Conceptual foundations of the unified theory of weak and electromagnetic interactions*†

Steven Weinberg

Lyman Laboratory of Physics, Harvard University and Harvard–Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138 U.S.A.

Our job in physics is to see things simply, to understand a great many complicated phenomena in a unified way, in terms of a few simple principles. At times, our efforts are illuminated by a brilliant experiment, such as the 1973 discovery of neutral current neutrino reactions. But even in the dark times between experimental breakthroughs, there always continues a steady evolution of theoretical ideas, leading almost imperceptibly to changes in previous beliefs. In this talk, I want to discuss the development of two lines of thought in theoretical physics. One of them is the slow growth in our understanding of symmetry, and in particular, broken or hidden symmetry. The other is the old struggle to come to terms with the infinities in quantum field theories. To a remarkable degree, our present detailed theories of elementary particle interactions can be understood deductively, as consequences of symmetry principles and of a principle of renormalizability which is invoked to deal with the infinities. I will also briefly describe how the convergence of these lines of thought led to my own work on the unification of weak and electromagnetic interactions. For the most part, my talk will center on my own gradual education in these matters, because that is one subject on which I can speak with some confidence. With rather less confidence, I will also try to look ahead, and suggest what role these lines of thought may play in the physics of the future.

Symmetry principles made their appearance in twentieth century physics in 1905 with Einstein's identification of the invariance group of space and time. With this as a precedent, symmetries took on a character in physicists' minds as *a priori* principles of universal validity, expressions of the simplicity of nature at its deepest level. So it was painfully difficult in the 1930s to realize that there are internal symmetries, such as isospin conservation,[1] having nothing to do with space and time, symmetries which are far from self-evident, and that only govern what are now called the strong interactions. The 1950s saw the discovery of another internal symmetry—the conservation of strangeness[2]—which is not obeyed by the weak interactions, and even one of the supposedly sacred symmetries of space-time—parity—was also found to be violated by weak interactions.[3] Instead of moving toward unity, physicists were learning that different interactions are apparently governed by quite different symmetries. Matters became yet more confusing with the recognition in the early 1960s of a symmetry group—the "eightfold way"—which is not even an exact symmetry of the strong interactions.[4]

These are all "global" symmetries, for which the symmetry transformations do not depend on position in space and time. It had been recognized[5] in the 1920s that quantum electrodynamics has another symmetry of a far more powerful kind, a "local" symmetry under transformations in which the electron field suffers a phase change that can vary freely from point to point in space-time, and the electromagnetic vector potential undergoes a corresponding gauge transformation. Today this would be called a U(1) gauge symmetry, because a simple phase change can be thought of as multiplication by a 1×1 unitary matrix. The extension to more complicated groups was made by Yang and Mills[6] in 1954 in a seminal paper in which they showed how to construct an SU(2) gauge theory of strong interactions. (The name "SU(2)" means that the group of symmetry transformations consists of 2×2 unitary matrices that are "special," in that they have determinant unity.) But here again it seemed that the symmetry, if real at all, would have to be approximate, because at least on a naive level gauge invariance requires that vector bosons like the photon would have to be massless, and it seemed obvious that the strong interactions are not mediated by massless particles. The old question remained: if symmetry principles are an expression of the simplicity of nature at its deepest level, then how can there be such a thing as an approximate symmetry? Is nature only approximately simple?

Sometime in 1960 or early 1961, I learned of an idea which had originated earlier in solid state physics and had been brought into particle physics by those like Heisenberg, Nambu, and Goldstone, who had worked in both areas. It was the idea of "broken symmetry," that the Hamiltonian and commutation relations of a quantum theory could possess an exact symmetry, and that the physical states might nevertheless not provide neat representations of the symmetry. In particular, a symmetry of the Hamiltonian might turn out to be not a symmetry of the vacuum.

As theorists sometimes do, I fell in love with this idea. But as often happens with love affairs, at first I was rather confused about its implications. I thought (as it turned out, wrongly) that the approximate symmetries—parity, isospin, strangeness, the eightfold way—might really be exact *a priori* symmetry principles, and that the observed violations of these symmetries might somehow be brought about by spontaneous symmetry breaking. It was therefore rather disturbing

*This lecture was delivered December 8, 1979, on the occasion of the presentation of the 1979 Nobel Prizes in Physics.
†This manuscript is referenced according to the author's own style, and not according to RMP's customary British reference system. We trust that this will not inconvenience the reader.

for me to hear of a result of Goldstone,[7] that in at least one simple case the spontaneous breakdown of a continuous symmetry like isospin would necessarily entail the existence of a massless spin zero particle—what would today be called a "Goldstone boson." It seemed obvious that there could not exist any new type of massless particle of this sort which would not already have been discovered.

I had long discussions of this problem with Goldstone at Madison in the summer of 1961, and then with Salam while I was his guest at Imperial College in 1961–62. The three of us soon were able to show that Goldstone bosons must in fact occur whenever a symmetry like isospin or strangeness is spontaneously broken, and that their masses then remain zero to all orders of perturbation theory. I remember being so discouraged by these zero masses that when we wrote our joint paper on the subject,[8] I added an epigraph to the paper to underscore the futility of supposing that anything could be explained in terms of a noninvariant vacuum state: it was Lear's retort to Cordelia, "Nothing will come of nothing: speak again." Of course, *The Physical Review* protected the purity of the physics literature, and removed the quote. Considering the future of the noninvariant vacuum in theoretical physics, it was just as well.

There was actually an exception to this proof, pointed out soon afterwards by Higgs, Kibble, and others.[9] They showed that if the broken symmetry is a local, gauge symmetry, like electromagnetic gauge invariance, then although the Goldstone bosons exist formally, and are in some sense real, they can be eliminated by a gauge transformation, so that they do not appear as physical particles. The missing Goldstone bosons appear instead as helicity zero states of the vector particles, which thereby acquire a mass.

I think that at the time physicists who heard about this exception generally regarded it as a technicality. This may have been because of a new development in theoretical physics, which suddenly seemed to change the role of Goldstone bosons from that of unwanted intruders to that of welcome friends.

In 1964 Adler and Weisberger[10] independently derived sum rules which gave the ratio g_A/g_V of axial-vector to vector coupling constants in beta decay in terms of pion–nucleon cross sections. One way of looking at their calculation (perhaps the most common way at the time) was as an analog to the old dipole sum rule in atomic physics: a complete set of hadronic states is inserted in the commutation relations of the axial vector currents. This is the approach memorialized in the name of "current algebra."[11] But there was another way of looking at the Adler–Weisberger sum rule. One could suppose that the strong interactions have an approximate symmetry, based on the group $SU(2)\times SU(2)$, and that this symmetry is spontaneously broken, giving rise among other things to the nucleon masses. The pion is then identified as (approximately) a Goldstone boson, with small but nonzero mass, an idea that goes back to Nambu.[12] Although the $SU(2)\times SU(2)$ symmetry is spontaneously broken, it still has a great deal of predictive power, but its predictions take the form of approximate formulas, which give the matrix elements for low energy pionic reactions. In this approach, the Adler–Weisberger sum rule is obtained by using the predicted pion nucleon scattering lengths in conjunction with a well-known sum rule,[13] which years earlier had been derived from the dispersion relations for pion–nucleon scattering.

In these calculations one is really using not only the fact that the strong interactions have a spontaneously broken approximate $SU(2)\times SU(2)$ symmetry, but also that the currents of this symmetry group are, up to an overall constant, to be identified with the vector and axial vector currents of beta decay. (With this assumption g_A/g_V gets into the picture through the Goldberger-Treiman relation,[14] which gives g_A/g_V in terms of the pion decay constant and the pion nucleon coupling.) Here, in this relation between the currents of the symmetries of the strong interactions and the physical currents of beta decay, there was a tantalizing hint of a deep connection between the weak interactions and the strong interactions. But this connection was not really understood for almost a decade.

I spent the years 1965–67 happily developing the implications of spontaneous symmetry breaking for the strong interactions.[15] It was this work that led to my 1967 paper on weak and electromagnetic unification. But before I come to that I have to go back in history and pick up one other line of thought, having to do with the problem of infinities in quantum field theory.

I believe that it was Oppenheimer and Waller in 1930[16] who independently first noted that quantum field theory when pushed beyond the lowest approximation yields ultraviolet divergent results for radiative self-energies. Professor Waller told me last night that when he described this result to Pauli, Pauli did not believe it. It must have seemed that these infinities would be a disaster for the quantum field theory that had just been developed by Heisenberg and Pauli in 1929–30. And indeed, these infinities did lead to a sense of discouragement about quantum field theory, and many attempts were made in the 1930s and early 1940s to find alternatives. The problem was solved (at least for quantum electrodynamics) after the war, by Feynman, Schwinger, and Tomonaga,[17] and Dyson.[19] It was found that all infinities disappear if one identifies the observed finite values of the electron mass and charge, not with the parameters m and e appearing in the Lagrangian, but with the electron mass and charge that are *calculated* from m and e, when one takes into account the fact that the electron and photon are always surrounded with clouds of virtual photons and electron–positron pairs.[18] Suddenly all sorts of calculations became possible, and gave results in spectacular agreement with experiment.

But even after this success, opinions differed as to the significance of the ultraviolet divergences in quantum field theory. Many thought—and some still do think—that what had been done was just to sweep the real problems under the rug. And it soon became clear that there was only a limited class of so-called "renormalizable" theories in which the infinities could be eliminated by absorbing them into a redefinition, or a "renormalization," of a finite number of physical parameters. (Roughly speaking, in renormalizable theories no coupling constants can have the dimensions of negative powers of mass. But every time we add a field or a space-time derivative to an interaction, we reduce the dimensionality of the associated coupling constant. So only a few simple types of interaction can be renormalizable.) In particular, the existing Fermi theory of weak interactions clearly was not renormalizable. (The Fermi coupling constant has the dimensions of $[\text{mass}]^{-2}$.) The sense of discouragement about quantum field theory persisted into the 1950s and 1960s.

I learned about renormalization theory as a graduate student, mostly by reading Dyson's papers.[19] From the beginning it seemed to me to be a wonderful thing that very few quantum field theories are renormalizable. Limitations of this sort are, after all, what we most *want;* not mathematical methods which can make sense out of an infinite variety of physically irrelevant theories, but methods which carry constraints, because these constraints may point the way toward the one true theory. In particular, I was very impressed by the fact that quantum electrodynamics could in a sense be *derived* from symmetry principles and the constraint of renormalizability; the only Lorentz invariant and gauge invariant renormalizable Lagrangian for photons and electrons is precisely the original Dirac Langrangian of QED. Of course, that is not the way Dirac came to his theory. He had the benefit of the information gleaned in centuries of experimentation on electromagnetism, and in order to fix the final form of his theory he relied on ideas of simplicity (specifically, on what is sometimes called minimal electromagnetic coupling). But we have to look ahead, to try to make theories of phenomena which have not been so well studied experimentally, and we may not be able to trust purely formal ideas of simplicity. I thought that renormalizability might be the key criterion, which also in a more general context would impose a precise kind of simplicity on our theories and help us to pick out the one true physical theory out of the infinite variety of conceivable quantum field theories. As I will explain later, I would say this a bit differently today, but I am more convinced than ever that the use of renormalizability as a constraint on our theories of the observed interactions is a good strategy. Filled with enthusiasm for renormalization theory, I wrote my Ph.D. thesis under Sam Treiman in 1957 on the use of a limited version of renormalizability to set constraints on the weak interactions,[20] and a little later I worked out a rather tough little theorem[21] which completed the proof by Dyson[19] and Salam[22] that ultraviolet divergences really do cancel out to all orders in nominally renormalizable theories. But none of this seemed to help with the important problem, of how to make a renormalizable theory of weak interactions.

Now, back to 1967. I had been considering the implications of the broken $SU(2)\times SU(2)$ symmetry of the strong interactions, and I thought of trying out the idea that perhaps the $SU(2)\times SU(2)$ symmetry was a "local," not merely a "global," symmetry. That is, the strong interactions might be described by something like a Yang–Mills theory, but in addition to the vector ρ mesons of the Yang–Mills theory, there would also be axial vector $A1$ mesons. To give the ρ meson a mass, it was necessary to insert a common ρ and $A1$ mass term in the Lagrangian, and the spontaneous breakdown of the $SU(2)\times SU(2)$ symmetry would then split the ρ and $A1$ by something like the Higgs mechanism, but since the theory would not be gauge invariant the pions would remain as physical Goldstone bosons. This theory gave an intriguing result, that the $A1/\rho$ mass ratio should be $\sqrt{2}$, and in trying to understand this result without relying on perturbation theory, I discovered certain sum rules, the "spectral function sum rules,"[23] which turned out to have a variety of other uses. But the $SU(2)\times SU(2)$ theory was not gauge invariant, and hence it could not be renormalizable,[24] so I was not too enthusiastic about it.[25] Of course, if I did not insert the ρ-$A1$ mass term in the Lagrangian, then the theory would be gauge invariant and renormalizable, and the $A1$ would be massive. But then there would be no pions and the ρ mesons would be massless, in obvious contradiction (to say the least) with observation.

At some point in the fall of 1967, I think while driving to my office at MIT, it occurred to me that I had been applying the right ideas to the wrong problem. It is not the ρ meson that is massless: it is the photon. And its partner is not the $A1$, but the massive intermediate bosons, which since the time of Yukawa had been suspected to be the mediators of the weak interactions. The weak and electromagnetic interactions could then be described[26] in a unified way in terms of an exact but spontaneously broken gauge symmetry. [Of course, not necessarily $SU(2)\times SU(2)$.] And this theory would be renormalizable like quantum electrodynamics because it is gauge invariant like quantum electrodynamics.

It was not difficult to develop a concrete model which embodied these ideas. I had little confidence then in my understanding of strong interactions, so I decided to concentrate on leptons. There are two left-handed electron-type leptons, the ν_{eL} and e_L, and one right-handed electron-type lepton, the e_R, so I started with the group $U(2)\times U(1)$: all unitary 2×2 matrices acting on the left-handed e-type leptons, together with all unitary 1×1 matrices acting on the right-handed e-type lepton. Breaking up $U(2)$ into unimodular transformations and phase transformations, one could say that the group was $SU(2)\times U(1)\times U(1)$. But then one of the $U(1)$'s could be identified with ordinary lepton number, and since lepton number appears to be conserved and there is no massless vector particle coupled to it, I decided to exclude it from the group. This left the four-parameter group $SU(2)\times U(1)$. The spontaneous breakdown of $SU(2)\times U(1)$ to the $U(1)$ of ordinary electromagnetic gauge invariance would give masses to three of the four vector gauge bosons: the charged bosons $W^{\pm}$, and a neutral boson that I called the Z^0. The fourth boson would automatically remain massless, and could be identified as the photon. Knowing the strength of the ordinary charged current weak interactions like beta decay which are mediated by $W^{\pm}$, the mass of the $W^{\pm}$ was then determined as about 40 GeV/$\sin\theta$, where θ is the γ-Z^0 mixing angle.

To go further, one had to make some hypothesis about the mechanism for the breakdown of $SU(2)\times U(1)$. The only kind of field in a renormalizable $SU(2)\times U(1)$ theory whose vacuum expectation values could give the electron a mass is a spin zero SU(2) doublet (ϕ^+, ϕ^0), so for simplicity I assumed that these were the only scalar fields in the theory. The mass of the Z^0 was then determined as about 80 GeV/sin2θ. This fixed the strength of the neutral current weak interactions. Indeed, just as in QED, once one decides on the menu of fields in the theory, all details of the theory are completely determined by symmetry principles and renormalizability, with just a few free parameters: the lepton charge and masses, the Fermi coupling constant of beta decay, the mixing angle θ, and the mass of the scalar particle. (It was of crucial importance to impose the constraint of renormalizability; otherwise weak interactions would receive contributions from $SU(2)\times U(1)$-invariant four-fermion couplings as well as from vector boson exchange, and the theory would lose most of its predictive power.) The naturalness of the whole theory is well demonstrated by the fact that much the same theory was independently developed[27] by Salam in 1968.

The next question now was renormalizability. The Feynman rules for Yang–Mills theories with unbroken gauge symmetries had been worked out[28] by deWitt, Faddeev, and Popov and others, and it was known that such theories are renormalizable. But in 1967 I did not know how to prove that this renormalizability was not spoiled by the spontaneous symmetry breaking. I worked on the problem on and off for several years, partly in collaboration with students,[29] but I made little progress. With hindsight, my main difficulty was that in quantizing the vector fields I adopted a gauge now known as the unitarity gauge[30]: this gauge has several wonderful advantages, it exhibits the true particle spectrum of the theory, but it has the disadvantage of making renormalizability totally obscure.

Finally, in 1971 't Hooft[31] showed in a beautiful paper how the problem could be solved. He invented a gauge, like the "Feynman gauge" in QED, in which the Feynman rules manifestly lead to only a finite number of types of ultraviolet divergence. It was also necessary to show that these infinities satisfied essentially the same constraints as the Lagrangian itself, so that they could be absorbed into a redefinition of the parameters of the theory. (This was plausible, but not easy to prove, because a gauge invariant theory can be quantized only after one has picked a specific gauge, so it is not obvious that the ultraviolet divergences satisfy the same gauge invariance constraints as the Lagrangian itself.) The proof was subsequently completed[32] by Lee and Zinn–Justin and by 't Hooft and Veltman. More recently, Becchi, Rouet, and Stora[33] have invented an ingenious method for carrying out this sort of proof, by using a global supersymmetry of gauge theories which is preserved even when we choose a specific gauge.

I have to admit that when I first saw 't Hooft's paper in 1971, I was not convinced that he had found the way to prove renormalizability. The trouble was not with 't Hooft, but with me: I was simply not familiar enough with the path integral formalism on which 't Hooft's work was based, and I wanted to see a derivation of the Feynman rules in 't Hooft's gauge from canonical quantization. That was soon supplied (for a limited class of gauge theories) by a paper of Ben Lee,[34] and after Lee's paper I was ready to regard the renormalizability of the unified theory as essentially proved.

By this time, many theoretical physicists were becoming convinced of the general approach that Salam and I had adopted: that is, the weak and electromagnetic interactions are governed by some group of exact local gauge symmetries; this group is spontaneously broken to U(1), giving mass to all the vector bosons except the photon; and the theory is renormalizable. What was not so clear was that our specific simple model was the one chosen by nature. That, of course, was a matter for experiment to decide.

It was obvious even back in 1967 that the best way to test the theory would be by searching for neutral current weak interactions, mediated by the neutral intermediate vector boson, the Z^0. Of course, the possibility of neutral currents was nothing new. There had been speculations[35] about possible neutral currents as far back as 1937 by Gamow and Teller, Kemmer, and Wentzel, and again in 1958 by Bludman and Leite–Lopes. Attempts at a unified weak and electromagnetic theory had been made[36] by Glashow and Salam and Ward in the early 1960s, and these had neutral currents with many of the features that Salam and I encountered in developing the 1967–68 theory. But since one of the predictions of our theory was a value for the mass of the Z^0, it made a definite prediction of the strength of the neutral currents. More important, now we had a comprehensive quantum field theory of the weak and electromagnetic interactions that was physically and mathematically satisfactory in the same sense as quantum electrodynamics—a theory that treated photons and intermediate vector bosons on the same footing, that was based on an exact symmetry principle, and that allowed one to carry calculations to any desired degree of accuracy. To test this theory, it had now become urgent to settle the question of the existence of the neutral currents.

Late in 1971, I carried out a study of the experimental possibilities.[37] The results were striking. Previous experiments had set upper bounds on the rates of neutral current processes which were rather low, and many people had received the impression that neutral currents were pretty well ruled out, but I found that in fact the 1967–68 theory *predicted* quite low rates, low enough in fact to have escaped clear detection up to that time. For instance, experiments[38] a few years earlier had found an upper bound of 0.12 ± 0.06 on the ratio of a neutral current process, the elastic scattering of muon neutrinos by protons, to the corresponding charged current process, in which a muon is produced. I found a predicted ratio of 0.15 to 0.25, depending on the value of the Z^0-γ mixing angle θ. So there was every reason to look a little harder.

As everyone knows, neutral currents were finally discovered[39] in 1973. There followed years of careful experimental study on the detailed properties of the neutral currents. It would take me too far from my subject to survey these experiments,[40] so I will just say that they have confirmed the 1967–68 theory with steadily improving precision for neutrino–nucleon and neutrino–electron neutral current reactions, and since the remarkable SLAC–Yale experiment[41] last year, for the electron–nucleon neutral currents as well.

This is all very nice. But I must say that I would not have been too disturbed if it had turned out that the correct theory was based on some other spontaneously broken gauge group, with very different neutral currents. One possibility was a clever SU(2) theory proposed in 1972 by Georgi and Glashow,[42] which has no neutral currents at all. The important thing to me was the idea of an exact spontaneously broken gauge symmetry, which connects the weak and electromagnetic interactions, and allows these interactions to be renormalizable. Of this I was convinced, if only because it fitted my conception of the way that nature ought to be.

There were two other relevant theoretical developments in the early 1970s, before the discovery of neutral currents, that I must mention here. One is the important work of Glashow, Iliopoulos, and Maiani on the charmed quark.[43] Their work provided a solution to what otherwise would have been a serious problem, that of neutral strangeness changing currents. I leave this topic for Professor Glashow's talk. The other theoretical development has to do specifically with the strong interactions, but it will take us back to one of the themes of my talk, the theme of symmetry.

In 1973, Politzer and Gross and Wilczek discovered[44] a remarkable property of Yang–Mills theories which they called "asymptotic freedom"—the effective coupling constant[45] decreases to zero as the characteristic energy of a process goes to infinity. It seemed that this might explain the experimental fact that the nucleon behaves in high energy deep inelastic electron scattering as if it consists of essentially free quarks.[46] But there was a problem. In order to give masses to the vector bosons in a gauge theory of strong interactions one would want to include strongly interacting scalar fields, and these would generally destroy asymptotic freedom. Another difficulty, one that particularly bothered me, was that in a unified theory of weak and electromagnetic interactions the fundamental weak coupling is of the same order as the electronic charge, e, so the effects of virtual intermediate vector bosons would introduce much too large violations of parity and strangeness conservation, of order $1/137$, into the strong interactions of the scalars with each other and with the quarks.[47] At some point in the spring of 1973 it occurred to me (and independently to Gross and Wilczek) that one could do away with strongly interacting scalar fields altogether, allowing the strong interaction gauge symmetry to remain unbroken so that the vector bosons, or "gluons," are massless, and relying on the increase of the strong forces with increasing distance to explain why quarks as well as the massless gluons are not seen in the laboratory.[48] Assuming no strongly interacting scalars, three "colors" of quarks (as indicated by earlier work of several authors[49]), and an SU(3) gauge group, one then had a specific theory of strong interactions, the theory now generally known as quantum chromodynamics.

Experiments since then have increasingly confirmed QCD as the correct theory of strong interactions. What concerns me here, though, is its impact on our understanding of symmetry principles. Once again, the constraints of gauge invariance and renormalizability proved enormously powerful. These constraints force the Lagrangian to be so simple, that the strong interactions in QCD must conserve strangeness, charge conjugation, and (apart from problems[50] having to do with instantons) parity. One does not have to assume these symmetries as *a priori* principles; there is simply no way that the Lagrangian can be complicated enough to violate them. With one additional assumption, that the u and d quarks have relatively small masses, the strong interactions must also satisfy the approximate $SU(2)\times SU(2)$ symmetry of current algebra, which when spontaneously broken leaves us with isospin. If the s quark mass is also not too large, then one gets the whole eightfold way as an approximate symmetry of the strong interactions. And the breaking of this $SU(3)\times SU(3)$ symmetry by quark masses has just the $(3,\bar{3})+(\bar{3},3)$ form required to account for the pion–pion scattering lengths[15] and the Gell-Mann Okubo mass formulas. Furthermore, with weak and electromagnetic interactions also described by a gauge theory, the weak currents are necessarily just the currents associated with these strong interaction symmetries. In other words, pretty much the whole pattern of approximate symmetries of strong, weak, and electromagnetic interactions that puzzled us so much in the 1950s and 1960s now stands explained as a simple consequence of strong, weak, and electromagnetic gauge invariance, plus renormalizability. Internal symmetry is now at the point where space-time symmetry was in Einstein's day. All the approximate internal symmetries are explained dynamically. On a fundamental level, there are no approximate or partial symmetries; there are only exact symmetries which govern all interactions.

I now want to look ahead a bit, and comment on the possible future development of the ideas of symmetry and renormalizability.

We are still confronted with the question whether the scalar particles that are responsible for the spontaneous breakdown of the electroweak gauge symmetry $SU(2)\times U(1)$ are really elementary. If they are, then spin zero semiweakly decaying "Higgs bosons" should be found at energies comparable with those needed to produce the intermediate vector bosons. On the other hand, it may be that the scalars are composites.[51] The Higgs bosons would then be indistinct broad states at very high mass, analogous to the possible s-wave enhancement in $\pi-\pi$ scattering. There would probably also exist lighter, more slowly decaying, scalar particles of a rather different type, known as pseudo-Goldstone bosons.[52] And there would have to exist a new class of "extra strong" interactions[53] to provide the binding force, extra strong in the sense that asymptotic freedom sets in not at a few hundred MeV, as in QCD, but at a few hundred GeV. This "extra strong" force would be felt by new families of fermions, and would give these fermions masses of the order of several hundred GeV. We shall see.

Of the four (now three) types of interactions, only gravity has resisted incorporation into a renormalizable quantum field theory. This may just mean that we are not being clever enough in our mathematical treatment of general relativity. But there is another possibility that seems to me quite plausible. The constant of gravity defines a unit of energy known as the Planck energy, about 10^{19} GeV. This is the energy at which gravitation becomes effectively a strong interaction, so that at this energy one can no longer ignore its ultraviolet divergences. It may be that there is a whole world of new physics with unsuspected degrees of freedom at these enormous energies, and that general relativity does not provide an adequate framework for understanding the physics of these superhigh energy degrees of freedom. When we explore gravitation or other ordinary phenomena, with particle masses and energies no greater than a TeV or so, we may be learning only about an "effective" field theory; that is, one in which superheavy degrees of freedom do not explicitly appear, but the coupling parameters implicitly represent sums over these hidden degrees of freedom.

To see if this makes sense, let us suppose it is true, and ask what kinds of interactions we would expect on this basis to find at ordinary energy. By "integrating out" the superhigh energy degrees of freedom in a fundamental theory, we generally encounter a very complicated effective field theory—so complicated, in fact, that it contains *all* interactions allowed by symmetry principles. But where dimensional analysis tells us that a coupling constant is a certain power of some mass, that mass is likely to be a typical superheavy mass, such as 10^{19} GeV. The infinite variety of nonrenormalizable interactions in the effective theory have coupling constants with the dimensionality of negative powers of mass, so their effects are suppressed at ordinary energies by powers of energy divided by superheavy masses. Thus the only interactions that we can detect at ordinary energies are those that are renormalizable in the usual sense, plus any nonrenormalizable interactions that produce effects which, although tiny, are somehow exotic enough to be seen.

One way that a very weak interaction could be detected is for it to be coherent and of long range, so that it can add up and have macroscopic effects. It has been shown[54] that the only particles which could produce such forces are massless particles of spin 0, 1, or 2. And furthermore, Lorentz invariance alone is enough to show that the long-range interactions of any particle of mass zero and spin 2 must be governed by general relativity.[55] Thus from this point of view we should not be too surprised that gravitation is the only interaction discovered so far that does not seem to be described by a renormalizable field theory—it is almost the only superweak interaction that *could* have been detected. And we should not be surprised to find that gravity is well described by general relativity at macroscopic scales, even if we do not think that general relativity applies at 10^{19} GeV.

Nonrenormalizable effective interactions may also be detected if they violate otherwise exact conservation laws. The leading candidates for violation are baryon and lepton conservation. It is a remarkable consequence of the SU(3) and SU(2)× U(1) gauge symmetries of strong, weak, and electromagnetic interactions, that all renormalizable interactions among known particles automatically conserve baryon and lepton number. Thus, the fact that ordinary matter seems pretty stable, that proton decay has not been seen, should not lead us to the conclusion that baryon and lepton conservation are fundamental conservation laws. To the accuracy with which they have been verified, baryon and lepton conservation can be explained as dynamical consequences of other symmetries, in the same way that strangeness conservation has been explained within QCD. But superheavy particles may exist, and these particles may have unusual SU(3) or SU(2)× U(1) transformation properties, and in this case, there is no reason why their interactions should conserve baryon or lepton number. I doubt that they would. Indeed, the fact that the universe seems to contain an excess of baryons over antibaryons should lead us to suspect that baryon nonconserving processes have actually occurred. If effects of a tiny nonconservation of baryon or lepton number such as proton decay or neutrino masses are discovered experimentally, we will then be left with gauge symmetries as the only true internal symmetries of nature, a conclusion that I would regard as most satisfactory.

The idea of a new scale of superheavy masses has arisen in another way.[56] If any sort of "grand unification" of strong and electroweak gauge couplings is to be possible, then one would expect all of the SU(3) and SU(2)× U(1) gauge coupling constants to be of comparable magnitude. (In particular, if SU(3) and SU(2) × U(1) are subgroups of a larger simple group, then the ratios of the squared couplings are fixed as rational numbers of order unity.[57]) But this appears in contradiction with the obvious fact that the strong interactions are stronger than the weak and electromagnetic interactions. In 1974 Georgi, Quinn, and I suggested that the grand unification scale, at which the couplings are all comparable, is at an enormous energy, and that the reason that the strong coupling is so much larger than the electroweak couplings at ordinary energies is that QCD is asymptotically free, so that its effective coupling constant rises slowly as the energy drops from the grand unification scale to ordinary values. The change of the strong couplings is very slow (like $1/\sqrt{\ln E}$) so the grand unification scale must be enormous. We found that for a fairly large class of theories the grand unification scale comes out to be in the neighborhood of 10^{16} GeV, an energy not all that different from the Planck energy of 10^{19} GeV. The nucleon lifetime is very difficult to estimate accurately, but we gave a representative value of 10^{32} years, which may be accessible experimentally in a few years. (These estimates have been improved in more detailed calculations by several authors.)[58] We also calculated a value for the mixing parameter $\sin^2\theta$ of about 0.2, not far from the present experimental value[40] of 0.23 ± 0.01. It will be an important task for future experiments on neutral currents to improve the precision with which $\sin^2\theta$ is known, to see if it really agrees with this prediction.

In a grand unified theory, in order for elementary scalar particles to be available to produce the spontaneous breakdown of the electroweak gauge symmetry at a few hundred GeV, it is necessary for such particles to escape getting superlarge masses from the spontaneous breakdown of the grand unified gauge group. There is nothing impossible in this, but I have not been able to think of any reason why it should happen. (The problem may be related to the old mystery of why quantum corrections do not produce an enormous cosmological constant; in both cases, one is concerned with an anomalously small "super-renormalizable" term in the effective Lagrangian which has to be adjusted to be zero. In the case of the cosmological constant, the adjustment must be precise to some fifty decimal places.) With elementary scalars of small or zero mass, enormous ratios of symmetry breaking scales can arise quite naturally.[59] On the other hand, if there are no elementary scalars which escape getting superlarge masses from the breakdown of the grand unified gauge group, then as I have already mentioned, there must be extra strong forces to bind the composite Goldstone and Higgs bosons that are associated with the spontaneous breakdown of $SU(2)\times U(1)$. Such forces can occur rather naturally in grand unified theories. To take one example, suppose that the grand gauge group breaks, not into $SU(3)\times SU(2)\times U(1)$, but into $SU(4)\times SU(3)\times SU(2)\times U(1)$. Since $SU(4)$ is a bigger group than $SU(3)$, its coupling constant rises with decreasing energy more rapidly than the QCD coupling, so the $SU(4)$ force becomes strong at a much higher energy than the few hundred MeV at which the QCD force becomes strong. Ordinary quarks and leptons would be neutral under $SU(4)$, so they would not feel this force, but other fermions might carry $SU(4)$ quantum numbers, and so get rather large masses. One can even imagine a sequence of increasingly large subgroups of the grand gauge group, which would fill in the vast energy range up to 10^{15} or 10^{19} GeV with particle masses that are produced by these successively strong interactions.

If there are elementary scalars whose vacuum expectation values are responsible for the masses of ordinary quarks and leptons, then these masses can be affected in order α by radiative corrections involving the superheavy vector bosons of the grand gauge group, and it will probably be impossible to explain the value of quantities like m_e/m_μ without a complete grand unified theory. On the other hand, if there are no such elementary scalars, then almost all the details of the grand unified theory are forgotten by the effective field theory that describes physics at ordinary energies, and it ought to be possible to calculate quark and lepton masses purely in terms of processes at accessible energies. Unfortunately, no one so far has been able to see how in this way anything resembling the observed pattern of masses could arise.[60]

Putting aside all these uncertainties, suppose that there is a truly fundamental theory, characterized by an energy scale of order 10^{16} to 10^{19} GeV, at which strong, electroweak, and gravitational interactions are all united. It might be a conventional renormalizable quantum field theory, but at the moment, if we include gravity, we do not see how this is possible. (I leave the topic of supersymmetry and supergravity for Professor Salam's talk.) But if it is not renormalizable, what then determines the infinite set of coupling constants that are needed to absorb all the ultraviolet divergences of the theory?

I think the answer must lie in the fact that the quantum field theory, which was born just fifty years ago from the marriage of quantum mechanics with relativity, is a beautiful but not a very robust child. As Landau and Källén recognized long ago, quantum field theory at superhigh energies is susceptible to all sorts of diseases—tachyons, ghosts, etc.—and it needs special medicine to survive. One way that a quantum field theory can avoid these diseases is to be renormalizable and asymptotically free, but there are other possibilities. For instance, even an infinite set of coupling constants may approach a nonzero fixed point as the energy at which they are measured goes to infinity. However, to require this behavior generally imposes so many constraints on the couplings that there are only a finite number of free parameters left[61]—just as for theories that are renormalizable in the usual sense. Thus, one way or another, I think that quantum field theory is going to go on being very stubborn, refusing to allow us to describe all but a small number of possible worlds, among which, we hope, is ours.

I suppose that I tend to be optimistic about the future of physics. And nothing makes me more optimistic than the discovery of broken symmetries. In the seventh book of *The Republic*, Plato describes prisoners who are chained in a cave and can see only shadows that things outside cast on the cave wall. When released from the cave at first their eyes hurt, and for a while they think that the shadows they saw in the cave are more real than the objects they now see. But eventually their vision clears, and they can understand how beautiful the real world is. We are in such a cave, imprisoned by the limitations on the sorts of experiments we can do. In particular, we can study matter only at relatively low temperatures, where symmetries are likely to be spontaneously broken, so that nature does not appear very simple or unified. We have not been able to get out of this cave, but by looking long and hard at the shadows on the cave wall, we can at least make out the shapes of symmetries, which though broken, are exact principles governing all phenomena, expressions of the beauty of the world outside.

It has only been possible here to give references to a very small part of the literature on the subjects discussed in this talk. Additional references can be found in the following reviews: E. S. Abers and B. W. Lee, "Gauge Theories" (Phys. Rep. C **9**, No. 1, 1973); W. Marciano and H. Pagels, "Quantum Chromodynamics" (Phys. Rep. C **36**, No. 3, 1978); J. C. Taylor, *Gauge Theories of Weak Interactions* (Cambridge University, 1976).

REFERENCES

[1]M. A. Tuve, N. Heydenberg, and L. R. Hafstad, Phys. Rev. **50**, 806 (1936); G. Breit, E. V. Condon, and R. D. Present, Phys. Rev. **50**, 825 (1936); G. Breit and E. Feenberg, Phys. Rev. **50**, 850 (1936).
[2]M. Gell-Mann, Phys. Rev. **92**, 833 (1953); T. Nakano and K. Nishijima, Prog. Theor. Phys. **10**, 581 (1955).
[3]T. D. Lee and C. N. Yang, Phys. Rev. **104**, 254 (1956); C. S. Wu *et al.*, Phys. Rev. **105**, 1413 (1957); R. Garwin, L. Lederman, and M. Weinrich, Phys. Rev. **105**, 1415 (1957); J. I. Friedman and V. L. Telegdi, Phys. Rev. **105**, 1681 (1957).

[4]M. Gell-Mann, Cal. Tech. Synchotron Laboratory Report CTSL–20 (1961), unpublished; Y. Ne'eman, Nucl. Phys. **26**, 222 (1961).

[5]V. Fock, Z. Phys. **39**, 226 (1927); H. Weyl, Z. Phys. **56**, 330 (1929). The name "gauge invariance" is based on an analogy with the earlier speculations of H. Weyl, in *Raum, Zeit, Materie*, 3rd edition (Springer, 1920). Also see F. London, Z. Phys. **42**, 375 (1927). (This history has been reviewed by C. N. Yang in a talk at City College, 1977.)

[6]C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954).

[7]J. Goldstone, Nuovo Cimento **19**, 154 (1961).

[8]J. Goldstone, A. Salam, and S. Weinberg, Phys. Rev. **127**, 965 (1962).

[9]P. W. Higgs, Phys. Lett. **12**, 132 (1964); **13**, 508 (1964); Phys. Rev. **145**, 1156 (1966); T. W. B. Kibble, Phys. Rev. **155**, 1554 (1967); G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964); F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964). Also see P. W. Anderson, Phys. Rev. **130**, 439 (1963).

[10]S. L. Adler, Phys. Rev. Lett. **14**, 1051 (1965); Phys. Rev. **140**, B736 (1965); W. I. Weisberger, Phys. Rev. Lett. **14**, 1047 (1965); Phys. Rev. **143**, 1302 (1966).

[11]M. Gell-Mann, Physics **1**, 63 (1964).

[12]Y. Nambu and G. Jona-Lasinio, Phys. Rev. **122**, 345 (1961); **124**, 246 (1961); Y. Nambu and D. Lurie, Phys. Rev. **125**, 1429 (1962); Y. Nambu and E. Shrauner, Phys. Rev. **128**, 862 (1962). Also see M. Gell-Mann and M. Lévy, Nuovo Cimento **16**, 705 (1960).

[13]M. L. Goldberger, H. Miyazawa, and R. Oehme, Phys. Rev. **99**, 986 (1955).

[14]M. L. Goldberger and S. B. Treiman, Phys. Rev. **111**, 354 (1958).

[15]S. Weinberg, Phys. Rev. Lett. **16**, 879 (1966); **17**, 336 (1966); **17**, 616 (1966); **18**, 188 (1967); Phys. Rev. **166**, 1568 (1967).

[16]J. R. Oppenheimer, Phys. Rev. **35**, 461 (1930); I. Waller, Z. Phys. **59**, 168 (1930); *ibid*. **62**, 673 (1930).

[17]R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948); Phys. Rev. **74**, 939, 1430 (1948); **76**, 749, 769 (1949); **80**, 440 (1950); J. Schwinger, Phys. Rev. **73**, 146 (1948); **74**, 1439 (1948); **75**, 651 (1949); **76**, 790 (1949); **82**, 664, 914 (1951); **91**, 713 (1953); Proc. Nat. Acad. Sci. USA **37**, 452 (1951); S. Tomonaga, Prog. Theor. Phys. (Japan) **1**, 27 (1946); Z. Koba, T. Tati, and S. Tomonaga, *ibid*. **2**, 101 (1947); S. Kanazawa and S. Tomonaga, *ibid*. **3**, 276 (1948); Z. Koba and S. Tomonaga, *ibid*. **3**, 290 (1948).

[18]There had been earlier suggestions that infinities could be eliminated from quantum field theories in this way, by V. F. Weisskopf, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd. **15** (6) 1936, especially p. 34 and pp. 5–6; H. Kramers (unpublished).

[19]F. J. Dyson, Phys. Rev. **75**, 486, 1736 (1949).

[20]S. Weinberg, Phys. Rev. **106**, 1301 (1957).

[21]S. Weinberg, Phys. Rev. **118**, 838 (1960).

[22]A. Salam, Phys. Rev. **82**, 217 (1951); **84**, 426 (1951).

[23]S. Weinberg, Phys. Rev. Lett. **18**, 507 (1967).

[24]For the nonrenormalizability of theories with intrinsically broken gauge symmetries, see A. Komar and A. Salam, Nucl. Phys. **21**, 624 (1960); H. Umezawa and S. Kamefuchi, Nucl. Phys. **23**, 399 (1961); S. Kamefuchi, L. O'Raifeartaigh, and A. Salam, Nucl. Phys. **28**, 529 (1961); A. Salam, Phys. Rev. **127**, 331 (1962); M. Veltman, Nucl. Phys. B **7**, 637 (1968); B **21**, 288 (1970); D. Boulware, Ann. Phys. (N.Y.) **56**, 140 (1970).

[25]This work was briefly reported in reference 23, footnote 7.

[26]S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[27]A. Salam, in *Elementary Particle Physics* (Nobel Symposium No. 8), edited by N. Svartholm (Almquist and Wiksell, Stockholm, 1968), p. 367.

[28]B. deWitt, Phys. Rev. Lett. **12**, 742 (1964); Phys. Rev. **162**, 1195 (1967); L. D. Faddeev and V. N. Popov, Phys. Lett. **B25**, 29 (1967). Also see R. P. Feynman, Acta Phys. Pol. **24**, 697 (1963); S. Mandelstam, Phys. Rev. **175**, 1580, 1604 (1968).

[29]See L. Stuller, MIT Ph.D. Thesis (1971), unpublished.

[30]My work with the unitarity gauge was reported in S. Weinberg, Phys. Rev. Lett. **27**, 1688 (1971), and described in more detail in S. Weinberg, Phys. Rev. D **7**, 1068 (1973).

[31]G. 't Hooft, Nucl. Phys. B **35**, 167 (1971).

[32]B. W. Lee and J. Zinn-Justin, Phys. Rev. D **5**, 3121, 3137, 3155 (1972); G. 't Hooft and M. Veltman, Nucl. Phys. B **44**, 189 (1972); B **50**, 318 (1972). There still remained the problem of possible Adler–Bell–Jackiw anomalies, but these nicely cancelled; see D. J. Gross and R. Jackiw, Phys. Rev. D **6**, 477 (1972), and C. Bouchiat, J. Iliopoulos, and Ph. Meyer, Phys. Lett. **38B**, 519 (1972).

[33]C. Beechi, A. Rouet, and R. Stora, Commun. Math. Phys. **42**, 127 (1975).

[34]B. W. Lee, Phys. Rev. D **5**, 823 (1972).

[35]G. Gamow and E. Teller, Phys. Rev. **51**, 288 (1937); N. Kemmer, Phys. Rev. **52**, 906 (1937); G. Wentzel, Helv. Phys. Acta **10**, 108 (1937); S. Bludman, Nuovo Cimento **9**, 433 (1958); J. Leite-Lopes, Nucl. Phys. **8**, 234 (1958).

[36]S. L. Glashow, Nucl. Phys. **22**, 519 (1961); A. Salam and J. C. Ward, Phys. Lett. **13**, 168 (1964).

[37]S. Weinberg, Phys. Rev. **5**, 1412 (1972).

[38]D. C. Cundy *et al*., Phys. Lett. **31B**, 478 (1970).

[39]The first published discovery of neutral currents was at the Gargamelle Bubble Chamber at CERN: F. J. Hasert *et al*., Phys. Lett. B **46**, 121, 138 (1973). Also see P. Musset, J. Phys. (Paris) 11/12 T34 (1973). Muonless events were seen at about the same time by the HPWF group at Fermilab, but when publication of their paper was delayed, they took the opportunity to rebuild their detector, and then did not at first find the same neutral current signal. The HPWF group published evidence for neutral currents in A. Benvenuti *et al*., Phys. Rev. Lett. **32**, 800 (1974).

[40]For a survey of the data see C. Baltay, *Proceedings of the 19th International Conference on High Energy Physics*, Tokyo, 1978. For theoretical analyses, see L. F. Abbott and R. M. Barnett, Phys. Rev. D **19**, 3230 (1979); P. Langacker, J. E. Kim, M. Levine, H. H. Williams, and D. P. Sidhu, "Neutrino '79" Conference; and earlier references cited therein.

[41]C. Y. Prescott *et al*., Phys. Lett. B **77**, 347 (1978).

[42]S. L. Glashow and H. L. Georgi, Phys. Rev. Lett. **28**, 1494 (1972). Also see J. Schwinger, Ann. Phys. (N.Y.) **2**, 407 (1957).

[43]S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970). This paper was cited in reference 37 as providing a possible solution to the problem of strangeness changing neutral currents. However, at that time I was skeptical about the quark model, so in the calculations of reference 37 baryons were incorporated in the theory by taking the protons and neutrons to form an SU(2) doublet, with strange particles simply ignored.

[44]H. D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973); D. J. Gross and F. Wilczek, Phys. Rev. Lett. **30**, 1343 (1973).

[45]Energy dependent effective coupling constants were introduced by M. Gell-Mann and F. E. Low, Phys. Rev. **95**, 1300 (1954).

[46]E. D. Bloom *et al*., Phys. Rev. Lett. **23**, 930 (1969); M. Breidenbach *et al*., Phys. Rev. Lett. **23**, 935 (1969).

[47]S. Weinberg, Phys. Rev. D **8**, 605 (1973).

[48]D. J. Gross and F. Wilczek, Phys. Rev. D **8**, 3633 (1973); S. Weinberg, Phys. Rev. Lett. **31**, 494 (1973). A similar idea had been proposed before the discovery of asymptotic freedom by H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. B **47**, 365 (1973).

[49]O. W. Greenberg, Phys. Rev. Lett. **13**, 598 (1964); M. Y. Han and Y. Nambu, Phys. Rev. **139**, B1006 (1965); W. A. Bardeen, H. Fritzsch, and M. Gell-Mann, in *Scale and Conformal Symmetry in Hadron Physics*, edited by R. Gatto (Wiley, 1973), p. 139; etc.

[50] G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976).

[51] Such "dynamical" mechanisms for spontaneous symmetry breaking were first discussed by Y. Nambu and G. Jona-Lasinio, Phys. Rev. **122**, 345 (1961); J. Schwinger, Phys. Rev. **125**, 397 (1962); **128**, 2425 (1962); and in the context of modern gauge theories by R. Jackiw and K. Johnson, Phys. Rev. D **8**, 2386 (1973); J. M. Cornwall and R. E. Norton, Phys. Rev. D **8**, 3338 (1973). The implications of dynamical symmetry breaking have been considered by S. Weinberg, Phys. Rev. D **13**, 974 (1976); D **19**, 1277 (1979); L. Susskind, Phys. Rev. D **20**, 2619 (1979).

[52] S. Weinberg, reference 51. The possibility of pseudo-Goldstone bosons was originally noted in a different context by S. Weinberg, Phys. Rev. Lett. **29**, 1698 (1972).

[53] S. Weinberg, reference 51. Models involving such interactions have also been discussed by L. Susskind, reference 51.

[54] S. Weinberg, Phys. Rev. **135**, B1049 (1964).

[55] S. Weinberg, Phys. Lett. **9**, 357 (1964); Phys. Rev. B **138**, 988 (1965); *Lectures in Particles and Field Theory*, edited by S. Deser and K. Ford (Prentice-Hall, 1965), p. 988; and reference 54. The program of deriving general relativity from quantum mechanics and special relativity was completed by D. Boulware and S. Deser, Ann. Phys. (N.Y.) **89**, 173 (1975). I understand that similar ideas were developed by R. Feynman in unpublished lectures at Cal. Tech.

[56] H. Georgi, H. Quinn, and S. Weinberg, Phys. Rev. Lett. **33**, 451 (1974).

[57] An example of a simple gauge group for weak and electromagnetic interactions (for which $\sin^2\theta=\frac{1}{4}$) was given by S. Weinberg, Phys. Rev. D **5**, 1962 (1972). There are a number of specific models of weak, electromagnetic, and strong interactions based on simple gauge groups, including those of J. C. Pati and A. Salam, Phys. Rev. D **10**, 275 (1974); H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32**, 438 (1974); H. Georgi, in *Particles and Fields* (American Institute of Physics, 1975); H. Fritzsch and P. Minkowski, Ann. Phys. (N.Y.) **93**, 193 (1975); H. Georgi and D. V. Nanopoulos, Phys. Lett. B **82**, 392 (1979); F. Gürsey, P. Ramond, and P. Sikivie, Phys. Lett. B **60**, 177 (1975); F. Gürsey and P. Sikivie, Phys. Rev. Lett. **36**, 775 (1976); P. Ramond, Nucl. Phys. B **110**, 214 (1976); etc.; all these violate baryon and lepton conservation.

[58] A. Buras, J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, Nucl. Phys. B **135**, 66 (1978); D. Ross, Nucl. Phys. B **140**, 1 (1978); W. J. Marciano, Phys. Rev. D **20**, 274 (1979); T. Goldman and D. Ross, CALT 68-704, to be published; C. Jarlskog and F. J. Yndurain, CERN preprint, to be published; M. Machacek, Harvard preprint HUTP-79/AO21, to be published in Nucl. Phys.; S. Weinberg, paper in preparation. The phenemonology of nucleon decay has been discussed in general terms by S. Weinberg, Phys. Rev. Lett. **43**, 1566 (1979); F. Wilczek and A. Zee, Phys. Rev. Lett. **43**, 1571 (1979).

[59] E. Gildener and S. Weinberg, Phys. Rev. D **13**, 3333 (1976); S. Weinberg, Phys. Lett. B **82**, 387 (1979). In general in this case there should exist at least one scalar particle with physical mass of order 10 GeV. The spontaneous symmetry breaking in models with zero bare scalar mass was first considered by S. Coleman and E. Weinberg, Phys. Rev. D **7**, 1888 (1973).

[60] This problem has been studied recently by S. Dimopoulos and L. Susskind, Nucl. Phys. B **155**, 237 (1979); E. Eichten and K. Lane, Phys. Lett., to be published; S. Weinberg, unpublished.

[61] S. Weinberg, in *General Relativity - An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University, 1979), Chap. 16.