\mu e^{-2in(\tau-\sigma)} + \bar{\alpha}_n^\mu e^{-2in(\tau+\sigma)} \right] \quad (2.12)$$

for the closed string. In both (2.11) and (2.12), the term $q^\mu + p^\mu \tau$ describes the center of mass motion of the string. The other modes describe its internal motions (vibrations). Note that the closed string contains twice as many modes as the open string corresponding to the left and right moving waves on the string.

To incorporate the constraints (2.5), we note that for the open string ($\alpha_0^\mu = p^\mu$)

$$\dot{x}^\mu \pm x'^\mu = \sum_{n=-\infty}^{\infty} \alpha_n^\mu e^{-in(\tau \pm \sigma)} \quad (2.13)$$

and therefore (2.5) is equivalent to

$$(\dot{x} \pm x')^2 = 2 \sum_{m=-\infty}^{\infty} L_m e^{-im(\tau \pm \sigma)} = 0 \quad (2.14)$$

where we have defined the Fourier modes

$$L_m = \frac{1}{2} \sum_{n=-\infty}^{\infty} \alpha_{m-n}^\mu \alpha_{n\mu}. \quad (2.15)$$

Thus, (2.5) holds if $L_m = 0$ for all m . For the closed string, there is another $\bar{L}_m$ associated with the $\bar{\alpha}$ -modes. It is instructive to analyze the L_0 constraint a little further; we have

$$L_0 = \frac{1}{2} p^2 + \sum_{n=1}^{\infty} \alpha_{-n}^\mu \alpha_{n\mu} = 0 \quad (2.16)$$

---

⁶ We put $2\alpha' = 1$ from now on.

This is the string analog of (2.3); owing to the infinitely many internal excitations, the mass of the string can assume infinitely many values such that a particular mass is associated with a particular vibrational excitation of the string. From (2.16) it might appear that $M^2 = -p^2$ is not positive due to the indefiniteness of the Minkowski metric. However, this is a gauge artifact as we will see below.

Instead of just imposing the constraints $L_m=0$ on the system, one may alternatively eliminate all unphysical degrees of freedom. For this purpose, one introduces light cone coordinates

$$x^\pm = \frac{1}{\sqrt{2}} (x^0 \pm x^{D-1}) , \quad x^i (i = 1, \dots, D-2) \quad (2.17)$$

where x^i are the transverse coordinates. One can then show (see e.g. Scherk, 1975) that the residual gauge invariance of (2.5) is entirely fixed by putting

$$x^+(\sigma, \tau) = q^+ + p^+ \tau \quad (2.18)$$

i.e. gauging to zero all the α_m^+ excitations. Substituting (2.18) into (2.5), we can solve for $x^-(\sigma, \tau)$

$$x^-(\sigma, \tau) = q^- + p^- \tau + i \sum_{m \neq 0} \frac{1}{m} \alpha_m^- \cos m \sigma e^{-im\tau} \quad (2.19)$$

where the α^- 's are now expressed as functions of the transverse excitations

$$\alpha_m^- = \frac{1}{2p_+} \sum_{n=-\infty}^{\infty} \alpha_{m-n}^i \alpha_n^i . \quad (2.20)$$

Observe that, up to the factor p_+^{-1} , α_m^- is just like L_m in (2.15) but with the sum ranging only over transverse indices. Furthermore putting $\alpha_m^+ = 0$ in (2.16) we see that (2.16) now reads

$$M^2 = -p^2 = 2 \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i \quad (2.16')$$

which is manifestly positive. All physical degrees of freedom are now contained in the transverse oscillations (and the center of mass coordinates and momenta). To understand why this is true one must investigate the theory in more detail but even without doing so, the analogy with electrodynamics may be helpful. There, the electromagnetic potential or photon field $A^\mu(x)$ has four components. In momentum space, one may decompose these components into time-like, longitudinal and transverse ones with respect to the momentum four-vector of the photon. Owing to gauge invariance, only the transverse components of the electromagnetic potential carry physical information while the other components are gauge degrees of freedom. In string theories, the conformal transformations take over the role of gauge invariance in that they may be used to eliminate the timelike and longitudinal components of the string. This analogy shows that the importance of conformal invariance in string theories can not be overemphasized.

3. QUANTIZATION, CRITICAL DIMENSION AND SPECTRUM

String theories possess some very remarkable properties which reveal themselves only after quantization. The most remarkable one is that string theories can be consistently quantized only in certain critical dimensions (recall that there were no such restrictions for the relativistic point particle which can be quantized in any dimension). Furthermore, in the bosonic string theories, quantization forces the groundstate of the string to be a tachyon - a particle of imaginary mass that travels faster than light. Obviously, neither of these properties was especially welcome to the physicists who tried to describe hadronic physics with string theories, but nowadays, with a completely changed perspective, one tends to view these features as virtues rather than as shortcomings of the theory. Moreover, in superstring theories, the tachyon disappears and the critical dimension is lowered.

The most powerful approach to quantize string theories is through functional integral methods (Polyakov, 1981). This method is well suited to compute higher order "radiative" corrections but it is technically demanding and requires an intimate knowledge of advanced mathematics such as Riemann surface theory. For this reason, we will not dwell on this topic but rather stick to the more conventional approach (which also has its pitfalls!). In the foregoing section, it has been explained that the relativistic string behaves in many ways like an ordinary violin string. Apart from the constraint (2.5), it satisfies a free two-dimensional wave equation (2.10) which can be easily solved, see (2.11) and (2.12). The motion is described through the modes α_n^μ (or α_m^μ and $\bar{\alpha}_n^\mu$) which are just ordinary harmonic oscillators. It is therefore not surprising that, after a canonical treatment (see e.g. Scherk, 1975), the quantized string is a collection of infinitely many harmonic oscillators. More precisely, the modes α_m^μ become creation operators for $m < 0$ and annihilation operators for $m > 0$ which are subject to the commutation relations

$$[\alpha_m^\mu, \alpha_n^\nu] = m \delta_{m+n, 0} \eta^{\mu\nu}, \quad (\alpha_m^\mu)^\dagger = \alpha_{-m}^\mu \quad (3.1)$$

The center of mass coordinates and momenta obey the usual commutation relations

$$[q^\mu, p^\nu] = i \eta^{\mu\nu} \quad (3.2)$$

The Fock space H of the theory is the product of the single harmonic oscillator Hilbert spaces and consists of all states of the form

$$\psi = \prod_r \alpha_{-r}^\mu |0, k\rangle \quad (3.3)$$

The groundstate $|0, k\rangle$ has momentum k and is annihilated by all oscillators α_m^μ with $m > 1$. Consequently, all other expressions now become "operator-valued". For instance, (2.15) is now an operator in H . However, some care must be exercised because, unlike the classical expressions, these operators may cease

to be well-defined. In particular, the sum over the oscillator vacuum energies diverges like $\frac{1}{2} \sum_{n=1}^{\infty} n$ as one can easily verify by computing the vacuum expectation value of L_0 . To avoid this problem, one modifies all potentially ill-defined operators by moving the annihilation operators to the right; this amounts to subtracting off all infinities. This procedure is referred to as "normal ordering" and denoted by semicolons. For instance, we have

$$L_0 = \frac{1}{2} \sum_{n=-\infty}^{\infty} : \alpha_{-n}^{\mu} \alpha_{n\mu} : = \frac{1}{2} p^2 + \sum_{n=1}^{\infty} \alpha_{-n}^{\mu} \alpha_{n\mu} \quad (3.4)$$

such that $\langle 0 | L_0 | 0 \rangle$ is now well-defined. The normal ordering leads to a very important modification in the algebra of the L_m -operators which now reads (a derivation of the extra term is given in Scherk, 1975)

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{D}{12} m(m^2 - 1) \delta_{m+n, 0} \quad (3.5)$$

The "central term" in (3.5) may be viewed as an "anomaly": with it, the L_m 's no longer generate the algebra of conformal transformations (1.2). The new contribution in (3.5) is the source of all the peculiarities that distinguish the quantized string from its classical counterpart. To restore conformal invariance, we are forced to a particular value of D and to a tachyonic ground state.

An obvious problem is already raised by the relation (3.1). Choosing the time-like excitation α_m^0 and remembering $\eta^{00} = -1$, we can easily calculate the norm of the state $\alpha_{-m}^0 |0, k\rangle$ ($m > 0$)

$$\langle 0, k | \alpha_m^0 \alpha_{-m}^0 | 0, k \rangle = \langle 0, k | [\alpha_m^0, \alpha_{-m}^0] | 0, k \rangle = -m < 0 \quad (3.6)$$

which is negative! This result is incompatible with the usual lore of quantum mechanics where the norm of a state is interpreted as a probability which should be positive. We must therefore devise a method to get rid of these "negative norm states". The clue to the solution is conformal invariance, and it is analogous to the solution of a related problem in quantum electrodynamics, see also the remarks at the end of section 2. There, the timelike component of the photon leads to a negative norm state but this state can be eliminated by a gauge transformation. One can do this either by imposing a gauge condition on physical states, in which case one has to prove that no negative norm states are left, or by going to a light cone gauge which contains only the transverse photon components. In the second case, there are evidently no negative norm states, but one must show by explicit computation that Lorentz invariance is not violated. These two ways of eliminating unphysical states have their analogs in string theory but here, the framework is much more restrictive.

Let us first discuss the method of defining the physical states by constraints. In the classical theory, we have the constraint (2.5) which is equivalent to $L_m = 0$ for all m . It is easy to see, however, that we cannot impose L_m to vanish on the physical states for all m . Namely, inserting (3.5) between

two physical states would lead to a contradiction immediately because of the "anomaly". Rather, as in the Gupta-Bleuler formulation of quantum electrodynamics, one must relax this condition by imposing this requirement only for "positive frequency" operators, that is

$$L_m |\text{phys}\rangle = 0 \quad \text{for } m > 1 \quad (3.7)$$

For L_0 , one must allow for an extra shift

$$[L_0 - \alpha(0)] |\text{phys}\rangle = 0 \quad (3.8)$$

where the intercept $\alpha(0)$ must be determined from the consistency requirement. Because of $L_{-m} = L_m^\dagger$, (3.7) implies

$$\langle \text{phys} | L_m | \text{phys} \rangle = 0 \quad \text{for all } m \neq 0 \quad (3.9)$$

so, in the classical limit, we recover (2.5).

The problem is now the following. By imposing (3.7) and (3.8), we single out a subspace H_{phys} of the full Hilbert space H spanned by the states (3.3). Under what conditions can one prove that H_{phys} is free of negative norm states? The answer to this question can only be given after a lengthy argument which we will not reproduce here (Brower, 1972; Goddard and Thorn, 1972; Goddard et al. 1973). It turns out that things work out only if

$$D = 26, \quad \alpha(0) = 1 \quad (3.10)$$

Assuming $\alpha(0) = 1$, one can make the following plausibility argument for the emergence of the number $D = 26$. Consider the following state for arbitrary D

$$\Psi = [\alpha_{-1}^\mu \alpha_{-1\mu} + A k_\mu \alpha_{-2}^\mu + B (k_\mu \alpha_{-1}^\mu)^2] |0, k\rangle \quad (3.11)$$

and impose the physical state constraints (3.7) and (3.8) on Ψ . It is actually sufficient to consider only the operators

$$\begin{aligned} L_0 &= \frac{1}{2} p^2 + \alpha_{-1} \alpha_1 + \alpha_{-2} \alpha_2 + \dots \\ L_1 &= p \alpha_1 + \alpha_{-1} \alpha_2 + \dots \\ L_2 &= \frac{1}{2} \alpha_1^2 + p \alpha_2 + \dots \end{aligned} \quad (3.12)$$

as the higher mode oscillators annihilate the state (3.11), and $L_3 = [L_2, L_1]$ etc. $(L_0 - 1)\Psi = 0$ leads to $k^2 = -2$, while the L_1 and L_2 constraints lead to

$$A = \frac{D-1}{5}, \quad B = \frac{D+4}{10} \quad (3.13)$$

The norm of the state Ψ for arbitrary D is

$$\langle \Psi | \Psi \rangle = \frac{2}{25} (26 - D)(D - 1) \quad (3.14)$$

Thus, $\langle \Psi | \Psi \rangle < 0$ for $D > 26$ in which case H_{phys} contains negative norm states and there is no hope of consistently quantizing the theory. For $D=26$, $|\Psi\rangle$ is a

zero norm state which does not correspond to a physical excitation (like the state with equally many timelike and longitudinal photons in quantum electrodynamics) and a consistent quantum theory exists. For $D < 26$, a consistent theory may exist but would contain extra states. Although the above argument proves the inconsistency for $D > 26$, it is, of course, not the whole story, but we hope at least that it gives the reader an idea as to where the number 26 comes from.

An alternative approach to quantization is by solving the constraints first as in (2.19) and expressing everything through the transverse oscillators α_m^i . In this case, all operators which were responsible for the occurrence of negative norm states have disappeared and unitarity is manifest. On the other hand, manifest Lorentz invariance has been lost and one must check explicitly whether it can be restored. After a tedious calculation, one recovers the conditions (3.10) and therefore the two approaches are entirely equivalent (Goddard et al., 1973).

The light-cone gauge is actually somewhat more convenient to describe the physical spectrum of string theories as it contains no unphysical operators. Taking into account the shift by $\alpha(0) = 1$, the quantum analog of the mass formula (2.16') is

$$\frac{1}{2}M^2 = \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i - 1. \quad (3.15)$$

Unlike in (2.16'), where M^2 varies continuously, M^2 has only integer values in the quantum theory. The lowest state $|0, k\rangle$ has no oscillator excitations, and therefore $M^2 = -1$. The next state is $\alpha_{-1}^i |0, k\rangle$, which has $M^2 = 0$. Since this state has only transverse excitations, it is like the photon. Continuing in this fashion, one obtains the following states ordered according to increasing mass

TABLE I. Open String Spectrum

$M^2 = -2$	$ 0, k\rangle$	Tachyon
$M^2 = 0$	$\alpha_{-1}^i 0, k\rangle$	"Photon"
$M^2 = 2$	$\alpha_{-1}^i \alpha_{-1}^j 0, k\rangle$ $\alpha_{-2}^i 0, k\rangle$	Massive "Spin-2" Excitation
$M^2 = 4$	$\alpha_{-1}^i \alpha_{-1}^j \alpha_{-1}^k 0, k\rangle$ $\alpha_{-1}^i \alpha_{-2}^j 0, k\rangle$ $\alpha_{-3}^i 0, k\rangle$	

etc.

The spectrum of the closed string can be analyzed in a similar fashion. We have already mentioned that there are twice as many oscillators in this case. The condition that there should be no distinguished point on the closed string leads to an additional constraint on the physical states, namely (in the light-cone gauge)

$$\sum_{n=1}^{\infty} (\alpha_{-n}^i \alpha_n^i - \bar{\alpha}_{-n}^i \bar{\alpha}_n^i) |\text{phys}\rangle = 0 \quad (3.16)$$

i.e. the number of unbarred and barred excitations must be the same. The mass formula for the closed string is

$$M^2 = \sum_{n=1}^{\infty} (\alpha_{-n}^i \alpha_n^i + \bar{\alpha}_{-n}^i \bar{\alpha}_n^i) - 2 \quad (3.17)$$

and the lowest state is therefore again a tachyon. Because of (3.16) neither of the states $\alpha_{-1}^i |0, k\rangle$ or $\bar{\alpha}_{-1}^i |0, k\rangle$ belongs to the physical spectrum. The next state is therefore

$$\psi^{ij} = \alpha_{-1}^i \bar{\alpha}_{-1}^j |0, k\rangle \quad (3.18)$$

which is massless because of (3.17). One can decompose ψ^{ij} into irreducible parts according to

$$\psi^{ij} = \psi_1^{(ij)} + \delta^{ij} \psi_2 + \psi_3^{[ij]} \quad (3.19)$$

where $\psi_1^{(ij)}$ is symmetric and traceless in (ij) , and $\psi_3^{[ij]}$ is antisymmetric in $[ij]$. A symmetric traceless two-index tensor describes a spin-2 particle (at least in 4 dimensions). It was this coincidence that inspired Scherk and Schwarz (1974) to make the identification

$$\psi_1^{(ij)} = \text{"Graviton"}. \quad (3.20)$$

(The state ψ_2 is referred to as "dilaton".) One of the remarkable things about string theory is that the existence of this "graviton" is a prediction rather than an input: even if one starts with open strings, which contain only a "photon", the "graviton" arises as an intermediate state. One is therefore inevitably forced to include gravity in the unification; there is no consistent string theory without gravity!

Before passing on, we make two further comments on the open string spectrum of table I (similar remarks apply to the closed string spectrum). Although we have not explained the notion of "spin" in 26 dimensions, it is evident from the table that there is a correlation between M^2 and the "spin". In fact, a plot reveals that the states lie on so-called "Regge-trajectories" see Fig.5. below. (There are also many "daughter trajectories" which are not shown.)

Fig.5. Open String spectrum.

Such "Regge-trajectories" were in fact observed in the sixties in the form of mesonic and baryonic resonances. In the modern interpretation, of course, the higher excited states on these trajectories have masses of the order of 10^{19} GeV and are therefore unobservable.

A second noteworthy feature of the string spectrum is the enormous increase in the number of states as one goes to higher and higher values of M^2 . One can show that the number of states $n(M)$ at a given mass level M^2 asymptotically grows like

$$n(M) \propto \text{const.} \left(\frac{M}{M_0} \right)^\alpha \exp \left(\frac{M}{M_0} \right) \quad (3.21)$$

where α depends on the dimension and M_0 is related to the fundamental scale $(\alpha')^{-1/2}$ of the theory (i.e. 1 GeV for hadronic physics and 10^{19} GeV for gravitational physics). Inserting (3.21) into the usual formula for the free energy⁷

$$F(T) = -kT \log \left[\sum_M n(M) \exp \left(- \frac{E(M)}{kT} \right) \right] \quad (3.22)$$

we see that the sum diverges at the critical temperature

$$kT_{\text{crit}} = M_0 \quad (3.23)$$

This result indicates that, at this temperature, a phase transition takes place. A natural interpretation is that, at $T = T_{\text{crit}}$, the string "breaks up"⁸

⁷ Actually, this formula should also include an integral over all 26-dimensional momenta, but its omission does not affect our main conclusion.

⁸ In fact, in the old days, the interpretation was slightly different. It was assumed that when one tries to heat hadronic matter beyond T_{crit} , it cooled down again by "boiling off" mesons, etc. Since therefore T_{crit} is a sort of ultimate temperature, it was sometimes referred to as "hell's temperature" (Hagedorn, 1968).

into its constituents, e.g. quarks and gluons, which then form a plasma for $T > T_{crit}$. At present, one does not understand what happens when a gravitational string is heated beyond T_{crit} and what its constituents could be. Perhaps, these are indications of a theory beyond string theory.

4. SPINNING STRINGS, SUPERSTRINGS AND HETEROTIC STRINGS

There is one obvious defect of the string theory discussed so far: it describes only bosons. In 1971, Ramond, and Neveu and Schwarz proposed new models in an attempt to remedy this defect. Both models share the feature that, in addition to the string coordinate $X^\mu(\sigma, \tau)$, they contain its fermionic partner $\lambda^\mu(\sigma, \tau)$ whose modes satisfy anticommutation relations rather than commutation relations. One can visualize this by attaching a fermion to every point on the world sheet. The mode expansion of the fermionic field $\lambda^\mu(\sigma, \tau)$ is quite analogous to (2.11) and (2.12) and is given by ⁹

$$\lambda^\mu(\sigma \pm \tau) = \sum_{n \in \mathbb{Z}} d_n^\mu e^{-in(\sigma \pm \tau)} \quad (4.1)$$

in the Ramond sector and

$$\lambda^\mu(\sigma \pm \tau) = \sum_{r \in \mathbb{Z} + \frac{1}{2}} b_r^\mu e^{-ir(\sigma \pm \tau)} \quad (4.2)$$

in the Neveu-Schwarz sector. A significant difference between (4.1) and (4.2) is that the oscillators in (4.1) are integer-moded ($n = \dots, -2, -1, 0, 1, 2, \dots$) whereas they are half-integer moded ($r = \dots, -3/2, -1/2, 1/2, 3/2, \dots$) in (4.2). The anticommutation relations are

$$\begin{aligned} \{d_m^\mu, d_n^\nu\} &= \delta_{m+n, 0} \eta^{\mu\nu} \\ \{b_r^\mu, b_s^\nu\} &= \delta_{r+s, 0} \eta^{\mu\nu} \end{aligned} \quad (4.3)$$

For $m=n=0$, we obtain $\{d_0^\mu, d_0^\nu\} = \eta^{\mu\nu}$ which means that d_0^μ is like a γ -matrix: it implies that the groundstate is a fermion in space-time. On the other hand, no such peculiarity occurs in the Neveu-Schwarz sector whose groundstate is a boson. Since both d_m^μ and b_r^μ behave like vectors under space-time Lorentz transformations, the spacetime statistics is not changed by the action of the oscillators on the groundstate, and therefore the Ramond spectrum contains only fermions and the Neveu-Schwarz spectrum contains only bosons. It may seem paradoxical at first sight that, although we started with two dimensional fermionic operators, the resulting spectrum in the embedding spacetime may consist of either bosons or fermions. However, it is well-known among the experts (Coleman, 1975; Mandelstam, 1975) that, in two dimensions, bosons and fermions

---

⁹ To be precise, λ^μ is a two-dimensional fermion and thus should have two components. The Dirac equation in two dimensions and the boundary conditions imply that the upper (lower) component are given by the same function $f(\sigma + \tau)$ (or $f(\sigma - \tau)$), and therefore we write out only one component, see Scherk (1975).

are equivalent, and therefore the two-dimensional statistics implies nothing about the statistics in space-time. From the above remarks, it follows that the space-time statistics depends only on the "modedness" of the oscillators, or, more precisely, on the boundary conditions obeyed by λ^μ !

To study the spectrum of these models, we switch again to the light-cone gauge. The mass formulas for the open spinning strings are given by

$$\frac{1}{2}M^2 = \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i + \sum_{n=1}^{\infty} n d_{-n}^i d_n^i \quad (4.4)$$

in the Ramond sector and by

$$\frac{1}{2}M^2 = \sum_{n=1}^{\infty} \alpha_{-n}^i \alpha_n^i + \sum_{r=1/2}^{\infty} r b_{-r}^i b_r^i - \frac{1}{2} \quad (4.5)$$

in the Neveu-Schwarz sector. In the Ramond sector the groundstate is a fermion which furthermore obeys the Dirac equation (we will not derive this here, see e.g. Scherk (1975) for a more detailed discussion); this groundstate has no further excitations and is therefore massless by (4.4). The groundstate in the Neveu-Schwarz sector has $M^2 = -1/2$ and is again a tachyon. The mass values are obtained by considerations similar to the ones that led to (3.10). In addition, the value of the critical dimension is lowered to

$$D = 10 \quad (4.6)$$

for spinning strings (again, the calculation required to prove this is very tedious and will not be reproduced here, see e.g. Green and Schwarz, 1981).

The excited states are simply obtained by acting with the oscillators on the ground state. In the Neveu-Schwarz sector, this procedure leads to the following table

TABLE 2. Neveu-Schwarz Spectrum

$M^2 = -1$	$ 0, k\rangle$	Tachyon
$M^2 = 0$	$b_{-\frac{1}{2}}^i 0, k\rangle$	"Photon"
$M^2 = +1$	$b_{-\frac{1}{2}}^i b_{-\frac{1}{2}}^i 0, k\rangle$ $\alpha_{-1}^i 0, k\rangle$	Massive "Spin-2" Excitation
$M^2 = +2$	$b_{-\frac{1}{2}}^i b_{-\frac{1}{2}}^j b_{-\frac{1}{2}}^k 0, k\rangle$ $b_{-\frac{1}{2}}^i \alpha_{-1}^j 0, k\rangle$ $b_{-\frac{3}{2}}^i 0, k\rangle$	

etc.

In the Ramond-sector, we obtain

TABLE 3. Ramond Spectrum*

$M^2 = 0$	$ 0, k\rangle$	Massless Fermion
$M^2 = 2$	$\alpha_{-1}^i 0, k\rangle$ $d_{-1}^i 0, k\rangle$	Massive "Spin- $\frac{3}{2}$ " Excitation

etc.

* $|0, k\rangle$ is a spinor wave function.

It was not until 1976 that it was recognized that, by combining these two spectra in a suitable way, a supersymmetric spectrum could be obtained in ten dimensions i.e. a spectrum containing equally many bosons and fermions at each mass-level (Gliozzi, Scherk and Olive, 1976). The clue was the elimination of half of the states in each sector. Obviously, we must eliminate all states of half integer M^2 in the Neveu-Schwarz sector since these cannot have fermionic partners in the Ramond sector where M^2 assumes only integer values. This is equivalent to discarding all states created by an even number of b-oscillators; observe that the troublesome tachyon is also eliminated in this way! The truncation to states with an odd number of b-oscillators is implemented by the projection operator

$$P_{NS} = \frac{1}{2} [1 - (-1)^{\sum_{r=1}^{\infty} b_{-r}^i b_r^i}]. \quad (4.7)$$

Although it is not so obvious, a similar truncation is needed in the Ramond sector which is accomplished by means of the projection operator

$$P_R = \frac{1}{2} [1 - \gamma^5 (-1)^{\sum_{n=1}^{\infty} d_{-n}^i d_n^i}] \quad (4.8)$$

where γ^5 is the ten-dimensional analog of the usual γ^5 matrix. The supersymmetry of this truncated system at the massless level is easily checked: a "Majorana-Weyl-spinor" in $D=10$ has eight real components which match with the eight components of the state $b_{-1/2}^i |0, k\rangle$. To demonstrate that there is an equal number of bosons and fermions at each mass level requires the following "aequatio identica satis abstrusa"

$$\frac{1}{2q} \left\{ \prod_{n=1}^{\infty} (1 + q^{2n-1})^8 - \prod_{n=1}^{\infty} (1 - q^{2n-1})^8 \right\} = 8 \prod_{n=1}^{\infty} (1 + q^{2n})^8 \quad (4.9)$$

which was already known to the German mathematician Jacobi in 1829! The above relation provides an example for the connection between string theories and rather deep mathematical results.

It may seem somewhat awkward to describe the superstring by a truncation via (4.7) and (4.8) and to have to rely on identities such as (4.9) to make the supersymmetry of the spectrum explicit. It was this circumstance and the desi-

re to explore superstring theory further that prompted Green and Schwarz in 1981 to develop the so-called "new formalism". Instead of gluing together the Neveu-Schwarz and the Ramond sector and projecting out half of the states in both sectors, they replaced the oscillators b_r^i and d_n^i by a single set of integer moded anticommuting oscillators $S_m^{\dot{a}}$ which carry a spinor index $\dot{a}=1,\dots,8$ rather than a vector index i and obey the anticommuting relations (Green and Schwarz, 1981)

$$\{S_m^{\dot{a}}, S_n^{\dot{b}}\} = \delta^{\dot{a}\dot{b}} \delta_{m+n,0}. \quad (4.10)$$

Because $S^{\dot{a}}$ belongs to a spinor representation 8_C of $SO(8)$, $S_m^{\dot{a}}$ transforms as a space-time spinor under transverse Lorentz rotations, and there are no apparent "paradoxes" any more with the space-time statistics of the states. A drawback of this new formalism is that it works so far only in the light cone gauge. The question of how to extend superstring theory "off-shell" is presently under investigation by many groups.

It is quite straightforward to work out the spectrum of the superstring by means of the new formalism. Remembering that there is no tachyon any more, we can write the massless groundstate of the open superstring as

$$\begin{aligned} |i\rangle & \in 8_V \quad (\text{vector representation of } SO(8)) \\ |a\rangle & \in 8_S \quad (\text{spinor representation of } SO(8)) \end{aligned} \quad (4.11)$$

where both indices i and a assume the values $1,\dots,8$ and 8_V and 8_S are the usual designations for the eight-dimensional vector and spinor representations of $SO(8)$. One could alternatively assign the spinor to the conjugate spinor representation of $SO(8)$ which is denoted by 8_C or $|a\rangle^{10}$. The excited states are obtained by acting on either $|i\rangle$ or $|a\rangle$ with either α_{-m}^i or $S_{-m}^{\dot{b}}$. The supersymmetry is now manifest; for example, the first excited level contains the states

$$\begin{aligned} 128 \text{ Bosons: } & \alpha_{-1}^i |j\rangle \quad \text{and} \quad S_{-1}^{\dot{a}} |b\rangle \\ 128 \text{ Fermions: } & \alpha_{-1}^i |b\rangle \quad \text{and} \quad S_{-1}^{\dot{a}} |j\rangle. \end{aligned} \quad (4.12)$$

One of the accomplishments of Green and Schwarz (1981) was the demonstration that one could construct the generators of the full Lorentz group $SO(1,9)$ in ten dimensions out of just the transverse oscillators α_m^i , $S_m^{\dot{a}}$ and the center of mass coordinates and momenta.

Closed superstrings can be constructed in complete analogy with the closed bosonic string discussed in section 3. One simply has to double the number of oscillators and the groundstate, so we now have bosonic oscillators α_m^i , $\bar{\alpha}_m^i$.

---

¹⁰ The group $SO(8)$ is unique in that it has three inequivalent eight-dimensional representations. This property is referred to as "triality" (see e.g. Slansky, 1981).

The groundstates are now obtained by decomposing the products

$$\begin{aligned} \text{Bosons } |i\rangle_L \otimes |j\rangle_R \quad \text{and} \quad |a\rangle_L \otimes |b\rangle_R \\ \text{Fermions } |i\rangle_L \otimes |b\rangle_R \quad \text{and} \quad |a\rangle_L \otimes |j\rangle_R \end{aligned} \quad (4.13)$$

into irreducible components. In particular, (4.13) contains the following states

$$\text{"Graviton"} = \text{symmetric traceless part of } |i\rangle_L \otimes |j\rangle_R \quad (4.14)$$

$$\text{"Gravitino"} = \gamma\text{-traceless parts of } |i\rangle_L \otimes |b\rangle_R \quad \text{and} \quad |a\rangle_L \otimes |j\rangle_R$$

Thus, we have two gravitinos and, since the massless states form a supermultiplet, it is no surprise that the full set of states (4.13) coincides with an $N=2$ supergravity multiplet in ten dimensions. Hence, in the same way as there is no closed string theory without gravity, there is no closed superstring theory without supergravity!

We have already mentioned that one may alternatively assign the groundstate spinor to the conjugate representation 8_c of $SO(8)$. In (4.13), both the left and right-moving spinors belong to the 8_s representation and the resulting theory is called "type II B". To get the so-called "type II A" theory, one must assign these spinors to different representations, and in this case the groundstates are obtained from the products

$$\begin{aligned} \text{Bosons } |i\rangle_L \otimes |j\rangle_R \quad \text{and} \quad |a\rangle_L \otimes |\dot{b}\rangle_R \\ \text{Fermions } |i\rangle_L \otimes |\dot{b}\rangle_R \quad \text{and} \quad |a\rangle_L \otimes |j\rangle_R \end{aligned} \quad (4.13')$$

Again, one gets an $N=2$ supergravity multiplet in this way.

Both the "type II A" and the "type II B" superstring theory are one-loop finite and free of anomalies and therefore good candidates for a unified theory. However, there is another type of string theories with these properties, namely the heterotic string theories (Gross et al., 1984). These are hybrids of the bosonic string in 26 dimensions and the superstring in ten dimensions. The most general solution to the free wave equation (2.10) is a superposition of left-moving and right-moving modes which only depend on $\sigma - \tau$ and $\sigma + \tau$, respectively, as is also evident from the closed string mode expansion (2.12). Since (2.10) contains no interactions, these do not interfere with each other and may therefore be chosen independently. The basic idea is now to take the left moving sector to be a superstring with states $(8_v)_L$ and $(8_s)_L$ and the right moving sector to be a bosonic string and to obtain the states of the full theory by a multiplication analogous to (4.13) and (4.14) but now with one half of the string in ten dimensions and the other in 26 dimensions. Absurd as though it may appear at first sight this idea does work! The crucial ingredient that makes it work is a "compactification" of the 26-dimensional part by which the momentum components k^I with $11 \leq I \leq 26$ become discrete. They are then no longer interpreted as momenta but rather as internal symmetry labels

such as isospin and strangeness quantum numbers. It is in this way that an internal symmetry is generated out of a purely bosonic theory which contains no internal symmetry (Frenkel and Kac, 1980; Goddard and Olive, 1984). Consistency¹¹ then forces these symmetry groups to be either $SO(32)$ or $E_8 \times E_8$ in accordance with the previously found restrictions from anomaly cancellations.

5. INTERACTING STRINGS

So far we have described free string theories. Knowing the spectrum, one would also like to calculate scattering amplitudes and other quantities of interest. To do so, one must develop a formalism for interacting strings. In this section, we will very sketchily explain how this can be done, mostly by drawing pictures. It is a rather demanding task to translate these pictorial representations of interacting strings into some kind of calculational scheme, and any attempt at a more detailed explanation would go far beyond the limitations of this article. The interested reader is referred to (Mandelstam, 1973; 1974; Cremmer and Gervais, 1974) for further details of the formalism. The essential result is that string interactions are very restricted and almost unique.

To understand the basic idea, it is useful to go back once more to the relativistic point particle which was discussed in section 2. Its interactions can be very simply represented as splitting of world-lines as in Fig. 6 below.

Fig.6. Interactions of the relativistic point particle.

It is very important that, for the relativistic point particle, there are essentially no limitations on such interactions: the world line may split at any of its point and branch off into arbitrarily many new world-lines. The mathematical description of such an interaction is through a vertex

$$V(x_0^\mu(\tau_0), x_1^\mu(\tau_0), \dots, x_n^\mu(\tau_0)) = \delta(x_0^\mu(\tau_0) - x_1^\mu(\tau_0)) \dots \delta(x_0^\mu(\tau_0) - x_n^\mu(\tau_0)) \quad (5.1)$$

where the world line of an incoming particle (parametrized by $x_0^\mu(\tau)$) splits up into the world lines of n particles (parametrized by $x_1^\mu(\tau), \dots, x_n^\mu(\tau)$) at the

---

¹¹ i.e. modular invariance (only for experts).

interaction time $\tau=\tau_0$. A "radiative correction" is obtained by splitting a world-line and joining the pieces at a later time, see Fig. 7.. In fact, these pictures are nothing but ordinary Feynman diagrams, and the knowledgeable reader will recall that it is not completely straightforward to translate these pictures into mathematically well-defined expressions.

Fig.7. A one-loop correction for the relativistic point particle.

At this level of the discussion, the interactions of strings are quite similar to the point particle interactions. Strings interact by touching at one point and joining into a single string: for the open string, the interaction point is always the boundary point while for the closed string the point of contact is arbitrary. These processes are depicted in Fig.8. below.

Fig.8. Interactions of open and closed strings.

A rather important point here is that although these are interactions between extended objects, the interaction itself is local: the instantaneous interaction takes place at one point only. The locality postulate rules out processes such as in Fig.9.

Fig.9. A forbidden interaction.

One can now associate a mathematically well-defined "vertex-operator" with such an interaction; it is essentially a string overlap δ -function analogous to (5.1). In terms of the individual states of the string theory (parts of which are shown in the tables), one gets infinitely many point particle interactions whose complexity increases with increasing "spin". It is especially instructive to calculate the three-point interactions between the massless excitations of the open and closed bosonic strings, respectively. These point particle interaction vertices turn out to coincide with the "three-gluon" vertex of Yang Mills theory (Neveu and Scherk, 1972) in the case of the open string and with the "three-graviton" vertex of Einstein's general relativity theory in the case of the closed string (Scherk and Schwarz, 1974). This means that

- (i) Ordinary Yang Mills theories are contained in the open bosonic string theory, and
- (ii) Einstein's relativity theory is contained in the closed bosonic string theory.

In a sense, one could have foreseen this result: the only consistent theories of massless particles of spin-1 and spin-2, respectively, are Yang-Mills theories and Einstein's general relativity, respectively (the gauge-invariance is absolutely necessary to eliminate unwanted helicity states). Nonetheless, the reader should pause at this point to appreciate the implications of this result. After all, the massless states are only a tiny part of the whole string spectrum, and one may therefore anticipate the existence of a much bigger symmetry which contains either ordinary gauge symmetries or general coordinate invariance as "the tip of the iceberg". It is one of the most fascinating problems of string theory what this huge symmetry might be and whether there is a generalization of the principle of equivalence that encompasses the postulates of general relativity.

Similar remarks apply to superstrings whose interactions are also given by overlap δ -functions (Mandelstam, 1974; Green and Schwarz, 1983). However, these are now harder to visualize and we will therefore refrain from drawing further diagrams. As before, one may calculate the point like interactions between massless particles. The result is that

- (iii) Ordinary supersymmetric Yang Mills theories are contained in the open superstrings, and
- (iv) Supergravity is contained in the closed superstring theory.

It is now obvious why superstring theory has completely absorbed the once thriving field of supergravity ¹².

---

¹² To be sure, there is one supergravity theory that does not fit into string theory, namely the maximally extended d=11 supergravity (Cremmer, Julia and Scherk, 1978).

Finally, radiative corrections can be discussed along similar lines. They correspond to first splitting and rejoining strings, see Fig.10.

Fig.10. A one-loop correction for the closed bosonic string.

The number of loops is equal to the number of holes in the world-surface of the string. Possible divergences appear when the diameter of such a hole shrinks to zero or when the surface is "pinched"; this is somewhat analogous to the divergences that appear in Feynman graphs when a loop shrinks to a point. It is conjectured, although not proven so far that, in contrast to point particle theories, the one-loop finiteness of string theories implies their finiteness to all orders of perturbation theory.

6. OUTLOOK

Up to now, we have concentrated on the basic features of string theory, namely those that would be included in any introductory treatment of the subject. However, there are many more advanced topics and, of course, many open problems. In this section we will try to give the reader a flavor of what these are but naturally our review will be incomplete. The areas of current research can be roughly divided into two parts. The first consists of attempts to extract physically testable predictions from superstring theory while the second centers on the underlying principles of string theory. Let us begin with the first.

As already mentioned, the theory currently thought to be most promising is the heterotic string with gauge group $E_8 \times E_8$ (Gross et al., 1985). Although there exist other versions of the heterotic string this theory is particularly attractive for phenomenology. The group E_8 is the largest of the exceptional Lie-groups and is big enough to accommodate all known symmetries and particles; for this reason, it has already been considered for grand unification several years ago. Furthermore, the $E_8 \times E_8$ theory has chiral fermions and is free of anomalies (this is also true for the other heterotic theories as well as for the type II B theory). Thus, it offers the possibility of getting chiral fermions in four dimensions in the desired representations of $SU(3) \times SU(2) \times U(1)$.

The way this is achieved in practice is related to the way in which the ten-dimensional theory is compactified to four dimensions. In the process of compactification, six dimensions are curled up to an "internal" manifold whose size is so small as to make it inaccessible to present day experiments (e.g. with diameter of the order of 10^{-33} cm). The number of chiral fermions which emerge in such a compactification is related to topological properties of the internal manifold, i.e. the number of its "holes" and "handles". This is an example of how qualitative features of our low energy world, such as the number of generations, may be linked to topological rather than metrical properties of a unified theory.

In a currently favored scenario (Candelas et al., 1985) the compactification occurs on a "Calabi-Yau manifold"¹³ and the gauge group $E_6 \times E_6$ is broken according to

$$E_6 \times E_6' \longrightarrow G \times E_6' \quad (6.1)$$

where the residual gauge group G is a subgroup of E_6 . All observed particles (quarks, leptons, etc.) transform under G whereas the particles associated with E_6' are almost completely decoupled from our known universe as they couple only gravitationally. The E_6' particles constitute a "shadow world" (thus we may be sitting in the middle of a "shadow mountain" without noticing it!). This is interesting, because invisible "shadow matter" may account for the dark matter whose origin is still not understood by astrophysicists. The observable group G must still be further broken to the standard low energy group $SU(3)_C \times SU(2)_L \times U(1)_Y$, and it is here that things get murky. Since the actual dynamics of the theory is unknown, one has to make many assumptions at this point, and the outcome of any calculation depends to a great extent on the assumptions that were put in at the beginning. A second problem is that the compactification on Calabi-Yau spaces is not unique; the number of solutions is astronomical and one can obtain almost any number of chiral generations depending on the topology of the Calabi-Yau manifold. One would rather prefer to have a unique solution to describe the compactification to our four-dimensional world. Another problem is that compactification to four dimensions is in no way preferential in superstring theories (unlike in the case of $d=11$ supergravity where four dimensions are preferred (Freund and Rubin, 1980)). It seems obvious that the solution to these problems will require a lot more work.

We next turn to the second area of problems having to do with questions of principle. String theories possess many "miraculous" properties which were

---

¹³ These manifolds are mathematically rather intricate (and interesting) objects, but there is no room here to discuss them in further detail.

usually found after long and arduous calculations especially in the early days of the subject. For instance, why does the massless state in the closed string theory behave like a graviton? We know that the only consistent theory of a massless spin-2 theory is Einstein's theory of general relativity, so even if we start with a massless free spin-2 field and try to make it interact, we must eventually bring in the full apparatus of Riemannian geometry. Of course, Riemannian geometry and the principle of equivalence were Einstein's points of departure, and it was only realized afterwards that the graviton was a massless spin-2 particle. However, in string theories we lack both the analog of Riemannian geometry and a generalized principle of equivalence, and so we must start from the other end. An indication that this may be the "wrong" end to start from is that until now we are only able to describe the string motion in a fixed (not necessarily flat) spacetime background while the string itself contains the seeds of curved spacetimes with nontrivial topology and should therefore be described in a much more general way. It seems therefore that in order to properly describe strings one must dissolve the very notion of spacetime in the same way that quantum mechanics does away with the notion of trajectory of an electron around the hydrogen atom. These conceptual problems are presently at the focus of research but is not clear how long it will take to solve them.

One attempt in this direction is covariant string field theory (see e.g. West (1986) for a recent review). The purpose here is to exhibit the invariances explicitly which generalize ordinary gauge invariance and invariance under general coordinate transformations. To illustrate the basic idea we introduce a "string functional" which associates a field with every string excitation according to

$$\psi = [\phi(x) + A_\mu(x)\alpha_{-1}^\mu + \dots]|0\rangle. \quad (6.2)$$

The physical state constraint $L_1\psi = 0$ can be easily evaluated using $L_1 = p_\nu\alpha_1^\nu + \dots$ and the basic commutator (3.1).

$$0 = L_1\psi = (p_\nu\alpha_1^\nu + \dots)(\phi + A_\mu\alpha_{-1}^\mu + \dots)|0\rangle = (p^\mu A_\mu + \dots)|0\rangle. \quad (6.3)$$

Hence, (6.3) implies the Landau gauge condition

$$\partial^\mu A_\mu(x) = 0. \quad (6.4)$$

We can release this gauge condition by introducing a gauge invariance associated with L_{-1} . To do so, we define a "gauge parameter string functional"

$$\Omega = [w(x) + w_\mu(x)\alpha_{-1}^\mu + \dots]|0\rangle. \quad (6.5)$$

Using $L_{-1} = p_\nu\alpha_{-1}^\nu + \dots$, we see that

$$\begin{aligned}\delta\Phi &= (\delta\phi + \delta A_\mu \alpha_{-1}^\mu + \dots)|0\rangle = L_{-1}\Omega = (p_\nu \alpha_{-1}^\nu + \dots)(w + w_\mu \alpha_{-1}^\mu + \dots)|0\rangle \\ &= (p_\mu w \alpha_{-1}^\mu + \dots)|0\rangle\end{aligned}\quad (6.6)$$

contains the transformation rule

$$\delta A_\mu = i\partial_\mu w \quad (6.7)$$

which is just the ordinary gauge transformation of the electromagnetic potential! From (6.6), we also see that $L_{-1}\Omega$ contains further transformations for the higher level fields in the expansion (6.2) and therefore an infinite tower of symmetries. But there is even more symmetry because there will be similar transformations for all L_{-m} with $m \geq 1$. This explicitly shows the "explosion" of symmetry in string theories which was alluded to in the introduction.

The main task is now to work out the fully gauge invariant action, i.e. the string analog of $F_{\mu\nu}F^{\mu\nu}$ with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, first at the free level and then for the interacting theory (which should in particular contain the three- and four-gluon vertices), and to repeat this exercise for superstring theory. A great deal of progress has been made during the last year, although it is probably too early to tell whether the conceptual breakthrough can be achieved in this way. However, apart from such considerations, one may anticipate that the formalism will be eventually useful in studying higher loop corrections and in finding classical and/or nonperturbative solutions to string theory.

A further question of considerable interest is why there are already so many superstring theories ¹⁴ that are fully consistent at the one-loop level (and presumably beyond) where one would be enough, and whether these theories are related. In (Freund, 1985; Casher et al., 1985; Englert, Nicolai and Schellekens, 1986), it has been suggested that all consistent superstring theories are just spontaneously broken versions of the purely bosonic D=26 string theory which should therefore be viewed as the "Urtheorie". In fact, it has been established there that superstrings are contained in the bosonic string but the question as to the dynamical origin of this symmetry breaking remains open. Again, much work is needed to make progress.

Finally, we should not close our eyes on the possibility that the final string theory may not yet have been found or that there exists a theory "beyond superstrings". While efforts in this direction have not borne fruit so far one may safely predict that the coming years will have some surprises in store which may change the course of theoretical high energy physics and our perception of it in unexpected ways.

---

¹⁴ About ten at the time of writing.

REFERENCES

Early Review Articles

Alessandrini, V., D. Amati, M. le Bellac and D. Olive (1971).

Phys. Rep. 1C, 269.

Mandelstam, S. (1974). Phys. Rep. 13C, 259.

Rebbi, C. (1974). Phys. Rep. 12C, 1.

Veneziano, G. (1974). Phys. Rep. 9C, 199.

Scherk, J. (1975). Rev. Mod. Phys. 47, 123.

Schwarz, J. H. (1973). Phys. Rep. 8C, 269.

References

Alvarez-Gaumé, L. and E. Witten (1983). Nucl. Phys. B234, 269.

Brower, R. C. (1972). Phys. Rev. DG, 1655.

Candelas, P., G. T. Horowitz, A. Strominger and E. Witten (1985).

Nucl. Phys. B258, 46.

Casher, A., F. Englert, H. Nicolai and A. Taormina Phys. Lett. 162B, 121.

Coleman, S. (1975). Phys. Rev. D11, 2088.

Cremmer, E. and J. L. Gervais (1974). Nucl. Phys. 876, 209.

Cremmer, E., B. Julia and J. Scherk (1978). Phys. Lett. 76B, 409.

Cremmer, E. and B. Julia (1979). Nucl. Phys. B159, 141.

Deser, S. and B. Zumino (1976). Phys. Lett. 62B, 335.

de Wit, B. and H. Nicolai (1982). Nucl. Phys. B208, 322.

Dirac, P. A. M. (1950). Can. J. Phys. 2, 129.

Englert, F., H. Nicolai and A. Schellekens (1986). Nucl. Phys. B274, 315.

Ferrara, S., D. Freedman and P. van Nieuwenhuizen (1976). Phys. Rev. D13, 3214

Frenkel, J. and V. Kac (1980). Inv.Math. 62, 23.

Freund, P. G. O. and M. A. Rubin (1980). Phys. Lett. 97B, 233.

Freund, P. G. O. (1985). Phys. Lett. 151B, 387.

Fritzsch, H., M. Gell-Mann and H. Leutwyler (1973). Phys. Lett. 47B, 365.

Gliozzi, F., J. Scherk and D. Olive (1976). Nucl. Phys. B122, 253.

Goddard, P. and C. B. Thorn (1972). Phys. Lett. 40B, 2; 235.

Goddard, P., J. Goldstone, C. Rebbi and C. B. Thorn (1973).

Nucl. Phys. B56, 109.

Goddard, P. and D. Olive (1984). In Vertex Operators in Mathematics and

Physics (Springer Verlag, 1984), 51.

Green, M. B. and J. H. Schwarz (1981). Nucl. Phys. B181, 502.

Green, M. B. and J. H. Schwarz (1982). Nucl. Phys. B198, 252; 441.

Green, M. B. and J. H. Schwarz (1983). Nucl. Phys. B218, 43.

Green, M. B. and J. H. Schwarz (1984). Phys. Lett. 149B, 117.

Gross, D. J., J. A. Harvey, E. Martinec and R. Rohm (1985).

Nucl. Phys. B256, 253.

- Hagedorn, R. (1968). Nuovo Cim. 56A, 1027.
- Mandelstam, S. (1973). Nucl. Phys. B64, 205.
- Mandelstam, S. (1974). Nucl. Phys. B69, 77.
- Mandelstam, S. (1975). Phys. Rev. D11, 3026.
- Nambu, Y. (1970), In Proceedings Int. Conf. on Symmetries and Quark Models (Gordon and Breach, 1970).
- Neveu, A. and J. H. Schwarz (1971). Nucl. Phys. B31, 86.
- Neveu, A. and J. Scherk (1972). Nucl. Phys. B36, 155.
- Polyakov, A. M. (1981). Phys. Lett. 103B, 207; 211.
- Ramond, P. (1971). Phys. Rev. D3, 2415.
- Scherk, J. (1975). Rev. Mod. Phys. 47, 123.
- Scherk, J. and J. H. Schwarz (1974). Nucl. Phys. B81, 118.
- Slansky, R. (1981). Phys. Rep. 79C, 1.
- Susskind, L. (1970). Nuovo Cimento 69A, 457.
- Thierry-Mieg, J. (1985). Phys. Lett. 156B, 199.
- Veneziano, G. (1968). Nuovo Cimento 57A, 190.
- Wess, J. and B. Zumino (1974). Nucl. Phys. B70, 39; Phys. Lett. 49B, 52.
- West, P. C. (1986). "Gauge Covariant String Field Theory",
CERN preprint TH 4460/86.

Recent Review Articles

- Schwarz, J. H. (1982). Phys. Rep. 89C, 223.
- Green, M. B. (1983). Surveys of High Energy Physics 3, 127.