# SUPERSTRINGS

**Considering the elementary building blocks of nature to be strings rather than point particles allows one to construct consistent quantum theories that unify gravity with the other known forces.**

John H. Schwarz

During the past three years many theoretical physicists have dedicated themselves to working on superstring theory. With varying degrees of conviction we believe that we have at hand for the first time many of the essential ingredients for an almost unique quantum theory that gives a unified description of all elementary particles and the forces between them. We also believe that this theory is free from the inconsistencies that have thwarted all previous attempts to construct a "unified field theory" that describes gravity together with the strong, weak and electromagnetic forces. In short, as some popular media like to put it, we may finally have "the theory of everything."

Some eminent physicists, however, view the emergence of superstring theory as a serious aberration that is corrupting a generation of students. They point out that there is no experimental evidence for superstrings currently, and claim that there never will be. Have we introduced a sinister new cult in the world of theoretical physics? Have many theorists been led astray by a few of us who wield excessive influence? Or could it be, as I would like to believe, that superstring theorists are developing new insights that will influence profoundly our view of the physical universe?

This article will present the case for superstrings and outline some of the obstacles that must still be overcome. A serious study of superstrings requires much modern mathematics. In fact, superstring theories are stimulating new developments in mathematics, and a new level of unification is emerging among particle theory, quantum gravity and certain branches of modern mathematics as a result of these developments. But we will settle here for a nonmathematical description of the results.

A string theory differs from a conventional quantum field theory in postulating that the elementary constituents of all matter are strings—one-dimensional curves—rather than point particles. The Planck length

$$L_{\rm P} = \left(\frac{\hbar G}{c^3}\right)^{1/2} \approx 1.6\times 10^{-33}\ \rm cm$$

and the Planck mass

$$M_{\rm P} = \left(\frac{\hbar c}{G}\right)^{1/2} \approx 2.2\times 10^{-5}\ {\rm g} \approx 1.2\times 10^{19}\ {\rm GeV}/c^2$$

characterize the size of these strings and their excitation energies, respectively. (See table 1.) Superstring theory differs from ordinary point-particle theories in important ways at these scales, but it may be well approximated by those theories at large distances or low energies. This extremely small length scale, or large energy scale, encourages the skeptics to claim that string theory will never be tested. If the energy available in laboratory experiments continues to increase by an order of magnitude every decade—a rate of increase that advances in accelerator technology have made possible over the past few decades—then we must admittedly wait for almost 200 years before we can directly study the Planck scale. But the Planck scale is necessarily the one relevant to a unified theory involving gravitation, so this objection would apply equally well to any other proposal for such a unification. In any case, as we will discuss, the theory should also have testable consequences at much lower energies.

## The standard model

A theoretical framework that describes all established experimental results in elementary-particle physics has been developed during the past 30 years; it is known in the trade as the standard model. To motivate our discussion of superstrings, let us review briefly some of the successes and limitations of the standard model.

Elementary-particle theorists have ignored the gravitational force between elementary particles until recently because it is very weak at the distances or energies explored experimentally. Theories of elementary parti-

**John H. Schwarz** is a professor of theoretical physics at the California Institute of Technology, Pasadena, California.

**Table 1. Characteristic scales**

| | Size cm | Excitation energy eV |
|---|---|---|
| Atoms | $10^{-8}$ | 10 |
| Nuclei | $10^{-13}$ | $10^{7}$ |
| Weak scale | $10^{-16}$ | $10^{11}$ |
| Planck scale | $10^{-33}$ | $10^{28}$ |

cles that ignore gravity must still be consistent with the principles of special relativity and of quantum theory. Quantum field theory is the mathematical framework that incorporates these two principles in a theoretical description of elementary particles regarded as point-like objects.

To be mathematically consistent and theoretically acceptable, a quantum field theory must be renormalizable and free from certain anomalies. Renormalizability means that the calculation of any physically measurable quantity in the theory gives a finite result in spite of the divergences that often appear when the quantity is expanded as an infinite series in the coupling constant. Anomalies in a field theory are terms that violate the conservation laws (or gauge invariance) when the theory is quantized. Many mathematically consistent quantum field theories are possible, and the standard model is a particular one selected on phenomenological grounds.

The standard model comprises particles of spin 0, $\frac{1}{2}$ and 1 only. Particles of each spin value play a specific role. Spin-1 particles transmit forces—gluons transmit the strong force, W and Z bosons the weak force and the photon the electromagnetic force. Quarks and leptons, the fundamental constituents of all ordinary matter, have spin $\frac{1}{2}$. Spin-0 particles, called Higgs scalars, induce spontaneous symmetry breaking (we will discuss this later).

Quantum field theories of a special type, called non-Abelian gauge theories or Yang–Mills theories, underlie the standard model. These theories are invariant under certain symmetry transformations of the basic field variables, whose magnitude may be chosen independently at each point in space–time. Lie groups provide a mathematical description of the continuous symmetry transformations in these theories. Sophus Lie, a Norwegian mathematician, defined the group theoretic properties of such continuous symmetry transformations in 1869, and Elie Cartan studied and classified them further. In the theory of Lie groups, a given finite symmetry transformation is built up by repeated application of infinitesimal transformations. A generator defines each infinitesimal transformation, and the number of linearly independent generators is called the dimension of the group. For example, SO($N$), the group of rotations in $N$ dimensions, is generated by infinitesimal rotations in planes described by pairs of axes; thus it has $N(N-1)/2$ generators. The group SO($N$) belongs to the "orthogonal" sequence—one of three infinite sequences that constitute the "classical" groups. The dimensions of groups in the other two sequences, called the unitary and symplectic sequences, are $N^2-1$ for $N$ larger than 1 and $N(2N+1)$ for $N$ larger than or equal to 1, respectively. These three classes and the five "exceptional" groups—$G_2$, $F_4$, $E_6$, $E_7$ and $E_8$, with dimensions 14, 52, 78, 133 and 248, respectively—are called simple Lie groups. There is in addition the one-dimensional Abelian group U(1), which describes the rotational symmetry of a circle.

Local fields, called Yang–Mills or gauge fields, associated with the generators of the Lie group define a gauge field theory. Spin-1 particles in the standard model are the quanta of these fields. It is possible to construct a gauge theory for any combination of simple Lie groups and U(1) factors. Quantum electrodynamics is a theory based on a U(1) symmetry. The particular choice employed in the standard model, SU(3)$\times$SU(2)$\times$U(1), is the minimal one consistent with the observed particles and couplings.

The gauge fields couple to the fields for quarks and leptons and transmit forces between them. The eight spin-1 fields called gluons, which are associated with the eight generators of the SU(3) subgroup, called the color group, describe the strong nuclear force, or color force, between quarks. The quarks come in three colors—red, green and blue (see table 2)—and the strong interaction transmitted between them by gluons is described by saying that the quarks form a three-dimensional representation of the color group: Quarks of different colors turn into one another under the action of the color force in exactly the same way that the components of a complex three-dimensional vector transform under the action of the group SU(3). On the other hand, leptons are color singlets—the lepton wavefunctions are unchanged by the color SU(3) transformations—and do not feel the color force. Table 2 lists the members of the "first family" of quarks and leptons, consisting of up and down quarks, the electron and the electron–neutrino. As the table indicates, two more families are known. There is no evidence for a fourth family, but this possibility is not yet definitively excluded. The SU(2)$\times$U(1) symmetry in the standard model gives a (partially) unified description of the electromagnetic and weak forces. The fields associated with its four generators correspond to the photon, which transmits the electromagnetic force, and the $W^{\pm}$ and $Z^0$, which carry the weak nuclear force. The group SU(2) does not act independently on the quark doublet in each family, but there is some "mixing" between families.

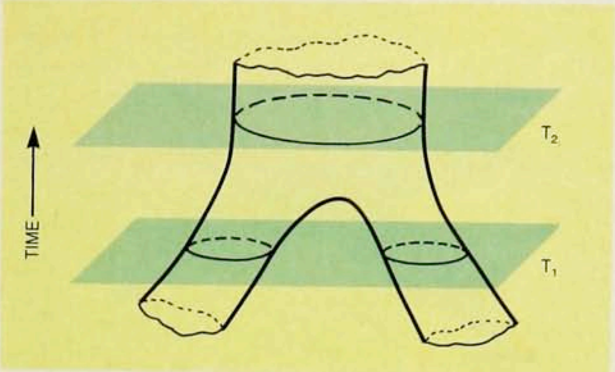
The electroweak transformation rules of the quarks and leptons are quite interesting. The fermions with "left-handed" polarization transform as doublets and feel the SU(2) force, whereas the ones with "right-handed" polarization do not, because they transform as singlets. They all feel the U(1) force. This is where parity violation is built into the theory. Altogether, counting polarizations, there are 12 quarks and 3 leptons in a family. Because the up quarks have charge $+\frac{2}{3}$ and the down quarks have charge $-\frac{1}{3}$, the total charge of the 12 quarks is $+2$ while the total charge of the 3 leptons is $-2$. That the sum is zero is important for ensuring that certain anomalies of the quantum theory cancel.

Each of the symmetries in a gauge theory represents an exact symmetry of the fundamental equations that describe the dynamics of various fields. However, it can happen that the solution with the lowest energy (the ground state or vacuum of the theory) does not have the full symmetry of the equations. In this case, one says that the symmetry is spontaneously broken. In the standard model, three of the 12 symmetries are spontaneously broken. Specifically, the eight generators of the SU(3) color group and one generator of the U(1) electromagnetic group correspond to unbroken symmetries. This U(1) generator corresponds to the photon; it is in fact a linear combination of a U(1) subgroup of the SU(2) factor and the U(1) factor itself, the Weinberg angle $\theta_W$ being a measure of this mixing. The gluons, the photon and, more generally, spin-1 particles that correspond to unbroken symmetry generators are massless, whereas the ones that correspond to spontaneously broken symmetries ($W^{\pm}$, $Z^0$) acquire mass. Higgs scalars—particles with spin 0—with suitable self-couplings and interactions with the other fields are introduced in the standard model to break the symmetries. Many theorists consider this an unaesthetic feature of the standard model. The construction gives a renormalizable quantum field theory, but it is disturbing that some "fine-tuning" of parameters is required to achieve the observed hierarchy of mass scales.

## Problems with the standard model

Impressive as the success of the standard model is, there is much that it does not explain. The choice of symmetry groups and representations is made on phenomenological grounds; the number of families is similarly chosen. Several coupling constants (especially those involving the Higgs fields), quark masses and mixing angles are parameters whose values may be freely adjusted to fit the facts. One hopes that many, or maybe even all, of these features are derivable from more fundamental principles. A nontrivial extension of the standard model—a deeper theory that reduces to the standard model at low energies—is therefore desirable. Moreover, the standard model is incomplete: It does not include gravity.

There have been attempts to combine general relativity, viewed as a classical field theory, with the standard model. But severe problems arise when the resulting system is interpreted as a quantum theory. In fact, any quantum mechanical description of gravity necessarily involves some very subtle conceptual issues. For example, conventional approaches to quantum theory require knowing whether two points have a space-like or time-like separation, but in a theory including gravity this is determined by the dynamics, and, furthermore, the answer is presumably only described by a probability amplitude. Another puzzling question is whether it is sensible to extend quantum mechanical notions to the entire universe—introducing a "wavefunction of the universe," for example, inevitably leads to the bizarre "many worlds" interpretation of quantum mechanics.




'Pants diagram' is a piece of world sheet representing the space-time history of two closed strings that join. There are two closed strings at time $T_1$, but only one at a later time $T_2$. **Figure 1**

There is also the disturbing possibility, pointed out by Stephen Hawking (Cambridge University), that black holes cause a loss of quantum phase coherence in the observable universe, in which case a density-matrix formalism would be required, instead of a pure state wavefunction.

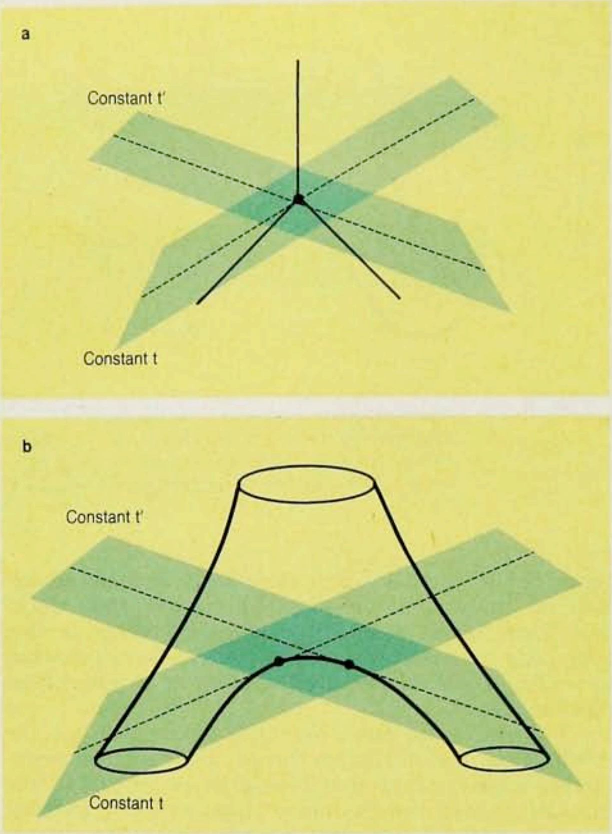
There are also more mundane issues that can be investigated in perturbation theory. For example, when general relativity is treated in isolation or coupled to the standard model, the Feynman diagrams that describe various quantum mechanical corrections give divergences that are not renormalizable. Although not every case has been studied completely, it is almost certain that all such theories are nonrenormalizable. From this I conclude that there is no consistent theory of quantum gravity that describes the elementary particles as points.

## Extensions of the standard model

Grand unification, supersymmetry, extra dimensions of space—these are some of the more significant theoretical proposals for going beyond the standard model. Each of these proposals can be incorporated in an ordinary gauge field theory, but they all fit naturally into string theories. I will briefly elaborate upon them in this section.

It seems to be a general principle that any new idea in elementary-particle physics implies new particles. (This is true for bad ideas as well as good ones.) Each of the three proposals to extend the standard model also implies the existence of new kinds of particles—axions, fractional electric charges and magnetic monopoles, for example. (Besides these, the string theory also predicts string excitations and shadow matter.) That none of these particles has been observed yet can be understood as a result of their either being too heavy to be produced in accelerators or else interacting too weakly to be observed. However, we may hope to observe some of them eventually.

Grand unification is a proposal for unifying the theoretical description of the strong force with that of the electroweak force. This is achieved by embedding the color SU(3) symmetry and the electroweak SU(2)×U(1) symmetry in a larger simple Lie group. This replaces the standard model with an asymptotically free theory that is better behaved at short distances. The smallest group that can accommodate the symmetries of the standard model is SU(5), and theories based on this symmetry have received the most attention. Larger groups such as SO(10), $E_6$, $E_7$ and $E_8$ have also been considered. Grand unification has




Space–time point at which two strings interact (b) is not unique but depends on the Lorentz frame; it is the point at which the time slice is tangent to the world sheet. In contrast, the point of interaction between two particles (a) is the same in all Lorentz frames. This difference is one reason why there is much more arbitrariness in the construction of theories for interacting particles than theories for interacting strings. **Figure 2**

had both successes and failures. The successes include a calculation of $\sin^2\theta_W$ to about 5% accuracy; grand unification has also allowed more elegant group theoretic assignments for classifying the quarks and leptons. The failures include the prediction of proton decay at a level that is now excluded experimentally and predictions for certain mass ratios of quarks and leptons that also do not agree with the measured values. However, that the successes of grand unification depend on rather general features, whereas its failures tend to depend sensitively on details of the particular model, gives some support to this proposal. Indeed, there is some evidence that in the context of string theory the successes survive but the failures do not.

Supersymmetry is another extension of the symmetry of the standard model—and one with profound implications. This symmetry has a somewhat different mathematical structure from that of ordinary Lie group symmetries, whose conserved charges are rotationally invariant (scalars). Supersymmetry charges, by contrast, transform under rotations like spin-$\frac{1}{2}$ particles (spinors). As a consequence, the irreducible representations used to describe particles contain different spins. In other words, a supersymmetry transformation "rotates" a particle of, say, integer spin into a partner particle of half-integer spin. When successive supersymmetry transformations are performed, the combined effect includes a space–time translation. It has been proved that supersymmetry is the only possible nontrivial extension of the Poincaré symmetry—translations, rotations and Lorentz boosts—of space-time. It is clearly of fundamental importance to learn whether nature uses supersymmetry.

Theoretical arguments suggest that divergence cancellations in supersymmetric theories could be helpful in resolving the fine-tuning problem alluded to earlier. There is no experimental evidence for supersymmetry yet. If the symmetry is relevant to nature, it must be spontaneously broken because unobserved partner particles certainly do not have the same masses as the known particles. Another reason for taking supersymmetry seriously is that it plays a central and inescapable role in string theory. However, our present level of understanding does allow for the possibility that it is completely broken at the Planck scale, in which case it would not be observable.

If supersymmetry is the correct solution to the fine-tuning problem, it should survive in some form down to the 100-GeV scale characteristic of weak interactions. Each known particle would then have one or more partner particles, with masses of the order of 100 GeV. If this is the case, some supersymmetry particles should be discovered soon, perhaps at the Fermilab collider. The names and spins of some of these hypothetical particles are listed in table 3. The experimental discovery of any of these would ensure a rich spectrum of new particles whose study could keep the proposed Superconducting Super Collider busy for a long time and would provide theorists with crucial guidance in sorting through the possible models.

I have listed in table 3 the spin-2 graviton and its partner, the spin-$\frac{3}{2}$ gravitino. The graviton, the quantum of gravity, is not part of the standard model. It interacts much too weakly at ordinary energies to be detected. However, we know that it must have spin 2 because the space–time metric in general relativity is a symmetric matrix. This distinguishes gravity from the other forces, which are mediated by spin-1 particles. Supersymmetric gravity—supergravity—theories, which contain one or more gravitinos, have been studied extensively during the past decade. Gravitinos are gauge particles for supersymmetry analogous to the spin-1 particles associated with Lie group symmetries.

The geometry of space–time is determined dynamically in general relativity. This is also true of string theory, which is a generalization of general relativity. In this context, it may be sensible to consider the possibility of extra dimensions of space; if the dynamics force the extra dimensions to curl up into a sufficiently small space, the resulting theory will not be in conflict with the observed three-dimensionality of the physical world. Remarkably, this idea goes back to the work in the 1920s of Theodor Kaluza and Oskar Klein, who suggested that a fifth dimension could be used to unify general relativity with electrodynamics. That scheme is no longer viable, but generalizations of the idea have been investigated in recent years as an outgrowth of work in string theory. These have been studied intensively in the context of various supergravity theories; an 11-dimensional version of supergravity was especially fashionable for a few years. The preferred dimensionality in superstring theories is ten—nine space and one time—so that six spatial dimensions should curl up, or "compactify."

## Three superstring theories

String theory has had an interesting history, but this is not the place to discuss it in detail. Suffice it to say that ideas for a string theory first appeared in physics in the late 1960s in attempts by Gabriele Veneziano (now at CERN), Yoichiro Nambu (University of Chicago) and many others to explain the physical origin of some mathematical features of strongly interacting particles, or hadrons. But these string theories predicted a massless spin-2 particle that had no relevance to hadron physics. The late Joël Scherk (Ecole Normale Supérieure, Paris) and I proposed in 1974 that this difficulty could be turned into a virtue by using strings of a size on the scale of the Planck length to describe gravity in unification with the other forces. This suggestion, we then thought, could overcome the divergences that plagued all attempts to develop a quantum theory of gravity. But interest in string theories—especially as candidates for a theory of hadrons—declined in the mid-1970s as the SU(3) color theory and the standard model became successful in explaining most experimental data. The current excitement started with the discovery—a mathematical discovery, one may call it—by Michael Green (Queen Mary College, London) and myself in 1984 that a particular string theory with space–time supersymmetry, and hence called superstring theory, is free from anomalies in ten dimensions only when the internal symmetry group is SO(32) (see PHYSICS TODAY, July 1985, page 17). I will discuss only superstring theories in what follows.

I am often asked why we stop at strings and why we do not consider objects with more than one dimension. It is extremely difficult to formulate a theory of elementary extended objects that is consistent with the usual requirements—such as unitarity and causality—of quantum theory. In the case of strings there appear to be a few schemes that are consistent. It is not known whether there are any at all for objects that have more than one dimension, such as two-dimensional membranes, but it would be a surprise if there were. The existence of string theories depends on special features that do not generalize to higher-dimensional objects.

It is remarkable that the particle spectra of all classical solutions of the known string theories each contain exactly one massless spin-2 graviton. Moreover, this graviton interacts in accord with the dictates of general covariance, which implies that general relativity gives a correct description at low energies. And the Planck length—the length scale characteristic of strings—arises naturally when we require the gravitational coupling to have the usual Newtonian value. This is the feature that led Scherk and me to propose that superstrings might be relevant to quantum gravity.

Strings can occur in two distinct topologies called open and closed. Open strings are line segments with free ends, whereas closed strings are loops (with the topology of a circle) and no free ends. In some theories strings have an intrinsic orientation (representable by an arrow). The various quantum mechanical excitations (normal modes) of the string for each solution of a particular string theory are interpreted as giving a spectrum of elementary particles. The excitations may involve rotational and vibrational degrees of freedom of the string or excitations of the various "internal" degrees of freedom that reside on it. The internal degrees of freedom arise from Lie group symmetries, supersymmetry and so forth. In string theory, one has a unified view of the rich world of elementary particles as different modes of a single fundamental string. String states that have masses much smaller than the Planck mass are finite in number and should correspond to observable particles. There are also an infinite number of modes with masses on the order of or larger than the Planck mass that are probably not observable. In general, they are unstable and decay into the light modes, although there could be some with magnetic charge, fractional electric charge or some other exotic property that are stable. Since we are unlikely to be able to make such superheavy particles, it would be possible to observe them only if they already exist in sufficient numbers as remnants of the Big Bang.

### Table 2. Quarks and leptons

| | Quarks color triplets | | Leptons color singlets | |
|---|---|---|---|---|
| Weak doublets (left-handed) | $\begin{pmatrix} u_R u_G u_B \\ d_R d_G d_B \end{pmatrix}$ | | $\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}$ | |
| Weak singlets (right-handed) | $(u_R u_G u_B)$ $(d_R d_G d_B)$ | | $e^-$ | |
| **Families** | | | | |
| 1 | u | d | e | $\nu_e$ |
| 2 | c | s | $\mu$ | $\nu_\mu$ |
| 3 | t | b | $\tau$ | $\nu_\tau$ |
| 4 | | ? | | |

R, red; G, green; B, blue.

Three consistent superstring theories are known. The type I theory is based on unoriented strings that may be open or closed. The other two theories are based on oriented closed strings that differ in internal symmetry; one of these is referred to as the type II superstring theory and the other as the heterotic string theory. (The developers of the heterotic theory point out that "heterosis" means increased vigor due to crossbreeding. This theory combines features of the superstring theory and the old bosonic string theory.)

The three theories are completely free of adjustable dimensionless parameters or any other arbitrariness. Thus, aside from this threefold choice, there is a completely unique theory that consistently incorporates quantum gravity. Of course, additional theories may still be found. It is also possible that the list will shrink. This could happen if two of the theories are shown to be equivalent or if one of them is found to be inconsistent. For example, if the type I theory turns out to be inconsistent and the other two theories are equivalent, we would be left with a unique theory to explain all fundamental physics.

But it is not enough to know the right theory. *Solutions* to equations, and not the equations themselves, provide a mathematical description of natural phenomena. To start with, we would want to know the quantum state of lowest energy and the low-lying excited states in any theory. It can happen that a theory has many possible vacuum configurations, or ground states. In that case one must make an arbitrary (phenomenological) choice to describe the experimental data—perhaps even adjusting a number of parameters—despite the underlying theory's uniqueness. We are faced with precisely this problem in string theory. A large number of solutions, all theoretically acceptable, seem possible. It would be disappointing, to say the least, if the appropriate solution must be chosen phenomenologically. Thus many string theorists specu-

### Table 3. Supersymmetry particles

| Particle (spin) | Supersymmetry partner (spin) |
|---|---|
| Gluon (1) | Gluino (½) |
| Photon (1) | Photino (½) |
| W (1) | Wino (½) |
| Quark (½) | Squark (0) |
| Lepton (½) | Slepton (0) |
| Graviton (2) | Gravitino (3/2) |

## Table 4. Ten-dimensional solutions

| Theory | Symmetry group | Number of supersymmetry generators |
|---|---|---|
| Type I | SO(32) | 1 |
| Type II | — | 2* |
| Type II | — | 2** |
| Heterotic | $E_8 \times E_8$ | 1 |
| Heterotic | SO(32) | 1 |
| Heterotic | SO(16)×SO(16) | 0 |

* The generators differ in handedness.
** The generators have the same handedness.

late that all but one or a few of the solutions will turn out to be inconsistent or unstable under a more thorough analysis that does not depend on perturbative expansions. As far as I can tell, this is only wishful thinking. However, optimistic conjectures have turned out to be correct on many previous occasions in this discipline. That mathematical "miracles" continue to turn up suggests that some fundamental features of string theory are not yet well understood.

In a theory of gravity, characterization of the vacuum configuration includes knowing the geometry of space-time. Can we derive the geometry of four-dimensional Minkowski space or of a realistic cosmology from superstring theories? There are perturbative classical solutions to each of the three theories for any space–time dimensionality less than or equal to ten. Thus the dimensionality of space–time is properly regarded as a property of the solution and not of the theory itself. Many of the solutions with dimensionality less than ten can be interpreted as having a ten-dimensional space–time manifold in which $10 - D$ spatial dimensions form a compact space $K$, so that altogether the space–time is a direct product of $D$-dimensional Minkowski space and $K$. However, there are other classes of solutions with fewer than ten dimensions that do admit such an interpretation.

The case $D = 10$ is special in that it is the largest value possible for any solution. One could say that string theory "predicts" that the dimensionality of space–time cannot be more than ten, but this is not a terribly enlightening statement. Though it would be much more satisfying to know why space–time has four dimensions, we have not achieved such an understanding yet. Indeed, each of the theories admits ten-dimensional solutions that are consistent as far as we can tell. These solutions (listed in table 4) are certainly not realistic, but they do seem to be of fundamental importance from a theoretical point of view. It is a real challenge to find a good theoretical reason to exclude them as potential vacuum configurations.

For fewer than ten dimensions, the number of vacuum configurations is much larger; enumerating them and identifying those that could be realistic has become something of an industry. At the moment the heterotic theory seems to offer the best prospects for realistic solutions, but it is not out of the question that the type I or type II superstring theories could also yield phenomenologically viable solutions.

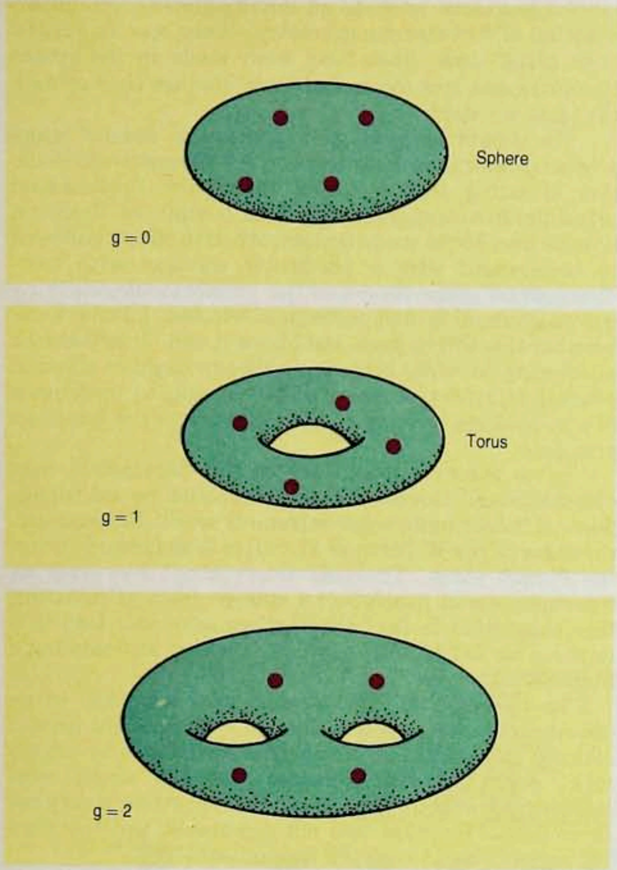
Will string theory ever be tested? It seems to me that there are several promising possibilities. First, the theory should enable us to calculate the properties of elementary particles at ordinary energies. A great deal of particle physics data should be calculable if the theory is unique and its solutions do not permit too much freedom. Some simple examples suggest that "low energy" phenomena should not be especially difficult to extract. Second, some particles with mass on the order of the Planck mass that were formed early in the Big Bang may have survived to the present epoch as observable stable entities. Magnetic monopoles could be one example. Characteristic features of superstring theory may also be required for an understanding of the cosmology of the very early universe. Our present understanding of string theory is not sufficient to allow definitive predictions in this regard, but with all the brainpower being brought to bear, there is no reason to be pessimistic about the eventual testability of the theory. As Edward Witten (Princeton University) recently noted, general relativity gave rise to various predictions that seemed quite hopeless to verify when they were made. Neutron stars, black holes, gravitational radiation and gravitational lenses may be counted among these predictions—and there is substantial observational evidence now for all of them.

## String interactions

Interactions between point particles are represented by Feynman diagrams in the perturbation expansion treatment of a quantum field theory. A trajectory in space-time, called the world line, describes the history of a particle's motion; world lines meet and bifurcate to represent the interactions the particle undergoes. The sum of the contributions associated with all allowed diagrams having the chosen initial and final states gives the complete interaction amplitude for those states. In particular, the diagrams must include all possible interactions appropriate to the theory in question. The diagrams can be classified by their topological properties; their contribution for a particular topology is given by a finite-dimensional integral. The integrals usually diverge, but there is a well-defined prescription for extracting finite results unambiguously in renormalizable theories.

String interactions can be formulated in an analogous manner. The space–time trajectory of a string is a two-dimensional surface called the world sheet. Feynman diagrams are then two-dimensional surfaces with specific incoming and outgoing strings, and are once again classified by their topology. The possible world sheet topologies are more limited in the type II and heterotic theories than in the type I theory. In the following I will therefore consider only the type II and the heterotic string theories. (The basic ideas are essentially the same in the type I theory.)

The type II and heterotic string theories each have a single fundamental interaction. It can be depicted by a portion of the world sheet, called the "pants diagram." (See figure 1.) When a plane representing a time slice at time $T_1$ intersects the diagram, one sees two closed strings. Intersecting the surface with a time slice at time $T_2$ reveals just one closed string. Clearly, at intermediate times the two closed strings approached each other,




**Feynman diagrams for strings** are two-dimensional surfaces. The diagrams are classified by the genus of, or the number of handles on, the surface. Points on the surfaces are topologically equivalent to tubes extending out to infinity and represent external strings. The genus $g$ of the surface also counts the number of loops, or the powers of $\hbar$, in the perturbative expansion of the theory. In type II and heterotic string theories, there is only one diagram at each order in $\hbar$. **Figure 3**

touched and joined. The reverse process in which one closed string splits to give two is also allowed.

The pants diagram describes an interaction that differs in fundamental respects from interactions in point-particle theories. A point-particle vertex and the pants diagram are drawn in figure 2. At what space–time point, we ask in each case, does the interaction that turns two particles into one take place? We can represent the time slices corresponding to two observers in distinct Lorentz frames by lines of constant $t$ or $t'$. The interaction in the point-particle theory occurs at a definite space–time point that all observers will identify unambiguously; in the string case, on the other hand, the interaction occurs at the point where the time slice is tangent to the surface, and this differs from one observer to another.

The interactions in figures 2a and 2b clearly differ fundamentally. The "manifold" of lines in the point-particle case is singular at the junction. Arbitrary choices are possible in the association of interactions with such vertices. This is part of the reason why ordinary quantum field theory has so much freedom in its construction. The string world sheet is a smooth manifold with no preferred points. That it describes interaction is purely a consequence of the topology of the surface. The nature of the interaction is therefore completely determined by the structure of the free theory, with none of the arbitrariness that exists in the point-particle case.

One may describe string world sheets as Riemann surfaces using techniques of complex analysis. This means that one can use complex coordinates $z$ and $\bar{z}$. A fundamental feature of string theory is that world sheets related by a conformal mapping $z \to f(z)$ are regarded as equivalent. Thus only surfaces that are conformally inequivalent need be included in performing the sum over distinct geometries. Fortunately the conformally inequivalent geometries for each topology can be characterized by a finite number of parameters, and thus the Feynman integrals are finite-dimensional.

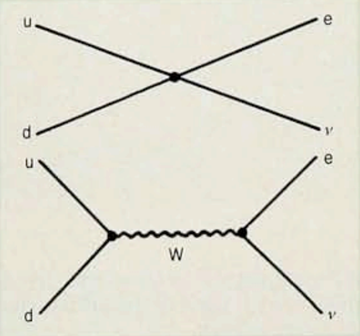
The topological classification of the Feynman diagrams is especially simple in the type II and heterotic theories. A single integer, the genus $g$, which is the number of handles on the surface (see figure 3), characterizes the world sheets. One can represent the external strings as points on the surface (shown as dots in the figure) because they are conformally equivalent to tubes extending off to infinity. The genus corresponds to the number of loops, or the power of $\hbar$, in the perturbation expansion. It is remarkable that there is just one diagram at each order of the perturbation expansion, especially as the number of them in ordinary quantum field theory is very large indeed.

The convergence properties of the integrals associated with diagrams in the string theory are also much better. The properties of multiloop ($g > 1$) amplitudes are not yet fully understood. The analysis involves various sophisticated issues at the frontiers of the theory of Riemann surfaces, algebraic geometry and maybe even number theory. However, it seems that all the divergences are of kinds that are well understood and must inevitably be present. For example, infrared divergences are expected in any sensible theory containing massless particles in four dimensions. We know how to deal with them. Similarly, divergences in amplitudes calculated with massive external string states are traced to the fact that the mass is shifted by the interactions. Other divergences, associated with so-called dilaton tadpoles, are also well understood and harmless. Most importantly, the kinds of divergences that result in parameters becoming arbitrary in renormalized quantum field theories or in amplitudes becoming completely undefined in nonrenormalizable field theories have no counterparts in string theory. (See the box on page 40.)

## Remaining challenges

Superstring theory is developing at a breathtaking pace as more and more clever people join in the enterprise. The project is indeed an enormously ambitious one, and many formidable hurdles remain. Determining a complete list

## A renormalizable quantum theory of gravity?




It may seem paradoxical that quantum corrections to general relativity give nonrenormalizable divergences, whereas string theory, which agrees with general relativity at low energies, is nonsingular. The essential reason can be traced to effects at the Planck scale that are present in string theory but not in general relativity. In particular, there is an infinite spectrum of string modes corresponding to particles with masses on the order of or greater than the Planck mass. These states contribute as virtual particles in scattering processes to produce subtle patterns of cancellations that soften the high-momentum (ultraviolet) behavior of the Feynman integrals.

This cancellation phenomenon is somewhat analogous to one that has played an important role in the recent history of particle physics. Before the discovery of the electroweak theory based on the SU(2)×U(1) symmetry, the "four-fermion theory" was developed to describe, for example, the neutron beta decay

$$n \rightarrow p + e^- + \bar{\nu}_e$$

The four Fermi fields in this theory were taken to interact at a point. One of the major problems with the four-fermion theory was that it was not renormalizable and quantum mechanical corrections could not be calculated for it. There is no four-fermion interaction in the electroweak theory. Instead, beta decay is described by the exchange of a virtual W particle, as depicted in the lower diagram. The two descriptions agree quite accurately, since the energy in the decay is much less than the W mass. However, the electroweak theory is renormalizable, so quantum mechanical corrections are well defined in it. String theory modifies the multi-graviton interactions in general relativity in an analogous fashion.

of consistent string theories is one hurdle I have already mentioned: Three theories are known, but it would be extremely nice if the number could be reduced to one. Then we could argue that there is a unique consistent theory that accounts for all of fundamental physics.

Development of fundamental principles for superstring theories and a more geometric formulation for them has attracted a great deal of effort. String theory has had a peculiar history, to say the least. It is instructive to compare its development with that of general relativity. In the case of relativity, Einstein began by formulating certain far-reaching principles—the equivalence principle and general covariance—then found their proper mathematical embodiment in the language of Riemannian geometry. This led to dynamical equations and experimental predictions, many of which have been tested and verified. In string theory, we have not yet identified the fundamental principles that generalize the equivalence principle and general coordinate invariance. These must surely exist, because general relativity is a low-energy (long-distance) approximation to string theory. Whatever these principles may be, they are likely to require a new kind of geometry, perhaps an infinite-dimensional generalization of Riemannian geometry. Some specific suggestions along these lines have been made in the recent literature, but it is too early to say whether they contain the ideas we want.

We should be in a good position to answer many profound questions once the correct geometric formulation of string theory, incorporating the fundamental principles in a comprehensible form, is achieved. It should then be possible to study nonperturbative effects and even to understand why a particular solution with four-dimensional space–time and the phenomenologically required symmetries and particles is selected. I do not know whether this will happen, but I hope it will. It will also be interesting to study how string theory modifies classical general relativity at short distances and to investigate how it resolves some of the profound issues of quantum gravity.

In a theory without adjustable parameters, any dimensionless number in nature should be calculable. Some of these numbers are extremely small. For example, the mass of the W boson is 17 orders of magnitude below the Planck mass. Theorists worry about how such an extremely small number can emerge from calculations. One suggestion is that the mathematics will lead to a formula for $\log(M_W/M_P)$, which is not so intimidating a number.

In the case of the cosmological constant, whose dimensionless value is less than $10^{-120}$, we might hope to identify a symmetry principle that forces it to be exactly zero. Some theorists consider this the single most challenging problem in physics. Prior to string theory the cosmological constant was not calculable, and therefore the problem could not even be studied.

I am quite confident that we are closing in on a unique fundamental theory of nature. But it is unrealistic to expect too much too soon. It will probably take a few decades of hard work to obtain a satisfactory understanding of what string theory is really all about. Settling this question will certainly require advances in mathematics, but the experimental results in the next 10–20 years also are likely to play an important role in shaping our ideas.

* * *

*I am grateful to Patricia Schwarz for suggesting improvements to this article.*

## References


1. J. H. Schwarz, ed., *Superstrings—The First 15 Years of Superstring Theory*, World Scientific, Singapore (1985). M. B. Green, J. H. Schwarz, E. Witten, *Superstring Theory*, Cambridge U. P., Cambridge, England (1987).
2. P. Ramond, Phys. Rev. D **3**, 2415 (1971). A. Neveu, J. H. Schwarz, Nucl. Phys. B **31**, 86 (1971).
3. J. Scherk, J. H. Schwarz, Nucl. Phys. B **81**, 118 (1974).
4. F. Gliozzi, J. Scherk, D. Olive, Nucl. Phys. B **122**, 253 (1977).
5. M. B. Green, J. H. Schwarz, Phys. Lett. B **109**, 444 (1982); Phys. Lett. B **122**, 143 (1983); Phys. Lett. B **136**, 367 (1984); Phys. Lett. B **149**, 117 (1984); Phys. Lett. B **151**, 21 (1985).
6. D. J. Gross, J. A. Harvey, E. Martinec, R. Rohm, Nucl. Phys. B **256**, 253 (1985); Nucl. Phys. B **267**, 75 (1986).
7. P. Candelas, G. Horowitz, A. Strominger, E. Witten, Nucl. Phys. B **258**, 46 (1985). E. Witten, Nucl. Phys. B **258**, 75 (1985).
8. D. Friedan, E. Martinec, S. Shenker, Nucl. Phys. B **271**, 93 (1986).
9. E. Witten, Nucl. Phys. B **268**, 253 (1986); Nucl. Phys. B **276**, 291 (1986).
10. Less technical treatments can be found in M. B. Green, Sci. Am., September 1986, p. 48; G. Taubes, Discover, November 1986, p. 34; M. Kaku, J. Trainer, *Beyond Einstein*, Bantam, New York (1987). ■