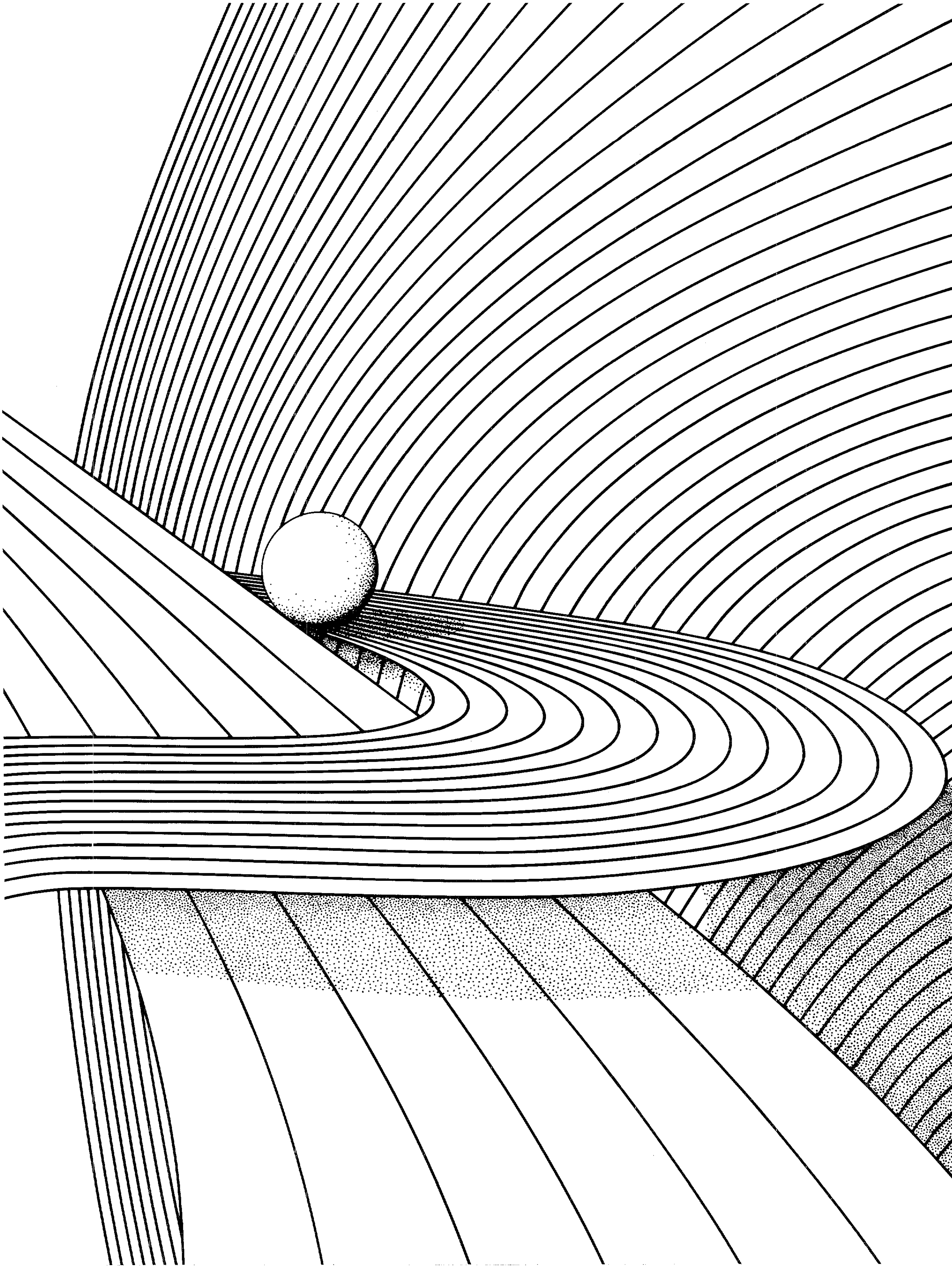
an essay on the role of supergravity in the search for unification

Toward a Unified Theory

by Richard C. Slansky



All throughout his history man has wanted to know the dimensions of his world and his place in it. Before the advent of scientific instruments the universe did not seem very large or complicated. Anything too small to detect with the naked eye was not known, and the few visible stars might almost be touched if only there were a higher hill nearby.

Today, with high-energy particle accelerators the frontier has been pushed down to distance intervals as small as 10^{-16} centimeter and with super telescopes to cosmological distances. These explorations have revealed a multifaceted universe; at first glance its diversity appears too complicated to be described in any unified manner. Nevertheless, it has been possible to incorporate the immense variety of experimental data into a small number of quantum field theories that describe four basic interactions—weak, strong, electromagnetic, and gravitational. Their mathematical formulations are similar in that each one can be derived from a local symmetry. This similarity has inspired hope for even greater progress: perhaps an extension of the present theoretical framework will provide a single unified description of all natural phenomena.

This dream of unification has recurred again and again, and there have been many successes: Maxwell's unification of electricity and magnetism; Einstein's unification of gravitational phenomena with the geometry of space-time; the quantum-mechanical unification of Newtonian mechanics with the wave-like behavior of matter; the quantum-mechanical generalization of electrodynamics; and finally the recent unification of electromagnetism with the weak force. Each of these advances is a crucial component of the present efforts to seek a more complete physical theory.

Before the successes of the past inspire too much optimism, it is important to note that a unified theory will require an unprecedented extrapolation. The present optimism is generated by the discovery of theories successful at describing phenomena that take place over distance intervals of order 10^{-16} centimeter or larger. These theories may be valid to much shorter distances, but that remains to be tested experimentally. A fully unified theory will have to include gravity and therefore will probably have to describe spatial structures as small as 10^{-33} centimeter, the fundamental length (determined by Newton's gravitational constant) in the theory of gravity. History suggests cause for further caution: the record shows many failures resulting from attempts to unify the wrong, too few, or too many physical phenomena. The end of the 19th century saw a huge but unsuccessful effort to unify the description of all Nature with thermodynamics. Since the second law of thermodynamics cannot be derived from Newtonian mechanics, some physicists felt it must have the most fundamental significance and sought to derive the rest of physics from it. Then came a period of belief in the combined use of Maxwell's electrodynamics and Newton's mechanics to explain all natural phenomena. This effort was also doomed to failure: not only did these theories lack consistency (Newton's equations are consistent with particles traveling faster than the speed of light, whereas the Lorentz invariant equations of Maxwell are not), but also new experimental results were emerging that implied the quantum structure of matter. Further into this century came the celebrated effort by Einstein to formulate a unified field theory of gravity and electromagnetism. His failure notwithstanding, the mathematical form of his classical theory has many remarkable similarities to the modern efforts to unify all known fundamental interactions. We must be wary that our reliance on quantum field theory and local symmetry may be similarly misdirected, although we suppose here that it is not.

Two questions will be the central themes of this essay. First, should we believe that the theories known today are the correct components of a truly unified theory? The component theories are now so broadly accepted that they have become known as the "standard model." They include the electroweak theory, which gives a unified description of quantum electrodynamics (QED) and the weak interactions, and quantum chromodynamics (QCD), which is an attractive candidate theory for the strong interactions. We will argue that, although Einstein's theory of gravity (also called general relativity) has a somewhat different status among physical theories, it should also be included in the standard model. If it is, then the standard model incorporates all observed physical phenomena—from the shortest distance intervals probed at the highest energy accelerators to the longest distances seen by modern telescopes. However, despite its experimental successes, the standard model remains unsatisfying; among its shortcomings is the presence of a large number of arbitrary constants that require explanations. It remains to be seen whether the next level of unification will provide just a few insights into the standard model or will unify all natural phenomena.

The second question examined in this essay is twofold: What are the possible strategies for generalizing and extending the standard model, and how nearly do models based on these strategies describe Nature? A central problem of theoretical physics is to identify the features of a theory that should be abstracted, extended, modified, or generalized. From among the bewildering array of theories, speculations, and ideas that have grown from the standard model, we will describe several that are currently attracting much attention.

We focus on two extensions of established concepts. The first is called supersymmetry; it enlarges the usual space-time symmetries of field theory, namely, Poincaré invariance, to include a symmetry among the bosons (particles of integer spin) and fermions (particles of half-odd integer spin). One of the intriguing features of supersymmetry is that it can be extended to include internal symmetries (see Note 2 in "Lecture Notes—From Simple Field Theories to the Standard Model). In the standard model internal local symmetries play a crucial role, both for classifying elementary particles and for de-

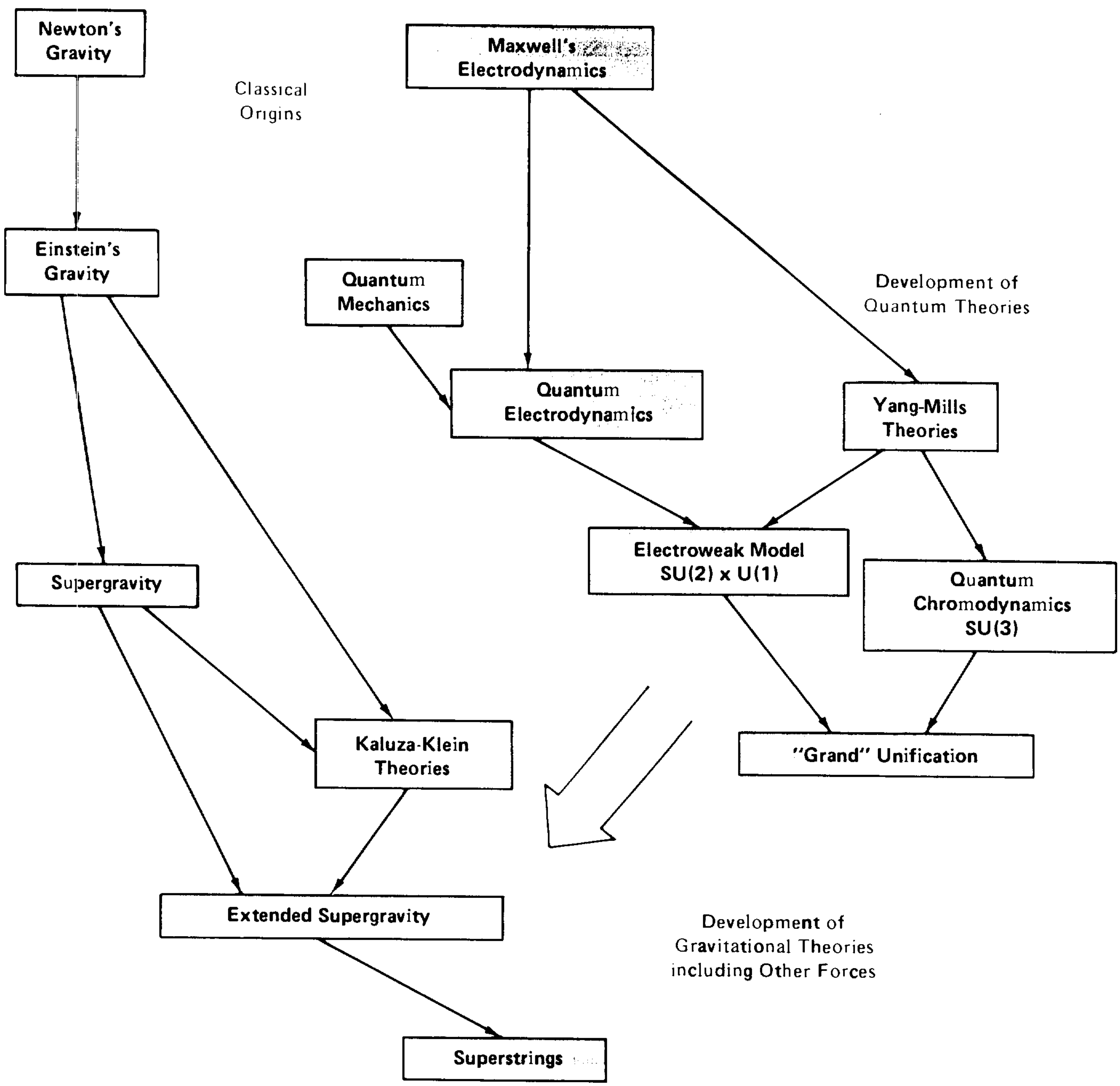


Fig. 1. Evolution of fundamental theories of Nature from the classical field theories of Newton and Maxwell to the grandest theoretical conjectures of today. The relationships among these theories are discussed in the text. Solid lines indicate a direct and well-established extension, or theoretical generalization. The wide arrow symbolizes the goal of present research, the unification of quantum field theories with gravity.

termining the form of the interactions among them. The electroweak theory is based on the internal local symmetry group SU(2) × U(1) (see Note 8) and quantum chromodynamics on the internal local symmetry group SU(3). Gravity is based on space-time symmetries: general coordinate invariance and local Poincaré symmetry. It is tempting to try to unify all these symmetries with supersymmetry.

Other important implications of supersymmetry are that it enlarges the scope of the classification schemes of the basic particles to include fields of different spins in the same multiplet, and it helps to solve some technical problems concerning large mass ratios that plague certain efforts to derive the standard model. Most significantly, if supersymmetry is made to be a local symmetry, then it automatically implies a theory of gravity, called supergravity, that is a generalization of Einstein's theory. Supergravity theories require the unification of gravity with other kinds of interactions, which may be, in some future version, the electroweak and strong interactions. The near successes of this approach are very encouraging.

The other major idea described here is the extension of the space-time manifold to more than four dimensions, the extra dimensions having, so far, escaped observation. This revolutionary idea implies that particles are grouped into larger symmetry multiplets and the basic interactions have a geometrical origin. Although the idea of extending space-time beyond four dimensions is not new, it becomes natural in the context of supergravity theories because these complicated theories in four dimensions may be derived from relatively simple-looking theories in higher dimensions.

We will follow these developments one step further to a generalization of the field concept: instead of depending on space-time, the fields may depend on paths in space-time. When this generalization is combined with supersymmetry, the resulting theory is called a superstring theory. (The whimsicality of the name is more than matched by the theory's complexity.) Superstring theories are encouraging because some of them reduce, in a certain limit, to the only supergravity theories that are likely to generalize the standard model. Moreover, whereas supergravity fails to give the standard model exactly, a superstring theory might succeed. It seems that superstring theories can be formulated only in ten dimensions.

Figure 1 provides a road map for this essay, which journeys from the origins of the standard model in classical theory to the extensions of the standard model in supergravity and superstrings. These extensions may provide extremely elegant ways to unify the standard model and are therefore attracting enormous theoretical interest. It must be cautioned, however, that at present no experimental evidence exists for supersymmetry or extra dimensions.

Review of the Standard Model

We now review the standard model with particular emphasis on its potential for being unified by a larger theory. Over the last several decades relativistic quantum field theories with local symmetry have succeeded in describing all the known interactions down to the smallest distances that have been explored experimentally, and they may be correct to much shorter distances.

Electrodynamics and Local Symmetry. Electrodynamics was the first theory with local symmetry. Maxwell's great unification of electricity and magnetism can be viewed as the discovery that electrodynamics is described by the simplest possible local symmetry, local phase invariance. Maxwell's addition of the displacement current to the field equations, which was made in order to insure conservation of the electromagnetic current, turns out to be equivalent to imposing local phase invariance on the Lagrangian of electrodynamics, although this idea did not emerge until the late 1920s.

A crucial feature of locally symmetric quantum field theories is this: typically, for each independent internal local symmetry there exists a gauge field and its corresponding particle, which is a vector boson (spin-1 particle) that mediates the interaction between particles. Quantum electrodynamics has just one independent local symmetry transformation, and the photon is the vector boson (or gauge particle) mediating the interaction between electrons or other charged particles. Furthermore, the local symmetry dictates the exact form of the interaction. The interaction Lagrangian must be of the form $eJ^{\mu}(x)A_{\mu}(x)$, where $J^{\mu}(x)$ is the current density of the charged particles and $A_{\mu}(x)$ is the field of the vector bosons. The coupling constant e is defined as the strength with which the vector boson interacts with the current. The hypothesis that all interactions are mediated by vector bosons or, equivalently, that they originate from local symmetries has been extended to the weak and then to the strong interactions.

Weak Interactions. Before the present understanding of weak interactions in terms of local symmetry, Fermi's 1934 phenomenological theory of the weak interactions had been used to interpret many data on nuclear beta decay. After it was modified to include parity violation, it contained all the crucial elements necessary to describe the low-energy weak interactions. His theory assumed that beta decay (e.g., $n \rightarrow p + e^- + \bar{\nu}_e$) takes place at a single space-time point. The form of the interaction amplitude is a product of two currents $J^{\mu}J_{\mu}$, where each current is a product of fermion fields, and $J^{\mu}J_{\mu}$ describes four fermion fields acting at the point of the beta-decay interaction. This amplitude, although yielding accurate predictions at low energies, is expected to fail at center-of-mass energies above 300 GeV, where it predicts cross sections that are larger than allowed by the general principles of quantum field theory.

The problem of making a consistent (renormalizable) quantum field theory to describe the weak interactions was not solved until the 1960s, when the electromagnetic and weak interactions were combined into a locally symmetric theory. As outlined in Fig.

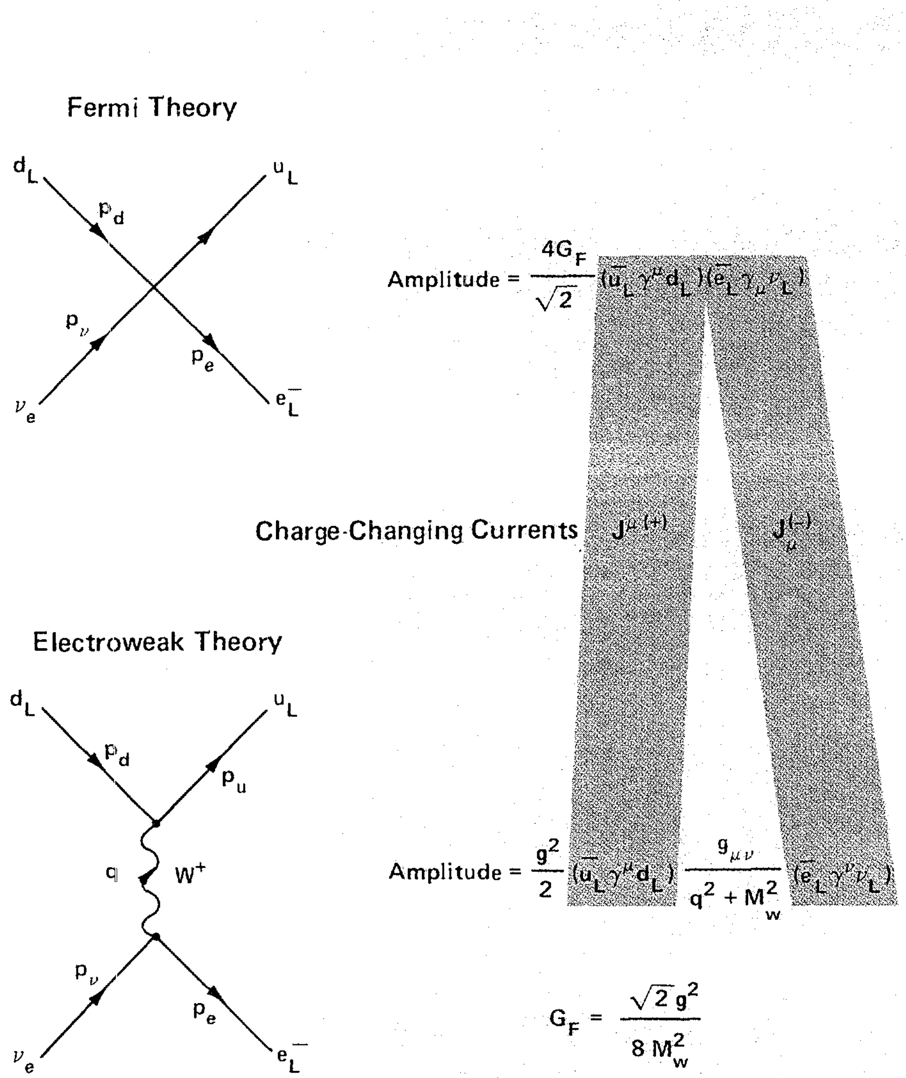


Fig. 2. Comparison of neutrino-quark charged-current scattering in the Fermi theory and the modern SU(2) × U(1) electroweak theory. (The bar indicates the Dirac conjugate.) The point interaction of the Fermi theory leads to an inconsistent quantum theory. The W^+ boson exchange in the electroweak theory spreads out the weak interactions, which then leads to a consistent (renormalizable) quantum field theory. $J_\mu^{(+)}$ and $J_\mu^{(-)}$ are the charge-raising and charge-lowering currents, respectively. The amplitudes given by the two theories are nearly equal as long as the square of the momentum transfer, $q^2 = (p_u - p_d)^2$, is much less than the square of the mass of the weak boson, M_W^2).

2, the vector bosons associated with the electroweak local symmetry serve to spread out the interaction of the Fermi theory in space-time in a way that makes the theory consistent. Technically, the major problem with the Fermi theory is that the Fermi coupling constant, G_F, is not dimensionless (G_F = (293 GeV)$^{-2}$), and therefore the Fermi theory is not a renormalizable quantum field theory. This means that removing the infinities from the theory strips it of all its predictive power.

In the gauge theory generalization of Fermi's theory, beta decay and other weak interactions are mediated by heavy weak vector bosons, so the basic interaction has the form $gW^{\mu}J_\mu$ and the current-current interaction looks pointlike only for energies much less than the rest energy of the weak bosons. (The coupling g is dimensionless, whereas G_F is a composite number that includes the masses of the weak vector bosons.) The theory has four independent local symmetries, including the phase symmetry that yields electrodynamics. The local symmetry group of the electroweak theory is SU(2) × U(1), where U(1) is the group of phase transformations, and SU(2) has the same structure as rotations in three dimensions. The one phase angle and the three independent angles of rotation in this theory imply the existence of four vector bosons, the photon plus three weak vector bosons, W^+, Z^0, and W^-. These four particles couple to the four SU(2) × U(1) currents and are responsible for the "electroweak" interactions.

The idea that all interactions must be derived from local symmetry may seem simple, but it was not at all obvious how to apply this idea to the weak (or the strong) interactions. Nor was it obvious that electrodynamics and the weak interactions should be part of the same local symmetry since, experimentally, the weak bosons and the photon do not share much in common: the photon has been known as a physical entity for nearly eighty years, but the weak vector bosons were not observed until late 1982 and early 1983 at the CERN proton-antiproton collider in the highest energy accelerator experiments ever performed; the mass of the photon is consistent with zero, whereas the weak vector bosons have huge masses (a little less than 100 GeV/c^2); electromagnetic interactions can take place over very large distances, whereas the weak interactions take place on a distance scale of about 10^{-16} centimeter; and finally, the photon has no electric charge, whereas the weak vector bosons carry the electric and weak charges of the electroweak interactions. Moreover, in the early days of gauge theories, it was generally believed, although incorrectly, that local symmetry of a Lagrangian implies masslessness for the vector bosons.

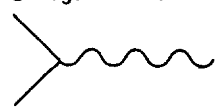
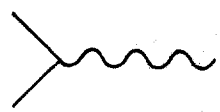
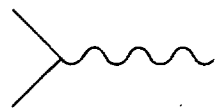
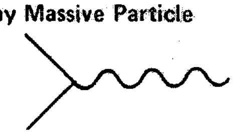
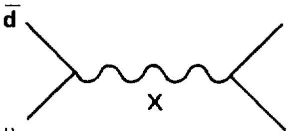
How can particles as different as the photon and the weak bosons possibly be unified by local symmetry? The answer is explained in detail in the Lecture Notes; we mention here merely that if the vacuum of a locally symmetric theory has a nonzero symmetry charge density due to the presence of a spinless field, then the vector boson associated with that symmetry acquires a mass. Solutions to the equations of motion in which the vacuum is not invariant under symmetry transformations are called spontaneously broken solutions, and the vector boson mass can be arbitrarily large without upsetting the symmetry of the Lagrangian.

In the electroweak theory spontaneous symmetry breaking separates the weak and electromagnetic interactions and is the most important mechanism for generating masses of the elementary particles. In the theories dicussed below, spontaneous symmetry breaking is often used to distinguish interactions that have been unified by extending symmetries (see Note 8).

The range of validity of the electroweak theory is an important issue, especially when considering extensions and generalizations to a theory of broader applicability. "Range of validity" refers to the energy (or distance) scale over which the predictions of a theory are valid. The old Fermi theory gives a good account of the weak interactions for energies less than 50 GeV, but at higher energies, where the effect of the weak bosons is to

Table 1

Review of fundamental interactions.

Interaction Example	Name	Local Symmetry
 Any Charged Particle; Any Charged Particle; Photon	Electromagnetic (QED)	U(1)
 Quark; Quark; Gluon	Strong (QCD)	SU(3)
 ν_e; e^-; W^+	Electroweak	SU(2) × U(1)
 Any Massive Particle; Any Massive Particle; Graviton	Gravity	Poincaré
 $\bar{d}$; u; e^+; u; X; (Proton Decay)	Conjectured Strong-Electroweak Unification	SU(5)

Local Symmetry: The generator of the electromagnetic U(1) is a linear combination of the generators of the electroweak U(1) and the diagonal generator of the electroweak SU(2). The general coordinate invariance of gravity permits several formulations of gravity in which different local symmetries can be emphasized.

Range of Force: The electromagnetic and gravitational forces fall off as $1/r^2$. Of course, the electromagnetic part of the electroweak force is long range.

Relative Strength at Low Energy: The strength of the strong interactions is extremely energy-dependent. At low energy hadronic amplitudes are typically 100 times stronger than electromagnetic amplitudes.

Number of Vector Bosons: The graviton can be viewed as the gauge particle for translations, and as a consequence it has a spin of 2. After all the symmetries of gravity are taken into account, the graviton is massless and has only two degrees of freedom with helicities (spin components) ±2.

Number of Vector Bosons	Range of Force	Relative Strength at Low Energy	Mass Scale
1 (photon)	Infinite	1/137	
8 (gluons)	10^{-13} cm	1	$\mu = 200\ \mathrm{MeV}/c^2$
4 (3 weak bosons, 1 photon)	10^{-15} cm (weak)	10^{-5}	$G_F^{-1/2} = 290\ \mathrm{GeV}/c^2$
(Graviton)	Infinite	10^{-38}	$G_N^{-1/2} = 1.2 \times 10^{10}\ \mathrm{GeV}/c^2$
24	10^{-29} cm	10^{-32}	$10^{15}\ \mathrm{GeV}/c^2$

Mass Scale: There is no universal definition of mass scale in particle physics. It is, however, possible to select a mass scale of physical significance for each of these theories. For example, in the electroweak and SU(5) theories the mass scale is associated with the spontaneous symmetry breaking. In both cases the vacuum value of a scalar field (which has dimensions of mass) has a nonzero value. In the weak interactions G_F is related directly to this vacuum value (see Fig. 2) and, at the same time, to the masses of the weak bosons. Similarly, the scale of the SU(5) model is related to the proton-decay rate and to the vacuum value of a different scalar field. In the Fermi theory G_F is the strength of the weak interaction in the same way that G_N is the strength of the gravitational interaction. However, in gravity theory, with its massless graviton, the origin of the large value of G_N is not well understood. (It might be related to a vacuum value but not in precisely the way that G_F is.) The QCD mass scale is defined in a completely different way. Aside from the quark masses, the classical QCD Lagrangian has no mass scales and no scalar fields. However, in quantum field theory the coupling of a gluon to a quark current depends on the momentum carried by the gluon, and this coupling is found to be large for momentum transfers below 200 MeV/c. It is thus customary to select $\mu = 200\ \mathrm{MeV}/c^2$ (where μ is the parameter governing the scale of asymptotic freedom) as the mass scale for QCD.

spread out the weak interactions in space-time, the Fermi theory fails. The electroweak theory remains a consistent quantum field theory at energies far above a few hundred GeV and reduces to the Fermi theory (with the modification for parity violation) at lower energies. Moreover, it correctly predicts the masses of the weak vector bosons. In fact, until experiment proves otherwise, there are no logical impediments to extending the electroweak theory to an energy scale as large as desired. Recall the example of electrodynamics and its quantum-mechanical generalization. As a theory of light in the mid-19th century, it could be tested to about 10^{-5} centimeter. How could it have been known that QED would still be valid for distance scales ten orders of magnitude smaller? Even today it is not known where quantum electrodynamics breaks down.

Strong Interactions. Quantum chromodynamics is the candidate theory of the strong interactions. It, too, is a quantum field theory based on a local symmetry; the symmetry, called color SU(3), has eight independent kinds of transformations, and so the strong interactions among the quark fields are mediated by eight vector bosons, called gluons. Apparently, the local symmetry of the strong interaction theory is not spontaneously broken. Although conceptually simpler, the absence of symmetry breaking makes it harder to extract experimental predictions. The exact SU(3) color symmetry may imply that the quarks and gluons, which carry the SU(3) color charge, can never be observed in isolation. There seem to be no simple relationships between the quark and gluon fields of the theory and the observed structure of hadrons (strongly interacting particles). The quark model of hadrons has not been rigorously derived from QCD.

One of the main clues that quantum chromodynamics is correct comes from the results of "deep" inelastic scattering experiments in which leptons are used to probe the structure of protons and neutrons at very short distance intervals. The theory predicts that at very high momentum transfers or, equivalently, at very short distances ($<10^{-13}$ centimeter) the quark and gluon fields that make up the nucleons have a direct and fundamental interpretation: they are almost noninteracting, point-like particles. Deep inelastic electron, muon, and neutrino experiments have tested the short-distance structure of protons and neutrons and have confirmed qualitatively this short-distance prediction of quantum chromodynamics. At relatively long distance intervals of 10^{-13} centimeter or greater, the theory must account for the existence of the observed hadrons, which are complicated composites of the quark and gluon fields. Until progress is made in deriving the list of hadrons from quantum chromodynamics, we will not know whether it is the correct theory of the strong interactions. This is a rather peculiar situation: the validity of QCD at energies above a few GeV is established (and there is no experimental or theoretical reason to limit the range of validity of the theory at even higher energies), but the long-distance (low-energy) structure of the theory, including the hadron spectrum, has not yet been calculated. Perhaps the huge computational effort now being devoted to testing the theory will resolve this question soon.

Gravity. Gravity theory (and by this is meant Einstein's theory of general relativity) should be added to the standard model, although it has a different status from the electroweak and strong theories. The energy scale at which gravity becomes strong, according to Einstein's (or Newton's) theory, is far above the electroweak scale: it is given by the Planck mass, which is defined as $(\hbar c/G_N)^{1/2}$, where G_N is Newton's gravitational constant, and is equal to 1.2×10^{19} GeV/c^2. (In quantum theories distance is inversely proportional to energy; the Planck mass corresponds to a length (the Planck length) of 1.6×10^{-33} centimeter.) Large mass scales are typically associated with small interaction rates, so gravity has a negligible effect on high-energy particle physics at present accelerator energies. The reason we feel the effect of this very weak interaction so readily in everyday life is that the graviton, which mediates the interaction, is massless and has long-range interactions like the photon. Moreover, the gravitational force has always been found to be attractive; matter in bulk cannot be "gravitationally neutral" in the way that it is typically electrically neutral.

At present there are no experimental reasons that compel us to include gravity in the standard model; present particle phenomenology is explained without it. Moreover, its theoretical standing is shaky, since all attempts to formulate Einstein's gravity as a consistent quantum field theory have failed. The problem is similar to that of the Fermi theory: Newton's constant has dimensions of (energy)$^{-2}$ so the theory is not renormalizable. However, like the Fermi theory, it is valid up to an energy that is a substantial fraction of its energy scale of 10^{19} GeV. This is the only known serious inconsistency in the standard model when gravity is included. Thus, including gravity in the standard model seems to pose many problems. Yet, there is a good reason to attempt this unification: there exist theoretical models (as we discuss later) that suggest that the electroweak and strong theories may cure the ills of gravitational theory, and unification with gravity may require a theory that predicts the phenomenological inputs of the electroweak and strong theories.

The mathematical structure of gravity theory provides another reason for its inclusion in the standard model. Like the other interactions, gravity is based on a local symmetry, the Poincaré symmetry, which includes Lorentz transformations and space-time translations. In this case, however, not all the generators of the symmetry group give rise to particles that mediate the gravitational interaction. In particular, Einstein's theory has no kinetic energy terms in the Lagrangian for the gauge fields corresponding to the six independent symmetries of the Lorentz group. The space-time translations have associated with them the gauge field called the graviton that mediates the gravitational interaction. The graviton field has a spin of 2 and is denoted by $e^a_\mu(x)$, where the vector index μ on the usual boson field is combined with the space-time translation index a to form a spin of 2. The metric tensor is, essentially, the square of $e^a_\mu(x)$. The massless graviton has two helicities (spin projections along the direction of motion) of values ± 2. In some ways these are merely technical differences, and gravity is like the other interactions. Nevertheless, these differences are crucial in the search for theories that unify gravity with the other interactions.

Summary. Let us summarize why the standard model including gravity may be the correct set of component theories of a truly unified theory.

- The standard model (with its phenomenologically motivated symmetries, choice of fields, and Lagrangian) correctly accounts for all elementary-particle data.
- The standard model contains no known mathematical inconsistencies up to an energy scale near 10^{19} GeV, and then only gravity gives difficulty.
- All components of the standard model have similar mathematical structures. Essentially, they are local gauge theories, which can be derived from a principle of local symmetry.
- There are no logical or phenomenological requirements that force the addition of further components to describe phenomena at scales greater than 10^{-16} centimeter. Thus, we are free to seek theories with a range of validity that may transcend the present experimental frontier.

We still have to cope with the huge extrapolation, by seventeen orders of magnitude, in energy scale necessary to include gravity in the theory. At best it appears reckless to begin the search for such a unification, in spite of the good luck historically with quantum electrodynamics. However, even if we ignore gravity, the energy scales encountered in attempts to unify just the electroweak and strong interactions are surprisingly close to the Planck mass. These more

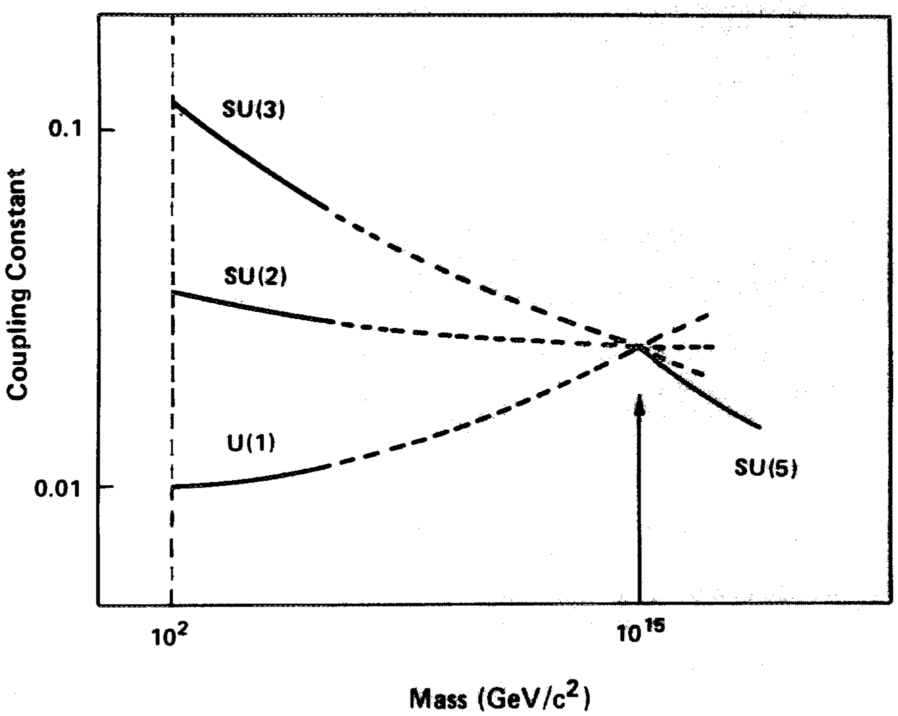


Fig. 3. Unification in the SU(5) model. The values of the SU(2), U(1), and SU(3) couplings in the SU(5) model are shown as functions of mass scale. These values are calculated using the renormalization group equations of quantum field theory. At the unification energy scale the proton-decay bosons begin to contribute to the renormalization group equations; at higher energies, the ratios track together along the solid curve. If the high-mass bosons were not included in the calculation, the couplings would follow the dashed curves.

modest efforts to unify the fundamental interactions may be an important step toward including gravity. Moreover, these efforts require the belief that local gauge theories are correct to distance intervals around 10^{-29} centimeter, and so they have made theorists more "comfortable" when considering the extrapolation to gravity, which is only four orders of magnitude further. Whether this outlook has been misleading remains to be seen. The components of the standard model are summarized in Table 1.

Electroweak-Strong Unification without Gravity

The SU(2) × U(1) × SU(3) local theory is a detailed phenomenological framework in which to analyze and correlate data on electroweak and strong interactions, but the choice of symmetry group, the charge assignments of the scalars and fermions, and the values of many masses and couplings must be deduced from experimental data. The problem is to find the simplest extension of this part of the standard model that also unifies (at least partially) the interactions, assignments, and parameters that must be put into it "by hand." Total success at unification is not required at this stage because the range of validity will be restricted by gravitational effects.

One extension is to a local symmetry group that includes SU(2) × U(1) × SU(3) and interrelates the transformations of the standard model by further internal symmetry transformations. The simplest example is the group SU(5), although most of the comments below also apply to other proposals for electroweak-strong unification. The SU(5) local symmetry implies new constraints on the fields and parameters in the theory. However, the theory also includes new interactions that mix the electroweak and strong quantum numbers; in SU(5) there are vector bosons that transform quarks to leptons and quarks to antiquarks. These vector bosons provide a mechanism for proton decay.

If the SU(5) local symmetry were exact, all the couplings of the vector bosons to the symmetry currents would be equal (or related by known factors), and consequently the proton decay rate would be near the weak decay rates. Spontaneous symmetry breaking of SU(5) is introduced into the theory to separate the electroweak and strong interactions from the other SU(5) interactions as well as to provide a huge mass for the vector bosons mediating proton decay and thereby reduce the predicted decay rate. To satisfy the experimental constraint that the proton lifetime be at least 10^{30} years, the masses of the heavy vector bosons in the SU(5) model must be at least 10^{14} GeV/c^2. Thus, experimental facts already determine that the electroweak-strong unification must introduce masses into the theory that are within a factor of 10^5 of the Planck mass.

It is possible to calculate the proton lifetime in the SU(5) model and similar unified models from the values of the couplings and masses of the particles in the theory. The couplings of the standard model (the two electroweak couplings and the strong coupling) have been measured in low-energy processes. Although the ratios of the couplings are predicted by SU(5), the symmetry values are accurate only at energies where SU(5) looks exact, which is at energies above the masses of the vector bosons mediating proton decay. In general, the strengths of the couplings depend on the mass scale at which they are measured. Consequently, the SU(5) ratios cannot be directly compared with the values measured at low energy. However, the renormalization group equations of field theory prescribe how they change with the mass scale. Specifically, the change of the coupling at a given mass scale depends only on all the elementary particles with masses less than that mass scale. Thus, as the mass scale is lowered below the mass of the proton-decay bosons, the latter must be omitted from the equations, so the ratios of the couplings change from the SU(5) values. If we assume that the only elementary fields contributing to the equations are the low-mass fields known experimentally and if the proton-decay bosons have a mass of 10^{14} GeV/c^2 (see Fig. 3), then the low-energy experimental ratios of the standard model couplings are predicted correctly by the renormalization group equations but the proton lifetime prediction is a little less than the experimental lower bound. However, adding a few more "low-mass" (say, less than 10^{12} GeV/c^2) particles to the equations lengthens the lifetime predictions, which can thereby be pushed well beyond the limit attainable in present-day experiments.

Thus, using the proton-lifetime bound directly and the standard model couplings at low mass scale, we have seen that electroweak-strong unification implies mass scales close to the scale where gravity must be included. Even if it turns out that the electroweak-strong unification is not exactly correct, it has encouraged the extrapolation of present theoretical ideas well beyond the energies available in present accelerators.

Electroweak-strong unified models such as SU(5) achieve only a partial unification. The vector bosons are fully unified in the sense that they and their interactions are determined by the choice of SU(5) as the local symmetry. However, this is only a partial unification. The choice of fermion and scalar multiplets and the choice of symmetry-breaking patterns are left to the discretion of the physicist, who makes his selections based on low-energy phenomenology. Thus, the "unification" in SU(5) (and related local symmetries) is far from complete, except for the vector bosons. (This suggests that theories in which all particles are more closely related to the vector bosons might remove some of the arbitrariness; this will prove to be the case for supergravity.)

In summary, strong experimental evidence for electroweak-strong unification, such as proton decay, would support the study of quantum field theories at energies just below the Planck mass. From the vantage of these theories, the electroweak and strong interactions should be the low-energy limit of the unifying theory, where "low energy" corresponds to the highest energies available at accelerators today! Only future experiments will help decide whether the standard model is a complete low-energy theory, or whether we are repeating the age-old error of omitting some low-energy interactions that are not yet discovered. Nevertheless, the quest for total unification of the laws of Nature is exciting enough that these words of caution are not sufficient to delay the search for theories incorporating gravity.

Toward Unification with Gravity

Let us suppose that the standard model including gravity is the correct set of theories to be unified. On the basis of the previous discussion, we also accept the hypothesis that quantum field theory with local symmetry is the correct theoretical framework for extrapolating physical theory to distances perhaps as small as the Planck length. Quantum field theory assumes a mathematical model of space-time called a manifold. On large scales a manifold can have many different topologies, but at short enough distance scale, a manifold always looks like a flat (Minkowski) space, with space and time infinitely divisible. This might not be the structure of space-time at very small distances, and the manifold model of space-time might fail. Nevertheless, all progress at unifying gravity and the other interactions described here is based on theories in which space-time is assumed to be a manifold.

Einstein's theory of gravity has fascinated physicists by its beauty, elegance, and correct predictions. Before examining efforts to extend the theory to include other interactions, let us review its structure. Gravity is a "geometrical" theory in the following sense. The shape or geometry of the manifold is determined by two types of tensors, called curvature and torsion, which can be constructed from the gravitational field. The Lagrangian of the gravitational field depends on the curvature tensor. In particular, Einstein's brilliant discovery was that the curvature scalar, which is obtained from the curvature tensor, is essentially a unique choice for the kinetic energy of the gravitational field. The gravitational field calculated from the equations of motion then determines the geometry of the space-time manifold. Particles travel along "straight lines" (or geodesics) in this space-time. For example, the orbits of the planets are geodesics of the space-time whose geometry is determined by the sun's gravitational field.

In Einstein's gravity all the remaining fields are called matter fields. The Lagrangian is a sum of two terms:

$$\mathscr{L} = \mathscr{L}_{\text{gravity}} + \mathscr{L}_{\text{matter}} , \qquad (1)$$

where the curvature scalar $\mathscr{L}_{\text{gravity}}$ is the kinetic energy of the graviton, and $\mathscr{L}_{\text{matter}}$ contains all the other fields and their interactions with the gravitational field. The interaction term in the Lagrangian, which couples the gravitational field (the metric tensor) to the energy-momentum tensor, has a form almost identical to the term that couples the electromagnetic field to the electromagnetic current. Newton's constant, which has dimensions of $(\text{mass})^{-2}$, appears in the ratio of the two terms in Eq. 1 as a coupling analogous to the Fermi coupling in the weak theory. This complicates the quantum generalization, just as it did in Fermi's weak interaction theory, and it is not possible to formulate a consistent quantum theory with Eq. 1. Actually, the situation is even worse, because $\mathscr{L}_{\text{gravity}}$ alone does not lead to a consistent quantum theory either, although the inconsistencies are not as bad as when $\mathscr{L}_{\text{matter}}$ is included.

This suggests that our efforts to unify gravity with the other interactions might solve the problems of gravity: perhaps we can join the matter fields together with the gravitational field in something like a curvature scalar and thereby eliminate $\mathscr{L}_{\text{matter}}$. In addition, generalizing the graviton field in this way might lead to a consistent (renormalizable) quantum theory of gravity. There are reasons to hope that the problem of finding a renormalizable theory of gravity is solved by superstrings, although the proof is far from complete. For now, we discuss the unification of the graviton with other fields without concern for renormalizability.

We will discuss several ways to find manifolds for which the curvature scalar depends on many fields, not just the gravitational

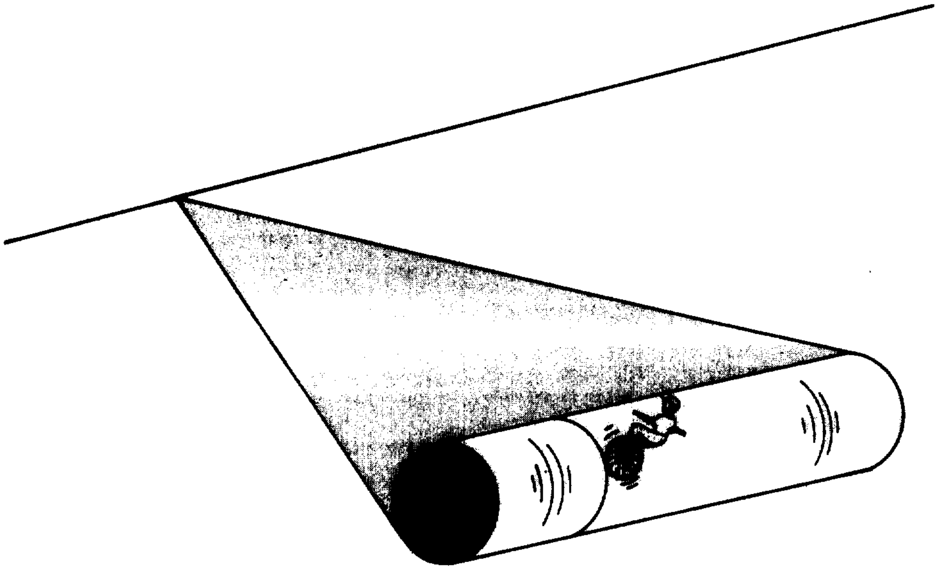


Fig. 4. Two-dimensional analogue of the vacuum geometry of a Kaluza-Klein theory. From great distances the geometry looks one-dimensional, but up close the second dimension, which is wound up in a circle, becomes visible. If space-time has more than four dimensions, then the extra dimensions could have escaped detection if each is wound into a circle whose radius is less than 10^{-16} centimeter.

field. This generally requires extending the 4-dimensional space-time manifold. The fields and manifold must satisfy many constraints before this can be done. All the efforts to unify gravity with the other interactions have been formulated in this way, but progress was not made until the role of spontaneous symmetry breaking was appreciated. As we now describe, it is crucial for the solutions of the theory to have less symmetry than the Lagrangian has.

In the standard model the generators of the space-time Poincaré symmetry commute with (are independent of) the generators of the internal symmetries of the electroweak and strong interactions. We might look for a local symmetry that interrelates the space-time and internal symmetries, just as SU(5) interrelates the electroweak and strong internal symmetries. Unfortunately, if this enlarged symmetry changes simultaneously the internal and space-time quantum numbers of several states of the same mass, then a theorem of quantum field theory requires the existence of an infinite number of particles of that mass. However, this seemingly catastrophic result does not prevent the unification of space-time and internal symmetries for two reasons: first, all symmetries of the Lagrangian need not be symmetries of the states because of spontaneous symmetry breaking; and second, the theorem does not apply to symmetries such as supersymmetry, with its anticommuting generators.

These two loopholes in the assumptions of the theorem have suggested two directions of research in the attempt to unify gravity with the other interactions. First, we might suppose that the dimensionality of space-time is greater than four, and that spontaneous symmetry breaking of the Poincaré invariance of this larger space separates 4-dimensional space-time from the other dimensions. The symmetries of the extra dimensions can then correspond to internal symmetries, and the symmetries of the states in four dimensions need not imply an unsatisfactory infinity of states. A second approach is to extend the Poincaré symmetry to supersymmetry, which then requires additional fermionic fields to accompany the graviton. A combination of these approaches leads to the most interesting theories.

Higher Dimensional Space-Time

If the dimensionality of space-time is greater than four, then the geometry of space-time must satisfy some strong observational constraints. In a 5-dimensional world the fourth spatial direction must be invisible to present experiments. This is possible if at each 4-dimensional space-time point the additional direction is a little circle, so that a tiny person traveling in the new direction would soon return to the starting point. Theories with this kind of vacuum geometry are generically called Kaluza-Klein theories.[1]

It is easy to visualize this geometry with a two-dimensional analogue, namely, a long pipe. The direction around the pipe is analogous to the extra dimension, and the location along the pipe is analogous to a location in 4-dimensional space-time. If the means for examining the structure of the pipe are too coarse to see distance intervals as small as its diameter, then the pipe appears 1-dimensional (Fig. 4). If the probe of the structure is sensitive to shorter distances, the pipe is a 2-dimensional structure with one dimension wound up into a circle.

The physically interesting solutions of Einstein's 4-dimensional gravity are those in which, if all the matter is removed, space-time is flat. The 4-dimensional space-time we see around us is flat to a good approximation; it takes an incredibly massive hunk of high-density (much greater than any density observed on the earth) matter to curve space. However, it might also be possible to construct a higher dimensional theory in which our 4-dimensional space-time remains flat in the absense of identifiable matter, and the extra dimensions are wound up into a "little ball." We must study the generalizations of Einstein's equations to see whether this can happen, and if it does, to find the geometry of the extra dimensions.

The Cosmological Constant Problem. Before we examine the generalizations of gravity in more detail, we must raise a problem that pervades all gravitational theories. Einstein's equations state that the Einstein tensor (which is derived from the curvature scalar in finding the equations of motion from the Lagrangian) is proportional to the energy-momentum tensor. If, in the absence of all matter and radiation, the energy-momentum tensor is zero, then Einstein's equations are solved by flat space-time and zero gravitational field. In 4-dimensional classical general relativity, the curvature of space-time and the gravitational field result from a nonzero energy-momentum tensor due to the presence of physical particles.

However, there are many small effects, such as other interactions and quantum effects, not included in classical general relativity, that can radically alter this simple picture. For example, recall that the electroweak theory is spontaneously broken, which means that the scalar field has a nonzero vacuum value and may contribute to the vacuum value of the energy-momentum tensor. If it does, the solution to the Einstein equations in vacuum is no longer flat space but a curved space in which the curvature increases with increasing vacuum energy. Thus, the constant value of the potential energy, which had no effect on the weak interactions, has a profound effect on gravity.

At first glance, we can solve this difficulty in a trivial manner: simply add a constant to the Lagrangian that cancels the vacuum energy, and the universe is saved. However, we may then wish to compute the quantum-mechanical corrections to the electroweak theory or add some additional fields to the theory; both may readjust the vacuum energy. For example, electroweak-strong unification and its quantum corrections will contribute to the vacuum energy. Almost all the details of the theory must be included in calculating the vacuum energy. So, we could repeatedly readjust the vacuum energy as we learn more about the theory, but it seems artificial to keep doing so unless we have a good theoretical reason. Moreover, the scale of the vacuum energy is set by the mass scale of the interactions. This is a dilemma. For example, the quantum corrections to the electroweak interactions contribute enough vacuum energy to wind up our 4-dimensional space-time into a tiny ball about 10^{-13} centimeter across, whereas the scale of the universe is more like 10^{28} centimeters. Thus, the observed value of the cosmological constant is smaller by a factor of 10^{82} than the value suggested by the standard model. Other contributions can make the theoretical value even larger. This problem has the innocuous-sounding name of "the cosmological constant problem." At present there are no principles from which we can impose a zero or nearly zero vacuum energy on the 4-dimensional part of the theory, although this problem has inspired much research effort. Without such a principle, we can safely say that the vacuum-energy prediction of the standard model is wrong. At best, the theory is not adequate to confront this problem.

If we switch now to the context of gravity theories in higher dimensions, the difficult question is not why the extra dimensions are wound up into a little ball, but why our 4-dimensional space-time is so nearly flat, since it would appear that a large cosmological constant is more natural than a small one. Also, it is remarkable that the vacuum energy winding the extra dimensions into a little ball is conceptually similar to the vacuum charge of a local symmetry providing a mass for the vector bosons. However, in the case of the vacuum geometry, we have no experimental data that bear on these speculations other than the remarkable flatness of our 4-dimensional space-time. The remaining discussion of unification with gravity must be conducted in ignorance of the solution to the cosmological constant problem.

Internal Symmetries from Extra Dimensions

The basic scheme for deriving local symmetries from higher dimensional gravity was pioneered by Kaluza and Klein[1] in the 1920s, before the weak and strong interactions were recognized as fundamental. Their attempts to unify gravity and electrodynamics in four dimensions start with pure gravity in five dimensions. They assumed that the vacuum geometry is flat 4-dimensional space-time with the fifth dimension a little loop of definite radius at each space-time point, just as in the pipe analogy of Fig. 4. The Lagrangian consists of the curvature scalar, constructed from the gravitational field in five dimensions with its five independent components. The relationship of a higher dimensional field to its 4-dimensional fields is summarized in Fig. 5 and the sidebar, "Fields and Spin in Higher Dimensions." The infinite spectrum in four dimensions includes the massless graviton (two helicity components of values ± 2), a massless vector boson (two helicity components of ± 1), a massless scalar field (one helicity component of 0), and an infinite series of massive spin-2 pyrgons of increasing masses. (The term "pyrgon" derives from πύργοσ, the Greek word for tower.) The Fourier expansion for each component of the gravitational field is identical to Eq. 1 of the sidebar. Since the extra dimension is a circle, its symmetry is a phase symmetry just as in electrodynamics.

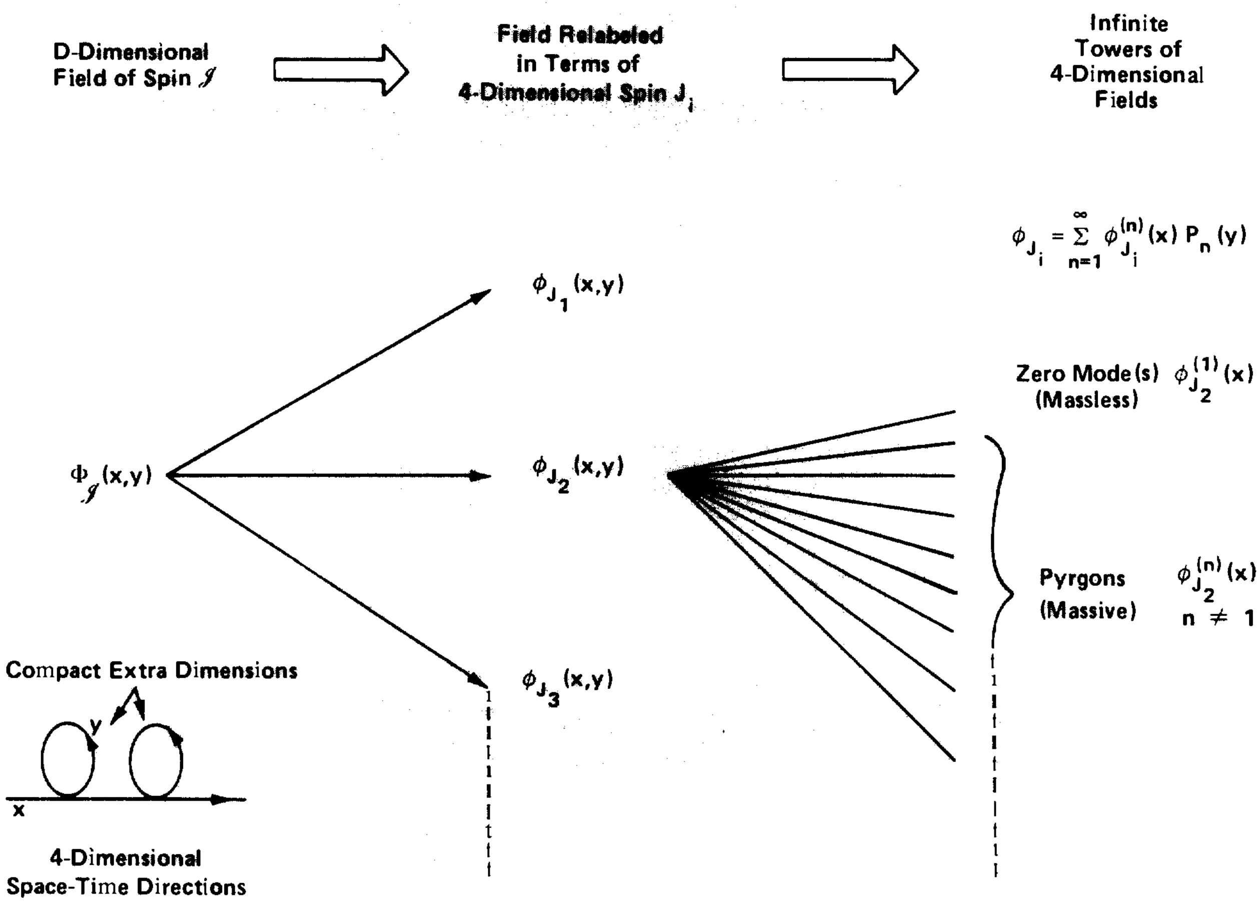


Fig. 5. A field in D dimensions unifies fields of different spins and masses in four dimensions. In step 1 the spin components of a single higher dimensional spin are resolved into several spins in four dimensions. (The total number of components remains constant.) Mathematically this is achieved by finding the spins J_1, J_2, ... in four dimensions that are contained in "spin-$\mathcal{J}$" of D dimensions. Step 2 is the harmonic expansion of the 4-dimensional spin components on the extra dimensions, which then resolves a single massless D-dimensional field into an infinite number of 4-dimensional fields of varying masses. When the 4-dimensional mass is zero, the corresponding 4-dimensional field is called a zero mode. The 4-dimensional fields with 4-dimensional mass form an infinite sequence of pyrgons.

The symmetry of this vacuum state is not the 5-dimensional Poincaré symmetry but the direct product of the 4-dimensional Poincaré group and a phase symmetry.

This skeletal theory should not be taken seriously, except as a basis for generalizing to more realistic theories. The zero modes (massless particles in four dimensions) are electrically neutral. Only the pyrgons carry electric charge. The interaction associated with the vector boson in four dimensions cannot be electrodynamics because there are no low-mass charged particles. (Adding fermions to the 5-dimensional theory does not help, because the resulting 4-dimensional fermions are all pyrgons, which cannot be low mass either.) Nevertheless, the hypothesis that all interactions are conse-

Fields and Spins in

Fields in Higher Dimensions. We describe here how to construct a field in higher dimensions and how such a field is related to fields in the 4-dimensional world in which we live. Higher dimensional fields unify an infinite number of 4-dimensional fields. A typical and simple example of this can be seen from a scalar field (a spin-0 field) in five dimensions. A scalar field has only one component, so it can be written as $\varphi(x,y)$, where x is the 4-dimensional space-time coordinate and y is the coordinate for the fifth dimension. We will assume that the fifth dimension is a little circle with radius R, where R is independent of x. (After this example, we examine the generalizations to more than five dimensions and to fields carrying nonzero spin in the higher dimensions.)

Functions on a circle can be expanded in a Fourier series; thus, the 5-dimensional scalar field can be written in the form

$$\varphi(x,y) = \sum_{n=-\infty}^{\infty} \varphi_n(x)\exp(iny/R)\,, \tag{1}$$

where n is an integer, and $\varphi_n(x)$ are 4-dimensional fields. The Fourier series satisfies the requirement that the field is single-valued in the extra dimension, since Eq. 1 has the same value at the identical points y and $y + 2\pi R$. Usually the wave equation of $\varphi(x,y)$ is a straightforward generalization of the 4-dimensional scalar wave equation (that is, the Klein-Gordon equation) in the limit that interactions can be ignored. The 5-dimensional Klein-Gordon equation for a massless 5-dimensional particle is

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 - \frac{\partial^2}{\partial y^2}\right)\varphi(x,y) = 0\,. \tag{2}$$

The presence of additional terms depends on the details of the Lagrangian, and we ignore them for the present description. It is a simple matter to substitute the Fourier expansion of Eq. 1 into Eq. 2 and use the orthogonality of the expansion functions $\exp(iny/R)$ to rewrite Eq. 2 as an infinite number of equations in four dimensions, one for each $\varphi_n(x)$:

$$\left[\frac{\partial^2}{\partial t^2} - \nabla^2 + \left(\frac{n}{R}\right)^2\right]\varphi_n(x) = 0\,. \tag{3}$$

Note the following very important point: for $n = 0$, Eq. 3 is the equation for a massless 4-dimensional scalar field, whereas for $n \neq 0$, Eq. 3 is the wave equation for a particle with mass $|n|/R$. The massless particle, or "zero mode," should correspond to a field observable in our world. The fields with nonzero mass are called "pyrgons," since they are on a "tower" of particles, one for each n. If R is near the Planck length (10^{-33} centimeter), then the pyrgons have masses on the order of the Planck mass. However, it is also possible that R can be much larger, say as large as 10^{-16} centimeter, without conflicting with experience.

The 4-dimensional form of the Lagrangian depends on an infinite number of fields and is very complicated to analyze. For many purposes it is helpful to truncate the theory, keeping a specially chosen set of fields. For example, 5-dimensional Einstein gravity is simplified by omitting all the pyrgons. This can be achieved by requiring that the fields do not depend on y, a procedure called "dimensional reduction." The dimensionally reduced theory should

quences of the symmetries of space-time is so attractive that efforts to generalize the Kaluza-Klein idea have been vigorously pursued. These theories require a more complete discussion of the possible candidate manifolds of the extra dimensions.

The geometry of the extra dimensions in the absence of matter is typically a space with a high degree of symmetry. Symmetry requires the existence of transformations in which the starting point looks like the point reached after the transformation. (For example, the environments surrounding each point on a sphere are identical.) Two of the most important examples are "group manifolds" and "coset spaces," which we briefly describe.

The tranformations of a continuous group are identified by N parameters, where N is the number of independent transformations in the group. For example, $N = 3$ for SU(2) and 8 for SU(3). These parameters are the coordinates of an N-dimensional manifold. If the vacuum values of fields are constant on the group manifold, then the vacuum solution is said to be symmetric.

Coset spaces have the symmetry of a group too, but the coordinates are labeled by a subset of the parameters of a group. For example, consider the space SO(3)/SO(2). In this example, SO(3) has three parameters, and SO(2) is the phase symmetry with one parameter, so the coset space SO(3)/SO(2) has three minus one, or two, dimensions. This space is called the 2-sphere, and it has the geometry of the surface of an ordinary sphere. Spheres can be generalized to any number of dimensions: the N-dimensional sphere is the coset space $[SO(N+1)]/SO(N)$. Many other cosets, or "ratios" of groups, make spaces with large symmetries. It is possible to find spaces with the symmetries of the electroweak and strong interactions. One such space is the group manifold SU(2) × U(1) × SU(3), which has twelve dimensions. More interesting is the lowest dimensional space with those symmetries, namely, the coset space $[SU(3) \times SU(2) \times U(1)]/[SU(2) \times U(1) \times U(1)]$, which has dimension $8 + 3 + 1 - 3 - 1 - 1 = 7$. (The SU(2) and the U(1)'s in the denominator differ from those in the numerator, so they cannot be "canceled.") Thus, one might hope that $(4 + 7 = 11)$-dimensional gravity would

Higher Dimensions

describe the low-energy limit of the theory.

The gravitational field can be generalized to higher (>5) dimensional manifolds, where the extra dimensions at each 4-dimensional space-time point form a little ball of finite volume. The mathematics requires a generalization of Fourier series to "harmonic" expansions on these spaces. Each field (or field component if it has spin) unifies an infinite set of pyrgons, and the series may also contain some zero modes. The terms in the series correspond to fields of increasing 4-dimensional mass, just as in the 5-dimensional example. The kinetic energy in the extra dimensions of each term in the series then corresponds to a mass in our space-time. The higher dimensional field quite generally describes mathematically an infinite number of 4-dimensional fields.

Spin in Higher Dimensions. The definition of spin in D dimensions depends on the D-dimensional Lorentz symmetry; 4-dimensional Lorentz symmetry is naturally embedded in the D-dimensional symmetry. Consequently a D-dimensional field of a specific spin unifies 4-dimensional fields with different spins.

Conceptually the description of D-dimensional spin is similar to that of spin in four dimensions. A massless particle of spin J in four dimensions has helicities $+J$ and $-J$ corresponding to the projections of spin along the direction of motion. These two helicities are singlet multiplets of the 1-dimensional rotations that leave unchanged the direction of a particle traveling at the speed of light. The group of 1-dimensional rotations is the phase symmetry SO(2), and this method for identifying the physical degrees of freedom is called the "light-cone classification." However, the situation is a little more complicated in five dimensions, where there are three directions orthogonal to the direction of the particle. Then the helicity symmetry becomes SO(3) (instead of SO(2)), and the spin multiplets in five dimensions group together sets of 4-dimensional helicity. For example, the graviton in five dimensions has five components. The SO(2) of four dimensions is contained in this SO(3) symmetry, and the 4-dimensional helicities of the 5-dimensional graviton are 2, 1, 0, −1, and −2.

Quite generally, the light-cone symmetry that leaves the direction of motion of a massless particle unchanged in D dimensions is $SO(D-2)$, and the D-dimensional helicity corresponds to the multiplets (or representations) of $SO(D-2)$. For example, the graviton has $D(D-3)/2$ independent degrees of freedom in D dimensions; thus the graviton in eleven dimensions belongs to a 44-component representation of SO(9). The SO(2) of the 4-dimensional helicity is inside the SO(9), so the forty-four components of the graviton in eleven dimensions carry labels of 4-dimensional helicity as follows: one component of helicity 2, seven of helicity 1, twenty-eight of helicity 0, seven of helicity −1 and one of helicity −2. (The components of the graviton in eleven dimensions then correspond to the graviton, seven massless vector bosons, and twenty-eight scalars in four dimensions.)

The analysis for massive particles in D dimensions proceeds in exactly the same way, except the helicity symmetry is the one that leaves a resting particle at rest. Thus, the massive helicity symmetry is $SO(D-1)$. (For example, SO(3) describes the spin of a massive particle in ordinary 4-dimensional space-time.) These results are summarized in Fig. 5 of the main text.

unify all known interactions.

It turns out that the 4-dimensional fields implied by the 11-dimensional gravitational field resemble the solution to the 5-dimensional Kaluza-Klein case, except that the gravitational field now corresponds to many more 4-dimensional fields. There are methods of dimensional reduction for group manifolds and coset spaces, and the zero modes include a vector boson for each symmetry of the extra dimensions. Thus, in the (4 + 7)-dimensional example mentioned above, there is a complete set of vector bosons for the standard model. At first sight this model appears to provide an attractive unification of all the interactions of the standard model; it explains the origins of the local symmetries of the standard model as space-time symmetries of gravity in eleven dimensions.

Unfortunately, this 11-dimensional Kaluza-Klein theory has some shortcomings. Even with the complete freedom consistent with quantum field theory to add fermions, it cannot account for the parity violation seen in the weak neutral-current interactions of the electron. Witten[1] has presented very general arguments that no 11-dimensional Kaluza-Klein theory will ever give the correct electroweak theory.

Supersymmetry and Gravity in Four Dimensions

We return from our excursion into higher dimensions and discuss extending gravity not by enlarging the space but rather by enlarging the symmetry. The local Poincaré symmetry of Einstein's gravity implies the massless spin-2 graviton; our present goal is to extend the Poincaré symmetry (without increasing the number of dimensions) so that additional fields are grouped together with the graviton. However, this cannot be achieved by an ordinary (Lie group) symmetry: the graviton is the only known elementary spin-2 field, and the local symmetries of the standard model are internal symmetries that group together particles of the same spin. Moreover, gravity has an exceptionally weak interaction, so if the graviton carries quantum numbers of symmetries similar to those of the standard model, it will interact too strongly. We can accommodate these facts if the graviton is a singlet under the internal symmetry, but then its multiplet in this new symmetry must include particles of other spins. Supersymmetry[2] is capable of fulfilling this requirement.

Four-Dimensional Supersymmetry. Supersymmetry is an extension of the Poincaré symmetry, which includes the six Lorentz generators $M_{\mu\nu}$ and four translations P_μ. The Poincaré generators are boson operators, so they can change the spin components of a massive field but not the total spin. The simplest version of supersymmetry adds fermionic generators Q_α to the Poincaré generators; Q_α transforms like a spin-½ field under Lorentz transformations. (The index α is a spinor index.) To satisfy the Pauli exclusion principle, fermionic operators in quantum field theory always satisfy anticommutation relations, and the supersymmetry generators are no exception. In the algebra the supersymmetry generators Q_α anticommute to yield a translation

$$\left\{Q_\alpha, \bar{Q}_\beta\right\} = \gamma^\mu_{\alpha\beta} P_\mu \,, \qquad (2)$$

where P_μ is the energy-momentum 4-vector and the $\gamma^\mu_{\alpha\beta}$ are matrix elements of the Dirac γ matrices.

The significance of the fermionic generators is that they change the spin of a state or field by $\pm$½; that is, supersymmetry unifies bosons and fermions. A multiplet of "simple" supersymmetry (a supersymmetry with one fermionic generator) in four dimensions is a pair of particles with spins J and $J -$ ½; the supersymmetry generators transform bosonic fields into fermionic fields and vice versa. The boson and fermion components are equal in number in all supersymmetry multiplets relevant to particle theories.

We can construct larger supersymmetries by adding more fermionic generators to the Poincaré symmetry. "N-extended" supersymmetry has N fermionic generators. By applying each generator to the state of spin J, we can lower the helicity up to N times. Thus, simple supersymmetry, which lowers the helicity just once, is called $N = 1$ supersymmetry. $N = 2$ supersymmetry can lower the helicity twice, and the $N = 2$ multiplets have spins J, $J -$ ½, and $J - 1$. There are twice as many $J -$ ½ states as J or $J - 1$, so that there are equal numbers of fermionic and bosonic states. The $N = 2$ multiplet is made up of two $N = 1$ multiplets: one with spins J and $J -$ ½ and the other with spins $J -$ ½ and $J - 1$.

In principle, this construction can be extended to any N, but in quantum field theory there appears to be a limit. There are serious difficulties in constructing simple field theories with spin 5/2 or higher. The largest extension with spin 2 or less has $N = 8$. In $N = 8$ extended supersymmetry, there is one state with helicity of 2, eight with 3/2, twenty-eight with 1, fifty-six with 1/2, seventy with 0, fifty-six with $-1/2$, twenty-eight with -1, eight with 3/2 and one with -2. This multiplet with 256 states will play an important role in the supersymmetric theories of gravity or supergravity discussed below. Table 2 shows the states of N-extended supersymmetry.

Theories with Supersymmetry. Rather ordinary-looking Lagrangians can have supersymmetry. For example, there is a Lagrangian with simple global supersymmetry in four dimensions with a single Majorana fermion, which has one component with helicity $+1/2$, one with helicity $-1/2$, and two spinless particles. Thus, there are two bosonic and two fermionic degrees of freedom. The supersymmetry not only requires the presence of both fermions and bosons in the Lagrangian but also restricts the types of interactions, requires that the mass parameters in the multiplet be equal, and relates some other parameters in the Lagrangian that would otherwise be unconstrained.

The model just described, the Wess-Zumino model,[3] is so simple that it can be written down easily in conventional field notation. However, more realistic supersymmetric Lagrangians take pages to write down. We will avoid this enormous complication and limit our discussion to the spectra of particles in the various theories.

Although supersymmetry may be an exact symmetry of the Lagrangian, it does not appear to be a symmetry of the world because the known elementary particles do not have supersymmetric partners. (The photon and a neutrino cannot form a supermultiplet because their low-energy interactions are different.) However, like ordinary symmetries, the supersymmetries of the Lagrangian do not have to be supersymmetries of the vacuum: supersymmetry can be spontaneously broken. The low-energy predictions of spontaneously broken supersymmetric models are discussed in "Supersymmetry at 100 GeV."

Local Supersymmetry and Supergravity. There is a curious gap in the spectrum of the spin values of the known elementary particles. Almost all spins less than or equal to 2 have significant roles in particle theory: spin-1 vector bosons are related to the local internal symmetries; the spin-2 graviton mediates the gravitational interaction; low-mass spin-½ fermions dominate low-energy phenomenology; and spinless fields provide the mechanism for spontaneous symmetry breaking. All these fields are crucial to the standard model, although there seems to be no relation among the fields of different spin. A spin of 3/2 is not required phenomenologically and is missing from the list. If the supersymmetry is made local, the resulting theory is supergravity, and the spin-2 graviton is accompanied by a "gravitino" with spin 3/2.

Local supersymmetry can be imposed on a theory in a fashion formally similar to the local symmetries of the standard model, except for the additional complications due to the fact that supersymmetry is a space-time symmetry. Extra gauge fields are required to compensate for derivatives of the space-time-dependent parameters, so, just as for ordinary symmetries, there is a gauge particle corresponding to each independent super-

Table 2

The fields of N-extended supergravity in four dimensions. Shown are the number of states of each helicity for each possible supermultiplet containing a graviton but with spin ≤ 2. Simple supergravity ($N = 1$) has a graviton and gravitino. $N = 4$ supergravity is the simplest theory with spinless particles. The overlap of the multiplets with the largest (+2) and smallest (−2) helicities gives rise to large additional symmetries in supergravity. $N = 7$ and $N = 8$ supergravities have the same list of helicities because particle-antiparticle symmetry implies that the $N = 7$ theory must have two multiplets (as for $N <$ 7), whereas $N = 8$ is the first and last case for which particle-antiparticle symmetry can be satisfied by a single multiplet.

	N						
Helicity	**1**	**2**	**3**	**4**	**5**	**6**	**7 or 8**
2	1	1	1	1	1	1	1
3/2	1	2	3	4	5	6	8
1		1	3	6	10	16	28
1/2			1	4	11	26	56
0				2	10	30	70
−1/2			1	4	11	26	56
−1		1	3	6	10	16	28
−3/2	1	2	3	4	5	6	8
−2	1	1	1	1	1	1	1
Total	4	8	16	32	64	128	256

symmetry transformation. However, the gauge particles associated with the supersymmetry generators must be fermions. Just as the graviton has spin 2 and is associated with the local translational symmetry, the gravitino has spin 3/2 and gauges the local supersymmetry. The graviton and gravitino form a simple ($N = 1$) supersymmetry multiplet. This theory is called simple supergravity and is interesting because it succeeds in unifying the graviton with another field.

The Lagrangian of simple supergravity[4] is an extension of Einstein's Lagrangian, and one recovers Einstein's theory when the gravitational interactions of the gravitino are ignored. This model must be generalized to a more realistic theory with vector bosons, spin-½ fermions, and spinless fields to be of much use in particle theory.

The generalization is to Lagrangians with extended local supersymmetry, where the largest spin is 2. The extension is extremely complicated. Nevertheless, without much work we can surmise some features of the extended theory. Table 2 shows the spectrum of particles in N-extended supergravity.

We start here with the largest extended supersymmetry and investigate whether it includes the electroweak and strong interactions. In $N = 8$ extended supergravity the spectrum is just the $N = 8$ supersymmetric multiplet of 256 helicity states discussed before. The massless particles formed from these states include one graviton, eight gravitinos, twenty-eight vector bosons, fifty-six fermions, and seventy spinless fields.

$N = 8$ supergravity[5] is an intriguing theory. (Actually, several different $N = 8$ supergravity Lagrangians can be constructed.) It has a remarkable set of internal symmetries, and the choice of theory depends on which of these symmetries have gauge particles associated with them. Nevertheless, supergravity theories are highly constrained and we can look for the standard model in each. We single out one of the most promising versions of the theory, describe its spectrum, and then indicate how close it comes to unifying the electroweak, strong, and gravitational interactions.

In the $N = 8$ supergravity of de Wit-Nicolai theory[6] the twenty-eight vector bosons gauge an SO(8) symmetry found by Cremmer and Julia.[5] Since the standard model needs just twelve vector bosons, twenty-eight would appear to be plenty. In the fermion sector, the eight gravitinos must have fairly large masses in order to have escaped detection. Thus, the local supersymmetry must be broken, and the gravitinos acquire masses by absorbing eight spin-½ fermions. This leaves 56 − 8 = 48 spin-½ fermion fields. For the quarks and leptons in the standard model, we need forty-five fields, so this number also is sufficient.

The next question is whether the quantum numbers of SO(8) correspond to the electroweak and strong quantum numbers and the spin-½ fermions to quarks and leptons. This is where the problems start: if we separate an SU(3) out of the SO(8) for QCD, then the only other independent interactions are two local phase symmetries of U(1) × U(1), which is not large enough to include the SU(2) × U(1) of the electroweak theory. The rest of the SO(8) currents mix the SU(3) and the two U(1)'s. Moreover, many of the fifty-six spin-½ fermion states (or forty-eight if the gravitinos are massive) have the wrong SU(3) quantum numbers to be quarks and leptons.[7] Finally, even if the quantum numbers for QCD were right and the electroweak local symmetry were present, the weak interactions could still not be accounted for. No mechanism in this theory can guarantee the almost purely axial weak neutral current of the electron. Thus this interpretation of $N = 8$ supergravity cannot be the ultimate theory. Nevertheless, this is a model of unification, although it gave the wrong sets of interactions and particles.

Perhaps the 256 fields do not correspond directly to the observable particles, but we need a more sophisticated analysis to find them. For example, there is a "hidden" local SU(8) symmetry, independent of the gauged SO(8) mentioned above, that could easily contain the electroweak and strong interactions. It is hidden in the sense that the Lagrangian does not contain the kinetic energy terms for the sixty-three vector bosons of SU(8). These sixty-three vector bosons are composites of the elementary supergravity fields, and it is possible that the quantum corrections will generate kinetic energy terms. Then the fields in the Lagrangian do not correspond to physical particles; instead the photon, electron, quarks, and so on, which look elementary on a distance scale of present experiments, are composite. Unfortunately, it has not been possible to work out a logical derivation of this kind of result for $N = 8$ supergravity.[8]

In summary, $N = 8$ supergravity may be correct, but we cannot see how the standard model follows from the Lagrangian. The basic fields seem rich enough in structure to account for the known interactions, but in detail they do not look exactly like the real world. Whether $N = 8$ supergravity is the wrong theory, or is the correct theory and we simply do not know how to interpret it, is not yet known.

Supergravity in Eleven Dimensions

The apparent phenomenological shortcomings of $N = 8$ supergravity have been known for some time, but its basic mathematical structure is so appealing that many theorists continue to work on it in hope that some variant will give the electroweak and strong interactions. One particularly interesting development is the generalization of $N = 8$ supergravity in four dimensions to simple ($N = 1$) supergravity in eleven dimensions.[9] This generalization combines the ideas of Kaluza-Klein theories with supersymmetry.

The formulation and dimensional reduction of simple supergravity in eleven dimensions requires most of the ideas already described. First we find the fields of 11-dimensional supergravity that correspond to the graviton and gravitino fields in four dimensions. Then we describe the components of each of the 11-dimensional fields. Finally, we use the harmonic expansion on the extra seven dimensions to identify the zero modes and pyrgons. For a certain geometry of the extra dimensions, the dimensionally reduced, 11-dimensional supergravity without pyrgons is $N = 8$ supergravity in four dimensions; for other geometries we find new theories. We now look at each of these steps in more detail.

In constructing the 11-dimensional fields, we begin by recalling that the helicity symmetry of a massless particle is SO(9) and the spin components are classified by the multiplets of SO(9). The multiplets of SO(9) are either fermionic or bosonic, which means that all the four-dimensional helicities are either integers (bosonic) or half-odd integers (fermionic) for all the components in a single multiplet. The generators independent of the SO(2) form an SO(7), which is the Lorentz group for the extra seven dimensions. Thus, the SO(9) multiplets can be expressed in terms of a sum of multiplets of SO(7) × SO(2), which makes it possible to reduce 11-dimensional spin to 4-dimensional spin.

The fields of 11-dimensional, $N = 1$ supergravity must contain the graviton and gravitino in four dimensions. We have already mentioned in the sidebar that the graviton in eleven dimensions has forty-four bosonic components. The smallest SO(9) multiplet of 11-dimensional spin that yields a helicity of 3/2 in four dimensions for the gravitinos has 128 components, eight components with helicity 3/2, fifty-six with 1/2, fifty-six with $-1/2$, and eight with $-3/2$. Since by supersymmetry the number of fermionic states is equal to the number of bosonic states, eighty-four bosonic components remain. It turns out that there is a single 11-dimensional spin with eighty-four components, and it is just the field needed to complete the $N = 1$ supersymmetry multiplet in eleven dimensions.

Thus, we have recovered the 256 components of $N = 8$ supergravity in terms of just three fields in eleven dimensions (see Table 3). The Lagrangian is much simpler in eleven dimensions than it is in four dimensions. The three fields are related to one another by supersymmetry transformations that are very similar to the simple supersymmetry transformations in four dimensions. Thus, in many ways the 11-dimensional theory is no more complicated than simple supergravity in four dimensions.

The dimensional reduction of the 11-dimensional supergravity, where the extra dimensions are a 7-torus, gives one version of $N = 8$ supergravity in four dimensions.[5] In this case each of the components is expanded in a sevenfold Fourier series, one series for each dimension just as in Eq. 1 in the sidebar, except that ny is replaced by $\Sigma n_i y_i$. The dimensional reduction consists of keeping only those fields that do not depend on any y_i, that is, just the 4-dimensional fields corresponding to $n_1 = n_2 = \ldots = n_7 = 0$. Thus, there is one zero mode (massless field in four dimensions) for each component. The pyrgons are the 4-dimensional fields with any $n_i \neq 0$, and these are omitted in the dimensional reduction.

The 11-dimensional theory has a simple Lagrangian, whereas the 4-dimensional, $N = 8$ Lagrangian takes pages to write down. In fact the $N = 8$ Lagrangian was first derived in this way.[5] It is easy to be impressed by a formalism in which everything looks simple. This is the first of several reasons to take seriously the proposal that the extra dimensions might be physical, not just a mathematical trick.

The seven extra dimensions of the 11-dimensional theory must be wound up into a little ball in order to escape detection. The

Table 3

The relation of simple ($N = 1$) supergravity in eleven dimensions and $N = 8$ supergravity in four dimensions. The 256 components of the massless fields of 11-dimensional, $N = 1$ supergravity fall into three n-member multiplets of SO(9). The members of these multiplets have definite helicities in four dimensions. The count of helicity states is given in terms of the size of SO(7) multiplets, where SO(7) is the Lorentz symmetry of the seven extra dimensions in the 11-dimensional theory.

	4-Dimensional Helicity								
n	**2**	**3/2**	**1**	**1/2**	**0**	**−1/2**	**−1**	**−3/2**	**−2**
44	1		7		1+27		7		1
84			21		7+35		21		
128		8		8+48		8+48		8	
Total	1	8	28	56	70	56	28	8	1

case described above assumes that the little ball is a 7-torus, which is the group manifold made of the product of seven phase symmetries. As a Kaluza-Klein theory, the seven vector bosons in the graviton (Table 3) gauge these seven symmetries. Since the twenty-eight vector bosons of $N = 8$ supergravity can be the gauge fields for a local SO(8), it is interesting to see if we can redo the dimensional reduction so that 11-dimensional supergravity is a Kaluza-Klein theory for SO(8), the de Wit-Nicolai theory. Indeed, this is possible. If the extra dimensions are assumed to be the 7-sphere, which is the coset space SO(8)/SO(7), the vector bosons do gauge SO(8).[10] This is, perhaps, the ultimate Kaluza-Klein theory, although it does not contain the standard model. The main difference between the 7-torus and coset spaces is that for coset spaces there is not necessarily a one-to-one correspondence between components and zero modes. Some components may have several zero modes, while others have none (recall Fig. 5).

There are other manifolds that solve the 11-dimensional supergravity equations, although we do not describe them here. The internal local symmetries are just those of the extra dimensions, and the fermions and bosons are unified by supersymmetry. Thus, 11-dimensional supergravity can be dimensionally reduced to one of several different 4-dimensional supergravity theories, and we can search through these theories for one that contains the standard model. Unfortunately, they all suffer phenomenological shortcomings.

Eleven-dimensional supergravity contains an additional error. In the solution where the seven extra dimensions are wound up in a little ball, our 4-dimensional world gets just as compacted: the cosmological constant is about 120 orders of magnitude larger than is observed experimentally.[11] This is the cosmological constant problem at its worst. Its solution may be a major breakthrough in the search for unification with gravity. Meanwhile, it would appear that supergravity has given us the worst prediction in the history of modern physics!

Superstrings

In view of its shortcomings, supergravity is apparently not the unified theory of all elementary particle interactions. In many ways it is close to solving the problem, but a theory that is correct in all respects has not been found. The weak interactions are not exactly right nor is the list of spin-½ fermions. There seems to be no good reason that the cosmological constant should be nearly or exactly zero as observed experimentally. The issue of the renormalizability of the quantum theory of gravity also remains unsolved. Supergravity improves the quantum structure of the theory in that the unwanted infinities are not as bad as in Einstein's theory with matter, but troubles still appear. Newton's constant is a fundamental parameter in the theory, and 4-fermion terms similar to those in Fermi's weak interaction theory are still present. In $N = 8$ supergravity, which is the best case, the perturbation solution to the quantum field theory is expected to break down eventually.

In spite of these difficulties we have reasons to be optimistic that supergravity is on the right track. It does unify gravity with some interactions and is almost a consistent quantum field theory. The line of generalization followed so far has led to theories that are enormous improvements, in a mathematical sense, over Einstein's gravity. It would seem reasonable to look for generalizations beyond supergravity.

Superstring theories may answer some of these questions. Just as the progress of supergravity was based on the systematic addition of fields to Einstein's gravity, superstring theory can also be viewed in terms of the systematic addition of fields to supergravity. Although the formulation of superstring theory looks quite different from the formulation of supergravity, this may be partially due to its historical origin.

Superstring theories were born from an early effort to find a theory of the strong interactions. They began as a very efficient means of understanding the long list of hadronic resonances. In particular, hadrons of high spin have been identified experimentally. It is interesting that sets of hadrons of different spins but the same internal quantum numbers can be grouped together into

"Regge trajectories." Figure 6 shows examples of Regge trajectories (plots of spin versus mass-squared) for the first few states of the Δ and N resonances; these resonances for hadrons of different spins fall along nearly straight lines. Such sequences appear to be general phenomena, and so, in the '60s and early '70s, a great effort was made to incorporate these results directly into a theory. The basic idea was to build a set of hadron amplitudes with rising Regge trajectories that satisfied several important constraints of quantum field theory, such as Lorentz invariance, crossing symmetry, the correct analytic properties, and factorization of resonance-pole residues.[12] Although the theory was a prescription for calculating the amplitudes, these constraints are true of quantum field theory and are necessary for the theory to make sense.

The constraints of field theory proved to be too much for this theory of hadrons. Something always went wrong. Some theories predicted particles with imaginary mass (tachyons) or particles produced with negative probability (ghosts), which could not be interpreted. Several theories had no logical difficulties, but they did not look like hadron theories. First of all, the consistency requirements forced them to be in ten dimensions rather than four. Moreover, they predicted massless particles with a spin of 2; no hadrons of this sort exist. These original superstring theories did not succeed in describing hadrons in any detail, but the solution of QCD may still be similar to one of them.

In 1974 Scherk and Schwarz[13] noted that the quantum amplitudes for the scattering of the massless spin-2 states in the superstring are the same as graviton-graviton scattering in the simplest approximation of Einstein's theory. They then boldly proposed throwing out the hadronic interpretation of the superstring and reinterpreting it as a fundamental theory of elementary particle interactions. It was easily found that superstrings are closely related to supergravity, since the states fall into supersymmetry multiplets and massless spin-2 particles are required.[14]

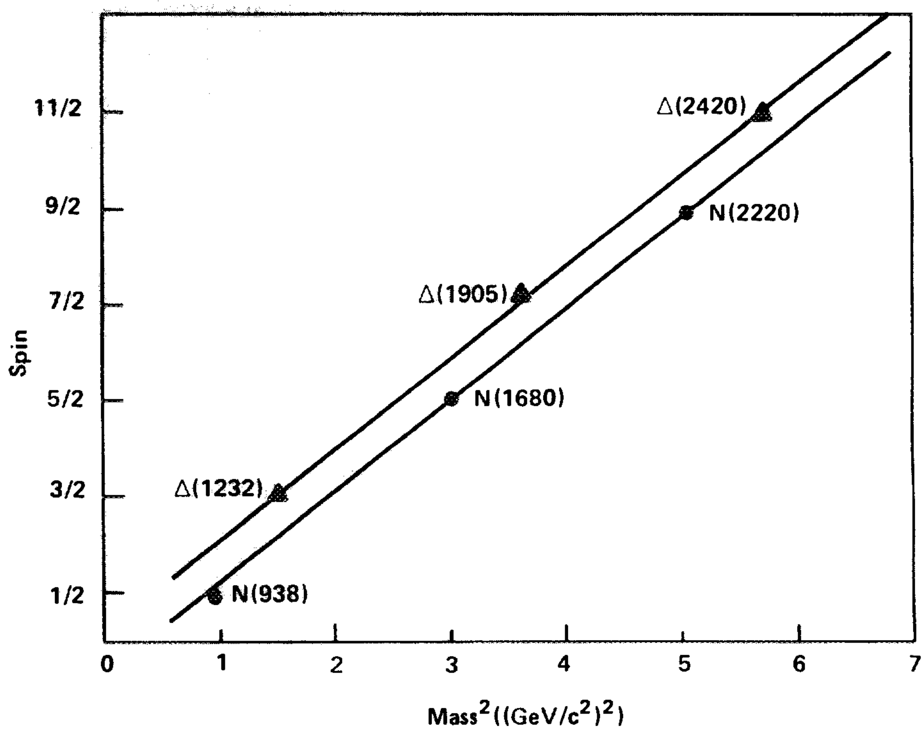


Fig. 6. Regge trajectories in hadron physics. The neutron and proton (N(938)) lie on a linearly rising Regge trajectory with other isospin-½ states: the N(1680) of spin 5/2, the N(2220) of spin 9/2, and so on. This fact can be interpreted as meaning that the N(1680), for example, looks like a nucleon except that the quarks are in an F wave rather than a P wave. Similarly the isospin-3/2 Δ resonance at 1232 MeV lies on a trajectory with other isospin-3/2 states of spins 7/2, 11/2, 15/2, and so on. The slope of the hadronic Regge trajectories is approximately (1 GeV/c^2)$^{-2}$. The slope of the superstring trajectories must be much smaller

The theoretical development of superstrings is not yet complete, and it is not possible to determine whether they will finally yield the truly unified theory of all interactions. They are the subject of intense research today. Our plan here is to present a qualitative description of superstrings and then to discuss the types and particle spectra of superstring theories.

Recent formulations of superstring theories are generalizations of quantum field theory.[15] The fields of an ordinary field theory, such as supergravity, depend on the space-time point at which the field is evaluated. The fields of superstring theory depend on paths in space-time. At each moment in time, the string traces out a path in space, and as time advances, the string propagates through space forming a surface called the "world sheet." Strings can be closed, like a rubber band, or open, like a broken rubber band. Theories of both types

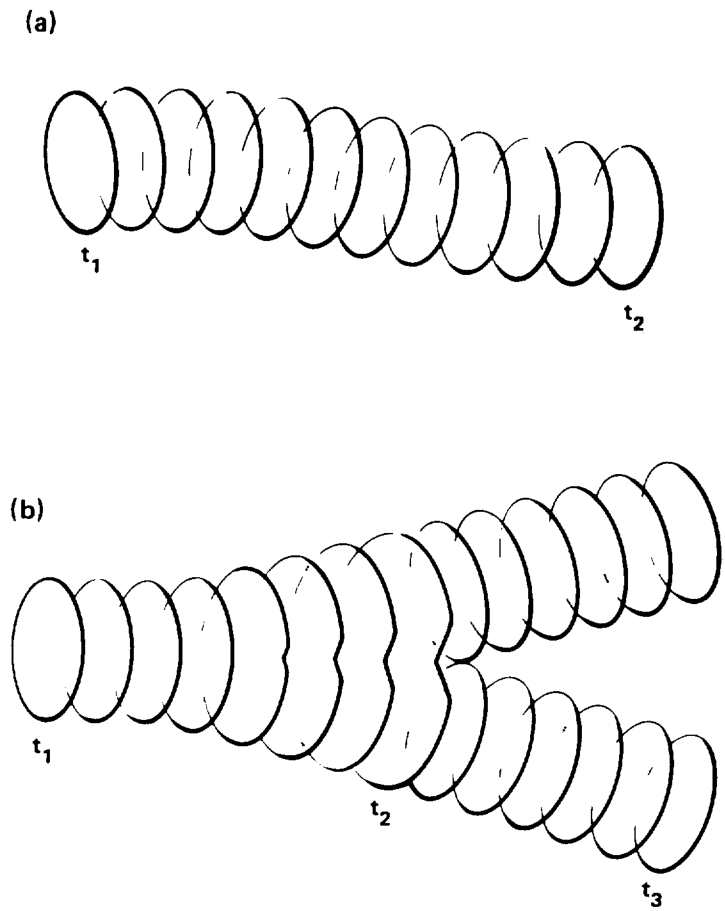


Fig. 7. Dynamics of closed strings. The figures show the string configurations at a sequence of times (in two dimensions instead of ten). In Fig. 7(a) a string in motion from times t_1 ***to*** t_2 ***traces out a world sheet. Figure 7(b) shows the three closed string interaction, where one string at*** t_1 ***undergoes a change of shape until it pinches off at a point at time*** t_2 ***(the interaction time). At time*** t_3 ***two strings are propagating away from the interaction region.***

are promising, but the graviton is always associated with closed strings.

Before analyzing the motion of a superstring, we must return to a discussion of space-time. Previously, we described extensions of space-time to more than four dimensions. In all those cases coordinates were numbers that satisfied the rules of ordinary arithmetic. Yet another extension of space-time, which is useful in supergravity and crucial in superstring theory, is the addition to space-time of "supercoordinates" that do not satisfy the rules of ordinary arithmetic. Instead, two supercoordinates θ_α and θ_β satisfy anticommutation relations $\theta_\alpha\theta_\beta + \theta_\beta\theta_\alpha = 0$, and consequently $\theta_\alpha\theta_\alpha$ (with no sum on α) $= 0$. Spaces with this kind of additional coordinate are called superspaces.[16]

At first encounter superspaces may appear to be somewhat silly constructions. Nevertheless, much of the apparatus of differential geometry of manifolds can be extended to superspaces, so applications in physics may exist. It is possible to define fields that depend on the coordinates of a superspace. Rather naturally, such fields are called superfields.

Let us apply this idea to supergravity, which is a field theory of both fermionic and bosonic fields. The supergravity fields can be further unified if they are written as a smaller number of superfields. Supergravity Lagrangians can then be written in terms of superfields; the earlier formulations are recovered by expanding the superfields in a power series in the supercoordinates. The anticommutation rule $\theta_\alpha\theta_\alpha = 0$ leads to a finite number of ordinary fields in this expansion.

The motion of a superstring is described by the motion of each space-time coordinate and supercoordinate along the string; thus the motion of the string traces out a "world sheet" in superspace. The full theory describes the motions and interactions of superstrings. In particular, Fig. 7 shows the basic form of the three closed superstring interactions. All other interactions of closed strings can be built up out of this one kind of interaction.[15] Needless to say, the existence of only one kind of fundamental interaction would severely restrict theories with only closed strings.

There is a direct connection between the quantum-mechanical states of the string and the elementary particle fields of the theory. The string, whether it is closed or open, is under tension. Whatever its source, this tension, rather than Newton's constant, defines the basic energy scale of the theory. To first approximation each point on the string has a force on it depending on this tension and the relative displacement between it and neighboring points on the string. The prob-

Table 4

Ground states of Type II superstrings. The 10-dimensional fields are listed according to the multiplets of the SO(8) light-cone symmetry. The 4-dimensional fields are listed in terms of helicity and multiplets of the SO(6) Lorentz group of the extra six dimensions.

	Helicity								
	2	3/2	1	1/2	0	−1/2	−1	−3/2	−2
Type IIA: Bosons									
1					1				
28			6		1 + 15		6		
35_v	1		6		1 + 20′		6		1
8_v			1		6		1		
56_v			15		$6 + 10 + \overline{10}$		15		
Type IIA: Fermions									
8_s				4		$\bar{4}$			
8_s				$\bar{4}$		4			
56_s		$\bar{4}$		4 + 20		$\bar{4} + \overline{20}$		4	
56_c		4		$\bar{4} + \overline{20}$		4 + 20		$\bar{4}$	
Type IIB: Bosons									
1 (twice)					1				
28 (twice)			6		1 + 15		6		
35_v	1		6		1 + 20′		6		1
35_c			$\overline{10}$		15		10		
Type IIB: Fermions									
8_s (twice)				4		$\bar{4}$			
56_s (twice)		$\bar{4}$		4 + 20		$\bar{4} + \overline{20}$		4	

lem of unravelling this infinite number of harmonic oscillators is one of the most famous problems of physics. The amplitudes of the Fourier expansion of the string displacement decouple the infinite set of harmonic oscillators into independent Fourier modes. These Fourier modes then correspond to the elementary-particle fields. The quantum-mechanical ground state of this infinite set of oscillators corresponds to the fields of 10-dimensional supergravity. Ten space-time dimensions are necessary to avoid tachyons and ghosts. The excited modes of the superstring then correspond to the new fields being added to supergravity.

The harmonic oscillator in three dimensions can provide insight into the qualitative features of the superstring. The maximum value of the spin of a state of the harmonic oscillator increases with the level of the excitation. Moreover, the energy necessary to reach a given level increases as the spring constant is increased. The superstring is similar. The higher the excitation of the string, the higher are the possible spin values (now in ten dimensions). The larger the string tension, the more massive are the states of an excited level.

The consistency requirements restrict superstring theories to two types. Type I theories have 10-dimensional $\mathcal{N} = 1$ supersymmetry and include both closed and open strings and five kinds of string interactions. Nothing more will be said here about Type I theories, although they are extremely interesting (see Refs. 14 and 15).

Type II theories have $\mathcal{N} = 2$ supersymmetry in ten dimensions and accommodate closed strings only. There are two $\mathcal{N} = 2$ supersymmetry multiplets in ten dimensions, and each corresponds to a Type II superstring theory. We will now describe these two superstring theories.

The Type IIA ground-state spectrum is the one that can be derived by dimensional reduction of simple supergravity in eleven dimensions to $\mathcal{N} = 2$ supergravity in ten dimensions. Thus, if we continue to reduce from ten to four dimensions with the hypothesis that the extra six dimensions form a 6-torus, we will obtain $\mathcal{N} = 8$ supergravity in four dimensions. The superstring theory adds both pyrgons and Regge recurrences to the 256 $\mathcal{N} = 8$ supergravity fields, but it has been possible (and often simpler) to investigate several aspects of supergravity directly from the superstring theory.

The classification of the excited 10-dimensional string states (or elementary fields of the theory) is complicated by the description of spin in ten dimensions. However, the analysis does not differ conceptually from the analysis of spin for 11-dimensional supergravity. The massless states, which form the ground state of the superstring, are classified by multiplets of SO(8), and the excitations of the string are massive fields in ten dimensions that belong to multiplets of SO(9). The ground-state fields of the Type IIA superstring are found in Table 4.

The Type IIB ground-state fields cannot be derived from 11-dimensional supergravity. Instead the theory has a useful phase symmetry in ten dimensions. The fields listed as occurring twice in Table 4 carry nonzero values of the quantum number associated with U(1). So far, the main application of the U(1) symmetry has been the

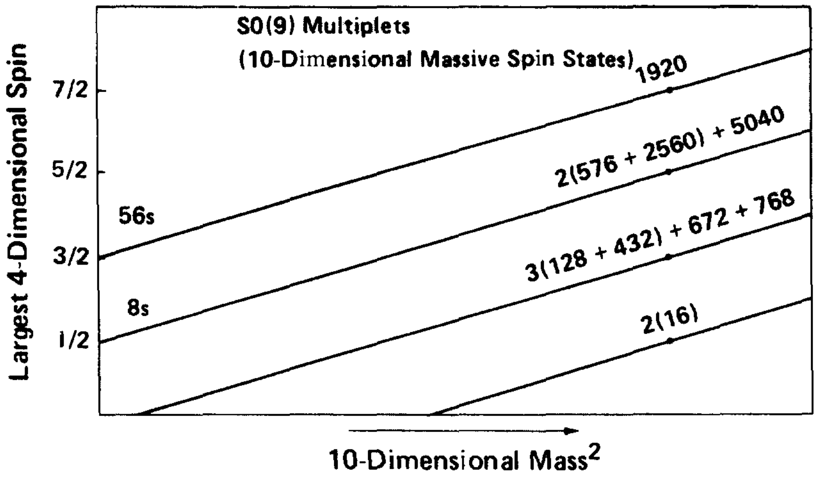


Fig. 8. The ground state and first Regge recurrence of fermionic states in the 10-dimensional Type IIB superstring theory. There are a total of 256 fermionic and bosonic states in the ground state. (The 56_s contains the gravitino.) The first excited states contain 65,536 component fields. Half of these are fermions. (Each representation of the fermions shown above appears twice.)

derivation of the equations of motion for the ground-state fields.[17] It will certainly have a crucial role in the future understanding of Type IIB superstrings.

The quantum-mechanical excitations of the superstring correspond to the Regge recurrences, which are massive in ten dimensions; they belong to multiplets of SO(9). Thus, it is possible to fill in a diagram similar to Fig. 6, although the huge number of states makes the results look complicated. We give a few results to illustrate the method.

The sets of Regge recurrences in Type IIA and IIB are identical. In Figure 8 we show the first recurrence of the fermion trajectories. (Note that only one-half of the 32,768 fermionic states of this mode are shown. The boson states are even messier.) The first excited level has a total of 65,536 states, and the next two excited levels have 5,308,416 and 235,929,600 states, respectively, counting both fermions and bosons. (Particle physicists seem to show little embarrassment these days over adding a few fields to a theory!)

The component fields in ten dimensions can now be expanded into 4-dimensional fields as was done in supergravity. Besides the zero modes and pyrgons associated with the ground states, there will be infinite ladders of pyrgon fields associated with each of the fields of the excited levels of the superstring.

The zero modes in four dimensions have been investigated only for the 6-torus; in this case all the zero modes come from the ground states. There is one zero mode for each component field, since the dimensional reduction is done as a 6-dimensional Fourier series on the 6-torus. The answers for other geometries are not yet known. It may be that many more fields become zero modes (or have nearly zero mass) in four dimensions when the dimensional reduction is studied for other spaces. An important problem is the analysis of superstrings on curved spaces, which has not yet been definitively studied.

Although not much progress has been made toward understanding the phenomenology of these superstring theories, there has been some formal progress. The theory described here may be a quantum theory of gravity. (It may take all those new fields to obtain a renormalizable theory.) Although local symmetries can be ruined by anomalies, Type II (and several Type I) superstrings satisfy the constraints. Also, the one-loop calculation is finite; there are no candidates for counter terms, so the theory may be finite. Of course, this promising result needs support from higher order calculations.

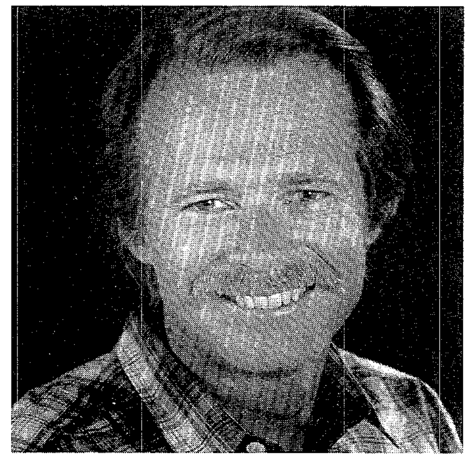
These results give some encouragement that superstrings may solve some long-standing problems in particle theory; whether they will lead to the ultimate unification of all interactions remains to be seen.

Postscript

The search for a unified theory may be likened to an old geography problem. Columbus sailed westward to reach India believing the world had no edge. By analogy, we are searching for a unified theory at shorter and shorter distance scales believing the microworld has no edge. Perhaps we are wrong and space-time is not continuous. Or perhaps we are only partly wrong, like Columbus, and will discover something new, but something consistent with what we already know. Then again, we may finally be right on course to a theory that unifies all Nature's interactions. ■

AUTHORS

Richard C. Slansky has a broad background in physics with more than a taste of metaphysics. He received a B.A. in physics from Harvard in 1962 and then spent the following year as a Rockefeller Fellow at Harvard Divinity School. Dick then attended the University of California, Berkeley, where he received his Ph.D. in physics in 1967. A two-year postdoctoral stint at the California Institute of Technology was followed by five years as Instructor and Assistant Professor at Yale University (1969-1974). Dick joined the Laboratory in 1974 as a Staff Member in the Elementary Particles and Field Theory group of the Theoretical Division, where his interests encompass phenomenology, high-energy physics, and the early universe.



References

1. For a modern description of Kaluza-Klein theories, see Edward Witten, *Nuclear Physics* B186(1981):412 and A. Salam and J. Strathdee, *Annals of Physics* 141(1982):316.

2. Two-dimensional supersymmetry was discovered in dual-resonance models by P. Ramond, *Physical Review D* 3(1971):2415 and by A. Neveu and J. H. Schwarz, *Nuclear Physics* B31(1971):86. Its four-dimensional form was discovered by Yu. A. Gol'fand and E. P. Likhtman, *Journal of Experimental and Theoretical Physics Letters* 13(1971):323.

3. J. Wess and B. Zumino, *Physical Letters* 49B(1974):52 and *Nuclear Physics* B70(1974):39.

4. Daniel Z. Freedman, P. van Nieuwenhuizen, and S. Farrara, *Physical Review D* 13(1976):3214; S. Deser and B. Zumino, *Physics Letters* 62B(1976):335; Daniel Z. Freedman and P. van Nieuwenhuizen, *Physical Review D* 14(1976):912.

5. E. Cremmer and B. Julia, *Physics Letters* 80B(1982):48 and *Nuclear Physics* B159(1979):141.

6. B. de Wit and H. Nicolai, *Physics Letters* 108B(1982):285 and *Nuclear Physics* B208(1982):323.

7. This shortage of appropriate low-mass particles was noted by M. Gell-Mann in a talk at the 1977 Spring Meeting of the American Physical Society.

8. J. Ellis, M. Gaillard, L. Maiani, and B. Zumino in *Unification of the Fundamental Particle Interactions*, S. Farrara, J. Ellis, and P. van Nieuwenhuizen, editors (New York:Plenum Press, 1980), p. 69.

9. E. Cremmer, B. Julia, and J. Scherk, *Physics Letters* 76B(1978):409. Actually, the $N = 8$ supergravity Lagrangian in four dimensions was first derived by dimensionally reducing the $N = 1$ supergravity Lagrangian in eleven dimensions to $N = 8$ supergravity in four dimensions.

10. M. J. Duff in *Supergravity 81*, S. Farrara and J. G. Taylor, editors (London: Cambridge University Press, 1982), p. 257.

11. Peter G.O. Freund and Mark A. Rubin, *Physics Letters* 97B(1983):233.

12. "Dual Models," *Physics Reports Reprint*, Vol. I, M. Jacob, editor (Amsterdam: North-Holland, 1974).

13. J. Scherk and John H. Schwarz, *Nuclear Physics* B81(1974):118.

14. For a history of this development and a list of references, see John H. Schwarz, *Physics Reports* 89(1982):223 and Michael B. Green, *Surveys in High Energy Physics* 3(1983):127.

15. M. B. Green and J. H. Schwarz, Caltech preprint CALT-68-1090, 1984.

16. For detailed textbook explanations of superspace, superfields, supersymmetry, and supergravity see S. James Gates, Jr., Marcus T. Grisaru, Martin Roček, and Warren Siegel, *Superspace: One Thousand and One Lessons in Supersymmetry* (Reading, Massachusetts: Benjamin/Cummings Publishing Co., Inc., 1983) and Julius Wess and Jonathan Bagger, *Supersymmetry and Supergravity* (Princeton, New Jersey:Princeton University Press, 1983).

17. John H. Schwarz, *Nuclear Physics* B226(1983):269; P. S. Howe and P. C. West, *Nuclear Physics* B238(1984):181.