# Superstring Theory

Michael B. Green,

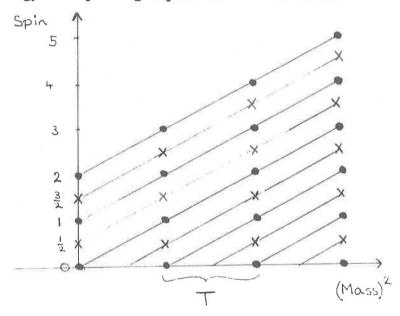
Dept. of Physics, Queen Mary College, University of London, Mile End Road, London E1 4NS U.K.

In this talk I shall review the structure of superstring theory and the reasons why particle physicists are enthusiastic about this somewhat novel approach to explaining the unity of the particles and their interactions. Needless to say the implications of these developments for cosmology must be very great, although the precise consequences are not yet clear. I shall make some very general comments about this but it is somewhat premature to draw any firm conclusions since the structure of these theories is not at all well understood. Detailed references to the subjects covered in this talk can be found in the bibliography at the end of the article.

String theories have a curious history since they were originally developed in the early 1970's in attempts to explain the structure of the strong interaction. In those days the strings were supposed to describe hadrons so that, for example, a meson was thought of as a quark and an antiquark tied together by a string. The need for such theories appeared to evaporate with the coming of QCD which provided an apparently satisfactory and more conventional field theory basis for the strong interactions. It was realized in the early 1970's that string theory does have the potential for describing Yang-Mills theory and gravity but in those days none of the available theories appeared to be consistent. However, with the incorporation of supersymmetry, the symmetry which relates fermions and bosons, these inconsistencies disappear. String theories of this type are called superstring theories.

It has been realized for a long time that the inconsistency of general relativity and quantum mechanics probably requires a radical modification of theories at short distances. This can be seen from a simple application of the Heisenberg uncertainty relation at the incredibly small distance scales around the Planck length ( $\sim 10^{-35}$  metres). At this scale the energy fluctuations are so great that according to general relativity microscopic virtual black holes form, in which case the structure of the theory ceases to have much meaning. In string theories this problem is avoided by a modification of the description of space-time which is no longer described in terms of a collection of points.

The conventional quantum field theories on which elementary particle theories are based, such as quantum electrodynamics and general relativity, describe the interactions of point particles such as photons, quarks, electrons, etc. By contrast, string theories describe the interaction of string-like relativistic quanta which have one-dimensional extension. The motion of a single string can be described in terms of the infinite number of normal modes which have unlimited frequencies,  $\nu_i$ . The separation between these frequencies depends on a parameter, the string tension T. Quantum mechanically, these modes become states of definite energy corresponding to particles of definite mass.



The spectrum of states in a typical superstring theory showing the  $(mass)^2$ and spin of the states. Bosons are indicated by dots and fermions by crosses. The massless states include the spin 2 graviton and the spin 1 gauge particle. The figure illustrates a typical spectrum that emerges from a superstring theory. The states lie on straight-line trajectories (which are known as Regge trajectories) in a plot of the angular momentum versus  $(mass)^2$ . There are an unlimited number of particles in the theory including particles of arbitrarily high spin. The fact that string theories consistently describe particles of arbitrary spin is one of the reasons for the importance of these theories. There is no consistent description of high-spin particles in conventional quantum field theories with a finite number of species of particles. In particular, the problems of quantum gravity, which describes the spin 2 graviton, have already been alluded to.

The need for a consistent description of high-spin particles is experimentally very evident in the data on *strong interaction* physics. Many strongly interacting particles – mesons and baryons – were discovered during the 1960's and found to lie on straight-line Regge trajectories in plots such as the one in the figure. Experimentally, the slope of these trajectories was found to be around 1  $Gev^2$ . There are no massless hadronic particle states so the intercept of the experimental trajectories is shifted from that of the figure. It was the similarity of the spectrum of the original string theory based on the Veneziano model (which was formulated in 1968) to the experimental hadronic data that motivated the original enthusiasm for string theory. However, the inability of string theories to reproduce the observed trajectory intercepts as well as the fact that the early string theories were inconsistent because they contained tachyons (particles with negative (mass)<sup>2</sup>) led to the demise of this application of string theory.

In its modern reincarnation string theory is supposed to encompass a unification of all the interactions. The figure shows that among the states of the string are massless states of spin 1 and spin 2. These are interpreted as Yang-Mills gauge particles and the graviton of general relativity. The string tension sets the dimensional scale of the theory. In a theory containing gravity this is the Planck scale. This means that the slope of the Reggetrajectories (*i.e.*, the inverse of the string tension) is tiny so that the first massive states have a mass of around

#### M.B. Green

the Planck mass  $(10^{19} Gev)$ . This is somewhat higher than the mass scale of the strong interactions! Another important feature is that superstring theories do not have the tachyon states which killed the original string theories.

The elementary particles that are observed in experiments (the quarks, leptons and gauge particles) are massless compared to the Planck scale and are expected to arise from among the superstring ground states. The massive string states are not directly observable but their existence is crucial for the consistency of the theory since they affect the very short distance (or high energy) behaviour of the theory, which is just the regime in which Einstein's theory has quantum mechanical problems.

One aspect of superstring theories which makes them most appealing as a fundamental description of physics is the fact that they very naturally incorporate the non-conservation of parity, or *chirality*, of observed physics. There are well-known problems, known as *chiral anomalies*, in incorporating chirality into quantum theories. Generically certain conservation laws (such as the conservation of electric charge) which are crucial for the consistency of the theory break down. The standard model of elementary particles (the union of QCD and the electroweak theory) is chiral and the fact that it must be free of anomalies provides experimentally observed constraints on the charges of the chiral quarks and leptons. In higher dimensions, such as the ten dimensions in which superstring theories are defined, there can also be anomalies in general relativity. Ensuring the absence of all anomalies is then much more difficult. The consequent restrictions on the possible theories lead to very few possibilities. The only possible ones contain unifications of all the forces and particles associated with symmetry groups which are subgroups of either SO(32) or  $E_8 \times E_8$ .

Since string theory contains general relativity as a low-energy approximation it must contain a principle that generalizes the geometrical priciple embodied in Einstein's theory. Unfortunately, such a principle is not yet understood – almost all the calculations are based on a formulation of the theory which is analogous to using a perturbative expansion of the Einstein theory around a particular fixed classical background geometry. Even at this level certain superstring theories avoid the familiar problems of quantum gravity such as infinities and anomalies. Requiring this consistencies selects only a very few possible theories in which the Yang-Mills theory associated with the forces other than gravity are unified in an almost unique manner. There are certainly further, unexplored, constraints that may well restrict the possible theories even further which raises the possibility that there is only one possible consistent quantum string theory. In order to make contact with physics such a theory has to explain a wealth of experimental data.

# The formulation of string theories

The dynamics of a classical relativistic string can be described by an action principle that generalizes that of a relativistic point particle. In the case of the particle the simplest Lorentz covariant action is simply the length of the world-line traversed as the particle moves through space-time. As a string moves through *D*-dimensional space-time it sweeps out a two-dimensional *world sheet*. Denoting the parameters of the world-sheet by  $\sigma, \tau$  the space-time coordinate of any point on the string at a given time,  $X^{\mu}(\sigma, \tau)$  ( $\mu = 0, \ldots, D - 1$ ), is a mapping of the sheet into *D*-dimensional space-time. The obvious generalization of the point particle action to the case of a string is to choose the action to be proportional to the area of the world-sheet

$$S = -T \int d\tau d\sigma \sqrt{\left(\dot{X} \cdot X'\right)^2 - \dot{X}^2 X'^2},\tag{1}$$

where  $\dot{X} = \partial X/\partial \tau$  and  $X' = \partial X/\partial \sigma$  and the world indices,  $\mu$ , are contracted by using the metric  $g^{\mu\nu}(X)$  of the space in which the world-sheet is embedded. The dimension of space-time is taken, in the first instance, to be arbitrary but is determined by consistency requirements to be discussed below. It is important for the interpretation of the theory that the area is a quantity that is reparametrization invariant, *i.e.*, it is invariant under arbitrary redefinitions of the parameters  $\sigma$  and  $\tau$ . This is general coordinate invariance in *two* dimensions. The action can be rewritten in a manner that makes this connection with two-dimensional general relativity very explicit by introducing a two-dimensional metric tensor  $h^{\alpha\beta}(\sigma,\tau)$  ( $\alpha,\beta=\tau,\sigma$ ) and interpreting the coordinates  $X^{\mu}(\sigma,\tau)$  as *D* fields on the world-sheet which transform as scalars under reparametrizations. This form of the action is given by

$$S = -\frac{T}{2} \int d\sigma d\tau \sqrt{h} h^{\alpha\beta} \partial_{\alpha} X^{\mu} \partial_{\beta} X^{\nu} g_{\mu\nu}(X), \qquad (2)$$

where  $h = -\det h_{\alpha\beta}$ . The two-dimensional metric tensor,  $h^{\alpha\beta}$ , is a sort of Lagrange multiplier. The two formulations, (1) and (2), are equivalent in the classical theory. This can be seen by deriving the  $h^{\alpha\beta}$  equations of motion by varying (2) with respect to  $h^{\alpha\beta}$ . These equations are constraint equations (they involve no derivatives of  $h^{\alpha\beta}$ ) and the solution for  $h^{\alpha\beta}$  can be substituted back into (2) which turns it into (1). In the quantum theory things are more subtle and it is very convenient to use the action (2) in which case many well-known geometric properties of two-dimensional surfaces can be applied. A particularly important feature of (2) is that in addition to obvious two-dimensional coordinate invariance the action is invariant under arbitrary rescalings of  $h^{\alpha\beta}$ 

$$h^{\alpha\beta} \to h^{\prime\alpha\beta} = e^{\phi(\sigma,\tau)} h^{\alpha\beta}.$$
 (3)

This amounts to an invariance of the action under two-dimensional conformal transformations. The conformal invariance of the theory is of paramount importance in understanding the consistency of string theory just as gauge invariance is a crucial aspect of the consistency of Maxwell's theory of electrodynamics. In that case the vector potential,  $A^{\mu}$ , describes four modes of oscillation including an obviously unphysical one in the time-like direction. The fact that this mode actually decouples from physical processes (along with the longitudinal component) is a result of the gauge symmetry of Maxwell's equations – we know that photons only have two independent modes of polarization. String theory has a potentially infinitely worse problem with time-like modes since the coordinate  $X^0(\sigma)$  gives an infinite number of time-like degrees of freedom, one for each point along the string (or, equivalently, one for each normal mode of  $X^0$ ). In fact string theory has an infinite number of gauge invariances, analogous to the single gauge symmetry of electrodynamics, which guarantee that all these unphysical modes decouple from physical processes. This enormous symmetry is encoded in the two-dimensional conformal invariance of the theory which is associated with an infinite dimensional algebra known as the Virasoro algebra.

This conformal symmetry is only a property of the theory under very special conditions. Most notably it only holds when the number of components of  $X^{\mu}$  in (2)(*i.e.*, the dimension of space-time) is restricted to specific values. For the Veneziano model this *critical* dimension is 26.

The superstring theories are constructed by including additional terms in the action (2) made out of fermion coordinates which are functions of  $\sigma$  and  $\tau$ . In this case the critical dimension is ten. In these critical dimensions only the transverse polarizations of the string are physical. There is a suggested procedure for dealing with string theories in dimensions lower than these critical values but it is only in the critical dimension that the physical modes are purely transverse, which corresponds most closely to the case of Yang-Mills theory. Furthermore it is at this point that the restrictions arise which give rise to an open-string spectrum which contains a massless vector state (a gauge boson) and a closedstring spectrum which includes the massless spin 2 state (the graviton).

# Quantum numbers

The original way of incorporating quantum numbers into string theory was to attach charges to the endpoints of open strings. Originally, these charges were required to lie in the fundamental representation of the groups SO(n), Sp(n) or U(n). It was then found that the only choice that is free of anomalies when the quantum loop amplitudes are considered is the group SO(32). In this case there are 496 massless gauge particles and their 496 massless fermion supersymmetric partners corresponding to the adjoint representation of the group.

A second symmetry group which is also known to give a cancellation of anomalies is the group  $E_8 \times E_8$ . This is not one of the groups that can be obtained from open strings. There is, however, a much more elegant way of incorporating quantum numbers into closed string theories by introducing a charge density on the world-sheet. This is used in the construction of the *heterotic* string theory. In that case the action (2) must have an additional term involving 32 world-sheet fermions that carry the internal quantum numbers. These fermions are chiral in the two-dimensional sense. Such a theory can only be made consistent if the gauge groups are  $E_8 \times E_8$  or Spin(32) (which has the same Lie algebra as SO(32)). In oder to investigate the consistency it is necessary to study Feynman diagrams for the scattering of strings. The world-sheets for the processes described by these Feynman diagrams are Riemann surfaces with nontrivial genus. The genus corresponds to the number of handles attached to the surface which is also the number of loops in the language of conventional Feynman diagrams. For the consistent theories, the Feynman diagrams are finite (at least at one loop and probably beyond). It is striking that the only theories that can be constructed in this way are precisely those with symmetry groups which were known to be free of anomalies and that these theories also give *finite* one-loop amplitudes. More recently a third theory has been added to the list of anomaly-free string theories. This one has a symmetry group  $O(16) \times O(16)$ which is a subgroup of both SO(32) and  $E_8 \times E_8$ . This theory, however, has divergent loop amplitudes.

It is interesting that the richest of the string theories, the heterotic string, has many properties in two dimensions (*i.e.*, when viewed as a field theory on the world-sheet) that we also want in the ten-dimensional theory, namely, coordinate invariance, supersymmetry, chirality and freedom from anomalies.

## The emergence of four-dimensional physics

We have seen that conformal invariance imposes strong restrictions arise even

when the background space-time is flat (i.e.,  $g^{\mu\nu} = \eta^{\mu\nu}$ , the *D*-dimensional Minkowski metric in (2)). Obviously, if string theory is to explain the world in which we live D-4 dimensions must be very curved so that in these dimensions the space-time metric is far from that of flat space. A theory defined by an action of the form of (2) is known as a two-dimensional nonlinear  $\sigma$  model. In this case the requirement of conformal invariance imposes very strong restrictions on the possible metrics of space-time. These constraints on the possible structure of the *D*-dimensional space-time geometries are quite unusual since they arise from a *two*-dimensional priciple, namely, that of world-sheet conformal invariance. There is no analogue of these constraints in conventional relativity based on the Einstein equations.

The condition that the  $\sigma$  model be conformally invariant can be translated into the condition that the renormalization group  $\beta$  functions of the theory vanish. These conditions had been studied in their own right before the connection with string theory had been appreciated. The metric  $g^{\alpha\beta}$  can be viewed as a coupling function for the  $\sigma$  model and there is a single associated renormalization group function  $\beta_{\mu\nu}^{g}$  which is a function only of the Riemann curvature made out of the metric tensor in the usual manner. The condition for conformal invariance is that this function should vanish. The  $\beta$  function can be calculated in perturbation theory in powers of 1/T giving the condition for conformal invariance

$$\beta_{\mu\nu}^{g} = R_{\mu\nu} + O(1/T) = 0, \tag{4}$$

where  $R_{\mu\nu}$  is the Ricci tensor (a contraction of the Riemann curvature). Remarkably, this is Einstein's equation together with corrections which are negligible at low energies. This demonstrates that string theory contains Einstein's equations at low energies. Of course, the series of terms in (4) is infinite, containing the structure of the complete string theory, with the higher terms providing crucial high-energy modifications of Einstein's equations which are needed for the consistency of the quantum theory.

#### M.B. Green

There can be other terms in the  $\sigma$  model in addition to (2), associated with background fields in addition to the metric. One example of such a term is one of the form

$$\int d\sigma d\tau \epsilon^{\alpha\beta} \partial_{\alpha} X^{\mu} \partial_{\beta} X^{\nu} B_{\mu\nu}(X), \tag{5}$$

where  $\epsilon^{\alpha\beta}$  is the two-dimensional Levi-Civita symbol and  $B_{\mu\nu}(X)$  is an antisymmetric tensor which is a new coupling which has its own renormalization group function,  $\beta^{B\mu\nu}$ . The condition for the theory to be conformally invariant now requires  $\beta^B = 0$  in addition to the vanishing of  $\beta^g$ . This condition is the equation of motion for the antisymmetric massless tensor field  $B_{\mu\nu}$  which is another of the massless states of the string. More generally, there are similar terms with a  $\sigma$  model coupling for each background field corresponding to the infinite number of component states of the string. In practise, most of the work on this subject considers terms in which the nonzero background fields correspond to the massless states of the string. There is a  $\beta$  function for each of these fields which is required to be zero if the theory is conformally invariant. These conditions on the  $\beta$  functions which express the two-dimensional conformal invariance are just the field equations for these massless fields in ten-dimensional space-time.

There has been a great deal of activity in trying to find solutions of the full equations for the vanishing of the  $\beta$  functions. There appears to be only one class of possible curved spaces which satisfy these equations. These spaces are known as *Calabi-Yau* spaces. The interesting case is when six dimensions are curved, leaving four flat dimensions. In this case the Calabi-Yau spaces are spaces of SU(3) holonomy (*i.e.*, the Riemann curvature,  $R_{\mu\nu}^{mn}$ , is an SU(3) matrix in the tangent-space indices, m, n) which have vanishing Ricci curvature. It is remarkable that these are precisely the spaces which had been conjectured to be required in order for the effective four-dimensional theory to have any chance of explaining observed four-dimensional physics.

Obtaining physics from Calabi–Yau spaces

There is still a long way to go before the standard model can be derived from the superstring theory. It is nevertheless interesting that there is a fairly clear scenario for how the huge symmetry groups associated with the ten-dimensional theory may break when the extra dimensions curl up and many realistic features emerge.

One procedure is to study the approximate equations that emerge from the low energy expansion of the superstring theory in terms of the massless fields of the theory. There are very strong topoological constraints which relate the Yang-Mills field strength,  $F_{\mu\nu}^{ab}$ , to the space-time curvature,  $R_{\mu\nu}^{mn}$ . These constraints suggest that the non-zero curvature in the six extra dimensions should be identified with the Yang-Mills field in these dimensions, R = F. However, as usual a non-zero value for  $\langle F \rangle$  leads to the breaking of the symmetry to a group which commutes with the subgroup in which the field strength lies. In the present case  $\langle F \rangle$  lies in a SU(3) subgroup of either of the large groups SO(32)or  $E_8 \times E_8$ .

Starting from the  $E_8 \times E_8$  group in ten dimensions (the SO(32) case seems not to be related to realistic physics) the effective four-dimensional theory has the smaller symmetry group  $E_6 \times E_8$  and is still supersymmetric. This acts a bit like an effective supersymmetric grand unified theory for the four-dimensional theory at high energy. The particle content includes fermions lying in 27's which are crucial for obtaining chiral physics. The number of generations of quarks and leptons is the number of 27's minus the number of  $\overline{27}$ 's. This is determined by a simple topological invariant of the internal six-dimensional Calabi-Yau space, namely its Euler characteristic. Since there are very many Calabi-Yau spaces an arbitrary choice must be made. Phenomenological considerations, including astrophysical abundances, suggest that there are, at most, four generations. This severely restricts the possible spaces of interest.

At this stage the symmetry group looks like one of the popular grand unified groups,  $E_6$ , multiplied by an extra  $E_8$  factor. This extra factor is associated with

a sector of matter which is neutral under the  $E_6$  charges and therefore is only observable via its gravitational interactions. If this exists in significant abundance it constitutes a novel form of dark matter. Unfortunately, it is premature to predict its properties.

The  $E_6$  grand unified group breaks to the standard model,  $SU(3) \times SU(2)_L \times U(1)$  at low energy. This can happen by a topological mechanism that is very different from the usual mechanism involving Higgs particles in the adjoint representation of  $E_6$ , which is just as well since the required Higgs particles are not present. The new mechanism makes use of the fact that those Calabi-Yau spaces which are plausibly relevant are not simply connected (*i.e.*,they have holes). As a result  $E_6$  flux loops can get trapped in the holes which leads to a breaking of the symmetry (but not supersymmetry). There are discrete choices for this breaking but the arbitrariness is much less than in the conventional Higgs mechanism. It is very easy to get the standard model from this breaking but almost inevitably there are extra symmetry factors such as  $U(1)^2$  or SU(2). This suggests that there are extra low energy gauge symmetries together with associated gauge particles.

Although the  $E_6$  factor acts like a GUT group there are crucial differences between this scheme and conventional GUT's. Most importantly, there is no sense in which the full  $E_6$  symmetry is restored at high energy. The topological mechanism that breaks  $E_6$  to the standard model cannot be continuously switched off. Symmetry restoration only arises at the exceedingly high energy scales at which the full  $E_8 \times E_8$  symmetry is apparent. In this way, although the gauge particles of  $E_6$  are unified at the compactifiation scale, the quarks and leptons may not be, which would avoid some unpleasant features of conventional  $E_6$  unification.

There are other plausible experimental consequences of this scheme which are only understood at a qualitative level but which have undoubtedly important consequences for cosmology. These include: The existence of at least one axion. The existence of Planck-mass, unconfined, fractionally-charged particles. The existence of Planck-mass, multiple-charged magnetic monopoles. In addition, superstring theory is totally different from conventional theories based on general relativity at temperatures around the Planck scale or distances of around the Planck distance. For example, it is intuitively plausible that there are no space-time singularities in superstring theory, thus avoiding the aesthetic problems in discussing classical black holes in Einstein's theory. Furthermore, it is clear that some of the folklore about the very early universe must be ammended. Presumably there was an era around the Planck time in which four dimensions expanded, leaving the other six dimensions effectively static. This would have also been the era in which string-like effects decoupled, inflation occurred, the  $E_8 \times E_8$  symmetry was broken and in which the seeds were sewn for generating the low-energy breakdown of supersymmetry and the low-energy mass scale of the observed particles. This is very speculative, but it provides the impetus for entirely new studies of these subjects.

The fact that the cosmological constant is so small is one of the most difficult facts to explain in any conventional quantum field theory. Its value is determined by the vacuum energy which is related to the various scales of symmetry breaking. In supersymmetric theories the cosmological constant vanishes. However, supersymmetry is not an exact symmetry of the observed world. Typically, after supersymmetry is broken the cosmological constant is predicted to have a value of order unity in units in which its experimental value is  $10^{-120}$ ! Clearly, a fundamental principle is missing. At the moment the string theory is not well enough understood to know if it embodies such a principle.

I have described this ingenious scheme, based on the low energy effective field theory obtained from the string, to illustrate that there is a plausible way in which observed physics can emerge from the apparently very abstract tendimensional starting point. The scheme has many difficulties and may ultimately be completely wrong. It is known, for example, that Calabi-Yau spaces are not exact solutions to the complete string theory but that there are solutions which are 'nearly' Calabi-Yau spaces in a sense which is not fully understood. There are also severe problems associated with the fact that perturbative calculations cannot be justified. Crucially missing from the scheme described above is an understanding of how supersymmetry is broken which is almost certainly a nonperturbative phenomenom. Associated with this the breaking of  $SU(2) \times U(1)$ is not understood together with the fact that the observed particles have nonzero masses which are tiny numbers when measured in Planck units. There are indications that these effects are due to a fermion condensate similar to that in a superconductor but there is little quantitative understanding of this.

# Other theoretical ideas

A very attractive alternative approach to the study of the curled-up, effectively four-dimensional theory is to study the complete string theory in backgrounds which are highly curved. It is not possible to take this to be a Calabi-Yau space since this is too complicated. However, it is possible to take the background to be an orbifold which is a singular limit of a Calabi-Yau space in which all the curvature is concentrated at points while the rest of the space is flat. These orbifolds are six-dimensional tori with a discrete isometry divided out. The singular nature of an orbifold is not a problem for the string theory while it would be in a conventional theory of gravity. Orbifolds can also be constructed which are not limits of ordinary manifolds. The fact that exact statements can be made about string theory in an orbifold background suggests that they may play a specially important rôle in the theory.

A major subject of present research is the quest for a formulation of a field theory of strings based on a geometrical principle. The description of strings in terms of a given bachground space is not satisfactory for a fundamental understanding of the theory. Just as in general relativity, we expect that string theory should have a geometrical formulation in which the particles (the strings) emerge as some kind of fluctuation in something like a metric field. Since string theory contains general relativity as a low energy approximation there must be a generalization of the principle of relativity to a larger principle based on a generalization of geometric ideas. It is natural, for example, to consider geometry on the space of configurations of strings (a much bigger space than space-time). Such a radical alteration of our notions of space-time is not apparent from the formulation of string theory in terms of two-dimensional world-sheet principles described above. Understanding such a principle might lead to a much deeper understanding of the predictions of the theory.

# Bibliography

The following are compilations of review articles, lectures and reprints.

Green, M. and Gross, D.I., editors, (1986) Workshop on Unified String Theories, 29 July - 16 August 1985 (World Scientific, Singapore).

Jacob, M., editor, (1974) Dual Theory. Physics Reports Reprint Volume I, (North-Holland, Amsterdam).

Schwarz, J.H., editor, (1985) Superstrings. A collection of reprints on string theory. (World Scientific, Singapore).