String Theory

John H. Schwarz

Traditional studies of the relativistic quantum physics of elementary particles assume that the particles can be described as mathematical points without any spatial extension whatsoever. This approach has had impressive success, but appears to break down at extremely high energies or short distances where gravitational forces become comparable in strength to the nuclear and electromagnetic forces that act between particles.

In 1974 the late Joel Scherk and I proposed overcoming this limitation by basing a unified description of all elementary particles and the forces that act among them on fundamental one-dimensional curves, called “strings”, rather than on point particles. String theories appear to be free from the inconsistencies that have plagued all previous attempts to construct a “unified field theory” that describes gravity together with the other forces. “Superstring” theories, which contain a special kind of symmetry called supersymmetry, show the most promise for giving realistic results.

Three String Theories

There are many consistent quantum field theories for point particles, although none contains gravity. It is extremely difficult, on the other hand, to formulate a theory of elementary extended objects that is consistent with the usual requirements of quantum theory. In the case of strings (one-dimensional elementary objects), there are a few schemes that appear to be consistent. It is not known whether there are any at all for objects of more than one dimension (such as two-dimensional membranes), but this appears to be very unlikely. The existence of string theories depends on special features that do not generalize to higher-dimensional objects.

Remarkably, every classical solution of each of the known string theories gives rise to a particle spectrum that contains 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. The characteristic length scale of the strings is the Planck length.[1] This is determined by requiring the gravitational coupling to have the usual Newtonian value.

Strings can occur in two distinct topologies called open and closed. Open strings are line segments with two free ends, whereas closed strings are loops with no free ends. Depending on the theory, strings may have an intrinsic orientation or not. For each solution of a particular string theory, there is a

[1] The Planck length is formed from the most fundamental constants – Planck’s constant ($\hbar$), the speed of light (c), and Newton’s constant (G). Specifically,

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

Table 1: Consistent string theories

Name	String topology	String orientation
Type I	Open and closed	Unoriented
Type II	Closed	Oriented
Heterotic	Closed	Oriented

corresponding spectrum of elementary particles given by the various quantum-mechanical excitations (normal modes) of the string. These include rotational and vibrational excitations as well as excitations of various "internal" degrees of freedom that can reside along the string. The internal degrees of freedom can describe 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 mass much below the Planck mass are finite in number and should correspond to observable particles. There are also an infinite number of modes with mass comparable to 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 or 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 only be possible to observe them if they already exist in sufficient number as remnants of the big bang.

At the present time there are three known consistent string theories (listed in Table 1). The type I superstring theory consists of both open and closed strings that are unoriented. The type II superstring theory and the heterotic string theory consist of oriented closed strings only. (They are distinguished by different internal degrees of freedom.) Each of the three theories is completely free of adjustable dimensionless parameters or any other arbitrariness. Thus, aside from this threefold choice, one has a completely unique theory that consistently incorporates quantum gravity.

To know the right theory is not enough. After all, nature is described by *solutions* of the equations. Normally, what is relevant is the quantum state of lowest energy (called the "vacuum") and low-lying excited states (which can be interpreted in terms of particle creation). It can happen that a theory has many different possible vacuum configurations. In this case one must make an arbitrary (phenomenological) choice to describe the experimental data – perhaps even adjusting a number of parameters – despite the fact that the underlying theory is itself unique. Precisely this problem seems to be facing us in the case of string theory. Despite the near uniqueness of the theory, there seems to be a very large number of possible solutions, any one of which is theoretically acceptable. If the choice can only be made phenomenologically, that would be disappointing. Thus, many string theorists speculate that when "nonperturbative" effects are properly understood, all but one (or a few) of the solutions will turn out to be inconsistent or unstable.

In a theory of gravity, the dynamics determines the geometry of space-time as

Table 2: $D = 10$ solutions

Theory	Symmetry group	Space-time supersymmetry
Type I	$SO(32)$	$N = 1$
Type II	–	$N = 2A$
Type II	–	$N = 2B$
Heterotic	$E_8 \times E_8$	$N = 1$
Heterotic	$SO(32)$	$N = 1$
Heterotic	$SO(16) \times SO(16)$	$N = 0$

part of the characterization of the vacuum configuration. We would like to derive the geometry of four-dimensional Minkowski space or a realistic cosmology. In fact, there are classical solutions to each of the three theories with any space-time dimension $D \leq 10$. Thus, the dimension of space-time is properly regarded as a property of the solution and not of the theory itself. Many of the solutions with $D < 10$ 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 $M_D \times K$ – a direct product of D-dimensional Minkowski space and K. However, there are other classes of solutions with $D < 10$ that do not seem to admit such an interpretation.

The case $D = 10$ is special in that it is the largest value possible. This is progress, but it would be much more satisfying to understand why $D = 4$ is necessary, which has not yet been achieved. Indeed, each of the theories admits solutions with $D = 10$, listed in Table 2, that are consistent so far as we can tell. These solutions 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 reason to exclude them as potential vacuum configurations. For $D < 10$ the number of vacuum configurations is much larger. It is a big industry to try to construct a complete list and identify the ones that could be realistic. 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.

Let us turn to the question of whether string theory can be tested. It seems to me that there are several promising possibilities. The first is to use it to calculate the properties of elementary particles at ordinary energies. After all, if the theory is unique and the solutions to the theory do not have too much freedom, then a great deal of particle physics data should be calculable. There is no reason that "low-energy" phenomena should be especially difficult to extract. A second possibility is that some Planck scale particles (such as magnetic monopoles) were formed early in the Big Bang and survive to the present epoch as observable stable entities. A related possibility is that characteristic features of superstring theory are required for a successful understanding of the cosmology of the very early universe. Our present understanding of string theory is not yet sufficient to make definitive testable predictions in any of these areas.

But with all the brainpower that is being brought to bear, the rate of progress is very impressive. Pessimism about eventual testability is probably unwarranted. (As Witten has noted, general relativity gave rise to various predictions that seemed quite hopeless to verify when they were made. These included neutron stars, black holes, gravitational radiation, and gravitational lenses. There is now substantial observational evidence for all of these.)

The Structure of Interaction

In the perturbation expansion treatment of quantum field theory, point-particle interactions are represented by Feynman diagrams. The history of the motion of a particle is a trajectory in space-time called the world line of the particle. Interactions are represented by joining or bifurcating world lines. The complete interaction amplitude for a given set of incoming particles and a given set of outgoing particles is given by a sum of contributions associated with all allowed diagrams with the chosen initial and final states. In particular, they must include all possible interactions appropriate to the theory in question. The diagrams can be classified by their topology and the contributions from diagrams of any particular topology is given by a finite-dimensional integral. The integrals usually diverge, but in renormalizable theories there is a well-defined prescription for extracting finite results unambiguously.

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 two-dimensional surfaces with specific incoming and outgoing strings, once again classified by their topology. The possible world-sheet topologies are more limited in the case of the type II and heterotic theories, which only contain closed oriented strings, than in the case of the type I theory. Therefore, the discussion that follows will be restricted to these theories. (The basic ideas are essentially the same in the type I theory. There are just some additional allowed topologies.)

The type II and heterotic string theories have a single fundamental interaction. It can be described by a portion of world sheet, called the "pants diagram", depicted in Fig. 1. When the diagram is intersected by a plane representing a time slice T_1, one sees two closed strings. Intersecting the surface by a time slice at time T_2 reveals just one closed string. Clearly, at intermediate times the two closed strings approached one another, touched, and joined. The reverse process in which two closed strings join to give one is also allowed.

The interaction structure described by the pants diagram differs in fundamental respects from interactions in point-particle theories. To explain the distinction, consider Fig. 2, where a point-particle vertex and the pants diagram are drawn. In each case, we can ask at what space-time point the interaction that turns two particles into one takes place. We can also represent the time slices corresponding to two observers in distinct Lorentz frames by the lines $t = \text{const}$, and $t' = \text{const}$. drawn in each case. In the point-particle theory there is a definite space-time point at which the interaction occurs that is unambiguously identified by all observers. In the string case, on the other hand, the

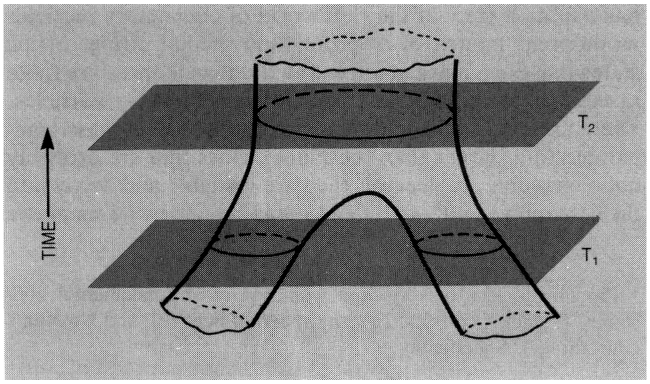

Fig. 1: Pants diagram.

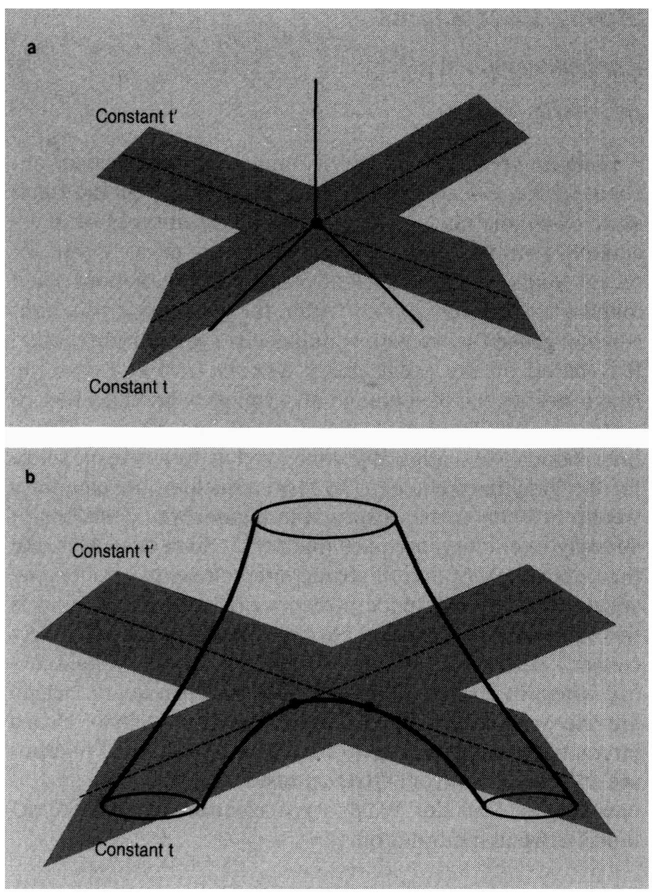

Fig. 2: Point-particle vertex (a) and pants diagram (b).

interaction point corresponds to the point at which the time slice is tangent to the surface, and this differs from one observer to another.

Clearly there is a fundamental difference between the interactions depicted in Figs. 2a and 2b. In the point-particle case (Fig. 2a) the "manifold" of lines is singular at the junction, which is a special point. 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 (Fig. 2b) is a smooth manifold with no preferred points. The fact that it describes interaction is purely a consequence of the topology of the surface. The nature of the interaction is completely determined by the structure of the free theory with none of the arbitrariness that exists in the point-particle case.

String world sheets are two-dimensional surfaces that can be described as Riemann surfaces using techniques of complex analysis. This means that (at least locally) one can use complex coordinates z and $\bar{z}$. A fundamental feature of string theory is that world sheets that can be related by a conformal mapping $z \rightarrow f(z)$ are regarded as equivalent. Thus, in performing the sum over distinct geometries only surfaces that are conformally inequivalent should be included. Fortunately, for each topology, the conformally inequivalent geometries 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 case of the type II and heterotic theories. The world sheets are characterized by a single integer, the genus, which is the number of holes in the surface. The external strings can be represented as points on the surface, since this is conformally equivalent to tubes extending off to infinity. The genus corresponds to the number of loops, i. e., 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.

Not only is the number of diagrams much less than in ordinary quantum field theory, but the convergence properties of the associated integrals are much better. The properties of multiloop (genus $g > 1$) amplitudes are not 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 appears extremely likely that the following is true: The only divergences that occur are ones of very well understood and inevitable origin. The types of divergences that result in parameters becoming arbitrary in renormalized quantum field theories, or amplitudes becoming completely undefined in nonrenormalizable field theories, have no counterparts in string theory.

How can it be that general relativity, interpreted as a quantum theory, has nonrenormalizable divergences, whereas string theory, which agrees with it 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 whose mass is of 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.

Remaining Challenges

A great deal of effort is being expended on the development of fundamental principles and a more geometric formulation. The history of string theory can be contrasted with that of general relativity. In that case, Einstein began by formulating certain far-reaching principles – the equivalence principle and general covariance – then finding 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, however, since general relativity is a low-energy (long distance) approximation to string theory. These principles, whatever they are, are likely to require a new kind of geometry, perhaps an infinite-dimensional generalization of Riemannian geometry, for their implementation. Some specific suggestions along these lines can be found in the recent literature, but our understanding is still far from complete.

Once the correct geometric formulation incorporating the fundamental principles of string theory in a comprehensible form is achieved, we should be in a good position to answer many profound questions. It should be possible to study nonperturbative effects and possibly understand why a particular solution with four-dimensional space-time and the phenomenologically required symmetries and particles is selected. It will also be interesting to study how string theory modifies classical general relativity at short distances. In particular, it would be interesting to know whether the singularity theorems are still valid. We could then also investigate how some of the profound issues of quantum gravity are resolved.

In a theory without adjustable parameters, any dimensionless number in nature should be calculable. Some of them are extremely small. For example, the cosmological constant, expressed in Planck units, is observed to be smaller than 10^{-120}. We might hope to identify a symmetry principle that forces it to be exactly zero, but none is known. 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 find it remarkable that there now seems to be a reasonable chance that we will find a unique fundamental theory of nature. Certainly, it would have been considered pure folly to express such a hope 10 years ago. I think it is unrealistic to expect too much too soon, however. It will probably take a few decades of hard work to obtain a satisfactory understanding of what string theory is really all about. This will require substantial advances in mathematics. Also, the experimental results that can be expected during the next 10–20 years are likely to play an important role in shaping our ideas.

See also: Supersymmetry and Supergravity.

Bibliography

P. C. W. Davies and J. Brown (eds.), *Superstrings: A Theory of Everything?* Cambridge University Press, Cambridge, 1988. (E)

M. Kaku and J. Trainer, *Beyond Einstein.* Bantam Books, New York, 1987. (E)

M. B. Green, "Superstrings", *Sci. Am.* **255**, 48–60 (September 1986). (I)

J. H. Schwarz, *Superstrings – The First 15 Years of Superstring Theory.* World Scientific, Singapore, 1985. Reprints and commentary in 2 vols. (A)

M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory*, 2 vols. Cambridge University Press, Cambridge, 1987. (A)

P. Ramond, "Dual theory for free fermions", *Phys. Rev. D* **3**, 2415 (1971); A. Neveu and J. H. Schwarz, "Factorizable dual model of pions", *Nucl. Phys.* **B31**, 86 (1971). (A)

J. Scherk and J. H. Schwarz, "Dual models for non-hadrons", *Nucl. Phys.* **B81**, 118 (1974). (A)

M. B. Green and J. H. Schwarz, "Anomaly cancellations in supersymmetric $D = 10$ gauge theory and superstring theory", *Phys. Lett.* **149B**, 117 (1984); "Infinity cancellations in $SO(32)$ superstring theory", *Phys. Lett.* **151B**, 21 (1985). (A)

D. J. Gross, J. A. Harvey, E. Martinec, and R. Rohm, "Heterotic string theory (I). The free heterotic string", *Nucl. Phys.* **B256**, 253 (1985); "Heterotic string theory (II). The interacting heterotic string", *Nucl. Phys.* **B267**, 75 (1986). (A)

P. Candelas, G. Horowitz, A. Strominger, and E. Witten, "Vacuum configurations for superstrings", *Nucl. Phys.* **B258**, 46 (1985); E. Witten, "Symmetry breaking patterns in superstring models", *Nucl. Phys.* **B258**, 75 (1985). (A)