String Theory and Nature

Michael Dine

Jerusalem Winter School on Theoretical Physics
Hebrew University
Jan 2003

In a polemic against string theory, published in the May 1986 issue of Physics Today, the Nobel prize winning physicist Sheldon Glashow, and Paul Ginsparg (now at Cornell), wrote that "The bad news is that years of intense effort by dozens of the best and the brightest have yielded not one verifiable prediction, nor should one soon be expected."

In the past 16 years, one could argue that Ginsparg and Glashow have been proven right. On the theoretical side, we have seen wonderful developments:

- Dualities: String-String, IIA-11 Dimensions, AdS/CFT,....

- We have understood some non-perturbative aspects of string theory, such as D-branes.

- There has been significant progress in understanding the quantum mechanics of black holes.

These developments have helped to propel some of the most interesting phenomenological proposals of the past few years, particularly the idea of large extra dimensions. But there is not yet a solid prediction one can point to.

In these lectures, I hope to explain why this is so – why string phenomenology is inherently difficult. But I also hope to outline some strategies for bringing string theory into contact with nature, both in accelerators and in the cosmos.

I will struggle to put together a coherent story. You will be the judges of whether it is, indeed, coherent. Even more difficult is the question of whether this story has elements of truth.

What is the Goal of String Phenomenology?
How will string theory be brought into contact with nature?

There are three points of view one might take on these questions:

- Solve the theory. Discover that it describes all of the phenomena of the Standard Model, and perhaps predicts a few others.– I am skeptical about this possibility. First, as we will see, string theory is inherently difficult. I am not optimistic that we can simply "solve it." Second, I think we already know that string theory has lots of perfectly good ground states which don't look anything like the real world. In fact, if anything, were the world not like it is, we would probably argue that nothing like it is possible.

- Examine large classes of string solutions. Find a solution which resembles the real world. Show that it replicates facts we know, and perhaps makes some predictions for future experiments.–This is probably the most widely adopted approach to string phenomenology. It has real promise, and in any case has stimulated much of our present understanding of the theory. The main difficulty, at present, is that this approach requires many assumptions, not only about the class of vacuum states, but about the stabilization of moduli (much more about moduli later). As we will see, it is unclear that the ground state of string theory should lie in any regime where weak coupling approximations are applicable, so it is not clear that one can calculate anything at all. Even granting that somehow the theory is stabilized at weak coupling, further assumptions about the stabilization of moduli are necessary, and it is unclear whether one is left with predictions at all.

- Examine broad classes of string solutions. Try to determine if there are any generic productions. This approach is not particularly popular, but it seems to me the most promising, at least at present.

In this last category are questions such as:

1. Is low energy supersymmetry a prediction of string theory?

2. If so, are there any general expectations for the soft breakings?

3. Are axions (as solutions of the strong CP problem) generic to string theory?

4. Is unification of couplings generic in string theory?

You can tell from the way I formulate these questions that I am partial to this viewpoint. Demonstrating that low energy supersymmetry is (or is not) a prediction of string theory is not something we know how to do at present. Success would be a triumph. Similarly, making some statement about the pattern of soft breakings (perhaps contingent on the discovery of supersymmetry) would be spectacular. While I am fond of this viewpoint, and this will be a theme through much of these lectures, my optimism is not widely shared. Taking, for example, your average physicist on the street (Ed Witten), there is an alternative view, that experiments will provide us guidance as to the correct way of thinking about a broad range of questions in string theory.

Plan of the Lectures

Lecture I. String Solutions

For any of the approaches I have enumerated here, it is necessary to have a survey of broad classes of string solutions. This will be the subject of the first lecture. It is impossible to survey, much less describe in any detail, all of the known string solutions in a lecture or two. So I will focus on a few. As an organizing principle, I will consider solutions with varying degrees of supersymmetry. So this lecture will include a brief introduction to supersymmetry. Then I will consider solutions with $N = 2, 4$ supersymmetry, followed by solutions with $N = 1$ supersymmetry. This will probably spill over into lecture II. Theories with $N = 0$ supersymmetry will be considered in lecture II and III.

Lecture II. Moduli in $N \leq 1$ Supersymmetry and Their Fate.

In this lecture we will begin by reviewing some aspects of the Standard Model, as well as some of its puzzles. We will discuss phenomenological approaches to low energy supersymmetry, with a view to understanding how these ideas might fit into string/M theory. Moduli seem to be an ubiquitous feature of string/M theory. This could represent some real feature of any quantum theory of gravity, or it could be a result of the fact that the existence of moduli, in any situation we know, is necessary to have any sort of systematic approximation procedure. In any case, understanding string phenomenology requires understanding the fate of the moduli. We will discuss computation of moduli potentials in $N = 1$ supersymmetry. These potentials are non-perturbative. They can be reliably calculated in certain regimes. However, as we will see, where they can be calculated, the system cannot describe nature. We will engage in some speculations about how the moduli might be fixed in a way which yields an approximate supersymmetry.

Lecture III. Features of $N = 0$ Theories.

To show that supersymmetry is a prediction of string theory (or not), we need to understand some features of $N = 0$ supersymmetry, and $N = 0$ supersymmetry. In some sense, we need to argue that non-supersymmetric strings don't make sense, or are in some way untypical or unacceptable. We won't succeed in doing this here, but we will consider some features of these theory, and especially of their moduli. We will see in lecture II, that with $N = 1$ supersymmetry, potentials are typically generated non-perturbatively. With $N = 0$ supersymmetry, they are typically generated, as we will see, *classically.* Indeed, in this lecture, we will understand, in some simple cases, the significance of tachyons in closed string theories, and their implications for the behavior of the moduli.

We will then go on and consider questions of the stability – and decay – of states in theories of gravity. We will review some aspects of the theory of false vacuum decay, taking into account some recent observations of Banks. Then we will consider bizarre type of vacuum instability discussed originally by Witten. We will argue that, if a non-supersymmetric string state describes the real world, this solution might have some relevance.

Lecture IV. Cosmology

Cosmology will be a crucial arena in which to test string theory, for obvious reasons. Some of these are conceptual (see Seiberg's lectures, but also here), some more phenomenological. Here I will talk a little bit about how inflation might fit into string theory. I will talk a bit about moduli. I will also explain why cosmology is a difficult problem, in the sense that even writing down string theories which describe toy cosmologies is hard.

This lecture will indicate some of the ways in which cosmology might provide insight into many of our basic questions about string phenomenology. I will conclude by suggesting one possible unified picture of string phenomenology, which solves some of the problems, both of phenomenology and cosmology.

Outline of Lecture I:

- A quick introduction to supersymmetry

- $N > 1$ Supersymmetry

- $N = 1$ SUSY – a survey

Introduction to Supersymmetry

(Good introductions provided in Appendices to Vol. 2 of Polchinski's text, Wess and Bagger, 1996 TASI lectures of Joe Lykken.)

Here we will for simplicity limit our considerations to four dimensions. We will occasionally remark about higher dimensions. We will also simplify our discussions by ignoring central charges. These are crucial for many issues, and discussed in the references mentioned.

Basic algebra:

$$\{Q^I_\alpha, Q^J_\beta\} = \delta^{IJ} \not{P}_{\alpha,\beta}. \qquad (1)$$

Here, $I, J = 1, \ldots N$, and $N = 1, 2, 4$ or 8. What is most relevant for our purposes is that the degree of supersymmetry constrains the possible multiplets, and also tightly constrains the form of the lagrangian which describes the system. These constraints provide great power for theoretical analysis. They also limit the possible relevance of these theories to the real world.

Because the mathematics of $N = 1$ supersymmetry is in some ways the simplest, it is convenient to describe these theories first.

N=1

The basic supersymmetry algebra is

$$\{Q_\alpha, Q^*_{\dot\beta}\} = 2\sigma_\mu P^\mu. \qquad (2)$$

We will first consider the case of global supersymmetry; later, we will consider the generalization to local supersymmetry. There are two irreducible representations of the supersymmetry algebra containing fields of spin less than or equal to one. These are the chiral and vector superfields. Chiral fields contain a Weyl spinor and a complex scalar; vector fields contain a Weyl spinor and a (massless) vector.

In superspace (using the conventions of [**?**]), a chiral superfield may be written as

$$\Phi(x,\theta) = A(x) + \sqrt{2}\theta\psi(x) + \theta^2 F + \dots \qquad (3)$$

Here A is the complex scalar, ψ the fermion, and F is an auxiliary field. Under a supersymmetry transformation with anticommuting parameter ζ, the component fields transform as

$$\delta A = \sqrt{2}\zeta\psi \qquad (4)$$

$$\delta\psi = \sqrt{2}\zeta F + \sqrt{2} i \sigma^\mu \bar{\zeta} \partial_\mu A \qquad \delta F = -\sqrt{2} i \partial_\mu \psi \sigma^\mu \bar{\zeta}. (5)$$

For vector superfields, the physical content is most transparent in a particular gauge (really a class of gauges) know as Wess-Zumino gauge. This gauge is analogous to the Coulomb gauge in QED. In that case, the gauge choice breaks manifest Lorentz invariance, but Lorentz invariance is still a property of physical amplitudes. Similarly, the choice of Wess-Zumino gauge breaks supersymmetry, but physical quantities are still supersymmetric. In this gauge, the vector superfield may be written as

$$\frac{V = -\theta\sigma^{\mu}\bar{\lambda}A_{\mu} + i\theta^2\overline{\theta}\overline{\lambda} - i\overline{\theta}^2\theta\lambda + 1}{2\theta^2\overline{\theta}^2 D.} \tag{6}$$

Here A_μ is the gauge field, λ_α is the gaugino, and D is an auxiliary field.

The analog of the gauge invariant field strength is a chiral field:

$$W_\alpha = -i\lambda_\alpha + \theta_\alpha D - i2(\sigma^\mu \bar{\sigma}^\nu \theta)_\alpha F_{\mu\nu} + \theta^2 \sigma^\mu_{\alpha\dot{\beta}} \partial_\mu \bar{\lambda}^{\dot{\beta}}. (7)$$

To construct an action with $N = 1$ supersymmetry, one starts with a set of chiral superfields, Φ^i, transforming in various representations of some gauge group $\mathcal{G}$. For each gauge generator, there is a vector superfield, V^a. The most general renormalizable lagrangian, written in superspace, is

$$\mathcal{L} = \sum_i \int d^4\theta \Phi_i^\dagger d^V \Phi_i + \sum_a \frac{1}{4g_a^2} \int d^2\theta W_\alpha^2 + c.c. + \int d^2\theta W(\Phi_i) + c.c. \quad (8)$$

Here $W(\Phi)$ is a holomorphic function of chiral superfields known as the superpotential.

In terms of the component fields, the lagrangian takes the form (again in Wess-Zumino gauge):

$$\mathcal{L} = \sum_i \left(|D\phi_i|^2 + i\psi_i\sigma^\mu D_\mu \psi_i^* + |F_i|^2\right) \qquad (9)$$

$$-\sum_a \frac{1}{4g_a^2}\left(\frac{F_{\mu\nu}^{a2} - i\lambda^a \not{D}\lambda^{a*} - 1}{2(D^a)^2}\right)$$

$$+i\sqrt{2}\sum_{ia} g^a \psi_i T^a \lambda^a \phi^* + c.c.$$

$$+\sum_{ij} \frac{1}{2}\frac{\partial^2 W}{\partial\phi^i\partial\phi^j}\psi^i\psi^j.$$

Here I have already changed notation, and have used ϕ_i to denote the scalar component of Φ. This is a common practice, and I will often use it where it will not lead to (too much) confusion. I have also solved for the auxiliary fields F_i and D_a using their equations of motion:

$$F_i = \frac{\partial W}{\partial \phi_i} \qquad D^a = g^a \sum_i \phi_i^* T^a \phi_i. \tag{10}$$

The first two lines are just the gauge invariant kinetic terms for the various fields, as well as potential terms for the scalars. The third line corresponds to Yukawa interactions of the gaugino and matter fields, with strength controlled by the strength of the gauge interactions. The last line yields fermion mass terms and Yukawa couplings among the various chiral fields.

It is often useful to consider effective lagrangians valid below some energy scale. In this case there is no restriction of renormalizability. On the other hand, one usually does want to expand the lagrangian in powers of momenta. It is a simple matter to generalize eq. (**??**) to include arbitrary terms with up to two derivatives:

$$\mathcal{L} = \sum_i \int d^4\theta K(\Phi^{i*}\Phi^i) + \sum_{ab} \int d^2\theta f_{ab}(\Phi) W_\alpha^a W^{\alpha b} + c.c. + \int d^2\theta W(\Phi_i) + c.c. \quad (11)$$

The functions W and f are holomorphic functions of the chiral fields (otherwise the last two terms are not supersymmetric); K is unrestricted. It is not hard to generalize the component lagrangian, and I will leave that as an exercise. Note that among the couplings are now terms:

$$\frac{1}{4}\mathrm{Re}\ f_{ab}(\phi_i)F^a_{\mu\nu}F^{\mu\nu b} + \frac{1}{4}\mathrm{Im}\ f_{ab}(\phi_i)F^a_{\mu\nu}\tilde{F}^{\mu\nu b} + \frac{\partial f_{ab}}{\partial \phi_i}F_i\lambda^a\lambda^b. \quad (12)$$

N=1 Supergravity

Once we include gravity, supersymmetry becomes a local symmetry. The gauge particle associated with this symmetry is the gravitino, a particle of spin 3/2. The basic supergravity multiplet consists of the graviton and the gravitino (two bosonic and two fermionic degrees of freedom). The effective lagrangian is still relatively easy to describe. We will focus on the bosonic terms:

First, the terms in the effective action with at most two derivatives or four fermions are completely specified by three functions:

1. The Kahler potential, $K(\phi,\phi^\dagger)$, a function of the chiral fields

2. The superpotential, $W(\phi)$, a holomorphic function of the chiral fields.

3. The gauge coupling functions, $f^a(\phi)$, which are also holomorphic functions of the chiral fields.

The lagrangian which follows from these can be found, for example, in [**?**, **?**]. Let us focus, first, on the scalar potential. This is given by

$$V = e^K \left[\left(\frac{\partial W}{\partial \phi_i} + \frac{\partial K}{\partial \phi_i} W \right) g^{i\bar{j}} \left(\frac{\partial W^*}{\partial \phi_j^*} + \frac{\partial K}{\partial \phi_j^*} W \right) - 3|W|^2 \right], \quad (13)$$

where

$$g_{i\bar{j}} = \frac{\partial^2 K}{\partial \phi_i \partial \phi_{\bar{j}}} \tag{14}$$

is the Kahler metric associated with the Kahler potential. In this equation, we have adopted units in which

$$G_N = \frac{1}{8\pi M^2}. \tag{15}$$

where $M \approx 2 \times 10^{18}$ GeV is the reduced Planck mass we encountered earlier.

$N=2$

N=4

If we do not include gravity, the basic multiplet consists of a vector field, three complex scalars, and four Weyl fermions. In $N=1$ language, there are three chiral superfields and one vector superfield in each multiplet. The potential for these fields arises from the D-terms and a superpotential. One can give expectation values to various scalars while leaving $V=0$. The symmetries of the model are sufficient to prevent any quantum corrections to the potential, perturbatively or non-perturbatively. This and other non-renormalization theorems play an important role in our understanding of the AdS-CFT correspondence. But they mean that these would be boring places to live.

The gravitational multiplet consists of the graviton, four gravitinos, a vector multiplet, and ***.

$N=4$ theories arise from compactification of the heterotic string or the Type I string on a six-torus; they also arise from compactification of the Type II theory on $K_3 \times T_2$ (and similarly of the eleven dimensional theory).

N=8

$N = 8$ theories are even more highly constrained. We will not even list the multiplets here, except to note that there are now eight gravitinos in the vector multiplet.

Lecture II

N=1 Theories and Their Properties
Outline

- Puzzles of the Standard Model

- Some successes of String Theory

- Some Illustrative String Models

- The Problem of Moduli

- Predicting Supersymmetry?

While the Standard Model is extraordinarily (depressingly?) successful, it has several shortcomings, (almost) all of which are addressed, in some degree, by string theory:

1. Inclusion of General Relativity (renormalizability, information puzzle,...)

2. Hierarchy problem

3. Generation Puzzle (Flavor)

4. Numerous Parameters

5. The Strong CP Problem

6. Cosmological Constant Problem

On the first problem, string theory has scored successes. The last problem is a severe one, and it is not clear that string theory has yielded much real progress, beyond the fact that, given (1), the problem is (almost) well posed.

What is the hierarchy problem?

Dimensional analysis: $m_H^2 \sim M_p^2$?

Feynman graphs of Standard Model are quadratically divergent:

It has long been argued that the hierarchy problem might be solved if nature is supersymmetric at low energies. Then the quadratic divergences are cancelled

Supersymmetry must be broken. These arguments suggest that the scale of this breaking should be not too different than the scale of weak interactions.

Some virtues of supersymmetry: One of the virtues of supersymmetry is that it can also explain, potentially, why the scale of supersymmetry breaking is hierarchically small compared to the scale of

Lecture III

N=0 Theories and The Connectedness of String Vacua

Outline

- Rohm's Model (or Scherk-Schwarz Compactification in String Theory). This is a model without supersymmetry. At large radius, it has no tachyons. At small radius, a tachyon appears. Near the critical radius, we can compute the effective action for the tachyon and moduli, and study their behavior. While the moduli disappear classically, we can only speculate on the overall structure of the theory.

- An N=0 Model due to Fabinger and Horava.

- Vacuum decay and the connectedness of strings. We review briefly the theory of false vacuum decay, with and without gravity. Gravity, as stressed by Banks, introduces new features which suggest that our simple-minded notions of metastable states connected by potentials may be misleading.

- Decay of the Kaluza-Klein vacuum. Here we review an analysis by Witten which suggests that the Kaluza-Klein vacuum is unstable. It is usually argued that, because a potential appears quantum mechanically for the moduli, this solution is irrelevant. But we will argue that this solution is potentially relevant if there are stable minima of the moduli potential with small or vanishing cosmological constant.

Various tachyonic instabilities in open string theory have been understood in recent years. Tachyons in closed string theory remain somewhat mysterious. In general, one has no technique to compute their potentials (nor is it clear that such a potential is meaningful), and so one does not know whether they correspond to rather familiar sorts of field theory instabilities, or represent some new phenomena. The difficulty in computing the tachyon potential is well known. The potential is a zero-momentum quantity, while the tachyon vertex operator is only defined on shell, far from zero momentum. One might almost hope that tachyons signal some sort of deeper disease of the theory. One might even speculate that such instabilities signal that non-supersymmetric string ground states are in some sense sick.

Rohm proposed a simple model of supersymmetry breaking in string theory. Take Type II theory compactified on a circle, and impose antiperiodic (Scherk-Scharz) boundary conditions on fermions. This breaks supersymmetry. It leads, at one loop, to a potential for the moduli, which Rohm computed. But for our purposes, what is more interesting is that, while for large radius there is not tachyon, a tachyon does appear at a finite value of the radius. In the RNS formulation, for example, the GSO projection is different in sectors with even and odd winding. Alternatively, in the Green-Schwarz formulation, the Green-Schwarz fermions are antiperiodic in the odd winding number sectors.

In general, the momenta are ($\alpha' = \frac{1}{2}$)

$$k_L = \frac{m}{2R} - nR; k_R = \frac{m}{2R} + nR. \qquad (16)$$

Thus, for example, in the sector with winding number one, the mass formula is ($\alpha' = \frac{1}{2}$):

$$L_o = k_L^2 - 1 = n^2 R^2 - 1. \qquad (17)$$

where in the last expression we have taken $m = 0$.

So at $R = 1$, a tachyon appears. At precisely this point, the would-be tachyon is massless, so we can calculate its potential using on-shell vertex operators. In the RNS formulation, for example, the tachyon vertex operator in the minus one ghost number picture is particularly simple:

$$V_{-1} = e^{2ik_L X_L - 2ik_R X_R} e^{ip\cdot x} e^{-\gamma} \qquad (18)$$

whereas in the ghost number zero picture it is

$$V_o = (p \cdot \psi - 2k_L \psi_5)(p \cdot \tilde{\psi} - 2k_R \tilde{\psi}_5) e^{2ik_L X_L - 2ik_R X_R} e^{ip\cdot x} \qquad (19)$$

Take the initial momenta, (k_L, k_R, p) to be $(1, -1, p_1), (-1, 1, p_2)$ and the final momenta to be $(-1, 1, p_3), (1, -1, p_4)$ (all momenta are defined to flow into the diagram). Then the scattering amplitude is obtained from

$$< V_{-1}(k_1, p_1, z_1) V_{-1}(k_2, p_2, z_2) V_0(k_3, p_3, z_3) V_0(k_4, p_4, z_4) \tag{20}$$

$$= \frac{|z_1 - z_4|^2}{|z_1 - z_2|^4} \frac{|z_2 - z_3|^2}{|z_2 - z_4|^2 |z_3 - z_4|^4 |z_1 - z_3|^2} \Pi |z_i - z_j|^{-\frac{p_i \cdot p_j}{2}} (-4 + s/2)^2 \tag{21}$$

where the last factor comes from the contraction of the fermions.

Taking $z_1 = \infty, z_2 = 0, z_3 = 1, z_4 = z$, the amplitude is

$$\mathcal{A} = c \int d^2z \frac{(1+s/8)^2}{|1-z|^{4+s/4}|z|^{2+t/4}}$$

$$= c' \frac{\Gamma(-t/8)\Gamma(-s/8)(1-u/8)\Gamma(1-u/8)}{\Gamma(-1+u/8)\Gamma(1+t/8)\Gamma(1+s/8)}$$

$$= c' \frac{u}{ts}(1-u/8)^2 \frac{\Gamma(1-t/8)\Gamma(1-s/8)\Gamma(1-u/8)}{\Gamma(1+u/8)\Gamma(1+t/8)\Gamma(1+s/8)}$$

Now we want to expand this in powers of momenta, in order to work out the effective action. The amplitude has singularities of the form $1/s$, $1/t$. We will keep terms up to order t, s. These are:

$$\mathcal{A} = \frac{u}{ts}(1+u/8)^2$$

$$= \frac{1}{2\alpha' t} + \frac{1}{2\alpha' s} - \frac{1}{2} + 2\alpha'[\frac{t+s+2s+2t}{64} + \frac{2t^2}{64s} + \frac{2s^2}{64t}].$$

The first term corresponds to radion exchange. The first term in the brackets corresponds to gauge boson exchange; the last two to radion and graviton exchange. To work out the graviton exchange piece, one works with the Feynman rules from the Einstein action, or calculates the three point coupling from the operator product expansion. The radion term is readily obtained from the three point coupling. This, again, can be obtained from the operator product expansion of two tachyons.

$$V_o(k_L, k_R, p)V_o(-k_L, -k_R, p') \sim \frac{4k_L k_R}{|z-z'|^2}[k_L^2 + \frac{p \cdot p'}{4}][k_R^2 + \frac{p \cdot p'}{4}]\partial X_L \bar{\partial} X_R e^{i(p+p')\cdot x} + \ldots \quad (22)$$

So from the radion exchange, we expect an additional contribution:

$$\frac{1}{2s}(1+s/8)^4+\frac{3}{64}(s+t)$$

so the terms match what we found above, plus a contact interaction. Subtracting off the contact interaction leaves the would-be quartic interaction. So, including the quartic interaction and the two derivative coupling of the radion to the tachyon gives the terms in the effective action:

$$\partial^2 R|T|^2+\frac{3}{8}|T|^4 \qquad (23)$$

In this case, the quartic coupling is negative. Alternatively, one can make a field redefinition,

$$R \to R - \frac{1}{2}T^2 \tag{24}$$

This yields

$$\mathcal{L}_{eff} = \frac{1}{4}T^4 + \partial_\mu T^2 \partial^\mu T^2. \tag{25}$$

This effective action, along with radion, gauge boson, and graviton exchanges reproduces the amplitude of eqn. **??** In this form, the quartic coupling is positive.

It is clear from this discussion that the sign of the quartic coupling is ambiguous. However, the low energy dynamics are not.

Let us work in the description with the positive quartic coupling, without the extra $\partial^2 R|T|^2$ coupling. Call σ the canonical field which measures the distance from the fixed point, i.e. $R = e^{\sigma}$. It is tempting to integrate out the tachyon by first minimizing its potential as a function of σ, and then viewing the resulting potential as a potential for the moduli – the dilaton and the radion. But this is not consistent. The resulting mass for σ, in particular, is of order one.

Instead, one can consider a cosmological solution of FRW type. In this case one finds, not surprisingly, that the system is driven to strong coupling and away from the tachyonic point (one also has the possibility of gravitational collapse in the far past or the far future).

So we have:

- In a small regime of the moduli space, we have calculated the tachyon potential. It appears to be a rather conventional sort of instability.

- The system, however, tends to evolve into a regime which we don't have calculational tools to study.

We are left to speculate. We can consider various sorts of dualities. T-duality doesn't help much. It relates the Rohm theory at small radius to a theory in ten dimensions which has a tachyon. S-duality is suggestive. We don't really know whether S-duality makes sense for these non-supersymmetric theories, but we might guess that it tells us, for fixed R, that as the coupling becomes very large, we are driven to a similar, weakly coupled theory. Trying to settle this requires further investigation. [The other possibility, alternative to finding a state with "negative cosmological constant", is that the potential is unbounded below; this seems unlikely, since it implies a peculiar behavior for the dual theory.]

Decay of the False Vacuum

I can't attempt to reproduce the beautiful lecture which Coleman gave on this subject years ago at Erice. I want to just outline the basic physics of vacuum decay. Consider a potential with two minima,

In quantum mechanics, one vacuum simply decays to the other. In field theory, this is not possible. Instead, vacuum decay proceeds by the formation of a bubble, the interior of which is "true vacuum," the exterior is false vacuum. In the semiclassical approximation, one finds the bubble by solving the Euclidean equations of motion for the field theory.

For a single scalar field, these are particularly simple. The lowest action solution can be shown to be O(4) symmetric (in four dimensions), so, for a single scalar field, $\phi = \phi(r^2)$. The Euclidean equation of motion is

$$\ddot{\phi} + \frac{3}{r}\dot{\phi} - V'(\phi) = 0. \tag{26}$$

Formally, this is the equation of motion for a particle in one dimension, with a time-dependent friction term. One looks for a solution which asymptotically reaches the false vacuum. The rate of bubble formation per unit time per unit volume is proportional to e^{-S}, the action for the "bounce."

In one approximation, the equations can be solved in a quite general form. But this won't concern us here. The typical solution looks like:

To understand what is going on in Minkowski space, one simply continues

$$r^2 = x_1^2 + x_2^2 + x_3^3 + x_4^2$$

to

$$\tau^2 = -x_o^2 + x_1^2 + x_2^2 + x_3^2.$$

Correspondingly, the transition region lies at

$$tau^2 = R^2.$$

In other words, the wall, after a microscopic time, expands at the speed of light.

Coleman, in a quite remarkable paper with De Luccia, subsequently considered this problem in the presence of gravity. Here the situation is strikingly different, since the asymptotic space time is different in the true and false vacua. Banks has recently revisited aspects of this problem.

One now has to allow for the non-trivial metric. Assuming $O(4)$ symmetry, this can be taken to have the form:

$$ds^2 = d\xi^2 + \rho(\xi)^2 d\Omega^2.$$

Here Ω is the metric of the three sphere. Now, if one has a scalar field, the relevant equations are:

$$\ddot{\phi} + \frac{3\rho'}{\rho} - V' = 0$$

$$\frac{\rho'^2 = 1 + \frac{1}{3}\kappa^2\rho^2(1}{2\phi'^2 - V.}$$

Coleman and De Luccia analyzed the thin wall case using these equations and found a number of remarkable results.

- For decay of De Sitter space to flat space, gravity enhances the decay rate.

- For decay of flat space to AdS, gravity suppresses the amplitude, and the amplitude vanishes for a critical value of the parameters.

The Minkowski continuation is somewhat more subtle in this case, but again, CD found some remarkable results. Perhaps most importantly, they found that even when decay to AdS seems allowed, the result is gravitational collapse.

Banks has looked at these issues recently, in light of developments in string theory. He has extended the CD results beyond the thin wall approximation, and argued that there may be no situation in which false vacuum decay makes sense. Indeed, he argues, based on these and other observations, that one cannot really think of such different states as being part of the same theory. This resolves a number of paradoxes connected with holography, which I won't review here.

***** more on Banks*****

Instead, I wish to turn to another type of vacuum decay, one which seems to distinguish supersymmetry and non-supersymmetric vacua in a decisive way.

The Decay of the Kaluza-Klein Vacuum

Some time ago, Witten pointed out an interesting feature of the simplest Kaluza-Klein theory: it is unstable. His reasoning goes as follows. The five dimensional Euclidean Einstein equations have a solution:

$$ds^2 = \frac{dr^2}{(1 - \frac{R^2}{r^2})} + r^2 d\Omega^2 + R^2(1 - \frac{R^2}{r^2})d\phi^2.$$

Here $0 < \phi < 2\pi$. The coefficient of $d\phi^2$ is the effective radius of the fifth dimension, i.e. it is the value of the radial dilaton field. Note that the singularity at $r = R$ is a coordinate singularity. Near this point, writing $r = R(1 + \sqrt{\frac{1}{2}}u^2)$ the metric becomes $du^2 + u^2 d\phi^2$.

Such a Euclidean solution is usually a signal of tunneling. Witten verified, in particular, that the small fluctuation operator for this solution has a single negative mode, yielding the requisite imaginary part.

He also explained, following Coleman and DeLuccia, what the decay corresponds to. To do this, it is necessary to find a suitable Minkowski continuation. Writing the metric of the three sphere as

$$d\Omega_3^2 = d\theta^2 + \sin^2(\theta) d\phi^2 + \sin^2(theta)\sin^2(\phi) d\psi^2)$$

one continues in the polar angle θ, $\theta \rightarrow i\tau + \frac{\pi}{2}$ and writes:

$$ds^2 = -d\tau^2 + sinh^2(\tau) d\Omega^2 + \frac{dr^2}{(1 - \frac{R^2}{r^2})} + (1 - \frac{R^2}{r^2}) d\phi^2.$$

As Witten says, here we “encounter a surprise.” This geometry describes a “bubble of nothing.” For times large compared to R, from the perspective of a four dimensional observer, space-time appears to end at the bubble wall. There is nothing obviously wrong about this space time. It is non-singular and geodesically complete, when viewed from five dimensions. But from the perspective of a four dimensional observer, the universe appears to end at the bubble wall.

This result may appear somewhat alarming. Its full significance is not completely clear since, unlike the conventional Coleman-Deluccia case analysis, we don’t know have a simple picture of what we are decaying to. The story is more complicated when we imagine bubbles forming in various regions of space-time and then colliding. As Witten says, it is likely that the result is “Gravitational Collapse.” (Horowitz and Maeda have recently generalized the solution).

One point stressed by Witten is that this solution distinguishes supersymmetric from non-supersymmetric compactifications. The point is quite simple. Near $R = 1$, we saw that the geometry is that of the plane in polar coordinates. As a result, fermions, in going around the origin (i.e. in letting $\phi \to \phi + 2\pi$), must pick up a factor of -1. At large R, these are just Scherk-Schwarz boundary conditions, i.e. they break supersymmetry. So the Witten solution is not admissable for supersymmetric compactifications. But it is admissable for non-supersymmetric ones!

In the case of the Rohm compactification, the Witten solution should be a good solution. It certainly solves the lowest order equations of motion, and there is no reason to think that one cannot generalize the solution to higher orders in the α' expansion.

The interior of the Witten bubble, i.e. the region $r \to R$, corresponds to shrinking the size of the compact dimension to a point. This happens smoothly. It is why, from a five dimensional perspective, the solution makes sense. We have seen, for the Rohm solution, that new issues arise as the radius becomes small. In particular, instabilities may appear. It is interesting that the Witten solution breaks the same Z_2 symmetry (which distinguishes even and odd windings) as the tachyon expectation value of the Rohm theory.

In any case, what is the relevance of this solution? After all, we have seen that perturbatively, there is a potential for the radion, and even classically, at small radius once we consider the instabilities. The lifetime for the bubble formation is

$$\Gamma = exp - \frac{\pi R^2}{4G}$$

which is very long for large R, compared to the time for motion of the field.

But this is not precisely what interests us. We are more interested in the possibility that the potential for the moduli has a minimum; for describing the world we see, this minimum should be at very small cosmological constant. In other words, the potential for the radius might be imagined to look something like:

We want to ask, in this case, whether a solution exists.

First, we should ask about T-duality. In other words, does a T-dual solution exist. (The Type II theory compactified in this way is T-dual, as we have remarked, to another theory. However, the bosonic part of the action is T-dual by itself).

To construct the solution, it is easiest to first reduce to four dimensions, Weyl-rescaling to the Einstein frame. It is in this form that T-duality is simple; it takes $R(x) \to \frac{1}{R(x)}$, while leaving the four dimensional metric alone.

The four dimensional metric is:

$$ds^2 = \frac{dr^2}{\sqrt{1 - R^2/r^2}} + r^2\sqrt{1 - R^2/r^2}d\Omega^2.$$

Performing the T-duality transformation and Weyl-rescaling back to five dimensions gives a new solution:

$$g_{\mu\nu} = dr^2 + r^2(1 - \frac{R^2}{r^2})d\Omega^2 + \frac{1}{(1 - \frac{R^2}{r^2})}d\phi^2 \quad g^{-2}(r) = g^{-2}\frac{1}{(1 - \frac{R^2}{r^2})}.$$

One can check that this is a solution to the five dimensional equations. There is also a non-trivial dilaton background, which solves the free scalar equation in the gravitational background.

So now consider the possibility that there is a minimum of the potential, say with zero cosmological constant. We want to ask whether a solution exists? To do this, we will consider the four dimensional equations for the bounce, just as do Coleman and DeLuccia. But we will require that as the Euclidean time tends to zero (corresponding to $r \to 1$), the solution tend to the Witten solution. In particular, this insures that the solution is, from a five dimensional perspective, non-singular.

Coleman and DeLuccia consider a scalar field, with lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - V(\phi). \tag{27}$$

They look for a solution with $O(4)$ symmetry, which interpolates between the false and true vacua. The symmetry dictates that the metric may be written in the form:

$$(ds)^2 = (d\xi)^2 + \rho(\xi)(d\Omega)^2. \tag{28}$$

The the Euclidean equations for ϕ and ρ are:

$$\phi'' + \frac{3\rho'}{\rho}\phi' = \frac{dU}{d\phi}, \tag{29}$$

and

$$\rho'^{\ 2} = 1 + \frac{1}{3}\kappa\rho^2(\frac{1}{2}\phi'^{\ 2} - U). \tag{30}$$

CD then consider an analog problem, where ξ plays the role of time. They require that at $\xi = 0$, the system starts off in (near) the true vacuum, while for $\xi \to \infty$ the system tends to the false vacuum. In particular, they require that the system start off with zero velocity.

How does the Witten solution look in this description? The coefficient of $d\phi^2$ in the metric we parameterize as the field $R(x)^2$ (as ϕ is defined, this field is dimensionless). In terms of the canonically normalized scalar, ϕ,

$$R = exp^{\frac{1}{\sqrt{3}}\sigma}. \quad (31)$$

For the Witten problem, the potential vanishes. We need also to Weyl rescale to four dimensions. In this case, the four dimensional Witten solution has the form:

$$ds^2 = \frac{dr^2}{\sqrt{1 - R^2/r^2}} + r^2\sqrt{1 - R^2/r^2}d\Omega^2. \quad (32)$$

To relate this to the CD equations, we need to solve

$$\frac{d\xi}{dr}) = \sqrt{1 - R^2/r^2}. \tag{33}$$

Then we can read off ρ from the metric. In particular, as $r \to 1$, we have

$$r = 1 + (3/42^(1/4)\xi)^(4/3)) \tag{34}$$

and

$$\rho = (r-1)^{1/4} \qquad \phi \approx \sqrt{3}\,\ln(r-1). \tag{35}$$

Note, in particular, that as $\xi \to 0$, the field and its derivative become large. In the analog mechanics problem, this corresponds to starting off with a big kick. It is a simple exercise with Mathematica to find the Witten solution from eqns. taking the limiting behavior above for the solution.

Now we can consider what happens when we introduce a potential. We require the same asymptotics for the coefficient of the $d\phi^2$ term relative to the dr^2 term, but the coefficient of the $d\Omega_3^2$ term is not fixed. Varying this coefficient changes slightly the "friction" in the analog problem at early times. This leads to a the possibility of "overshoot." By varying this term, one tries to find a solution which asymptotically approaches the "false vacuum."

In general, we find that the existence of a solution depends on the height of the potential. For a given functional form, as we increase the size of the potential, we find, not surprisingly, that the solution eventually disappears. The properties of these solutions are still under investigation. The implications I will leave for you to speculate on.

One point worthy of note: in this form, it is clear that decay from de Sitter space is possible. It seems that even anti-de Sitter space can decay in this way.