



Unified theory of elementary-particle forces

At sufficiently small distances, perhaps less than 10^{-29} cm, the weak, electromagnetic, and strong interactions appear to be no more than different components of the same fundamental force.

Howard Georgi and Sheldon L. Glashow

The notion of what are the "elementary" or structureless constituents of matter keeps changing as we are able to probe smaller and smaller distances with higher and higher energies. As long as we were limited by the energy available in chemical processes, the elementary particles were atoms; later they were protons, neutrons and electrons; currently we can smash matter into pieces sufficiently fine that quarks and leptons appear to be the elementary constituents of matter.

Quarks are the constituents of the

hadrons, the particles (such as protons or pions) that interact via the strong force. There is evidence for at least five types ("flavors") of quarks; they are generally labelled *u* (for up) *d* (for down), *s* (for strange), *c* (for charm) and *b* (for bottom or beauty). Further, each type of quark comes in three varieties ("colors"), generally labelled "red," "white" (or sometimes "green") and "blue." The strong interaction, which binds quarks into hadrons, is (we believe) an interaction in which the three colors play a role analogous to the charge in electrodynamics. The currently accepted theory of these interactions is quantum chromodynamics, or QCD for short.

At this time there is some evidence for six kinds of leptons: the electron, the muon and the tau, and the neu-

trinos associated with each. The leptons have no color and are almost completely unaffected by the chromodynamic interactions.

Except for the three neutrinos, quarks and leptons carry electric charges, and therefore interact electromagnetically. All the particles, including neutrinos, interact via the weak interaction. Particles that interact electromagnetically do not change their identity: except for possible particle-antiparticle pairs, the particles that emerge after an interaction have the same properties (charge, strangeness, and so on) as the particles that entered into the interaction. The weak interactions, on the other hand, do change particle identities. The classic example is beta decay in which a neutron decays into a proton, an electron

The authors are professors of physics at Harvard University. Sheldon Glashow shared the Nobel Prize for Physics with Abdus Salam and Steven Weinberg in 1979 for contributions to the unification of the weak and electromagnetic interactions.

A neutrino collides with a proton in a hydrogen bubble chamber at Brookhaven National Laboratory. Such events yield information on the nature of weak interactions.

and an antineutrino ($\bar{\nu}_e$); this occurs when one of the d quarks in the neutron decays into a u (which remains bound to the other d and u that were in the neutron) together with an electron and an antineutrino.

Unification

Quantum electrodynamics, quantum chromodynamics, and the currently accepted theory of weak interactions adequately describe the forces among the elementary particles down to 10^{-16} cm (about 1% of the proton radius), which is the shortest distance probed by today's accelerators. What happens at shorter distances we do not know. But we suspect that at distances of the order of 10^{-29} cm, all three interactions—along with others not yet observed—will be unified. That is, all interactions will have the same strength and the distinctions between quarks, antiquarks and leptons will disappear.

Although this proposed unification takes place at ridiculously small distances (or high energies), it has important consequences for the "low-energy" world of contemporary physics. One such consequence is a prediction of an angle that appears in the theory of weak interactions, called θ_w , the weak mixing angle. Unified theories predict $\sin^2\theta_w \approx 0.20$. When this prediction was first worked out, the best experimental value was $\sin^2\theta_w \approx 0.35$. But as the experiments have improved, the value has marched steadily down; the experimental value now is $\sin^2\theta_w = 0.23 \pm .02$, almost in agreement with the unified theory.

The most spectacular consequence of the unification has yet to be tested conclusively. It is that the proton itself (and thus all matter) is unstable: In a time on the order of 10^{31} years a proton can decay into a positron and a π^0 . The positron eventually annihilates an electron, producing gamma rays, and the π^0 decays quickly into gamma rays as well. The net result is that a hydrogen atom disappears into energy. Although the chance that any one proton decays in any one year is absurdly small, a macroscopic chunk of material has a very large number of protons. Experiments are now in preparation to look at very large samples of material to test the prediction.

Before explaining the unification we briefly review its components, QCD, QED and the weak interactions, in a language particularly suited to the unification, the language of gauge theories. For our examples we will chiefly

use the quarks u and d and the leptons e and ν . These are the constituents of all ordinary materials; the heavier quarks and leptons are of interest almost exclusively to particle physicists.

Electrodynamics

Quantum electrodynamics (QED) is the quantum theory of the electromagnetic interactions of charged point particles. It is an extremely successful theory in that its predictions have been verified experimentally to great precision, in part because it is characterized by a small dimensionless coupling constant. The concept of a "coupling constant" will be central to our idea of unification, so we will describe it in some detail.

The idea starts with Coulomb's law. In relativistic quantum theory it is natural to measure the products of charges (in electrostatic units) in units of $\hbar c$; it is also natural to make use of the fact that charges are quantized as integral multiples of the proton charge, e . Thus we write Coulomb's law as

$$F/\hbar c = \alpha Q Q' / r^2$$

where Q and Q' are integers. We have absorbed the unit of charge, e , into the dimensionless coupling constant

$$\alpha = e^2/\hbar c$$

whose experimental value is pretty nearly $1/137$. It is small enough to allow straightforward perturbation expansions in powers of α .

We should emphasize that while the quantization of electromagnetic charge is an experimental fact, quantum electrodynamics does not require it. The theory would still make sense if there were also a particle with electric charge $Q = \sqrt{2}$, for example. But in the unified theory we will describe below, charge quantization is automatic. Also note that the charge Q is an additive quantum number, and that it has opposite values for a particle and its antiparticle.

All the theories we will discuss are relativistic quantum field theories. In such a theory both the forces that act on particles and the particles themselves are represented by quantized fields. A field and its corresponding quanta represent dual aspects of matter: A field quantum behaves as a particle and the expectation value of the field is the wave function of the particle. We shall distinguish between the "particles" that correspond to the constituents of matter (these are the quarks and leptons—particles with spin $1/2$) and "forces" that determine the interactions between particles (these are associated with particles of integral spin, such as the photon).

The dynamics of the electromagnetic force are given by the vector potential, $A(x)$ —the "photon field." Quantum

electrodynamics is invariant under an extremely large group of transformations. We can add to the vector potential the gradient of any scalar function and simultaneously change the phase of all particle fields by an amount proportional to the scalar:

$$\begin{aligned} A(x) &\rightarrow A(x) + \nabla\theta(x) \\ \psi(x) &\rightarrow \psi(x)e^{ieQ\theta(x)/\hbar c} \end{aligned}$$

and this transformation leaves the theory unchanged. This symmetry is called a gauge invariance, specifically, a $U(1)$ gauge invariance. The $U(1)$ refers to the fact that the phase transformation is unitary (does not change the normalization of the wave function) and that the integer Q that appears in it is just a number, a 1×1 matrix.

The gauge invariance is such a powerful symmetry that it, together with the requirement that the coupling constant be dimensionless, completely determines the form of the interaction. Because of the symmetry, the photon field, $A(x)$, is coupled to each particle field with a strength eQ ; the symmetry also ensures that the photon is massless.

Because of the central role played by gauge invariance in the electromagnetic interactions, the photon field $A(x)$ is called a "gauge field" and the photon a "gauge particle."

Chromodynamics

The quark fields are three-component vectors in a "color space," the components corresponding to the colors red, white and blue. (We shall denote such "color vectors" by arrows: $\vec{q}$) The force between quarks is mediated by a set of "gluon fields" analogous to the photon field of electrodynamics) whose coupling strength is given by a set of color charges, T_a . These charges are 3×3 matrices (traceless and Hermitian) in color space. There are exactly eight independent such matrices, and there are eight different types of color charges T_a , and eight different types of gluon fields G_a . The eight matrices T_a are 3×3 analogs of the famous 2×2 Pauli matrices, σ_i ; to normalize them one generally requires that the trace of T_a^2 should be one half. The color charges are also additive, and particles and antiparticles have opposite color charges.

When a quark is shaken it can emit any one of the eight gluons. The strength of the interaction (analogous to eQ) between a quark and the gluon field G_a is $g_3 T_a$. Two of the matrices T_a , conventionally called T_3 and T_8 , are diagonal, so that quarks that emit gluons G_3 or G_8 do not change their color; the other gluon fields cause transitions between colors.

Quantum chromodynamics also has a gauge invariance. We can add to the gluon field a set of gradients (and simultaneously perform unitary trans-

formations on the matrices) and change the phases of the quark fields without changing the theory. Because the fields are matrices, the gauge transformations are more complex than they are in electrodynamics; for example, the quark fields transform with a matrix phase factor

$$\bar{q}(x) \rightarrow \exp[i g_3 \sum_a \theta_a(x) T_a] \bar{q}(x)$$

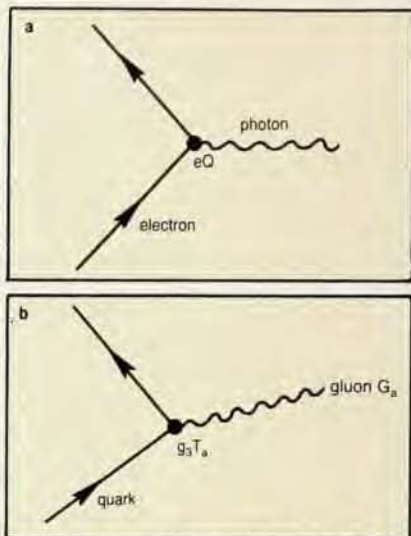
Because only two of the matrices are diagonal, the colors of quarks can be mixed up differently at each point in space—and all the physical results of the theory remain the same.

The net effect of the virtual exchange of all these gluons is to produce a force that tends to bind quarks together into systems that are color-neutral, that is, for which $\bar{q} T_a q$ is zero ($a = 1, \dots, 8$). The simplest such systems are quark-antiquark pairs (the π^+ , for example, consists of a u and a $\bar{d}$), and quark triplets (the proton, for example, contains $2u$'s and a d , one quark of each color); the color-neutrality of quark triplets reflects the tracelessness of the T_a matrices.

Actually, we think that the binding of quarks and gluons into color-neutral systems is absolute: The quarks and gluons are permanently confined within hadrons. There is thus no hope of ever seeing an isolated quark or gluon.

In addition to the color charges, the quarks have electromagnetic charges, Q . The u quark for example has $Q = 2/3$ and the d quark has $Q = -1/3$. Thus, the proton has total charge $Q = 1$ and the neutron (which is two d 's and a u) has total charge $Q = 0$. Notice that, though the quarks have fractional electric charge, the fractional charges are just right to give integral charge to the color-neutral systems of three quarks or a quark and an antiquark. This is another thing we would like our unified theory to explain. Unlike the photon, the gluons themselves must carry color charges. This is related to the fact that the color charges (being matrices) do not commute; the $SU(3)$ symmetry is non-Abelian (contains noncommuting elements) and simple (cannot be separated into commuting subsets). It is easy to prove that for simple groups the charges are quantized. Another way to arrive at the same conclusion is to realize that any system with color has the same color as some combination of quarks. Color charges are quantized because they must all be multiples of the color charges of the three quarks.

Because the color charges T_a are quantized, the dimensional coupling constant g_3 that appears, for example, in the gauge transformation measures a minimum coupling strength in QCD, thus playing a role analogous to that of e , the quantum of electromagnetic charge, in QED. The constant g_3 has



Fundamental interactions (a) in QED and (b) in QCD. In (a) an electron emits (or absorbs) a photon, changing its energy and momentum; the interaction strength is given by eQ (Q is the dimensionless charge, an integer). In (b) a quark emits (or absorbs) a gluon, changing its energy, momentum, and possibly its "color."

the same units as e , so that we can form the dimensionless combination

$$\alpha_3 = g_3^2 / 4\pi$$

Experiments show that $\alpha_3 \gg \alpha$: The color forces (which are "strong interactions" after all) are much stronger than electromagnetic forces.

Asymptotic freedom

The relativistic quantum mechanical vacuum is a lot more than just an empty space. It is seething with virtual particles. For example, the same electromagnetic interaction that allows an electron to emit a photon also allows an electron, a positron and a photon to appear out of nothing. Of course, these cannot be real particles because energy is not conserved, but they can exist as virtual particles for a short time consistent with the uncertainty principle.

If a real charged particle is added to this complicated vacuum, it can polarize the virtual electron-positron pairs. If the real particle has a positive charge, the virtual positive charges are pushed away from it slightly while the virtual negative charges are pulled toward it. The net result is that some of the virtual positive charge is pushed far away and the real charge is surrounded by a negatively charged vacuum.

Suppose now that we want to measure the charge on one of these particles. The standard way is to measure the electric field on a sphere of radius r around the charge and use Gauss's law to find the charge enclosed. But, because of the polarization of the vacuum, this charge will vary with r : It will

decrease as r increases because the sphere contains more and more negatively charged vacuum.

You might think that we could find the "bare" particle charge by going in very close, making r very small. But that doesn't work. The density of negative charge in the vacuum increases as you get closer to the real charge, and no matter how small you make the sphere, there is always a significant amount of negative charge in the enclosed vacuum (this can only happen if the "bare" positive charge is infinite and the density of negative charge in the vacuum goes to infinity as r goes to zero).

This sounds dangerous, but it is okay so long as the charge in any finite sphere is finite. After all, in any experiment, the charges are separated by some characteristic distance d . For example, in an atomic physics experiment a typical distance is $d \sim 10^{-8}$ cm. The appropriate "charge" to use for describing the electromagnetic interactions in such an experiment is the charge inside a sphere of radius d around the point charges.

This process of replacing the infinite bare charge with a finite charge measured at a given distance from the point charge is called "renormalization;" it works only for a specific class of field theories. We shall deal only with such "renormalizable" theories.

The renormalization does not affect charge quantization: The charge in a sphere of radius d around an electron is the same as the charge in a sphere of radius d around the muon (there are some effects due to the mass difference between μ and e , and the proton is a bit more complicated because of the strong interactions, but never mind). This means that the dimensionless charge Q still takes on integral values (except for quarks), but the coupling constant α now depends on d ; its value is $1/137$ at an atomic distance, $\sim 10^{-8}$ cm.

In QED, α is small, so the dependence of α on d is weak—usually negligible, in fact—but it has been measured. (It contributes to the Lamb shift, for example.) In QCD, where α_3 is much larger, this effect is very important: A color-charged particle polarizes the vacuum to a substantial degree. Furthermore, the gauge particles that mediate the color force themselves carry color charge; the cloud of virtual gluons that surrounds a color-charged particle thus effectively spreads out the charge of the particle. It turns out that this effect overwhelms the effect of vacuum polarization, so that a positive color charge is surrounded by a positively charged vacuum. Because the charge is spread out, the color charge within a sphere of radius r decreases (very slowly) to zero as r goes to zero. The color forces get weaker at short distances! This property of chromodynamics is called as-

ymptotic freedom. It was discovered by H. David Politzer at Harvard (now at Caltech) and by David Gross and Frank Wilczek at Princeton. Asymptotic freedom has been tested successfully in recent experiments that probe very short distances.

Weak interactions

The weak interactions are also described by a gauge theory. In this case the fundamental particles appear in doublets, so the invariance is an SU(2) symmetry. (This is the same symmetry group as the angular momentum of quantum mechanics.) The weak charges are the three independent 2×2 traceless, Hermitian matrices R_i (proportional to the Pauli matrices). Coupled to these three charges are three gauge particles, W (these are often called "intermediate vector bosons"). Again, because the charges are matrices, the emission or absorption of a W_i can change the identity of a particle. Thus, for example, an electron can emit a W_1 and turn into a neutrino.

The W_3 is electrically neutral, but the W_1 and W_2 carry electric charge. The fields corresponding to particles of definite charge are the combinations $W^\pm = W_1 \pm iW_2$.

The weak interactions have the strange property that they discriminate between particles of different handedness (that is, with spins directed along or against the direction of motion).

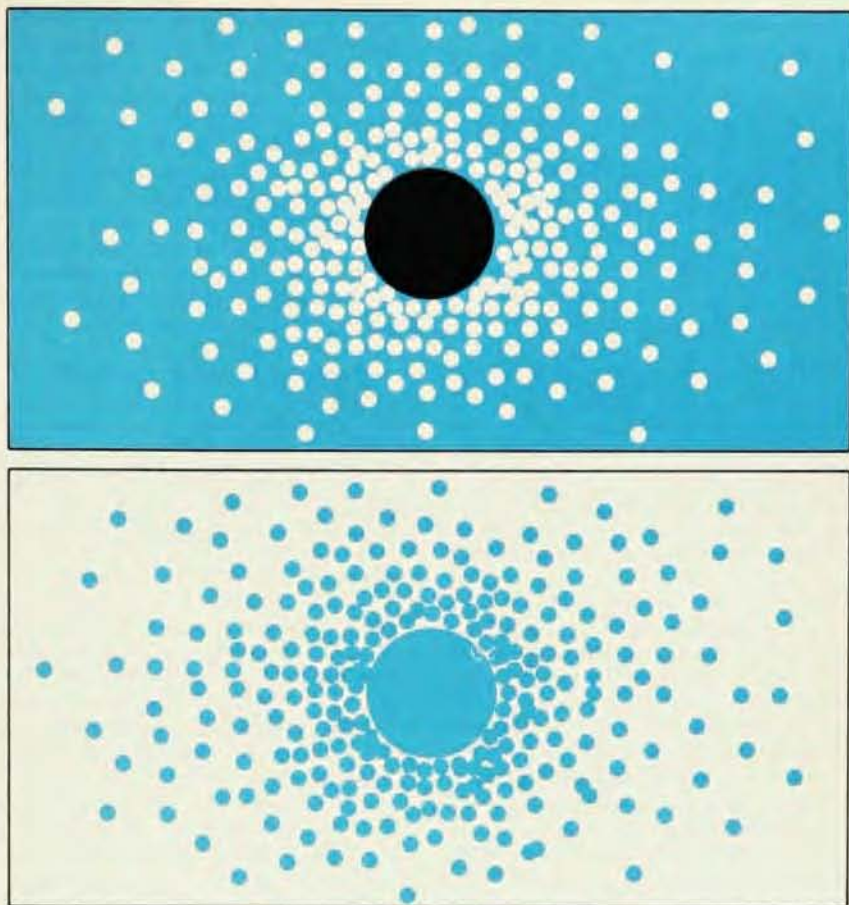
The handedness of a massive particle can of course be changed by bringing it to rest and accelerating it in the opposite direction without changing its spin. Thus, massive particles have both left-handed and right-handed components. However, the neutrinos apparently have zero rest mass; so like photons they always travel at the speed of light. (Recent experiments may lead to some changes in this picture—see PHYSICS TODAY, July, page 17). Because a neutrino cannot be stopped, its handedness never changes. In fact, experimentally only left-handed neutrinos and right-handed anti-neutrinos have been observed. (The antiparticle of a right-handed particle is left handed.) Their oppositely spinning counterparts are presumed not to exist.

Some of the quarks and leptons carry no weak charge at all (they are "singlets" under the SU(2) symmetry) while the same particles with oppositely directed spins are members of weak-charged doublets. The singlets include

$e^-_R, e^+_L, u_R, \bar{u}_L, d_R, \bar{d}_L$
and the corresponding doublets are

$\begin{matrix} \nu_L & e^+_R & u_L & \bar{d}_R \\ e^-_L & \bar{\nu}_R & d_L & \bar{u}_R \end{matrix}$

The subscripts indicate the left or right handedness of the particles. The strange, charmed, truth and beauty



Vacuum polarization. In QED (a) a point charge within the large spot is surrounded by a cloud of virtual charges of the opposite sign. In QCD (b) a point "charge" within the large spot is surrounded by virtual quarks and gluons with predominantly the same color as the central quark.

quarks and the leptons mu and tau and their associated neutrinos fall into similar singlets and doublets.

The weakly charged doublets couple, via one of the charge-matrices R_i , to the W particles. The exchange of a W^+ or W^- (or, of a W_1 or W_2) gives rise to the weak force observed in nuclear physics. The beta decay of a neutron, for example, involves a virtual W^- : a d quark emits a W^- , becoming a u quark; the W^- in turn decays into an electron and an antineutrino.

The dimensional coupling constant g_2 that appears in the gauge transformation has the dimensionless equivalent $\alpha_2 = g_2^2/\hbar c$, which is about $1/30$ at a distance of 10^{-16} cm.

SU(2) $\times$ U(1)

Because the W^+ and W^- carry electric charge, we expect the weak and electromagnetic interactions to be related. In fact, we can construct a relationship between the charges:

$$Q = R_3 + S$$

The charge S is the same for each component of a weak SU(2) multiplet. For particles that are singlets under weak SU(2), $R_3 = 0$ and S is just the

electric charge. For SU(2) doublets, $R_3 = \pm 1/2$ and S is the average electric charge of the doublet. For example, the right-handed electron has $R_3 = 0$, $S = -1$ and $Q = -1$; the left-handed electron neutrino and electron have $S = -1/2$ and, respectively, $R_3 = +1/2$ and $-1/2$ and $Q = 0$ and -1 .

As for the other charges we have discussed, S is associated with a gauge invariance. Particles for which $S = 0$ can emit a gauge particle, which we shall call V ; we shall denote the dimensionless coupling constant for this process α_1 . But because the charges Q, R_3 and S are related, the three gauge fields (or, equivalently, the particles γ, W_3 and V) are related as well. The photon field (the vector potential, A) is a linear combination of the fields W_3 and V :

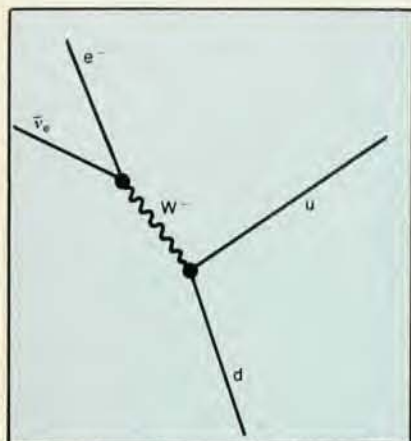
$$A = V \cos \theta_w + W_3 \sin \theta_w$$

There is, of course, another, independent linear combination of the V and W_3 fields; it is called the Z^0

$$Z^0 = W_3 \cos \theta_w - V \sin \theta_w$$

The parameter θ_w is called the weak mixing angle; its value is determined by the coupling strengths e, g_2 and g_1 .

Like the photon, the Z^0 has zero

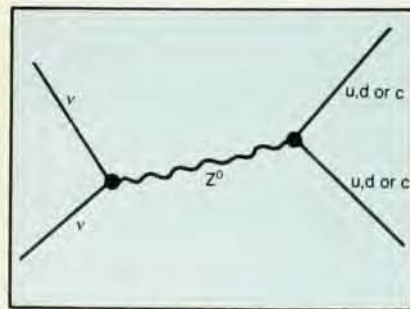


Beta decay. A d quark within a neutron emits a virtual W^- , becoming a u (the neutron thus turns into a proton). The virtual W^- in turn decays into an electron and an antineutrino.

electric charge and can be emitted without changing a particle's identity. But, unlike the photon, the Z^0 can be emitted and absorbed directly by neutrinos. The virtual exchange of Z^0 gives rise to a weak force between neutrinos and the other particles; this is called a "neutral-current" interaction because the exchanged particle is electrically neutral. In contrast, the classic weak interaction shown in the graph above is called a "charged-current" interaction. The neutral-current interaction is very different and much harder to observe because no change in particle identity is involved. Recent experiments, however, have in fact seen the results of weak neutral currents, thus confirming this aspect of the unified theory of weak and electromagnetic interactions. (See *PHYSICS TODAY*, November 1973, page 17, and September 1978, page 17). Because the theory combines an $SU(2)$ symmetry for some of its fields (the $W_\pm$) and a $U(1)$ symmetry for another (V), it is called an $SU(2) \times U(1)$ theory.

Spontaneous symmetry breakdown

Something crucial is missing from the above description of weak and electromagnetic interactions. The electro-



Neutral-current weak interactions have recently been detected experimentally. They involve the exchange of a virtual Z^0 particle.

magnetic force has an infinite range ($F \propto 1/r^2$), and is mediated by a particle (the γ) with zero mass. The weak interactions, however, have a very short range ($\sim 10^{-16}$ cm), and the exchanged particles, the W^+ , W^- and Z^0 must be massive—indeed about 100 times more massive than a proton. What has happened to gauge invariance, which tells us that the gauge particles have zero mass? The answer is that the underlying force law is symmetrical but the vacuum is not. The vacuum distinguishes the W^+ , W^- and Z^0 from the photon, making them heavy but leaving the photon massless. The structure of the vacuum breaks the $SU(2) \times U(1)$ gauge invariance and allows the gauge particles to develop mass. Such a situation is called a spontaneous symmetry breakdown.

A useful analogy is the breakdown of rotational invariance in a crystal. Consider, for example, a grain of salt. It is built up out of sodium ions and chloride ions. The forces between the ions are electromagnetic forces, which do not pick out any special directions in space. However, when the ions are packed together, they form a cube. If you lived inside a grain of salt, you would find that your space does have some special directions; the directions perpendicular to the faces of the cube, for example, have special properties. It is true that a "giant" living outside the cube of salt could pick it up and rotate it (and you with it) without affecting the physics of your world, but to you inside the cube, the rotational symmetry of the laws of physics would not be obvious. The rotational symmetry has been spontaneously broken by your environment, the salt crystal.

The relativistic quantum-mechanical vacuum, in which we all live, is like the salt crystal. The symmetry which is broken spontaneously is not rotational invariance, but the gauge invariance which, if unbroken, would make the W^+ , W^- , and W^0 interchangeable and massless. The vacuum picks out a direction in this "gauge space" so that the W^0 direction is singled out. In the process the W^+ , W^- , and the Z^0 (a particular combination of W^0 and V^0 , remember) get a mass, while the photon remains massless because the electromagnetic gauge invariance is not broken by the vacuum.

The gauge theory of the weak and electromagnetic interactions, with spontaneously broken $SU(2) \times U(1)$ symmetry, was worked out in the 1960's by Glashow at Copenhagen (now Harvard), Steven Weinberg at MIT (now also at Harvard) and Abdus Salam at Imperial College, London, and the International Centre for Theoretical Physics in Trieste. Glashow worked out the form of the gauge theory but did not

know how to give mass to the W and the Z . Weinberg and Salam worked out the effect of the spontaneous symmetry breakdown (using a mechanism developed earlier by Peter Higgs at the University of Edinburgh and T. W. B. Kibble at Imperial College) and produced a consistent theory. Glashow, Salam and Weinberg received the Nobel Prize for this work in 1979.

There is another aspect of the spontaneous breakdown of the gauge symmetry which can be understood in the crystal analog. It is "graininess". The salt crystal, after all, is grainy on an atomic scale 10^{-8} cm. Living inside the crystal, you might well distinguish three different domains of distance, associated with different physics:

- Distances much smaller than 10^{-8} cm. If you do an experiment that probes the structure of your world at distances much smaller than 10^{-8} cm, it doesn't matter very much that you are living in a crystal. You will find that the results of your experiment will be approximately rotationally invariant. This is because the nonuniform electric fields that mess up the crystal's rotation invariance are very weak compared with the kinds of fields you have to generate to probe subatomic distances. Once you are inside the atom, it is the constituents of the atom that matter, not how the atom is put together with other atoms in the crystal.
- Distances of the order of 10^{-8} cm. Here you will see all of the complicated interatomic forces that are responsible for packing the atoms into a crystal.
- Distances much larger than 10^{-8} cm. Here you will see the cubic structure of the crystal; there is no apparent rotational invariance that one can see at this scale. All that remains is the discrete symmetry group associated with the crystal structure.

Just as a crystal is grainy at an atomic scale, the vacuum is grainy at a smaller scale, on the order of 10^{-16} cm, associated with the spontaneous breakdown of the $SU(2) \times U(1)$ symmetry. Again, we can identify three distinct regions of distance scale:

- Distances much smaller than 10^{-16} cm. Here you will see the $SU(2) \times U(1)$ gauge structure of the world as an explicit (though approximate) symmetry. At such short distances, the masses of the W^+ , W^- , and Z^0 (which are around 100 GeV) are negligible compared to the energies required to do the experiment. (To probe down to $\Delta x \sim 10^{-16}$ cm requires an energy about 10^4 GeV.) Thus, the fact that the W and Z^0 are massive while the photon is not can be ignored. All the particles are light compared to the energy of the experiment.

- Distances of the order of 10^{-16} cm. Here your life will be very complicated. You will see the physics responsi-

ble for spontaneously breaking the $SU(2) \times U(1)$ symmetry. W 's and Z^0 's will show up in your experiments, but they will look very different from photons, because their rest energies are not small compared to the energies of your probes.

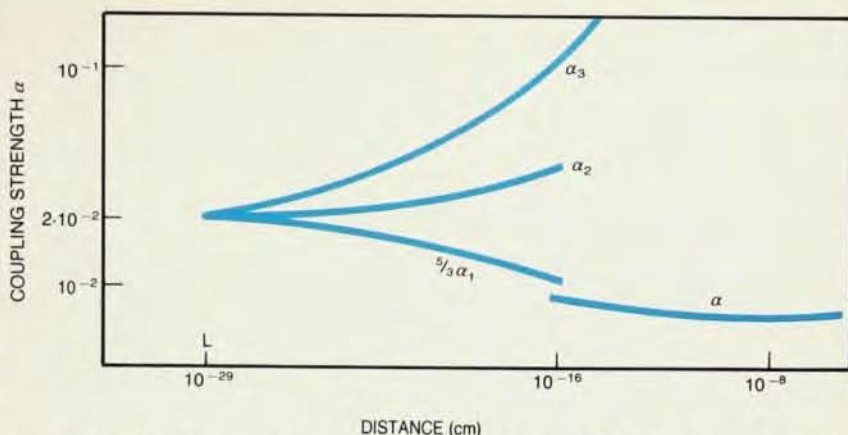
► Distances much greater than 10^{-16} cm. Here you will not see the $SU(2) \times U(1)$ gauge symmetry at all. You will not even see W 's and Z^0 's directly. You do not have a probe with high enough energy to produce them. You see the unbroken electromagnetic gauge invariance directly. But the heavier gauge particles only show up in the weak short-range interactions caused by their virtual exchange. These interactions would hardly be noticeable except that they do things the electromagnetic and strong interactions do not do. The W^+ and W^- exchanges change particle identities. These show up in beta decay. The Z^0 exchanges cause neutral-current interactions of neutrinos. And they all violate parity symmetry because the left- and right-handed components of the quarks and leptons interact differently.

This last region is the domain of contemporary particle-physics experiments; indeed, all of the short-range weak effects we have described have actually been observed. The fact that everything fits together as expected gives us confidence that the picture of weak interactions as a spontaneously broken $SU(2) \times U(1)$ gauge theory is correct, even though no one has seen a W or a Z^0 .

A unified theory

The $SU(2) \times U(1)$ gauge theory we have just described is a partial unification of the weak and electromagnetic interactions. It describes the charged-current and the neutral-current weak interactions and the electromagnetic interactions, which at first sight look very different as gauge interactions. But, although the theory is unified, there are still two different interactions involved, associated with the $SU(2)$ and the $U(1)$ groups. There are two dimensionless coupling constants: α_2 for the $SU(2)$ gauge particles (the three W 's) and α_1 for the $U(1)$ gauge particles (the V^0). Their values can be determined experimentally because α , the ordinary electromagnetic coupling constant is given by $1/\alpha = 1/\alpha_1 + 1/\alpha_2$ and the weak mixing angle θ_w is given by $\sin^2 \theta_w = \alpha/\alpha_2$.

Also, the electric charge is only partially quantized. Within an $SU(2)$ multiplet (such as the doublet e^-_L, ν_L) all charge differences are integers, because a particle in a multiplet can emit a $W^\pm$ and become another member of the multiplet. But the average charge of the multiplet is not quantized. The average charge is the $U(1)$ charge S ,



Coupling strengths vary with distance. At distances up to the "unification distance" L all three constants are the same (up to a group-theoretic factor of $5/3$ for α_1). The strong coupling, α_3 , becomes large at distances greater than 10^{-14} cm, signalling the breakdown of QCD perturbation theory due to quark confinement. At distances larger than 10^{-16} cm the spontaneous breakdown of $SU(2) \times U(1)$ collapses α_2 and α_1 into the single electromagnetic coupling α .

and, as in the case of the electromagnetic charge in the QED theory, there is no theoretical reason for it to be quantized. And finally, there is still no connection between quarks and leptons. But, now we have all the pieces to the puzzle, and we can fit them together into a unified theory.

If we merely combined the three gauge theories we would have a theory whose symmetry is $SU(3) \times SU(2) \times$

$U(1)$. Such a combination of symmetries can be components of an $SU(5)$ symmetry. This is a symmetry group whose fundamental representation has five elements (like the triplets of colored quarks or the weak doublets), and whose charges are 5×5 traceless Hermitian matrices.

The five-element vectors that form the basis of the group representation (analogous to the color-triplets of

The fundamental particles in $SU(5)$

Particle	Representation	"metacolor"	Q	T_3	T_8	R_3
d^i_R	5	a	-1/3	1/2	$1/2\sqrt{3}$	0
d^w_R	5	b	-1/3	-1/2	$1/2\sqrt{3}$	0
d^b_R	5	c	-1/3	0	$-1/\sqrt{3}$	0
e^+_R	5	d	1	0	0	1/2
ν_R	5	e	0	0	0	-1/2
u^i_L	10	a + d	2/3	1/2	$1/2\sqrt{3}$	1/2
u^w_L	10	b + d	2/3	-1/2	$1/2\sqrt{3}$	1/2
u^b_L	10	c + d	2/3	0	$-1/\sqrt{3}$	1/2
d^i_L	10	a + e	-1/3	1/2	$1/2\sqrt{3}$	-1/2
d^w_L	10	b + e	-1/3	-1/2	$1/2\sqrt{3}$	-1/2
d^b_L	10	c + e	-1/3	0	$-1/\sqrt{3}$	-1/2
e^+_L	10	d + e	1	0	0	0
$\bar{u}^i_L$	10	b + c	-2/3	-1/2	$-1/2\sqrt{3}$	0
$\bar{u}^w_L$	10	a + c	-2/3	1/2	$-1/2\sqrt{3}$	0
$\bar{u}^b_L$	10	a + b	-2/3	0	$1/\sqrt{3}$	0
$\bar{d}^i_L$	$\bar{5}$	a	1/3	-1/2	$-1/2\sqrt{3}$	0
$\bar{d}^w_L$	$\bar{5}$	b	1/3	1/2	$-1/2\sqrt{3}$	0
$\bar{d}^b_L$	$\bar{5}$	c	1/3	0	$1/\sqrt{3}$	0
e^-_L	$\bar{5}$	d	-1	0	0	-1/2
ν_L	$\bar{5}$	e	0	0	0	1/2
$\bar{u}^i_R$	$\bar{10}$	a + d	-2/3	-1/2	$-1/2\sqrt{3}$	-1/2
$\bar{u}^w_R$	$\bar{10}$	b + d	-2/3	1/2	$-1/2\sqrt{3}$	-1/2
$\bar{u}^b_R$	$\bar{10}$	c + d	-2/3	0	$1/\sqrt{3}$	-1/2
$\bar{d}^i_R$	$\bar{10}$	a + e	1/3	-1/2	$1/2\sqrt{3}$	1/2
$\bar{d}^w_R$	$\bar{10}$	b + e	1/3	1/2	$1/2\sqrt{3}$	1/2
$\bar{d}^b_R$	$\bar{10}$	c + e	1/3	0	$1/\sqrt{3}$	1/2
e^-_R	$\bar{10}$	d + e	-1	0	0	0
u^i_R	$\bar{10}$	b + c	2/3	1/2	$1/2\sqrt{3}$	0
u^w_R	$\bar{10}$	a + c	2/3	-1/2	$1/2\sqrt{3}$	0
u^b_R	$\bar{10}$	a + b	2/3	0	$1/\sqrt{3}$	0

quarks, for example) contain both quarks and leptons, but only of a single handedness. One of these, for instance, is

$$d_R^r \quad d_R^w \quad d_R^b \quad e_R^+ \quad \bar{\nu}_R$$

All of these are related by an SU(5) symmetry; the first three are related by the SU(3) color symmetry, the last two by the weak SU(2).

The twenty-four SU(5) matrices are coupled to twenty-four gauge fields. Eight of the matrices have elements only in the upper left 3×3 sub-matrix; these are coupled to the eight gluons. Four of the matrices are coupled to the $W^\pm, \gamma$ and Z^0 . The photon, for example couples to

$$Q = \begin{pmatrix} -1/3 & 0 & 0 & 0 & 0 \\ 0 & -1/3 & 0 & 0 & 0 \\ 0 & 0 & -1/3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

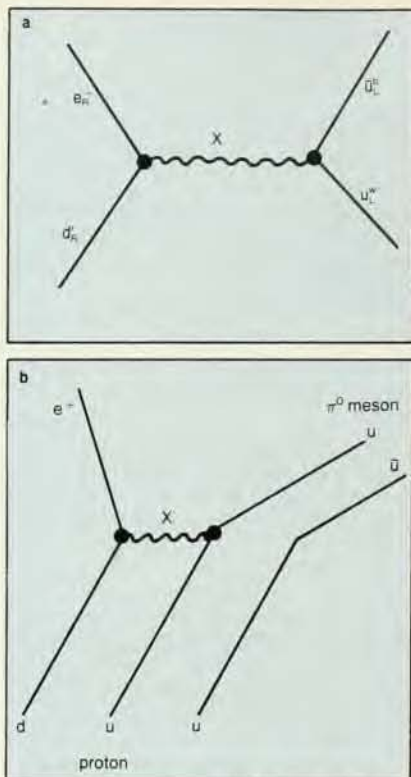
The W particles couple to charges that contain non-zero elements only in the lower right 2×2 sub-matrix.

The remaining twelve gauge particles cause transitions between the quarks and antileptons of the fundamental five. We shall call them X 's.

Because SU(5) is a simple group, all the charges—including, of course, Q —are quantized. We have thus finally arrived at a theoretical understanding of charge quantization.

At least, we have arrived at a reason for charge quantization of the right-handed d -quarks, positrons and antineutrinos—and, consequently, of their left-handed antiparticles. But something even better is in store for us. The next simplest family with an SU(5) gauge invariance can be built out of the five fundamental objects by forming ten pairs. (In an analogous way, a spin-zero state can be formed from pair of spin- $1/2$ states.) The ten states formed from pairs of right-handed states have all the right charges (color, electric and weak) to form the remaining left handed particles and antiparticles; those formed from pairs of left-handed states make up the remaining right-handed states. The process is illustrated in the table on the previous page. To clarify the way individual states are built up out of the fundamental 5 (and their antiparticles, called $\bar{5}$) we have denoted the analogs of the SU(3) "colors" (r, w, b)—the "metacolors," so to say—by letters $a-e$. The table also gives the values of the (diagonal) color-charges T_3 and T_8 and of the (diagonal) weak charge R_3 for each of the particles. (This is, of course, not in any sense a dynamical building process.)

In the language of group theory, the quintuplets 5 and $\bar{5}$ are the smallest irreducible representations of SU(5). The antisymmetric combination of two



Proton decay. (a) One of the superheavy gauge particles X can be exchanged when two quarks come within about 10^{-29} cm of each other. (b) When such an exchange occurs in a proton it decays into a positron and a pion.

5 's (or two $\bar{5}$'s) is a 10 (or a $\bar{10}$) and is the next-smallest representation of SU(5). These representations are all we need to describe all the quarks and leptons. The other leptons and quarks with other flavors also have the same SU(5) representations: The quarks s and c together with the leptons μ and ν_μ form another SU(5) "family," as do the quarks t and b with the τ and ν_τ . There is, in fact some astrophysical evidence that there are not very many such families. (See "Cosmology and elementary-particle physics" by Michael Turner and David Schramm, PHYSICS TODAY, September 1979, page 43.)

It is important to realize that this building process did not have to work. It represents the first, the simplest, and in some ways the most remarkable success of the SU(5) unification.

Providing a reason for charge quantization is, to us, one of the nicest aspects of the theory. Charge quantization follows simply from the building process. Even if more complicated SU(5) families exist, Q must still be quantized in multiples of $1/3$, because all charges must be sums of the charges of the simplest family. Furthermore, the fact that all observed systems have integral Q is connected with their color neutrality. The only way to "build" color-neutral systems

out of the simplest family is to use the particles $e^+ \bar{\nu}_R$ or $\bar{\nu}_R$ or the combination $d_R^r + d_R^w + d_R^b$, each of which has integral charge.

In SU(5) charge quantization is trivial because the charges in the 5 are "commensurate." Actually, however, the quantization of charge is somewhat deeper: It has a topological aspect and is related to the existence of magnetic monopoles in gauge theories that are based on simple groups, which in turn requires charge quantization.

This SU(5) unification was worked out by us at the end of 1973.

The unification scale

All of the charges in SU(5) are treated symmetrically, so in some sense, there ought to be only one coupling constant describing all of the interactions. But that is certainly not the situation at distances of the order of 10^{-16} cm and larger. There α_3 is greater than α_2 which, in turn, is greater than α_1 . For this reason, and for others to which we will come below, there must be another level of spontaneous symmetry breakdown, this one associated with the breaking of SU(5) down to SU(3) $\times$ SU(2) $\times$ U(1).

We must assume that the vacuum is grainy and structured at another scale L much smaller than the 10^{-16} cm associated with spontaneous breakdown of SU(2) $\times$ U(1). As before, there are three different regions of distance scale:

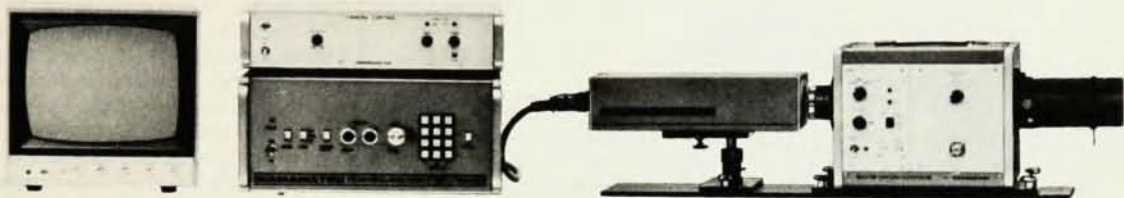
► Distances much less than L . Here you will see the SU(5) gauge invariance as an explicit, approximate symmetry. All the gauge particles, the photon, the gluons, the W 's, and Z^0 and the X 's are light compared to the energies required to probe these short distances.

► Distances of the order of L . At these distances you will see the complicated physics of the spontaneous breakdown of SU(5). The X 's will be produced, but will be much heavier than all the other gauge particles, with a mass of order $\hbar c/L$.

► Distances much greater than L . Here you will see only the SU(3) and the SU(2) $\times$ U(1) gauge invariance directly. The X 's are too heavy to be produced directly, but they do give rise to very weak, very short-range interactions.

We can now understand the disparity in coupling constants. In the first region, at very short distances where the SU(5) gauge symmetry is explicit, there is indeed only a single coupling constant describing all of the gauge interactions. But in the second region the symmetry is broken so that for larger distances the couplings diverge. Since the SU(3) coupling constant α_3 is the most asymptotically free, it increases faster than α_2 as the distance at which it is measured increases. The

WHICH STREAK CAMERA SYSTEM IS BEST FOR YOU?



10 PICOSECOND TIME RESOLUTION

TEMPORALDISPERSER MODEL C-979**

- Exclusive streak tube with built-in MCP*
- Dynamic range better than 1:100
- Trigger jitter less than ± 50 ps
- Streak speeds full scale 1, 2, 5, 10 ns fast plug-in, or 10, 20, 50, 100 ns slow plug in
- Easy to set up and get operating first day
- Outstanding record of dependable operation worldwide

DELIVERY 30-90 DAYS

2 PICOSECOND TIME RESOLUTION

TEMPORALDISPERSER MODEL C-1370
NEW

- Time resolution better than 2 ps
- Exclusive streak tube with built-in MCP*
- Trigger jitter less than ± 50 ps
- Broad spectral range 200~850 nm (with use of optional UV optics)
- Easy to operate
- Real time analysis

DELIVERY 9-12 MONTHS

X-RAY

TEMPORALDISPERSER MODEL C-1102X
NEW

- Complete with 5° bias custom flange to fit your vacuum chamber port.
- Time resolution better than 30 ps
- Trigger jitter less than ± 50 ps
- Exclusive streak tube with built-in MCP*
- Variable sweep speeds controllable at the front panel
- Optional SIT camera/temporal analyzer readout system available

DELIVERY 6-9 MONTHS

1 NANOSECOND TIME RESOLUTION

TEMPORALDISPERSER MODEL C-1155**
NEW

- Ideal for applications too slow for picosecond cameras
- Time resolution better than 1 ns
- Wide range of streak time from 100 ns to 50 μ s full screen in a single unit
- Exclusive streak tube with built-in MCP*
- Dynamic range better than 1:100
- Small size, light weight

DELIVERY 6 MONTHS

*Each of the four models features an exclusive Hamamatsu streak tube with built-in microchannel plate (MCP). Provides weak signal amplification greater than 3×10^3 , needs no external image intensifier.

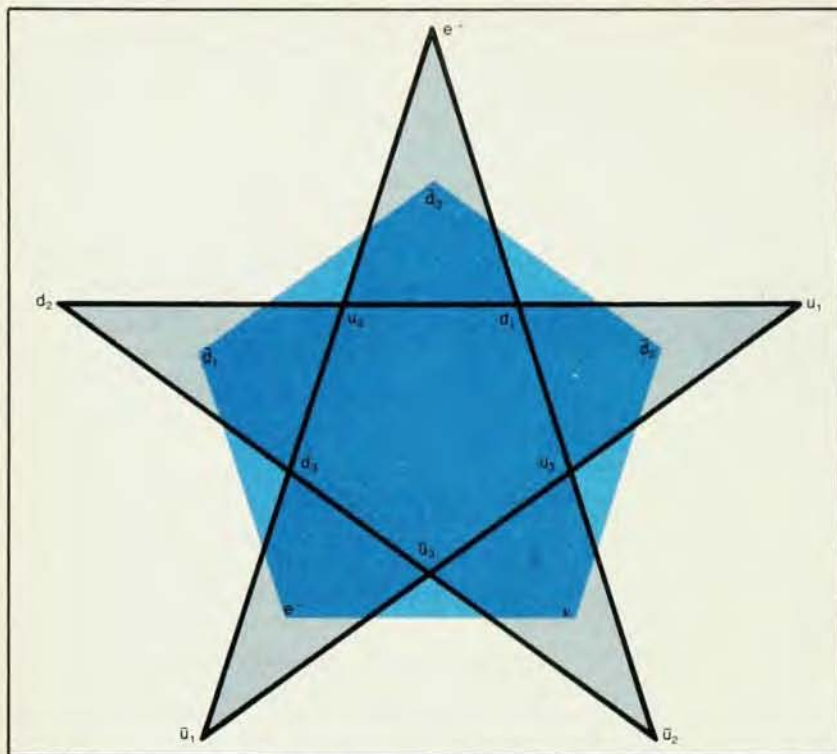
**Optional S-1 photocathode available for Models C-979 and C-1155. Extends spectral range to 180nm~1100nm.

CALL OR WRITE FOR DESCRIPTIVE LITERATURE

HAMAMATSU

HAMAMATSU CORPORATION • 420 SOUTH AVENUE • MIDDLESEX, NEW JERSEY 08846 • PHONE: (201) 469-6640
International Offices in Major Countries of Europe and Asia.

Circle No. 17 on Reader Service Card



Relationships among particles in SU(5). Each of the fundamental particles has some combination of the four independent charges and can be plotted as a point in a four-dimensional space. The diagram shows a two-dimensional projection of such a plot which preserves some of the symmetry of the four-dimensional original. The states shown are all left-handed particles. The pentagon represents the group of states that transforms according to the "5" representation of SU(5), while the five-pointed star and the inner pentagon represents the "10." The vertical axis is a linear combination of T_3 , R_3 and Q ; the horizontal axis is a combination of T_3 and R_3 .

U(1) coupling constant α_1 , on the other hand, is not asymptotically free at all (because the gauge particle does not carry the charge). So, α_1 actually decreases as the distance increases. This situation is depicted in the graph on the previous page.

The figure shows, qualitatively, that at distances much larger than L , the couplings have the right form, with α_3 greater than α_2 greater than α_1 . But we can do better. At distances smaller than 10^{-16} cm, all of the couplings are rather small, and their dependence on the distance at which they are measured is actually calculable. Then, if two of the couplings at 10^{-16} cm are known, they can be followed to shorter and shorter distances until they meet. This will be a distance of the order of L . Furthermore, one can then predict the values for distances larger than L . For example, one can use α (the combination of α_2 and α_1 that occurs in the electromagnetic interactions), which is known very well experimentally, and α_3 (the gluon coupling constant), which has been measured, but with rather large errors, and then estimate L and θ_w . The results are that L is about 10^{-29} cm and $\sin^2\theta_w$ is about 0.20.

This prediction was worked out at Harvard in 1974 by Georgi, Helen

Quinn (now at SLAC) and Weinberg for a large class of unified theories, including SU(5). At the time, it looked bad for unification because the experimental value of $\sin^2\theta_w$ was about 0.35.

This discrepancy between the predicted and observed values didn't bother the two of us at the time because we were not certain that the SU(3) and the SU(2) $\times$ U(1) were correct descriptions of physics at 10^{-16} cm, so we thought that SU(5) might have to be generalized to include new physics. However, in the years between 1974 and the present, the SU(3) and SU(2) $\times$ U(1) theories have passed a number of important experimental tests so that today we are much more confident in the underpinnings of the SU(5) theory. And, fortunately for the theory, the experimental value of $\sin^2\theta_w$ has moved down, and today is $0.23 \pm .02$, almost in agreement with the unification prediction.

Proton decay

The X particles in the SU(5) unified theory are very heavy because they are associated with a spontaneous symmetry breakdown at a very short distance. Their mass is about $\hbar c/L$, or 10^{15} times the mass of the proton. We can never produce them in accelera-

tors, but it may be possible to observe the effect of their virtual exchange. The interactions that are mediated by X-exchange are very weak because they have such a short range. But, like the ordinary weak interactions, these very weak interactions cause processes that could not occur at all without them.

The most interesting of these are interactions that change baryon number. Baryon number is defined as one third the number of quarks minus one third the number of antiquarks. The baryon number of a proton is one while the baryon number of a meson or a lepton is zero.

The SU(3) and the SU(2) $\times$ U(1) gauge interactions do not change baryon number: They conserve it, because their gauge particles do not cause transitions between quark and lepton or quark and antiquark. If baryon number were exactly conserved, the proton would have to be absolutely stable because it is the lightest particle with baryon number one. Because proton decays have not been observed (although some experimenters have tried to find it) it was quite reasonable to assume that baryon number is absolutely conserved, and most physicists did.

The processes that change baryon number, which are caused by the X particles in the SU(5) theory, can, however, lead to proton decays. One of the X particles, for example is coupled to "metacolors" a and d . A d_R quark can emit such an X change into an e^+ ; also, a u_L quark can absorb the same X and change into a $\bar{u}_L$. Exchanging an X can thus turn a pair of quarks, d and u , into a lepton and an antiquark, e^+ and $\bar{u}$. The baryon number of the system changes by one unit, from $2/3$ to $-1/3$.

The sketch on page 36 shows what happens if this process takes place inside a proton: It decays into a positron and a neutral pion. The decay is extremely rare. Because the X is so heavy, two quarks must come very close together to exchange it.

Quinn, Weinberg, and Georgi used the calculation of L to estimate the proton decay rate. This estimate has since been refined by many people, including Andrzej Buras, John Ellis, Mary Gaillard, and Demetres Nanopoulos, from CERN, Terry Goldman and Douglas Ross from Caltech and William Marciano from Rockefeller. The present estimate is that the decay rate is about one decay per proton every 10^{31} years.

Salam and Jogesh Pati of the University of Maryland independently, and in a different context, suggested that the proton might be unstable. Their model is based on a different version of the QCD theory of hadrons.

What is the chance of observing pro-

ton decay directly? How is it possible to measure a proton lifetime of about 10^{31} years when the age of the universe since the big bang is only about 10^{10} years? The answer is that you get yourself a lot of protons. For example, in 1000 tons of matter there are about 5×10^{32} protons and neutrons. We expect about 50 of these to decay each year. So you simply have to monitor everything that goes on inside 1000 tons of matter and distinguish proton and neutron decays from everything else, and you should detect baryon number changing processes.

Several groups are planning experiments on this scale (PHYSICS TODAY, January, page 17). The experiments are to be done underground to minimize the confusion caused by cosmic rays interacting in the sample of matter. One experiment is planned to take place in a salt mine near Cleveland, another in a silver mine in Utah, and yet another in an iron mine in Minnesota. On a slightly smaller scale are European experiments planned for tunnels under the Alps, an ongoing experiment in a gold mine in South Dakota, and an Indian-Japanese collaboration in the Kolar gold field.

Another very exciting idea has come out of the unification: Although we cannot build such a machine ourselves, there is one machine that may have directly probed the region of distances on the order of 10^{-29} cm—the big bang of the universe itself. About 10^{-40} sec after the singularity that signaled the beginning of things the universe had expanded to a radius comparable to L . In this unimaginable era, the breakdown of the unifying gauge invariance was just beginning to appear, and the interactions that change baryon number were as strong as anything else. It seems possible that they may have produced more baryons than antibaryons. If so, unification is the solution to an old puzzle in astrophysics: Why is the universe built out of matter rather than anti-matter?

This interesting speculation was first made by Motohiko Yoshimura of KEK in Japan, and was subsequently elaborated by many physicists including Ellis, Gaillard, Nanopoulos, Weinberg, Leonard Susskind from Stanford and Sam Treiman and Frank Wilczek from Princeton. They have shown that to produce an excess of baryons over antibaryons requires that the baryon-number changing interactions look different when they are run backwards in time. This condition is satisfied in the $SU(5)$ theory.

Astrophysical observations and experiments deep within the Earth may thus provide clues about the structure of the world at 10^{-29} cm, a distance so small we could never hope to probe it directly. □

MAGNETIC SHIELDING

**Material For PMT's,
Power Supplies, Transformers,
CRT's, Relays & other components:**

- CO-NETIC AA ALLOY —
High Permeability
.002" to .100" thick

EXCLUSIVE: Perfection Annealed —
No further anneal required if severe
forming is avoided.

- NETIC S3-6 ALLOY — High Saturation Induction.
.004" to .095" thick



MAGNETIC SHIELD DIVISION

PERFECTION MICA CO.
740 North Thomas Drive
Bensenville, Ill. 60106, USA
Phone 312 / 766-7800
TWX 910-256-4815

Send for NEW
Material, Application
and Fabrication
Guide MG-5

Circle No. 18 on Reader Service Card

pulsar

BOTHERED BY ARCS AND SPARKS?



USE SYSTEM-60 FOR DATA TRANSMISSION

pulsar associates, inc.

high voltage electronics...

11491 Sorrento Valley Road, San Diego, CA 92121
(714) 455-5933 TELEX 181768 PULSAR INC SDG

Circle No. 19 on Reader Service Card