Theoretical Motivation for Studying Superstrings

Nathan Berkovits
Instituto de Física Teórica, Universidade Estadual Paulista
Rua Pamplona 145, 01405-900, São Paulo, SP, Brazil

Received on 16 January, 2004

I try to explain why many high-energy theoretical physicists are studying superstrings.

There are various motivations for studying superstring theory, both mathematical and physical. Since I am a physicist, I will only mention the physical motivations. When string theory was discovered in the early 1970's, it was orignally intended to be a model for describing strong interactions. The basic discovery was that by extending the pointlike nature of particles to one-dimensional extended objects called strings, one could obtain S-matrix scattering amplitudes for the fundamental particles which contained many of the properties found in scattering experiments of mesons. The action for string theory is proportional to the area of a two-dimensional worldsheet, as opposed to the action for point-particles which is based on the length of a one-dimensional worldline.

Amazingly, the masses and coupling constants of the fundamental particles in string theory are not inputs in the theory, but are instead fixed by consistency requirements such as Lorentz invariance and unitarity. In fact, unlike theories based on point particles, string theory not only predicts the masses of the fundamental particles, but also predicts the dimension of spacetime. In the simplest string theory, this dimension turns out to be 26, rather than the experimentally observed spacetime dimension of 4. However, it is possible to 'compactify' all but four of the dimensions to small circles, in which case only four-dimensional spacetime is observable at low energies.

For open string theory (where particles are represented by one-dimensional objects with two ends), the particle spectrum contains a massless 'gluon', as well as an infinite number of massive particles whose masses and spins sit on 'Reggae trajectories'. These Reggae trajectories of massive particles are welcome for describing strong interactions since they are needed for producing scattering amplitudes with the properties seen in experiments. Unfortunately, string theory also predicts fundamental particles which are not needed for describing strong interactions. One of these particles is tachyonic, i.e. its $(\text{mass})^2$ is negative implying that it travels faster than the speed of light. The presence of such a particle makes the vacuum unstable, which is not acceptable in a physical theory.

The resolution of this tachyon problem was found in a series of remarkable discoveries which led to the concept of supersymmetry, a symmetry relating bosonic and fermionic particles. The first discovery was the existence of a new consistent string theory whose spacetime dimension turns out to be 10 rather than 26. The second discovery was that the action for this new string theory depends on a two-dimensional worldsheet containing both bosonic and fermionic parameters, and the action is invariant under a worldsheet supersymmetry which transforms the bosonic and fermionc parameters into each other. The third discovery was that, after performing a projection operation which removes half the particles but leaves a unitary S-matrix, the particle spectrum and interactions of this 'superstring' theory are invariant under a ten-dimensional spacetime-supersymmetry which transforms bosons into fermions. This projection operation removes the problematic tachyon from the spectrum but leaves the massless gluons, as well as an infinite number of massive particles. Superstring theory also contains fermionic counterparts to the gluon (called the gluino), as well as an infinite number of massive fermions.

Another particle which survives the projection operation is a massless spin-two particle called the graviton (as well as its fermionic counterpart, the massless spin-3/2 particle called the gravitino). Although this massless spin-2 particle comes from closed string theory (where particles are represented by one-dimensional circles), unitarity implies that the two ends of an open string can join to form a closed string, so these massless spin-two particles are produced in the scattering of gluons. Since the only consistent interactions of massless spin-two particles are gravitational interactions, string theory 'predicts' the existence of gravity. Therefore, without prior intention, superstring theory was found to give a unified description of Yang-Mills and gravitational interactions.

Since the energy scale of gravitational interactions is much larger than the energy scale of strong interactions, a unification of these interactions implies that the massive particles predicted by superstring theory contain masses of the order of the Planck mass (about 10^{19} GeV), and are therefore unrelated to meson particles found in experiments. So the original motivation for using string theory as a model for strong interactions is no longer viable, assuming that one interprets the massless spin-two particle as the graviton of general relativity. Instead, superstring theory can be used as a model for a unified theory which includes all four of the standard interactions: gravitational, strong, weak, and electromagnetic (the last three are described by a spontaneously broken Yang-Mills theory).

The usual obstacle to constructing a quantum unified theory (or even a quantum theory of gravity) is that the Einstein-Hilbert action for general relativity is non-renormalizable. This is easily seen from the fact that the gravitational coupling constant (Newton's constant) is dimensionful, unlike the coupling constant of Yang-Mills theory. So for a scattering amplitude of three gravitons at L loop-order, power counting arguments imply that the amplitude diverges like Λ^{2L} where Λ is the cutoff. The only way to remove this divergence is if there is some miraculous cancellation of Feynmann diagrams.

One way to cancel divergences in Feynmann diagrams is to introduce fermions into the theory with the same interactions and masses as the bosons. Since internal loops of fermions contribute with an extra minus sign as compared with internal loops of bosons, there is a possibility of cancellations. If a theory is supersymmetrized (i.e. fermions are introduced in such a manner that the theory is symmetric under a transformation which exchanges the bosons and fermions), then the above conditions are satisfied. The supersymmetrization of gravity is called supergravity, and for a few years, it was hoped that such a theory might be free of non-renormalizable divergences. However, it was later realized that even after supersymmetrizing gravity to a theory with the maximum number of supersymmetries (which is called N=8 supergravity), the non-renormalizable divergences are still present.

As already mentioned, the fundamental particles of superstring theory include the graviton and the gravitino (like supergravity), but also include an infinite set of massive bosons and fermions. It turns out that after including the contributions of the infinite massive particles, the non-renormalizable divergences in the loop amplitudes completely cancel each other out. Although the explicit proof of the preceding statement is rather technical, there are various 'handwaving' arguments which are convincing. One of these arguments involves the nature of superstring interactions which are 'smoother' than the interactions of point-particles. For example, the three-point diagram for point-particles has a vertex where the three external point-particles coincide. But the three-point diagram for closed strings is like a pair of pants, where the two cuffs and the waist are the external strings. Unlike the vertex in a point-particle diagram, there is no singular point on a pair of pants.

So superstring theory provides a consistent theory of quantum gravity which, unlike all other attempts, does not suffer from non-renormalizable divergences. However, it requires an infinite set of massive particles which are unobservable in any foreseeable experiment. In addition, the theory includes a set of massless particles such as the gluons and gluinos of super-Yang-Mills and also a scalar massless boson called the dilaton. If superstring theory really describes nature (and is not just a model for a unified quantum theory of gravity and Yang-Mills), these massless particles must become the leptons, quarks, and gluons of the standard model where the masses of the above particles come from spontaneous symmetry breaking. One important unsolved problem in superstring theory is that it is very difficult to give a mass to the dilaton in a natural way, so one needs to explain why noone has observed massless scalars in experiments.

Although superstring theory is the only candidate for a renormalizable quantum theory of gravity, only a few researchers worked in this field between 1975 (when it was realized that string theory could not serve as a model for strong interactions) and 1985. One reason for the lack of interest was that there appeared to be different versions of superstring theory (called Type I, Type IIA and Type IIB), none of which resembled very closely the structure of the standard model. In the Type I theory, the gauge group for super-Yang-Mills was thought to be arbitrary, and in the Type IIA and Type IIB theories, the gauge group had to be abelian. However, in 1985, it was learned that absence of anomalies restricted the gauge group of the Type I theory to be $SO(32)/Z_2$. Although this gauge group is not very interesting for phenomenology, it was soon realized that there is another type of superstring theory, called the 'heterotic' superstring (since it combines features of the bosonic string and superstring), which has two possible gauge groups: $SO(32)/Z_2$ or $E_8 \times E_8$ (E_8 is one of the exceptional groups). The $E_8 \times E_8$ version of the heterotic superstring was very attractive for phenomenologists since it is easy to construct grand unified theories starting from the exceptional subgroup E_6.

For this reason, the next five years attracted many researchers into the field of superstring theory. However, it was soon clear that without understanding non-perturbative effects, superstring theory would not be able to give explicit predictions for a grand unified model (other than vague predictions, such as supersymmetry at a suitably high energy scale). The problem was that four-dimensional physics depends crucially on the type of compactification which is used to reduce from ten to four dimensions. Although there is a symmetry called T-duality which relates some compactifications in superstring theory, there is a large class of compactifications which are not related by any symmetry. In principle, the type of compactification is determined dynamically, however, the selection of the correct compactification scheme requires non-perturbative information. So, for this reason, any researchers left the field of string theory after 1989 to work in other areas such as supercollider phenomenology.

Recently, it has been learned that many non-perturbative features of four-dimensional supersymmetric Yang-Mills theories can be understood without performing explicit instanton computations. Although this had been conjectured in 1977 for N=4 super-Yang-Mills, the conjecture was treated skeptically until 1994 when convincing evidence was presented for the case of N=2 super-Yang-Mills. One of these non-perturbative features is an 'S-duality' symmetry which relates the super-Yang-Mills theory at large values of the coupling constant with a super-Yang-Mills theory at small values of the coupling constant. For N=4 super-Yang-Mills, S-duality maps the theory at strong coupling into the same theory at weak coupling, while for N=2 super-Yang-Mills, S-duality maps the theory at strong coupling into a different theory at weak coupling.

These S-duality symmetries are also believed to be present in superstrings and relate superstring theory at large values of the coupling constant with a theory at small values of the coupling constant. S-duality maps the Type IIB superstring at strong coupling into the same Type IIB superstring at weak coupling, and maps the Type I superstring at strong/weak coupling into the heterotic superstring at weak/strong coupling with gauge group $SO(32)/Z_2$.

There is also believed to a duality symmetry which maps the Type IIA superstring at strong coupling into a new eleven-dimensional theory called M-theory, and which maps the heterotic superstring with gauge group $E_8 \times E_8$ at strong coupling into a version of M-theory with boundaries. M-theory is known to contain the massless particle of eleven-dimensional supergravity (which is the maximum possible dimension for supergravity) as well as massive particles which are still not understood. It is believed to be related to a theory constructed from two-dimensional extended objects called membranes (as opposed to the one-dimensional extended objects called strings).

So by studying the perturbative regime of superstring theory where the coupling constant is small, one can use S-duality symmetry to obtain non-perturbative information where the coupling constant is large. Furthermore, duality symmetries relate the five different superstring theories, suggesting that these five theories can be understood as perturbative vacua of some unique underlying non-perturbative theory which would be the 'Theory of Everything'. This has attracted renewed interest in superstring theory, and there is optimism that by studying M-theory, one will gain a greater understanding of duality symmetries. However, the problem of getting explicit predictions out of superstring theory is probably still far from being resolved. Although S-duality symmetries may help in understanding superstring theory at very small and very large values of the coupling constants, it is not clear if it will be possible to extrapolate these results to the physically interesting values of the coupling constants which is somewhere between the two extremes.

Acknowledgements

I would like to thank the organizers of this meeting for a very enjoyable conference. This work was partially supported by CNPq grant number 300256/94-9, FAPESP grant 99/12763-0, and Pronex grant 66.2002/1998-9.