Elegant Connections in Physics:

A Cursory Glance at the "Standard Model" of Elementary Particle Physics

— by Dwight E. Neuenschwander (Second from left.)

As the Society of Physics Students (SPS) anticipates the 2008 Sigma Pi Sigma Congress at Fermi National Accelerator Laboratory, we might contemplate some of the concepts that keep Fermilab physicists (and their competitors) awake at night. If everything in the universe is built ultimately out of irreducible particles, then the fundamental forces in nature must be understandable as operating between those particles. This article attempts to offer an overview of the so-called "standard model" of elementary particle physics. I assume the background such as one encounters in courses in "Modern Physics," elementary quantum mechanics, and electrodynamics. I offer this primer because this jargon-laden topic forms a complicated subject filled with long chains of abstract reasoning and caveats galore. But it is guided by a small number of powerful fundamental principles, such as conservation laws and local gauge invariance. It never ceases to amaze me how such abstract reasoning about objects that we cannot directly see, makes predictions that are rendered observable through still more chains of reasoning expressed through technology.

If this primer at least makes some sense of the jargon then it will have served a useful purose. It is organized into five sections:

I. The Cast of Characters and Their Environment
II. Fermion Currents
III. The Charges
IV. Radiation Particles: Vector Bosons
V. Unification, Mass, and the Higgs Boson.

I. The Cast of Characters and Their Environment

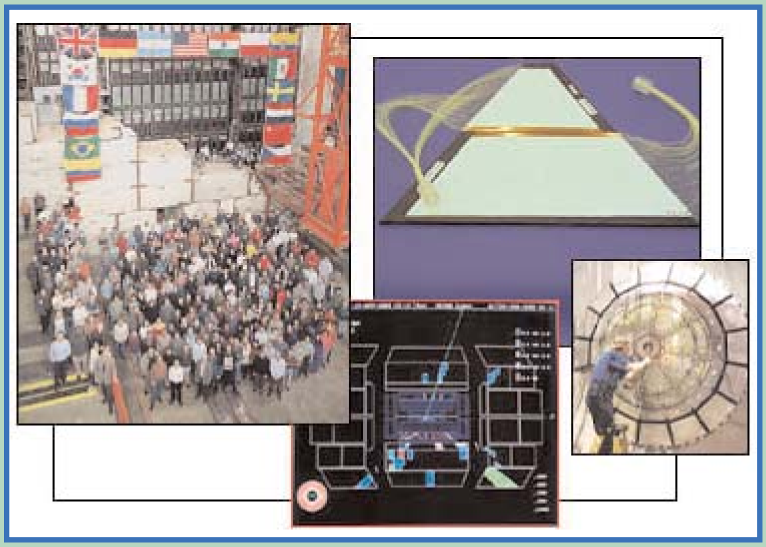
Matter is held together, and signals are transmitted from one event to another, ultimately by the interactions of elementary particles. At this level we know of 3+1 fundamental forces. I say "3+1" because the "1" is gravitation, that is negligible between elementary particles until we get to distance scales ~10^{-33} cm or smaller; hence we ignore gravitation here; and we do not yet have a self-consistent quantum theory of gravity anyway. The three fundamental forces about which we know something meaningful at the level of elementary particles are:

(1) electromagnetism that binds electrons to the nuclei of atoms and makes light;

(2) the "weak" force whose most prominent manifestation is beta decay; and

(3) the "strong" (or "chromodynamic") force that holds protons and neutrons together. Despite profound differences between these forces, they can be described in a common vocabulary. Its main ingredients are:

(1) "matter particles";
(2) the "charges" they carry; and
(3) "radiation" particles exchanged between the charged matter particles.



The DZero Experiment at Fermilab

Photo credit: Fermilab.

The DØ Experiment consists of a worldwide collaboration of scientists conducting research on the fundamental nature of matter. The experiment is located at the world's premier high-energy accelerator, the Tevatron Collider, at the Fermi National Accelerator Laboratory (Fermilab) in Batavia, IL, USA. The research is focused on precise studies of interactions of protons and antiprotons at the highest available energies. It involves an intense search for subatomic clues that reveal the character of the building blocks of the universe.

Charges and Coupling Constants

Each fundamental force corresponds to a characteristic charge carried by particles that participate in the force. The charge can be factored into a "coupling constant" and a dimensionless quantized coefficient. For instance, the electric potential energy between two electrons is proportional to e^2, illustrating that the coupling constant for the electromagnetic force is the electric charge $e = 1.6\times10^{-19}$ C. Particles that carry electric charge do so in quantized amounts qe where, for example, $q = -1$ for the electron and $q = +1$ for the positron.

The first theory of the weak interaction was articulated by Enrico Fermi in 1934[1] in the context of beta decay,

$$n \to p + e^- + \bar{\nu}_e \,,$$

where e^- denotes the electron and $\bar{\nu}_e$ an electron-type antineutrino. In writing the interaction term, Fermi introduced a constant G_F, to be fit to data. The fit gave $G_F = 1.16\times10^{-5}\ \text{GeV}^{-2}$, the "Fermi constant"; therefore $\sqrt{G_F}$ can be taken as the weak force coupling constant.

The strong coupling g *varies* with the energy of the interacting particles, as all coupling constants do to some extent. At very high energies (small distances) g becomes quite small, the particles that carry strong charge become approximately free ("asymptotic freedom"). In tems of momentum transfer quarks in hadrons are nearly

(continued on page 2)

free, but it is a high energy density region. Conversely, at low energies g becomes large and the particles carrying strong charge are bound together with maximum strength ("infrared slavery")—but those topics are another story.

The coupling constants for each of the fundamental forces can be converted into a dimensionless strength parameter (a "fine structure constant") for mutual comparison, through the reduced Planck's constant $\hbar$ and the speed of light in vacuum c:

Electromagnetic: $e^2 / \hbar c \approx 1/137$ (at zero momentum transfer)

Weak: $G_F / \hbar c \sim 10^{-7}$

Strong (at the Tevatron): $g^2 / \hbar c \approx 0.11$

For completeness, if G denotes Newton's gravitational constant and m the electron mass, then $Gm^2 / \hbar c \approx 10^{-45}$.

Bosons and Fermions

Every elementary particle has an intrinsic angular momentum, or "spin" **S**. Indeed, one useful definition of "elementary particle" is "a state of definite mass and spin."[2] In quantum theory, to have a spin vector **S** means that there exists a quantum number S such that $|\mathbf{S}|^2 = S(S+1)\hbar^2$. When projected onto a designated z-axis there are $2S+1$ values (the "multiplicity") of S_z, which differ by integer steps from $-S$ to $+S$. The quantum number S itself may take on the values 0, ½, 1, 3/2, 2, ... Particles with integer values of S are "bosons" (from Bose-Einstein statistics: there is no limit on how many identical bosons can be in the same quantum state) and particles with half-odd-integer values of S are "fermions" (from Fermi-Dirac statistics: no two identical fermions can be in the same quantum state).[3] A spin-½ particle has two possible values of $S_z = \pm½$. In specifying a particle's spin state, the spin quantum number S_z is to the unit of angular momentum $\hbar$ as q is to the unit of charge: e and $\hbar$ are the *fundamental* constants; q and S_z assign respectively how many units of charge or spin angular momentum a particle carries.

The elementary particles are the quanta of their respective *fields*. The familiar photon provides the prototypical example: the photon is the quantum of the electromagnetic field. A field extends through all of space and time, and in general varies with the spacetime coordinates. Scalar fields (one number specifies the field value at each spacetime event) have quanta that carry spin 0. There are vector fields, for which the number of *components* to be specified at each event equals the number of spacetime *coordinates*. The quanta of a vector field have spin 1 (e.g., the electromagnetic field's quanta are spin-1 photons). There are also "spinor" fields. Their quanta are the spin-½ fermions. Scalar fields, spinor fields, and vector fields (and tensor fields, with spin 2 quanta, such as the conjectured graviton, which does not concern us here) have distinctive Lorentz transformation properties; for example, vector fields transform the same as do the coordinates themselves; scalar fields are invariant; spinor fields have some distinctive Lorentz transformation properties of their own.[4]

The Fundamental Fermions

Over the last century, physics has shown us that matter is made of atoms whose architecture features electrons swarming about a nucleus that in turn consists of protons and neutrons. The nuclear constituents are not elementary, but have internal parts. So far as we can tell, the elementary constituients of matter are all spin-½ particles called "leptons" and "quarks."[5]

Ordinary familiar matter is made of the stable, lowest-mass, "first generation" leptons and quarks. The first generation leptons are the electron and its neutrino. The first generation quarks are denoted u and d, for their whimsical names "up" and "down." The electron carries $q = -1$, the neutrino $q = 0$, the u quark has $q = +2/3$ and the d quark $q = -1/3$. A proton's quark composition is *uud*, and the neutron's *udd*. The second and third generations of quarks and leptons echo the first generation in all their charges,[6] but have higher masses and are therefore less stable than their first-generation cousins. The second generation leptons are the muon and its neutrino; the second-generation quarks are called c and s (for "charm" and "strange" with their own peculiar quantum numbers for specialized reactions). The third generation includes the tau lepton and its neutrino, and the t ("top") and b ("bottom") quark. The muon's mass is about 207 electron masses, the tau lepton ~ 1000 electron masses. The u and d masses are around 3-10 electron masses; the heaviest quarks, t and b, have masses ~ 170-400 proton masses. "Mass" is not well-defined for a quark because quarks have never been observed in isolation. They are always interacting inside "hadrons," the strongly-interacting quark composites (three quarks for "baryons" fermions like the proton and neutron; and a quark-antiquark pair for integer-spin "mesons"). Thus the u and d masses are *not* precisely one-third of the proton mass, because the proton's *uud* composition is only the *net* quark composition. In the background there is a continual "noise" of particle-antiparticle annihilation, gluons being exchanged,...but I am getting ahead of my story. The electric charges of the quarks and leptons are listed in Table 1.

	1ST GENERATION	2ND	3RD
LEPTONS			
$q = -1$	e	μ	τ
$q = 0$	ν_e	ν_μ	ν_T
QUARKS			
$q = +2/3$	u	c	t
$q = -1/3$	d	s	b

Table 1: Generations and electic charges of the fundamantal fermions.

When we say "particle" we also have in mind its corresponding antiparticle as well. A particle and its antiparticle have the same mass but opposite quantum numbers. For instance, the antiparticle of the electron is the positron. We do not need to think of particles and antiparticles as different objects in the theory. Taking a page from circuit analysis, where positive charges flowing from left to right equals negative charges flowing from right to left, when we see an antiparticle moving through the laboratory, we can *represent* it on a spacetime diagram as a particle moving backwards in time. For instance, consider electron-positron annihilation into two photons:

$$e^+ + e^- \rightarrow 2\gamma$$

The situation as seen in the laboratory has an electron emit the first photon at some point A (see Fig. 1a), then annihilate with the positron into the second photon at point B. To write the theory in these terms requires us to think of the positron and the electron as *distinct species*. But a tremendous simplification results if, instead, we think of the positron as an *electron moving backwards in time*! The annihilation process can then be *represented* in the spacetime diagram (and in the equations of the theory) as the adventures of a *single* electron (Fig. 1b): it moves *forward* in time and emits a photon at point A. The electron continues forward in time to point B, where it emits the second photon and recoils *backwards* in time away from point B. Any "time slice"

(continued on page 3)

before point B sees both the forward-moving electron and the "backwards-moving" electron, the latter showing up in the laboratory as a forward-moving positron. (In the theory one *does* have to sum over all diagrams that have the same initial and final states, and the same number of electron-photon vertices. Thus one must consider the processes of Figs. 1b and 1c together.)

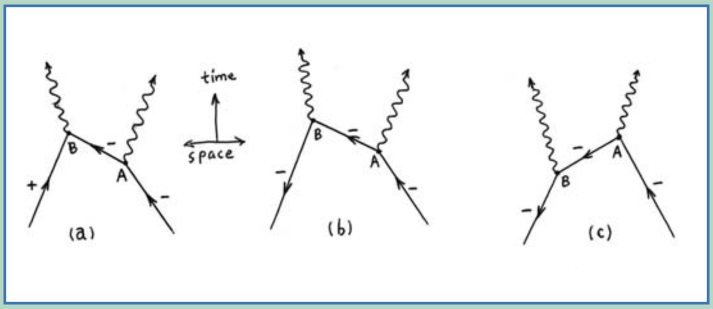


Fig. 1 (a) An electron and a positron annihilate to produce two photons (two photons are required to conserve momentum). (b) The same process with a continuous electron current. The positron is *represented* as the electron moving backwards in time. Thus positrons "are" electrons, and the theory does not have to treat them as distinct entities. (c) The order of emission of a photon by the electron, and the emission of the photon by annihilation, can be reversed. When calculating the cross-sections for electron-positron annihilation, both possibilities have to be included.

All quarks and leptons enjoy the weak force; therefore they all carry "weak charge." The quarks—but not the leptons—participate in the strong force; therefore the "strong charge" of all leptons equals zero, while all quarks carry non-zero strong charge. All the quarks and half of the leptons interact electromagnetically; the neutrinos sit this one out.

Charge Spaces

To participate in a fundamental force, the fermion must carry a non-zero value of the quantized conserved "charge" that characterizes the fundamental force.

Begin with a set of *N* fermion species that participate in a given interaction. Each of these fermions therefore carries a nonzero value of the necessary charge. Let the particles 1, 2,..., *N* be described by the quantum mechanical wave functions

$$\psi_1 , \psi_2 , \psi_3 , ..., \psi_N \qquad (1a)$$

which can be arranged into a vector denoted ψ. This vector exists in "charge space" for the respective interaction:

$$\psi \equiv \begin{matrix} \psi_1 \\ \psi_2 \\ . \\ . \\ \psi_n \end{matrix} \qquad (1b)$$

To form charge space scalars we need the row vector $\psi^\dagger$ that corresponds to the column vector ψ:

$$\psi^\dagger \equiv (\psi_1^* , \psi_2^* , \psi_3^* , ..., \psi_n^*) \qquad (2)$$

where * denotes complex conjugate.

Whenever the particle participates in the interaction for which it has the "ticket" of the necessary charge, its identity may stay the same or it may turn into another fermion from the same charge family. We can write a matrix of transitions between this family of fermions (Fig. 2). Suppose we put the *initial* fermion states in rows and the *final* fermion states in columns. Let the matrix elements $A(i\rightarrow j)$ denote the "radiation particle" emitted *necessary to make the reaction proceed*,

$$\psi_i \rightarrow \psi_j + A(i\rightarrow j).$$

If particle ψ_j does not carry the same charge as particle ψ_i, then $A(i\rightarrow j)$ must carry the descrepancy. These transitions form a kind of matrix:

	ψ_1	ψ_2 ..	ψ_n
ψ_1	$A(1\rightarrow 1)$	$A(1\rightarrow 2)$	$A(1\rightarrow N)$
ψ_2	$A(2\rightarrow 1)$	$A(2\rightarrow 2)$	$A(2\rightarrow N)$
.	...	...	...
.	...	...	...
ψ_n	$A(N\rightarrow 1)$	$A(N\rightarrow 2)$	$A(N\rightarrow N)$

Fig. 2. The matrix of transitions in "charge space."

Such transformations are shown in the "vertex" of Fig. 3a. By extension, two separated fermions may interact with one another when one of them emits a radiation particle which subsequently gets absorbed by another fermion, as in Fig. 3b. Each vertex will be proportional to the coupling constant (g, say), and thus the two-fermion interaction will be proportional to g^2. The radiation particle in the *exchange* forms an intermediate state. Of course, one can envision more complicated diagrams with more exchanges of radiation particles, including diagrams in which a fermion emits and absorbs the same radiation particle itself. Diagrams also exist for processes where a radiation particle is emitted into the final state.

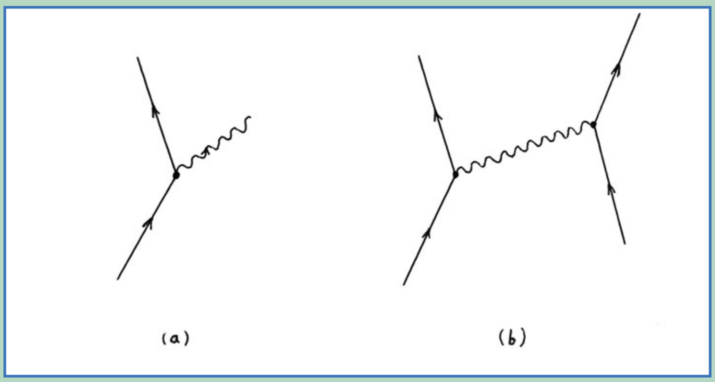


Fig. 3. (a) The fundamental vertex for the emission or absorption of a radiation particle. (b) Two fermions interact by exchanging a radiation particle. If the absorption does not occur, the radiation particle propagates away (e.g., photons as light).

Let us describe the vertex transition in a way that allows some kind of analysis. In going from a set of "initial" fermion states represented by the column vector ψ in charge space, to the "final" state represented by a new column vector ψ', we can postulate an operator or matrix Γ that turns the original ψ into the new ψ':

$$\psi' = \Gamma \psi \qquad (3)$$

or, by components for individual fermion species (repeated indices summed),

(continued on page 4)

$$\psi'_i = \Gamma_{ij}\, \psi_j \,. \tag{4}$$

Notice that by the rules of transposing a matrix,

$$\psi'^{\dagger} = \psi'^{\dagger}\; \Gamma^{\dagger} \,. \tag{5}$$

So long as the fermion options in the charge space still exist, the probability of their existence must remain unchanged (recall the fermion currents are continuous). Interpreting ψ as a quantum wave function for the fermion states, this means that $\psi'^{\dagger}\,\psi'$ must equal $\psi^{\dagger}\psi$. This, in turn, requires $\Gamma^{\dagger} = \Gamma^{-1}$, the transpose-complex conjugate of Γ equals the multiplicative inverse of Γ, and one says that Γ is "unitary."

Any unitary operator can always be written as a Hermitian operator Λ raised into a complex exponential, where "Hermitian" means $\Lambda^{\dagger} = \Lambda$, and raising an operator into an exponential is interpreted in terms of a Taylor series:

$$\begin{aligned} \Gamma &\equiv 1 + i\Lambda + \tfrac{1}{2}(i\Lambda)^2 + \ldots \\ &= \exp(i\Lambda) \end{aligned} \tag{6}$$

Because the vector ψ in charge space has N components, matrix representations of Γ and Λ will be $N{\times}N$ matrices, and therefore have N^2 components. If we require the determinant of Γ to be unity then the trace of Λ vanishes; this gives one constraint among the matrix elements, leaving N^2-1 *independent* components. Such transformation matrices are elements of the "special unitary group in N dimensions," SU(N). Let the transformation be paramaterized by a set of N^2-1 continuous, real parameters ε_k, k = 1,2,..., N^2-1. We therefore assume that, to first order in ε

$$\Lambda(\varepsilon) = \tfrac{1}{2}\varepsilon_k\, \lambda^k \equiv \varepsilon \cdot \lambda/2 \tag{7}$$

(the ½ is traditional) where each of the "generators" $\tfrac{1}{2}\lambda^k$ has *constant* matrix elements. Part of our responsibility in trying to understand elementary particle physics is to discern the "best" way to group the particles into their respective "charge spaces" (with their transformation matrices) for each force.

Spacetime and Relativity

The charges define an "internal space" that exists among the family of charged fermions. The fermions and the radiation particles move through external spacetime, and they do so at very high energy in cosmic rays and accelerators. Therefore their kinematic and dynamics must be described by Special Relativity, which we review briefly here, at least to fix some notations for later use.

The "invariant spacetime interval" forms the central working tool of Relativity. Consider two events in spacetime: Event 1 occurs at the coordinates (t, x, y, z) in the lab frame, and Event 2 at coordinates (t+dt, x+dx, y+dy, z+dz). An observer riding aboard a rocket that moves with velocity relative to the lab frame records these same events with rocket-frame coordinates (t', x', y', z').[7] The invariance of the speed of light between all inertial frames means that, in general, time and distances are not separately invariant, $dt' \neq dt$ and $dx' \neq dx$. However, a *combination* of space and time displacements is "invariant," having the same numerical value in both frames. That quantity is the "spacetime interval," which in Special Relativity says

$$c^2dt'^2 - (dx'^2 + dy'^2 + dz'^2) = c^2dt^2 - (dx^2 + dy^2 + dz^2) \tag{8}$$

where c denotes the speed of light in a vacuum. Henceforth, I shall absorb the speed of light into the time coordinate so that dt measures time in meters.

An observer using in the frame for which two events occur at the same place records proper time $d\tau$. By virtue of the Invariance of the Spacetime Interval the numerical value of the Interval for those events equals proper time squared:

$$d\tau^2 = dt^2 - (dx^2 + dy^2 + dz^2) \tag{9}$$

Whatever the coordinates (x^0, x^1, x^2, x^3) being used, the Spacetime Interval $d\tau^2$ may be written out component by component as the sum

$$d\tau^2 = g_{\mu\nu}\, dx^{\mu}\, dx^{\nu} \tag{10}$$

(sum repeated indices), where the elements of the matrix $\{g_{\mu\nu}\}$ form the "metric tensor." In general, they convert *coordinate displacements* into physical *lengths*. For rectangular spatial coordinates the metric tensor components are especially simple:

$$g_{tt} = 1, \quad g_{xx} = g_{yy} = g_{zz} = -1 \tag{11}$$

with all off-diagonal terms zero. The multiplicative inverse of the metric tensor has components denoted $g^{\mu\nu}$, that we will need later.

Consider a particle of mass m moving freely, without interactions. Its four-dimensional relativistic momentum relative to, say, the Lab Frame, may be computed as the particle's mass times the coordinate displacement per *proper* time:

$$\begin{aligned} p^{\mu} &= m\, dx^{\mu}/d\tau \\ &= m\,(dx^{\mu}/dt)(dt/d\tau) \,. \end{aligned} \tag{12}$$

From the Interval we find $dt/d\tau = (1-v^2)^{-1/2} \equiv \gamma$, where $v^{\mu} = dx^{\mu}/dt$ denotes the particle's coordinate velocity in this frame. The time component of the four-dimensional spacetime momentum ("4-momentum") is the particle's energy,

$$p^0 = m\, dt/d\tau = m\,\gamma \equiv E \tag{13}$$

which may also be written $E = m+K$ to define the particle's kinetic energy K. The space components gives the relativistic generalization of familiar "3-momentum,"

$$\mathbf{p} = m\,\mathbf{v}\gamma \,. \tag{14}$$

The energy and momentum components form a "4-vector" in spacetime,

$$p^{\mu} = (E, \mathbf{p}) \tag{15a}$$

Defining $p_{\mu} = g_{\mu\nu}\, p^{\nu}$ gives the spacetime conjugate vector

$$p_{\mu} = (E, -\mathbf{p}) \tag{15b}$$

so that the scalar product of 4-dimensional vectors in spacetime may be formed:

$$p_{\mu}\, p^{\mu} = E^2 - p^2 \tag{16}$$

which as a scalar in spacetime is invariant between inertial frames. Using the expressions for E and $\mathbf{p}$ give

$$E^2 - p^2 = m^2 \tag{17}$$

in units where c = 1.

Four-dimensional vectors such as the energy-momentum four-vector are justified by the name "vector" because they have the same

(continued on page 5)

number of components as there are spacetime coordinates, and transform the same as the coordinates do under a Lorentz transformation. Later we will meet "pseudoscalars" and "pseudovectors" too. The distinction between scalars and vectors on the one hand, and pseudoscalars and pseudovectors on the other hand, concerns their transformation properties under an inversion of the coordinates, taking **r** to **–r** (an example of an "improper Lorentz transformation"). This will be crucial in assigning "weak charges" for the weak interaction.

Quantum Mechanics

The particle wave functions mentioned above evolve according to the principles of quantum mechanics. Non-relativistic quantum theory begins with the relation between a free particle's energy and momentum,

$$p^2/2m = E\ . \tag{18}$$

Multiply this from the right with the quantum wave function Ψ, and use the quantum mechanical algorithms

$$E \rightarrow \hbar\, i\, \partial_t \tag{19a}$$

$$\mathbf{p} \rightarrow -\hbar\ \ i\, \nabla \tag{19b}$$

to obtain the free-particle Schrödinger equation,

$$-(\hbar^2/2m)\ \partial^2\Psi/\partial x^2\ =\ -(\hbar/i)\ \partial\Psi/\partial x/\partial t\ . \tag{20}$$

Special Relativity insists that space and time be considered equal partners. However, the Schrödinger equation deals with space and time asymmetrically. The situation can be redeemed in two ways, by combining the quantum algorithms of Eqs. (19) with the *relativistic* energy-momentum relation of Eq. (17). One way, applicable to bosons, takes second derivatives of Ψ with respect to *both* space and time. The other way, which describes fermions, takes *first* derivatives of Ψ with respect to *both* space and time. To get started, let's return to the relativistic energy-momentum relation, Eq. (17), that applies to a free particle. Multiply it from the right by the wave function Ψ and apply the quantum algorithm for energy and momentum in terms of derivatives (Eqs. 19). Thereby do we obtain the Klein-Gordon equation,

$$\hbar^2\ (\nabla^2 - \partial_t{}^2)\Psi\ =\ m^2\ \Psi \tag{21}$$

which describes the evolution of boson wave functions. We will "factor" it in the next section to "derive" the quantum wave equation for fermions, the Dirac equation.

II. Fermion Currents

The fermions carry the charges that are the "tickets" to participating in a fundamental force. Those charges are conserved *locally*. That means if a charge initially exists inside some region $\mathfrak{R}$, for it to be found later outside of $\mathfrak{R}$ it had to move *through* the surface S that forms the boundary of $\mathfrak{R}$.[8] The rate at which the charge inside $\mathfrak{R}$ decreases must equal the flux of charged current through the surface S:

$$\int_S \mathbf{j}\cdot\mathbf{n}\ d\boldsymbol{A} = -\ d/dt \int_{\mathfrak{R}} \rho\ dV \tag{22}$$

where ρ denotes the charge density, and **j** the current density. In many instances **j** may be written in terms of the particle velocity **v** as $\rho\mathbf{v}$. In quantum mechanics, the probability density for locating a particle is given in terms of a wave function by its square, $\Psi^*\Psi$. For a charged fermion carrying some charge Q, the charge density will, we expect, be given by $Q\ \Psi^*\Psi$, and the current density should look like $Q\psi^*\mathbf{v}\psi$. This idea is basically correct, but the case of spin-½ fermions requires some other factors besides **v** to go between Ψ^* and Ψ, because of the spinor nature of these wave functions. Let's see what features emerge that are peculiar to wave functions of spin-½ fermions.

In non-relativistic quantum mechanics, the spin of a particle has to be put in by hand. But spin *emerges* from the Dirac equation. To see how that comes about, let's go back to the Klein-Gordon equation, Eq. (21), and factor its difference of squares into the square of a difference. To facilitate this, let us introduce a four-vector

$$\gamma^\mu = (\gamma^0, \gamma^1, \gamma^2, \gamma^3)\ =\ (\gamma^0, \boldsymbol{\gamma}) \tag{23}$$

and consider what happens if we multiply out the left-hand side of this expression:

$$\hbar^2\ (\nabla\cdot\gamma + \gamma^0\,\partial_t)^2\ \psi =\ m^2\ \psi\ , \tag{24}$$

where we require the quantities γ^μ to be such that, when squared, Eq. (24) recovers the Klein-Gordon equation for ψ. This requires the γs to be quantities for which

$$\gamma^\mu\,\gamma^\nu + \gamma^\nu\,\gamma^\mu = 2g^{\mu\nu}\ . \tag{25}$$

Clearly the γ^μ cannot be ordinary real or complex numbers; they are *hypercomplex numbers*. They have to be something that can be represented by matrices. The smallest dimensionality of these so-called Dirac gamma-matrices is 4×4.[9] A popular representation of the Dirac matrices looks like this:

$$\gamma^0 = \begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix} \qquad \gamma^k = \begin{matrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{matrix} \qquad k = 1,2,3 \tag{26}$$

where here 1 denotes the 2×2 unit matrix, 0 the 2×2 zero matrix, and σ^k the kth Pauli matrix, which will appear below (Eq. 36).

This factorization of the Klein-Gordon equation allows the possibility

$$i\hbar\ \ (\nabla\cdot\gamma + \gamma^0\,\partial_t)\psi = m\psi\ , \tag{27}$$

the Dirac equation. In succinct summation convention notation the Dirac equation for the free particle reads

$$i\hbar\ \ \gamma^\mu\ \ \partial_\mu\psi = m\ \psi \tag{28}$$

When one takes the non-relativistic limit of Eq. (27), the Schrödinger equation with *spin-½ matrices already present* emerges.[10] The Dirac equation *intrinsically* describes particles of spin ½. The four-entry Dirac γ-matrices keep track of the spin-up/down quantum numbers for both the electron *and* for its antiparticle, the positron.

The three "space component" gamma-matrices $\boldsymbol{\gamma} = (\gamma^1, \gamma^2, \gamma^3)$ correspond (with some manipulation) to the particle's velocity in the non-relativistic limit. The "probability density" and "probability current density" for a Dirac fermion is

$\bar{\psi}\psi$ a scalar

and

$\psi\,\gamma^\mu\,\psi$ a 4-vector,

(continued on page 6)

where

$$\bar{\psi} \equiv \psi^{\dagger}\gamma^{0} . \tag{29}$$

When one considers so-called "improper" Lorentz transformations that allow the coordinate axes to be inverted, then one must consider pseudoscalars and pseudovectors (also called "axial vectors"). Because these objects are important to the weak force, we will consider them below, when we discuss the various charges that are characteristic of the various forces.

III. The Charges

Electric charge

The charge for electrodyanamics is of course electric charge, which comes in values qe where q is quantized. In purely electromagnetic interactions, the electric charge stays with its fermion. The "radiation particle," the photon γ—so denoted because in high energy physics the photons are typically gamma rays—carries no electric charge itself; thus, we do not have "atoms" of pure light. Examples of electromagnetic interactions include electron-electron scattering,

$$e^- + e^- \rightarrow e^- + e^- ,$$

bremsstrahlung ("braking radiation") and Compton scattering,

$$e^- \rightarrow e^- + \gamma ,$$

and electron-positron annihilation,

$$e^- + e^+ \rightarrow 2\gamma$$

The transition matrix for the first generation particles, for the electromagnetic force, gives

	e	v	u	d
e	γ	0	0	0
v	0	0	0	0
u	0	0	γ	0
d	0	0	0	γ

The zero for, say, e-going-to-u, does *not* mean that the electron and up quark do not interact electromagnetically (rather, they *do*), the zero means the electromagnetic interaction does not *change* an electron into a u-quark. Therefore, for any fermion wave function ψ, where the particle has electric charge quantum number q, under the electromagnetic interaction only the *phase* of ψ can be changed:

$$\psi' = e^{iq\varepsilon}\psi \tag{30}$$

where $\varepsilon^* = \varepsilon$ may be a function of position but requires only a single number at every event in spacetime. This change in the wave function is merely multiplication by a complex number, the "unitary group of transformations in one dimension," U(1).

Weak Charge of the Leptons

Carrying spin ½, electrons exhibit two possible orientations for their spin vector's z-component when **S** is projected onto the particle's linear momentum **p**. If S_z points in the same direction as **p** ("spin up" relative to **p**) then one says that electron state is "right-handed," or one says the electron has "positive helicity." If the spin vector's z-component points opposite **p** ("spin down") then that electron state is called "left-handed" (Fig. 4b) or "negative helicity." Moving electrons have momentum, and all electrons have spin, and electrons are found in *both* of these "helicity" states.

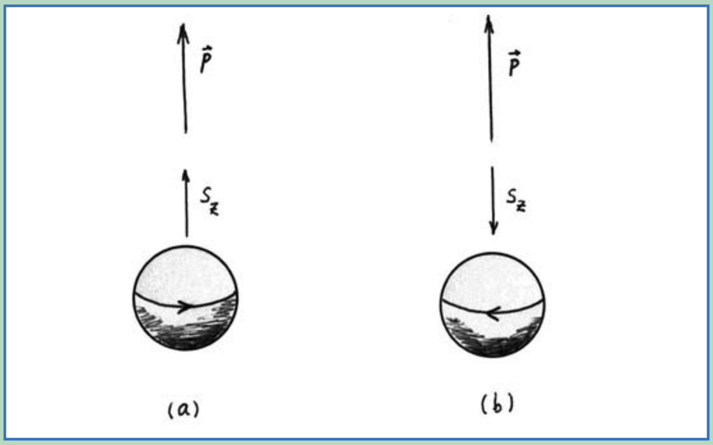


Fig. 4. (a) A "right-handed," or "positive helicity" fermion state. (b) A left-handed, or "negative helicity" fermion state.

However, neutrinos occur *only* in the left-handed state (antineutrinos are only right-handed), which offers the key for classifying the leptons into weak charge assignments. For the purposes of the weak interaction, it's advantageous to partition the fundamental fermions into "left handed" and "right-handed" multiplets. We can group the left-handed electron and the neutrino together in a two-component "left-handed doublet" (multiplicity = 2) arranged as the matrix,

$$\psi_L = \begin{matrix} v \\ e_L \end{matrix} \tag{31a}$$

where e_L denotes the wave function of the left-handed electron and v the neutrino's wave function. The right-handed electron remains by itself as a singlet (multiplicity = 1),

$$\psi_R = e_R . \tag{31b}$$

The left-handed doublet reminds us of a spin-½ itself, and we introduce a "weak isospin" quantum number τ, where $2\tau+1 = 2$ so that τ=½. Next we assign $\tau_z = +½$ to the neutrino and $\tau_z = –½$ to the left-handed electron. The right-handed electron "singlet" gets $\tau = 0$ and τ_z=0. The weak charges in units of $\sqrt{G_F}$ are tabulated in Table 2.

	τ	τ_z
v	½	+½
e_L	½	–½
e_R	0	0

Table 2. Lepton weak charges

(continued on page 7)

Let's look some more at that doublet, and write it as

$$\psi_L = \begin{pmatrix} a \\ b \end{pmatrix} = a\,|1\rangle + b\,|2\rangle \qquad (32)$$

where

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (33a)$$

denotes the "pure neutrino" basis vector in the space of all first-generation leptons, and

$$|2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad (33b)$$

denotes the "pure left-handed electron" basis vector. The set of all (first-generation) lepton states forms an abstract space spanned by the two basis unit vectors $|1\rangle$ and $|2\rangle$. When "asked" in an experiment whether it is a neutrino or a left-handed electron, the lepton has the probability $|a|^2$ of being a neutrino and the probablility $|b|^2$ of being a left-handed electron. Meanwhile, when it's "not asked," the lepton may be some mixture of neutrino and left-handed-electron. So long as the lepton exists, $|a|^2 + |b|^2 = 1$.

Just as a fermion with "spin up" can have its spin flipped to "spin down," so too can transitions exist between the neutrino and left-handed electron. When we form the transition matrix of the doublet of weakly interacting leptons that have weak isospin ½, we have to entertain the possibility of left-handed electrons converting into neutrinos and vice-versa, in reactions such as

$$e_L \rightarrow \nu + W^-$$

$$\nu \rightarrow e_L + W^+ .$$

The radiation particles $W^\pm$ carry ±1 unit of electric charge and ±1 unit of weak isospin. This offers the possibility of unifying the weak and electromagnetic forces, a major point to which we will return. In the meantime, that transition matrix, so far, might look something like this:

	ν	e_L
ν	0	W^+
e_L	W^-	0

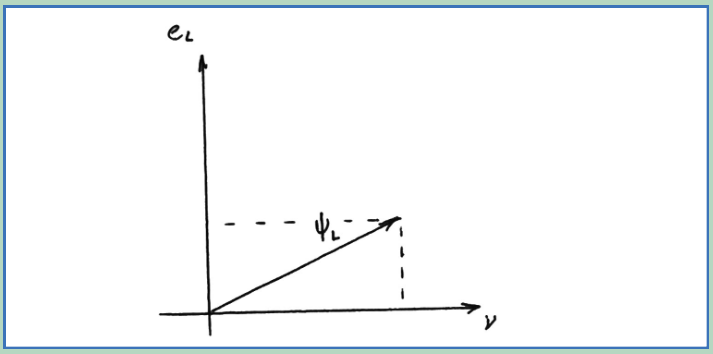


Fig. 5: "Weak isospin space," where the neutrino and left-handed electron are basis states for the generic first-generation lepton state. To change a neutrino into a left-handed electron rotates the isospin-space vector from $|1\rangle$ to $|2\rangle$.

Such transitions can be seen as rotations of the "state vector" in "weak isospin space" (Fig. 5). The language of elementary particle physics is that of transformations within such an "internal charge spaces."

Continuous Transformations of the Doublet

When we encounter states of spin-½ or weak isospin, we find ourselves dealing with doublets of states, and thus our transformation matrix will be 2×2. More formally, we are dealing with the SU(2) group of transformations, which has $2^2-1 = 3$ generators, 2×2 matrices with *constant* entries. Let us find them.

Because the matrix Λ of Eq. (16) is Hermitian, we may paramaterize it as

$$\Lambda = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \qquad (34a)$$

where A and C are real and B* denotes the complex conjugate of B. Requiring the trace to Λ to be zero sets C = –A. Therefore we can write

$$\Lambda = A\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix} \qquad (34b)$$

Writing $B = \alpha - i\beta$, we have

$$\Lambda = A\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \alpha\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \beta\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad (35)$$

Let $\varepsilon_1 = \alpha$, $\varepsilon_2 = \beta$, and $\varepsilon_3 = C$, and denote the 2×2 generators with σ's, famously known as the "Pauli spin matrices":

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (36)$$

Introducing the a coupling constant g that drives the transitions, and a traditional factor of ½ (recall the ε's are arbitrary so far anyway), we have shown that

$$\Lambda = \Lambda(\varepsilon) = \tfrac{1}{2}g\,[\varepsilon_1\sigma_1 + \varepsilon_2\sigma_2 + \varepsilon_3\sigma_3] . \qquad (37)$$

If we have any 2-dimensional charge space described by a doublet generically of the form

$$\psi = \begin{pmatrix} a \\ b \end{pmatrix} \qquad (38)$$

then σ_1 and σ_2 *scramble* a and b,

$$\sigma_1|\psi\rangle = \begin{pmatrix} b \\ a \end{pmatrix}, \qquad \sigma_2|\psi\rangle = \begin{pmatrix} ib \\ -ia \end{pmatrix} \qquad (39a)$$

while σ_3 *distinguishes* them, yielding eigenvalues as a diagonal matrix:

$$\sigma_3\psi = \begin{pmatrix} a \\ -b \end{pmatrix} \qquad (39b)$$

so that $g\,\sigma_3$ yields the charges as eigenvalues when operating on an eigenstate. Thus,

(continued on page 8)

$$g\,\sigma_3\,|1\rangle = g|1\rangle \text{ and } g\,\sigma_3\,|2\rangle = -g|2\rangle\ . \qquad (40)$$

In a two-dimensional charge space where one particle species has the opposite charge of the other, then we *assign* the charges as the eigenvalues of the diagonal generator.

We have seen that to conserve probability (or maintain "unitarity"), we must restrict our transformations according to

$$\psi' \approx \exp(\,½ig\,\boldsymbol{\lambda\cdot\varepsilon})\,\psi\ , \qquad (41)$$

or first order in the ε,

$$\psi' = \psi + ½ig\,\varepsilon_k\,\lambda^k\psi\ . \qquad (42)$$

The transformations form a *group*: each transformation has an inverse; an identity transformation exists; and successive transformations are equivalent to a single transformation given by the product of the individual ones. Because our group elements are operators in exponentials, we must consider the multiplicative properties of group elements. If A and B are operators, then

$$e^A e^B = \exp\{\,A + B + ½(A\,,\,B) + ...\,\} \qquad (43)$$

where the omitted terms are commutators of commutators. Two successive transformations, $\Lambda_1\Lambda_2$, involves commutators of the λ-matrices. If all the matrices commute then the group is said to be "Abelian" (after Norwegian mathematician Niels Henrik Abel). In particular, U(1) is Abelian. But if not all the matrices commute, then the group of transformations is said to be "non-Abelian." Such is the case with SU(N), where the commutator algebra is determined by the "structure constants" f^{abc} according to

$$(\lambda^a\,,\,\lambda^b\,) = 2i\,f^{abc}\,\lambda^c \qquad (44)$$

(sum over c). For example, in the case of SU(2), f^{abc} is the totally antisymmetric Levi-Ceviti symbol ε^{abc} which = 0 if any two superscripts are equal, but = +1 (–1) if abc is an even (odd) permutation of 1-2-3.

Quark Weak Charges

The quarks also are grouped, for the weak interaction, into weak isospin doublets and singlets. These assignments for quark weak charges come from spectroscopy and selection rules. With the quark wave functions denoted by the particle names, the weak isospin doublets for the first two generations are

$$\begin{matrix} u_L & c_L & t_L \\ (d_C)_L & (s_C)_L & (b_C)_L \end{matrix}$$

where

$$d_C \equiv d\cos\theta_C + s\sin\theta_C \qquad (45)$$
$$s_C \equiv s\cos\theta_C - d\sin\theta_C$$

and θ_C is the "Cabbibo angle," with $\sin\theta_C \approx 0.23$. The Cabbibo angle (fit to data) is necessary because the hadronic weak decays are systematically diminished compared to the leptonic weak decays. The quark right-handed states are weak isospin singlets,

$$u_R,\quad c_R,\quad (d_C)_R,\quad (s_C)_R\ .$$

The other generations also show similar mixing.

Chromodynamic Charge

The quarks are the *only* elementary fermions to participate in the strong force. All leptons carry zero strong charge. These chromodynamic charges are whimsically name after primary colors, so we speak of "color charge" (hence "chromodynamic"). To the *strong* force, there is no difference between u or d or quarks; those distinctions, called "quark flavor," are determined by their electromagnetic and weak interactions. No, the strong force couples only to color charge, and all flavors of quarks are alike in their color charge possibilities. To make sense of "color charge," let's consider an analogy to ordinary spin-½.

When writing the generic spin state χ of any spin-½ particle, one writes

$$\chi = \begin{matrix} a \\ b \end{matrix} \qquad (46)$$
$$= a\,|1\rangle + b\,|2\rangle$$

where

$$|1\rangle = \begin{matrix} 1 \\ 0 \end{matrix} \qquad (47a)$$

denotes the spin up basis vector in the space of all spin ½ states, and

$$|2\rangle = \begin{matrix} 0 \\ 1 \end{matrix} \qquad (47b)$$

denotes the spin down basis vector. When asked in an experiment whether the fermion has spin up or spin down, it replies with a probability $|a|^2$ of being spin up, and the probability $|b|^2$ of being spin down. But so long as the spinning particle exists, $|a|^2 + |b|^2 = 1$.

Color charge is analogous to spin, except there are three possibilities instead of two. The generic color state has three possible independent "directions" in "color space," so to specify the chromodynamics charge state of any quark requires a *triplet* of numbers, arranged into a vector in three-dimensional color charge space. Where spin had the two spin up and spin down eigenvectors, quantum chromodynamics has three color state eigenvectors called (arbitrarily) "red," "blue," and "green."

$$\psi = \begin{matrix} R \\ B \\ G \end{matrix}$$

$$= R\begin{matrix} 1 \\ 0 \\ 0 \end{matrix} + B\begin{matrix} 0 \\ 1 \\ 0 \end{matrix} + G\begin{matrix} 0 \\ 0 \\ 1 \end{matrix}$$

$$\equiv R\,|1\rangle + B\,|2\rangle + G\,|3\rangle\ . \qquad (48)$$

An arbitrary "color charge state" is represented by a vector in this abstract color space, as in Fig. 6.

(continued on page 9)

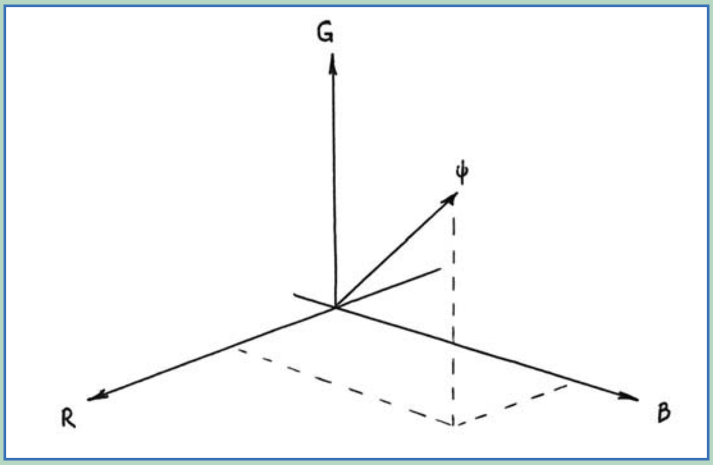


Fig. 6. An arbitrary vector in color space.

Incidentally, these amusing names for the color charge basis states are chosen (a) because physicists have a sense of humor, and (b) with sober reason, because of an analogy to color optics: as the three primary colors combined in equal proportions produce white light, similarly no state of matter has been found so far that exhibits *net* chromodynamic charge. Matter is color-neutral, which means that the quarks that make up protons and neutrons contain equal amounts of all three color charges. In particular, the color-charge mixture inside the proton ($p = uud$) is formed by $\psi_{\text{(proton color)}} = \varepsilon^{abc}\, \psi_a\, \psi_b\, \psi_c$ where a, b, c denote R, B, G and the ψ_a are quark fermion wave functions. Try it: the net color charge vanishes.

The appearance of the triplet of color states suggests that the group of color charge transformations,

$$\psi' = \exp(ig\, \tfrac{1}{2}\, \boldsymbol{\lambda}\cdot\boldsymbol{\varepsilon})\, \psi \approx \psi + ig\, \tfrac{1}{2}\, \varepsilon_k\, \lambda_k\, \psi\,, \qquad (49)$$

that turn one color state into another color state, would be SU(3). The group SU(3) has eight independent λ's, and if you go through the same kind of analysis whereby we derived the Pauli matrices of SU(2) above, you will find that two of the SU(3) matrices are diagonal. Proceeding as we did in before you can show that the eight λ-matrices of SU(3) are the following.

$$\lambda^1 = \begin{matrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \quad \lambda^2 = \begin{matrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \quad \lambda^3 = \begin{matrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{matrix}$$

$$\lambda^4 = \begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{matrix} \quad \lambda^5 = \begin{matrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{matrix}$$

$$\lambda^6 = \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{matrix} \quad \lambda^7 = \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{matrix}$$

$$\lambda^8 = \begin{matrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1/\sqrt{3} & 0 \\ 0 & 0 & -2/\sqrt{3} \end{matrix} \qquad (50)$$

It will be noticed that λ^1 and λ^2 mix R and B; λ^4 and λ^5 mix R and G; λ^6 and λ^7 mix B and G; and λ^3 and λ^8 distinguish R, B, and G, yielding two sets of eigenvalues for $|1\rangle$, $|2\rangle$, and $|3\rangle$.

Let ψ_R, ψ_B, and ψ_G be the wave functions of quarks that carry pure R, B, or G color charge respectively. The transition matrix for quantum chromodynamics is therefore

	ψ_R	ψ_B	ψ_G
ψ_R	$g(R\to R)$	$g(R\to B)$	$g(R\to G)$
ψ_B	$g(B\to R)$	$g(B\to B)$	$g(B\to G)$
ψ_G	$g(G\to R)$	$g(G\to B)$	$g(G\to G)$

where the $g(C_1\to C_2)$ are the "radiation particles" called "gluons" that convert a quark of color C_1 into a quark of color C_2. Because SU(3) is a non-Abelian group of transformations, the gluons carry color charge themselves, and therefore bound states of gluons-to-gluons ("glueballs") may exist.

"Vector" and "Axial Vector" Currents

A *parity transformation* changes the position vector **r** to **−r**, "inverting space" through the origin.[11] This is equivalent to a mirror reflection followed by rotating the image 180° about a normal to the mirror. Why should we care? If physical space were inverted, how could we tell? One might say that a parity transformation merely changes a right-handed coordinate system into a left-handed one. But coordinate systems are not part of nature, so it seems the fundamental laws of physics could cheerfully ignore a parity transformation. But with *interacting* particles in the game there may be more to parity than just the handedness of coordinate systems or the invariance properties of empty space. Whether elementary particles operate the same in their original and parity-transformed versions become an *experimental* question.

Here is an analogy to illustrate what I mean. The wheels of cars are held to the vechicle with lug nuts that you loosen and re-tighten to change a tire. Most lug nuts are right-handed, they tighten when rotated clockwise. However, some cars use left-handed lug nuts which tighten when they are rotated counter-clockwise. When the car is at rest, the handedness or "parity" of the lug nuts makes no difference: if properly tightened, right-handed and left-handed lug nuts hold the static wheel with equal security. One might say that "parity is a good symmetry" of a *static* car.

However, once the car gets set in motion then a preferred handedness reveals itself. If all the wheels are held on with right-handed threads, then the wheels on the driver's right side tend to tighten themselves even more as the car rolls forward, while the lug nuts on the left side of the car are more apt to eventually loosen. No longer does "lug nut parity" remain a good symmetry. To compensate for this, some manufacturers, e.g., Studebaker, make left-handed lug nuts for the wheels on the driver's left side.

The *rolling* of the wheels makes a difference to the torque maintained by the handedness of the lug nuts, even though on a static car, the handedness makes no difference. This parable illustrates that the *interactions* of the elementary particles may compromise some symmetries, even when "space itself" offers no reason to question the symmetry. In considering the motion of charged fermions, the charge current suggested by the Dirac equation is a *vector*. End of story? Not so fast!

Define a "parity operator" P that inverts space, so that for any function of the position vector **r**,

(continued on page 10)

$$P\, f(\mathbf{r}) = f(-\mathbf{r})\,. \qquad (51)$$

The reversal *of a reversal* restores the original situation:

$$P^2 f(\mathbf{r}) = f(\mathbf{r}) \qquad (52)$$

So if some function, say $g(\mathbf{r})$, is an eigenfunction of parity with eigenvalue π, which means $Pg(\mathbf{r}) = \pi g(\mathbf{r})$, then it follows by considering P^2 that $\pi = \pm 1$. A state with $\pi = +1$ is said to have "even parity," and a state with $\pi = -1$ has "odd parity." *Eigenstates* of P are *either* even *or* odd.

What effect does P have on observables? The position vector $\mathbf{r}$, and thus vectors derived from $\mathbf{r}$ by vector addition and scalar multiplication (e.g., velocity, momentum, acceleration, force) have odd parity because $P\mathbf{r} = -\mathbf{r}$. If $\mathbf{A}$ and $\mathbf{B}$ are odd-parity vectors then the dot product $\mathbf{A} \bullet \mathbf{B}$ has even parity: a "scalar." For the same $\mathbf{A}$ and $\mathbf{B}$, the vector product $\mathbf{A}\times\mathbf{B}$ also has even parity. Examples include angular momentum $\mathbf{r}\times\mathbf{p}$ and the magnetic field. Even-parity vectors are called "pseudovectors" or "axial vectors" to distinguish them from the odd-parity "vectors," which are also called "polar vectors." If $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are odd-parity vectors, then $\mathbf{C} \bullet (\mathbf{A}\times \mathbf{B})$ is an odd-parity scalar, a "pseudoscalar."

How do we determine if a given interaction conserves parity? Consider the electron in a hydrogen atom. When the electron sees only the proton's Coulomb potential, the Hamiltonian is

$$H = \mathbf{p}^2/2m - (1/4\pi\varepsilon_o)e^2/r\,, \qquad (53)$$

which has even parity. Include the spin-orbit interaction. Now the electron also sees a magnetic field $\mathbf{B}$ due to its motion relative to the proton, so that $\mathbf{B}$ is proportional to the electron's angular momentum $\mathbf{L}$. The electron's magnetic dipole moment $\boldsymbol{\mu}$ is proportional to its spin $\mathbf{S}$. Therefore the interaction energy $-\boldsymbol{\mu} \bullet \mathbf{B}$ is proportional to $\mathbf{L}\cdot\mathbf{S}$, a scalar because both $\mathbf{L}$ and $\mathbf{S}$ are axial vectors. Electrodynamics, so far as anyone can tell, conserves parity, because the electromagnetic Hamiltonian commutes with P.

But suppose the electromagnetic interaction did *not always* conserve parity; how would we know it? Imagine a universe where the electron-proton electromagnetic interaction in the hydrogen atom included an interaction that was proportional to $\mathbf{S}\cdot\mathbf{p}$, where $\mathbf{p} = m\mathbf{v}$, in addition to the spin-orbit interaction. Since it is the product of an axial vector and a polar vector, $\mathbf{S}\cdot\mathbf{p}$ is a pseudoscalar. Now the electromagnetic interaction would have a scalar *and* a pseudoscalar part. These interactions would each contribute something to the perturbative corrections to the spectrum of the electron's energy states, and that's how one could measure the scalar and the pseudoscalar contributions. And electrodynamics would not always conserve parity, because the parity operator and the total electron Hamiltonian would not commute. That means, like the lug nuts on a rolling car, some sense of orientation might be preferred in decays and scattering cross-sections.

But so far as anyone has ever been able to tell, electrodynamics does, in fact, conserve parity. So does the strong interaction. Their fermion currents seem to be entirely "vector" currents, given by the Dirac expressions

$$j^\mu \sim \bar{\psi}\,\gamma^\mu\,\psi \qquad (54)$$

which have the Lorentz transformation properties of a spacetime 4-vector.

The weak interaction offers a more compicated story.

Parity Violation in the Weak Interaction

The distinctive feature of the weak interaction is the participation of the neutrino or antineutrino, as in spontaneous neutron decay:

$$n \rightarrow p + e^- + \bar{\nu}_e\,.$$

More fundamentally, this is the weak interaction between quarks inside the proton (recall that $p = uud$):

$$d \rightarrow u + e^- + \bar{\nu}_e\,.$$

This reaction proceeds through an intermediate spin-1 particle we shall meet later, the W boson:

$$d \rightarrow u + W^+ \text{ then } W^+ \rightarrow e^- + \bar{\nu}_e\,.$$

Before 1956 everyone assumed that the weak interaction, like electrodynamics and the strong force, respects parity conservation.

That year, however, the community was confronted with the "τ–θ paradox." Two particle states, one called τ and the other called θ, had equal masses and lifetimes, and seemed to be the same particle—except that they decayed into states of opposite parity. The assumption of parity conservation was making the situation "puzzling," as T. D. Lee and C. N. Yang observed in their celebrated paper, "Question of Parity Conservation in Weak Interactions."[12] They realized that the situation would make sense if τ and θ were the same particle with two decay channels of opposite parity, which would mean that the weak interaction did not conserve parity. They wrote

...existing experiments do indicate parity conservation in strong and electromagnetic interactions to a high degree of accuracy, but that for the weak interactions... parity conservation is so far only an extrapolated hypothesis unsupported by experimental evidence.

By early 1957, the definitive experiment had been done, led by Chien Shiung Wu of Columbia University.[13] Wu and her colleagues worked with cobalt-60, which beta-decays to nickel-60 with a half-life of 5.2 yr. At low temperature the Co^{60} nuclei had their spins aligned by a strong magnetic field. The decay changes the nuclear spin by one unit. After the decay the Ni^{60} nuclei had their spins aligned in the opposite sense. Since the nuclear spin changed by one unit in the weak decay

$$Co^{60} \rightarrow Ni^{60} + e^- + \bar{\nu}_e\,,$$

and since the electron and antineutrino must move in opposite directions to conserve linear momentum, their orbital angular momentum vanishes. Therefore their spins must add up to 1, and point in the same direction to conserve angular momentum. This is illustrated in Fig. 7 below.

(continued on page 11)

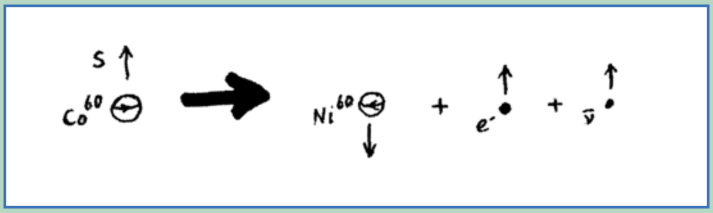


Fig. 7. The spins of the reaction $Co^{60} \rightarrow Ni^{60} + e^{-} + \bar{v}_e$.

Figure 8 shows the spin **S** and linear momentum **p** options for the electron and antineutrino emitted along the spin axis of the nucleus. I will denote these two situations as "state |1⟩" and "state |2⟩."

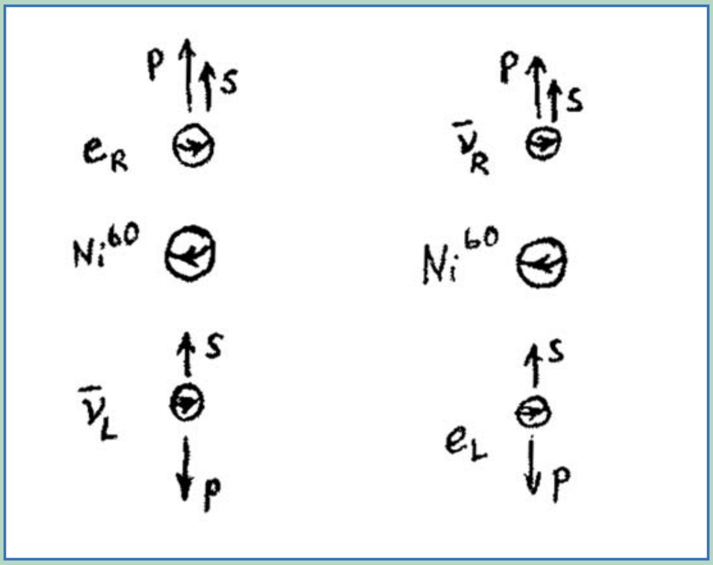


Fig. 8: State |1⟩ State |2⟩

A parity transformation is equivalent to reflecting through a mirror, followed by a rotation of 180° about a normal to the mirror. *It thus follows that state |2⟩ is a parity inversion of state |1⟩:*

$$P\,|1\rangle = |2\rangle \,. \tag{55}$$

To the extent that *both* states |1⟩ and |2⟩ occur, we will know to what degree parity is (or is not) conserved. In state |1⟩ the electrons emerge along the "north pole" of the original Co^{60} nucleus, but in state |2⟩ the electron emerges along the "south pole." Furthermore the electron in state |1⟩ is "right-handed," which means $\mathbf{S}\cdot\mathbf{p} > 0$. The antineutrino of state |1⟩, however, is "left-handed," because $\mathbf{S}\cdot\mathbf{p} < 0$. But in state |2⟩ the handedness situation is reversed. These outcomes are collected in the Table 3 below.[14]

	State |1⟩	State |2⟩
electron	R	L
antineutrino	L	R

Table 3: Handedness in the decay options of the Wu *et. al.* experiment.

If the weak interactions conserve parity, then states |1⟩ and |2⟩ should occur with equal probability. This probability can be measured by counting the number of electrons N_1 that emerge as described in state |1⟩ , and counting N_2, the number of electrons that emerge according to the description of state |2⟩. Define the parity measuring ratio R,

$$R \equiv \frac{N_1 - N_2}{N_1 + N_2} \,. \tag{56}$$

If parity is conserved then $N_1 = N_2$ and $R = 0$. If parity is "maximally violated" then $R = \pm 1$; specifically, if *all* the electrons emerge as in state |1⟩ we will have $R = 1$, and if they all emerge in state |2⟩ then $R = -1$. If the observed value of R were, say, 0.10, then Nature would allow about a 10 percent relative abundance for right-handed electrons and left-handed antineutrinos, and parity would be almost but not exactly conserved. The problem thus reduces to comparing the number of electrons emitted along opposite directions.

Wu's group measured $R = 1$ to within five percent. Parity is not conserved in the weak interaction. This result $R = 1$ requires that in the theory, the weak interaction currents include both the "vector" and "axial vector" terms, and that they be equally weighted. Not only is parity conservation violated, but it is *maximally* violated. Thus came the so-called "*V–A*" fermion current of the weak interaction.[15] How do we build an axial vector in terms of currents made of Dirac gamma matrices and fermion wave functions?

The textbooks [16] show that the quantity

$$\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{57}$$

is a pseudoscalar, that is, it maintains invariance under a proper Lorentz transformation, but changes sign under an improper transformation such as coordinate inversion. Therefore, the four-vector made of fermion wave functions and Dirac matrices γ^μ can now be made into an axial vector by replacing γ^μ with $\gamma^\mu\gamma_5$. For any fermion wave function we may always split the left-handed and right-handed helicity states by writing

$$\begin{aligned}\psi &= \tfrac{1}{2}(1 + \gamma_5)\,\psi + \tfrac{1}{2}(1 - \gamma_5)\,\psi \\ &= \psi_L + \psi_R \,.\end{aligned} \tag{58}$$

Therefore the "*V–A*" currents, proposed by George Sudarshan and Robert Marshak in 1958,[17] and independently in the same year by Murray Gell-Mann and Richard Feynman,[18] would become

$$j^\mu(\text{weak}) = \bar{\psi}\,\gamma^\mu\,(1 + \gamma_5)\,\psi \tag{59}$$

Now let's return to the other half of a fermion-fermion interaction, with two vertices connected by the propagation of a gauge boson.

Here we encounter one of the caveats I mentioned at the beginning. The weak interactions to not conserve parity (the space-inverted reactions are not as equally probable as the original), and they do not maintain invariance under charge conjugation (swapping antiparticles for particles). However, the weak interactions are *almost* invariant under the combined charge-conjugation-and-parity transformation. I say "almost" because the neutral kaon system does not quite respect this combined operation, the so-called "CP violation." CP violation allows the decays of unstable antiparticles to proceed ever-so-slightly slower than the decays of their corresponding particles. And it's a good thing, too: Had matter and antimatter always decayed symmetrically (equal rates), then in the early universe, the annihilations would have been complete and the universe a rather uninteresting

(continued on page 12)

swarm of photons! But I digress down one of those side paths that make this subject so very interesting.

IV. Radiation Particles

Fermions that carry a non-zero value of a charge have the "ticket" to interact with one another. Because signals cannot propagate instantaneously, the charged fermions interact by exchanging a radiation particle. The first fermion emits a radiation particle, and the second fermion absorbs it.

Mass and Range

The electromagnetic force has infinite range. In its simplest version of the Coulomb potential, the potential energy drops off as $1/r$; to go beyond the range of interaction the particles must get infinitely far apart. However, the weak and strong forces are very short-ranged, essentially the size of a nucleus, about 10^{-15} m, or less. What accounts for the difference in ranges? Suppose the radiation particle has mass M. From this mass and the fundamental constants $\hbar \sim 10^{-34}$ J s and $c \sim 10^8$ m/s we can form the length $\hbar/Mc \sim 10^{-42}$ kg m/M. Therefore to have infinite range the radiation particle for electrodynamics must have zero mass, and, indeed, the photon has zero mass.[19] To have a range on the order of 10^{-15} m the radiation particle of the weak interaction must have a mass ~ 100 times the mass of a proton.

Here appears another caveat. The photon is massless and electromagnetism has infinite range, and the gluon is also masseless, but the chromodynamic force has short range! The caveat here is called "color confinement," and it's a color charge screening effect. Confinement has yet to be proved rigorously from quantum chromodynamics (although asymptotic freedom has), although there are many confinement models that seem to work.

Why the Radiation Particles have Spin 1, Part I : The Lagrangians

The exchanged particles are spin-1 bosons, "vector bosons." Let us see why that must be so, and what are the consequences.

The Klein-Gordon equation describes the evolution of a boson wave function, and the Dirac equation controls the evolution of a fermion wave function. You may have encountered the concept of equations of motion being the logical consequence of requiring a "functional" or "action integral" to be a minimum. You may recall that the *fundamental* quantity is the *action J* which for a mechanical system takes the form of a definite integral,

$$J = \int_a^b L\ (x, v, t)\, dt \tag{60}$$

where L may be an explicit function of position $x(t)$, velocity $v = dx/dt$, and time t. L is called the "Lagrangian," and for a mechanical system L is the *difference* between the system's kinetic and potential energy. A functional like J is a machine into which one inserts a function $x(t)$; one then evaluates the definite integral and obtains the real number J. A main task in this business requires one to find the $x(t)$ that makes J a minimum. All the textbooks on this subject show how the $x(t)$ that will do the the trick must satisfy the "Euler-Lagrange equation,"

$$\partial L/\partial x = d/dt\,(\partial L/\partial v)\ . \tag{61}$$

Equations of motion such as $\mathbf{F} = m d\mathbf{v}/dt$ and $\mathbf{r}\times\mathbf{F} = d(\mathbf{r}\times\mathbf{p})/dt$ are the Euler-Lagrange equations for Lagrangians in generalized (not necessarily rectangular) coordinates.

The functional, Lagrangian, and Euler-Lagrange equation can be generalized to quantum fields. The functional becomes an integral over spacetime; the indepenent variable t gets replaced with the four spacetime variables; and the dependent variable x gets replaced with a quantum field, φ for bosons and ψ for fermions. To preserve unitarity, φ and φ^*, or ψ and $\bar{\psi}$, must always appear together in the combination $\varphi^*\varphi$ or $\bar{\psi}\psi$. A field and its conjugate are considered distinct fields.

For the scalar field φ, let φ_μ denote $\partial\varphi/\partial x^\mu$. The Euler-Lagrange equation for the field φ^* reads

$$\partial L/\partial\varphi^* = \partial_\mu\,(\partial L/\partial\varphi^*_\mu)\ . \tag{62}$$

Consider the following Lagrangian density (with $\hbar = 1$) for the non-interacting complex scalar field φ, and insert it into the Euler-Lagrange equation of Eq. (62):

$$L = (\partial_\mu\varphi^*)(\partial^\mu\varphi) - m^2\varphi^*\varphi \tag{63}$$

This yields the Klein-Gordon equation with the self-interaction,

$$\partial_\mu\,\partial^\mu\ \varphi = m^2\varphi \tag{64}$$

The Lagrangian for non-interacting Dirac fermions is

$$L = \psi^\dagger(\gamma^\mu i\,\partial_\mu - m)\psi \tag{65}$$

The Euler-Lagrange equation for $\bar{\psi}$ gives the Dirac equation for ψ,

$$i\gamma^\mu\ \partial_\mu\psi = m\psi \tag{66}$$

The Klein-Gordon and Dirac equations, with their Lagrangians, are written so far only for free particle quanta. But the real world has interactions. We are considering three fundamental interactions between elementary particles: the strong force, the weak force, and the electromagnetic force. All of these interactions exhibit "local gauge invariance." Let us look at the prototype of a "locally gauge invariant" theory, electrodynamics. After that we'll provide the next installment on why the radiation particles must be spin-1 bosons.

Electrodynamics and Gauge Invariance

Maxwell's equations for the electric and magnetic fields in free space produced by some charge of density ρ and a current density $\mathbf{j}$ are:

$$\nabla \bullet \mathbf{E} = \rho/\varepsilon_o \tag{67a}$$

$$\nabla\times\mathbf{B} = \mu_o\,\mathbf{j} + (1/c^2)\partial\mathbf{E}/\partial t \tag{67b}$$

$$\nabla \bullet \mathbf{B} = 0 \tag{67c}$$

$$\nabla\times\mathbf{E} = -\partial\mathbf{B}/\partial t \tag{67d}$$

where ε_o and μ_o are the electric permittivity and magnetic permeability of vacuum. The electric and magnetic fields are derivable from a vector potential $\mathbf{A}$ and a scalar potential V,

(continued on page 13)

$$\mathbf{B} = \nabla\times\mathbf{A} \,. \tag{68a}$$

and

$$E = -\nabla V - \partial\mathbf{A}/\partial t \tag{68b}$$

The Maxwell equations with the sources then become

$$\nabla^2\mathbf{A} - (1/c^2)\partial^2\mathbf{A}/\partial t^2 = -\mu_o\,\mathbf{j} + \nabla\,[\nabla\bullet\mathbf{A} + (1/c^2)\,\partial V/\partial t] \tag{69a}$$

and

$$\nabla^2 V + \partial/\partial t\,(\nabla\bullet\mathbf{A}) = -\rho/\varepsilon_o \tag{69b}$$

Let these potential components in spacetime be denoted $A^\mu = A^\mu(x^\nu)$. These are the components of four-vector in spacetime,

$$\{A^\mu\} = (V\,,\,\mathbf{A}) \tag{70}$$

Let's also introduce the four-vector current density for the sources,

$$\{j^\mu\} = (\rho,\,\mathbf{j}) \tag{71}$$

We can also define lower-index four-vectors $A_\mu = g_{\mu\nu}A^\mu$. Explicitly these are

$$\{A_\mu\} = (V,\,-\mathbf{A}) \tag{72}$$

For the derivatives with respect to spacetime coordinates we have

$$\{\partial_\mu\} = (\partial/\partial t\,,\,\nabla\,) \tag{73a}$$

and

$$\{\partial^\mu\} = (\partial/\partial t\,,\,-\nabla\,)\,. \tag{73b}$$

At this point it seems convenient to introduce the "Faraday tensor" with components $F^{\mu\nu}$, according to

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu\,. \tag{74}$$

Notice that the Faraday tensor is antisymmetric:

$$F^{\mu\nu} = -F^{\nu\mu} \tag{75}$$

In rectangular spatial coordinates the tensor's elements can be presented as the 4×4 matrix

$$\{F^{\mu\nu}\} = \begin{matrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_3 \\ E_3 & B_2 & -B_1 & 0 \end{matrix} \tag{76}$$

Eqs. (69) can be written collectively, with source terms, as

$$\partial_\mu F^{\mu\nu} = j^\nu\,. \tag{77}$$

At events in spacetime away from sources, we have "pure electrodynamics," controlled by the homogeneous equation

$$\partial_\mu F^{\mu\nu} = 0\,. \tag{78}$$

In terms of the vector potentials, this becomes a Klein-Gordon-like equation but with no mass term for the boson,

$$\ddot{y}\partial_\mu\;\partial^\mu A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0\,. \tag{79}$$

Because the potentials are related to the **E** and **B** fields by derivatives, we may perform a "gauge transformation" where, for any function χ,

$$A'^{\,\mu} = A^\mu - \partial^\mu\chi \tag{80}$$

or splitting out the time and space components,

$$\mathbf{A}' = \mathbf{A} + \nabla\,\chi \tag{81a}$$

$$V' = V - \partial\chi/\partial t\,. \tag{81b}$$

This transformation, for any χ, leaves $F'^{\mu\nu} = F^{\mu\nu}$, which is to say $\mathbf{E}' = \mathbf{E}$ and $\mathbf{B}' = \mathbf{B}$, an illustration of "gauge invariance."[20]

Suppose we now require that Eq. (77) be the Euler-Lagrange equation of some Lagrangian $L^{(A)}$ that is a function of the A^μ and their derivatives. The Lagrangian that does the trick for *pure electrodynamics* (no charges or currents)

$$\begin{aligned} L^{(A)} &= \tfrac{1}{2}\,(\mathbf{E}^2 - \mathbf{B}^2)\,. \\ &= -\tfrac{1}{4}\,F_{\mu\nu}F^{\mu\nu} \end{aligned} \tag{82}$$

Why Radiation Particles are Vector,
Part 2: The Covariant Derivative

Recall the real scalar field whose Lagrangian density is

$$L = (\partial_\mu\varphi^*)(\partial^\mu\varphi) - m^2\varphi^*\varphi - V(\varphi^*\varphi) \tag{83}$$

where a scalar field self-interaction potential V has been included. Suppose the field's phase is allowed to change by *different amounts* at different events, so that $\varepsilon = \varepsilon(x)$. The transformation for the field φ says:

$$\varphi' = \exp[\tfrac{1}{2}ig\varepsilon(x)]\varphi(x)\,. \tag{84}$$

When writing the new Lagrangian, everything stays the same as the original one except for the derivatives. They result in

$$\partial_\mu\varphi' = \exp(\tfrac{1}{2}ig\varepsilon)\partial_\mu\varphi + \tfrac{1}{2}ig\,(\partial_\mu\varepsilon\,)\varphi\,. \tag{85}$$

The first term represents the original kinetic energy term, but the second terms spoils the invariance. Instead of giving up on invariance under a local change of phase, we *enlarge the definition of the derivative* by introducing a new field A_μ, a vector in spacetime, and require *its* transformation to guarantee the invariance of this extended Lagrangian under a local gauge transformation. This kind of derivative is called the "covariant derivative," dignified by changing the ∂_μ notation to D_μ. To make this covariant derivative we add to the usual partial derivative a vector field with components $\{A_\mu\}$ that "hooks on" to φ with charge g, a scheme called "minimal coupling":

$$\partial_\mu\,\varphi \rightarrow D_\mu\varphi \equiv (\partial_\mu + \tfrac{1}{2}ig\,A_\mu)\varphi \tag{86}$$

Why are the A_μ components of a *vector* in spacetime? Because they have the same *Lorentz transformation* properties as does the gradient, the derivative with respect to spacetime coordinates. And we also require the covariant derivative to have the same *gauge transformation* rule as the field L itself," (Eq. 87 here).

$$(D_\mu\,\varphi)' = \exp[\tfrac{1}{2}\,ig\varepsilon(x)]\,[\,D_\mu\,\varphi(x)\,]\,. \tag{87}$$

(continued on page 14)

that determines the gauge transformation properties of the A_μ:

$$A'_\mu = A_\mu - \partial_\mu \varepsilon \quad (88)$$

Now we can write a Lagrangian that exhibits invariance under a local gauge transformation. It is

$$L = (D_\mu \varphi^*)(D^\mu \varphi) - m^2 \varphi^*\varphi - V(\varphi^*\varphi) \ldots \quad (89)$$

where ... denotes other terms that describe any other contributions that the field A_μ *itself* brings into the system (e.g., its own kinetic energy). When multiplying out $(D_\mu\varphi^*)(D^\mu\varphi)$ the Lagrangian takes the form

$$L = L^{(\varphi)} + j^\mu A_\mu + g^2 A_\mu A^\mu \varphi^*\varphi \quad (90)$$

where $L^{(\varphi)}$ is the original scalar field Lagrangian. The charged boson current takes the form

$$j^\mu \equiv ig\,(\varphi^{*\mu}\varphi - \varphi^*\varphi^\mu) \,. \quad (91)$$

The second and third terms in Eq. (90) describe interactions between the φ and A_μ fields, first order and second order in g respectively. Note that

$$j^\mu A_\mu = j^0 A_0 - \mathbf{j} \cdot \mathbf{A} \quad (92)$$

has the charge density ρ interacting with the electric potential A_0 and a current density **j** interacting with a vector potential **A**. It appears that we have "derived" the electrodynamic interactions of a charged scalar field, from the requirement of local gauge invariance under U(1)!

To describe the field φ of electric charge g coupled to the electromagnetic field A^μ, we merely invoke this algorithm:

(a) Write the Lagrangian for the φ field;

(b) Replace the partial derivatives ∂_μ with covariant derivatives D_μ, and

(c) Add the Lagrangian $L^{(A)}$ for gauge fields.

One has thereby "gauged" the transformation. So if our original Lagrangian for the charged particles reads

$$L^{(\text{original})} = (\partial_\mu\varphi^*)(\partial^\mu\varphi) - m^2 \varphi^*\varphi - V(\varphi^*\varphi) \,. \quad (93)$$

Then our algorithm says to change ∂ to D and add $L^{(A)}$:

$$L^{(\text{gauged})} = (D_\mu\varphi^*)(D^\mu\varphi) - m^2 \varphi^*\varphi - V(\varphi^*\varphi) + L^{(A)} \quad (94)$$

The matter-gauge field interactions are already contained in the covariant derivatives of φ.

This algorithm works the same way for fermion fields, such as the electron, with their Dirac Lagrangian.

The Field Tensor for Non-Abelian Fields

The electromagnetic transformations proceed through the group of transformations U(1), the weak force through SU(2), and the chromodynamic interaction through SU(3). Consider here the transformations of SU(N), which form a non-Abelian group:

$$\psi' = \exp(\tfrac{1}{2}\, ig\, \boldsymbol{\lambda} \cdot \boldsymbol{\varepsilon})\psi \,. \quad (95)$$

To first order in ε this means writing out the exponential's argument component by component,

$$\psi' \approx \psi + \tfrac{1}{2} ig\, \varepsilon_k \lambda^k \psi \quad (96)$$

where

$$(\lambda^a, \lambda^b) = 2i f^{abc} \lambda^c \,. \quad (97)$$

For the ath component of the charge space vector, we have

$$\psi'_a = \psi_a + \tfrac{1}{2}\, ig\, \varepsilon_k (\lambda^k)^{ab} \psi_b \,. \quad (98)$$

For instance, in the case of SU(N), when we introduce the vector gauge fields to make the kinetic energy term invariant, we have to introduce one for each generator; this means a set of vector fields $\{A^k_\mu\}$ ($k = 1,2,\ldots,N^2-1$).

The covariant derivative becomes now

$$D_\mu^{ab}\psi_b \equiv [\delta^{ab}\,\partial_\mu + \tfrac{1}{2} ig(\boldsymbol{\lambda}\cdot\mathbf{A}_\mu)^{ab}]\psi_b \,. \quad (99)$$

As before, we *require* the covariant derivative to have the same transformation rule as the field ψ_b itself, so that

$$(D^{ab}_\mu\, \psi_b)' = \exp[\tfrac{1}{2}\, ig\, \boldsymbol{\lambda}\cdot\boldsymbol{\varepsilon}]\,[\, D^{ab}_\mu\, \psi_b] \,. \quad (100)$$

When we work this out we are presented with the required transform property of the vector field A_μ:

$$A'^a_\mu = A^a_\mu - \partial_\mu \varepsilon^a + f^{abc} \varepsilon^b A^c_\mu \quad (101)$$

We have seen that a vector field's Lagrangian consists of $-\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu}$, but now the "Faraday tensor" picks up a term *quadratic* in the fields:

$$F^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu \,. \quad (102)$$

The term $g f^{abc} A^b_\mu A^c_\nu$ means the gauge fields interact, not only with the charged matter particles, but with one another directly.

Note two important points about the gauge fields, whether they are Abelian or non-Abelian:

(1) The gauge boson is the quantum of a spacetime *vector* field and has spin 1, because it has the same transformation properties as the gradient vector ∂_μ.

(2) If the gauge field quantum has mass M, then its mass term in the Lagrangian $L^{(A)}$ will be $\tfrac{1}{2} M^2 A_\nu A^\nu$, to give the correct term for mass in the Klein-Gordon equation. But such a term spoils the local gauge invariance. Therefore the gauge field must have zero mass, M=0.

This second point raises a dramatic red flag: the photons and gluons are massless, but the gauge bosons for the weak interaction have masses on the order of a hundred protons! This brings us to the Higgs boson and "hidden symmetries," to which we turn next.

In summary for now: gauge invariance *requires* the existence of a spin-1, massless gauge boson: the photon in the case of electrodynamics, the eight species of gluons in the case of quantum chromodynamics.[21] We will discuss the gauge bosons of the weak interactions in the context of the unification of the weak and electromagnetic interactions, which illustrates one more level of complexity: the Higgs spin-0 boson.

V. Unification, Mass, and the the Higgs Boson

When we wrote the transition matrix for the doublet of weakly interacting leptons, we had to entertain the possibility of left-handed

(continued on page 15)

electrons converting into neutrinos and vice-versa:

$e_L \rightarrow \nu_e + W^-$

where W (we can now say) denotes a gauge boson for the weak interaction. Because the fermion's weak isospin has changed from –½ to +½, and the change in electric charge has been from −1 to 0, this W will have to carry weak isospin quantum number −1 and electric charge quantum number −1 (it's the electric charge denoted in the W's superscript). Indeed, if the transformations of this left-handed doublet are supposed to be described by SU(2) then there will be *three W* particles. The other two possibilities for them are

$\nu_e \rightarrow e_L + W^+$

and

$e_L \rightarrow e_L + W^0$

which might also show up in

$\nu_e \rightarrow \nu_e + W^0$.

Because the left- and right-handed electron states can also interact electromagnetically, it might be tempting to identify the W^0 as the photon, but the neutrino does not interact electromagnetically so it can't be that simple. First, the weak interaction has a very short range, which means the weak force gauge bosons somehow acquire large masses. In contrast, electodynamics has infinite range, so the photon carries zero mass. However, such a model with the W^0 as the photon was proposed by Julian Schwinger in 1957.[22] Schwinger got around the mass problem in a clever but contrived way by introducing some auxiliary spin-0 particles with which the W's would interact. While there's nothing intrinsically wrong with that (indeed, the Higgs mechanism discussed below does the same), in Schwinger's model no prediction of the W masses could be made. A successful unification of the electromagnetic and weak interaction would *require* the $\nu_e \rightarrow \nu_e + W^0$ transition, a so-called "neutral current" event, which would allow neutrino-neutrino scattering when a second neutrino absorbs the W^0 emitted by the first. Such neutral current scatterings were modeled by Bludman in 1958, where again the W masses had to be put in by hand.[23] But having the same W^0 mediating interactions between neutrinos and left-polarized electrons assumes that the charged-current and neutral-current interactions have the same structure, and experiment shows this not to be the case. Furthermore if the W^0 is not the photon then nothing has been accomplished towards unifying two of the fundamental interactions. Unification attempts received a boost around 1964 with the "Higgs mechanism."

If elementary particle physics currently has an analog to the legendary quest for the Holy Grail, it must surely be the search for the Higgs boson. All the fundamental fermions and the massive gauge bosons are thought to begin intrinsically massless, but *acquire* their masses through interactions with the Higgs field. All this is done while preserving gauge invariance.

The Higgs field is a *postulated* scalar field. We can think of it as a relativistic analogy to some kind of "pressure" that exists throughout all of spacetime, but we don't need any fluid: one may view the Higgs field as a property of *spacetime itself*. The Higgs boson is the quantum of the Higgs field. Because the field is a scalar field, the Higgs boson has spin 0.

It must be understood that the Higgs field and its spin-0 boson are, to date, *conjectured*; an example of the species has yet to be seen cleanly in the laboratory. But its role in the theory of elementary particles ties everything together so elegantly and convincingly that, whether or not the Higgs mechanism is *literally* true, it offers an *effective* description that seems to approximate *something* real.

Finding an actual Higgs particle propagating with the right quantum numbers through a particle detector would cinch this important section in the foundation of physics.

Where does the name "Higgs" come from? As you might suppose, the Higgs boson acquired its name by repeated references in the literature to the author who first envisioned its role in the theory. While he was at the University of Edinburgh in 1964, Peter Higgs suggested the scenario that has became known as the "Higgs mechanism."[24]

The recurring major theme in the history of physics has been the quest to unify the interactions in Nature. Isaac Newton unified "terrerestrial" and "celestial" mechanics with his law of *universal* gravitation. Michael Faraday and James Maxwell unified optics, electricity, and magnetism into electrodynamics. Einstein unified energy with mass, time with space, gravitation with geometry, and spent the last two decades of his life trying (unsuccessfully) to unify electrodynamics with gravitation. The Higgs boson makes possible the unification of the weak interaction with electromagnetism, the "electroweak interaction."

We have seen how electric charge and weak isospin, by themselves, cannot *unify* these two forces. A useful suggestion was made by Sheldon Glashow in 1961[25], similar ideas were put forward in 1964 by Adabus Salam and J. C. Ward[26], and everything was pulled together by Steven Weinberg in 1967[27]. The idea is to give the particles another "charge," called "weak hypercharge" Y that *combines* electric charge and weak isospin. Weak hypercharge is defined, in terms of electric charge quantum numbers q and weak isospin τ_z according to

$$Y = 2(q - \tau_z). \tag{103}$$

Therefore the full set of quantum numbers for the familiar weakly and electromagnetically interacting leptons are listed in Table 4.

	q	τ_z	Y
ν_e	0	+½	−1
e_L	−1	–½	−1
e_R	−1	0	–2

Table 4. Lepton electroweak charges.

Recall that the leptonic states group themselves into the two-component doublet for the left-handed leptons,

$$\psi_L = \begin{matrix} \nu_e \\ e_L \end{matrix} \tag{104a}$$

which transforms under SU(2). There is also the right-handed singlet

$$\psi_R = e_L\ . \tag{104b}$$

(continued on page 16)

Let us suppose all the fermion masses are initially set to zero. The fermions will acquire their masses, as will some of the gauge bosons, through their interactions with the Higgs field. With only these massless fermions, the mass term $m\,\psi_R{}^{\dagger}\gamma^0\psi_R$ disappears from the fermion Lagrangian density, leaving

$$L^{(\psi)} = \psi_L{}^{\dagger}\gamma^0\,\gamma^\mu\,\partial_\mu\,\psi_L + \psi_R{}^{\dagger}\gamma^0\,\gamma^\mu\,\partial_\mu\,\psi_R \tag{105}$$

Let us now postulate a Higgs field, or rather two Higgs fields represented with a doublet

$$\phi = \begin{pmatrix} \phi^a \\ \phi^b \end{pmatrix} . \tag{106}$$

Let this Higgs field intereact *with itself* through some self-interaction potential energy V, which is a function of $\phi^\dagger\phi$. The potential that has the features we need later on takes the form $V = h\,(\phi^\dagger\phi)^2$ where h is a constant to be determined. As a *real* scalar field, the Higgs Lagrangian is now

$$L^{(\phi)} = \tfrac{1}{2}\,(\partial_\mu\,\phi^\dagger)(\partial^\mu\,\phi) - M^2\,\phi^\dagger\phi - h\,(\phi^\dagger\phi)^2 . \tag{107}$$

(the kinetic energy term requires the ½ when the scalars are not complex), where M denotes the Higgs boson mass, another constant to be determined.

Next, we allow the fermions to interact with the Higgs fields through coupling G_e, yet another constant to be fit to data. The simplest coupling has $\bar\psi_L\phi\,\psi_R$, and we must include its conjugate $\bar\psi_R\,\phi^\dagger\,\psi_L$. Note the change in fermion helicity. Now the coupled fermion-Higgs Lagrangian density is

$$\begin{aligned} L^{(\psi-\phi)} = {} & \bar\psi_L\,\gamma^\mu\partial_\mu\,\psi_L + \bar\psi_R\,\gamma^\mu\partial_\mu\,\psi_R \\ & + \tfrac{1}{2}\,(\partial_\mu\phi^\dagger)(\partial^\mu\phi) - M^2\,\phi^\dagger\phi - h\,(\phi^\dagger\phi)^2 \\ & - G_e(\bar\psi_L\,\phi\psi_R + \bar\psi_R\,\phi^\dagger\,\psi_L) . \end{aligned} \tag{108}$$

The Higgs-fermion interactions can be represented with the spacetime diagram vertex of Fig. 9, where a fermion comes in, emits (or absorbs) a Higgs boson, and a fermion of the opposite helicity goes out.

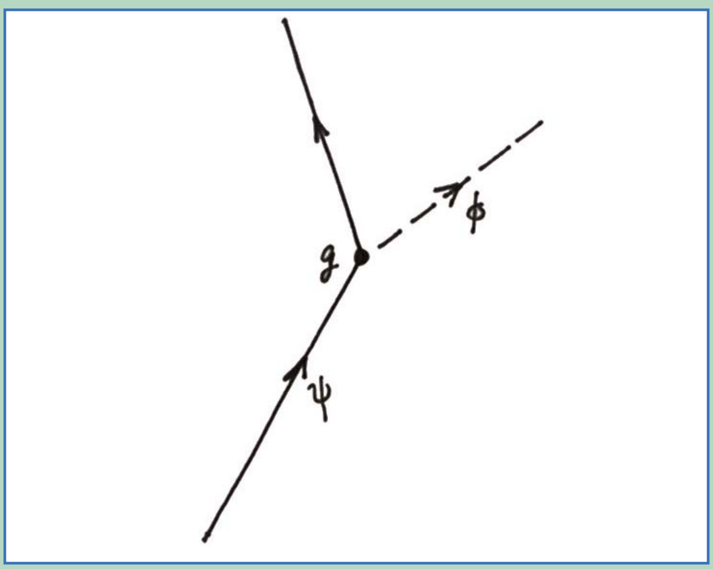


Fig. 9. The vertex diagram of the Higgs-fermion interaction.

Consider the "vacuum expectation value" $\langle\phi\rangle$ of the Higgs field. The vacuum expectation value is the average value of the Higgs scalar field in its ground state througout all of space. Suppose furthermore that this vacuum expectation value written out for both components of the Higgs doublet becomes

$$\langle\phi\rangle = \begin{pmatrix} \langle\phi^a\rangle \\ \langle\phi^b\rangle \end{pmatrix} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix} \tag{109}$$

where λ is a *constant*. By *invoking* this value we implement so-called "spontaneous symmetry breaking." "Spontaneous" means the vacuum expectation value is set nonzero by *assumption*.[28] The term "symmetry breaking" means the transformation symmetry of the Lagrangian is spoiled by this nonzero vacuum expectation value. In other words the vacuum state does not exhibit the symmetry of the original Lagrangian.

Notice what will happen when we now replace ϕ by $\langle\phi\rangle$ in the Lagrangian density. As you may anticipate, say from Fig. 9, what *was* a fermion-Higgs vertex will become a mass term for the fermion. After properly doing all the matrix multiplication on the doublets and their conjugates, and noting that $\langle\phi\rangle$ is a constant, we obtain:

$$\begin{aligned} L^{(\psi-\phi)}\big|_\lambda = {} & \bar\psi_L\,\gamma_\mu\partial_\mu\,\psi_L + \bar\psi_R\,\gamma^\mu\,\partial_\mu\,\psi_R \\ & - (M^2 + h)\,\lambda^2 - G_e\lambda\bar\psi_R\,e_L . \end{aligned} \tag{110}$$

Notice that the term $M^2\,\langle\phi^b\rangle^2$ is *not* present, which is indistinguishable from setting the mass of the ϕ^b boson equal to zero. This is an example of a so-called "Goldstone boson," an awkward development because no massless particles of spin zero are known to exist.

But we have not yet invoked gauge invariance with its covariant derivative. We have two charges to "gauge:" weak isospin, that transforms under SU(2), and weak hypercharge, that transforms under U(1).

Let us gauge weak isospin with a set of three gauge boson fields $\{A^k{}_\mu\}$, $k = 1, 2, 3$; coupled to weak isospin via coupling constant g. Let this gauge boson's non-Abelian Faraday tensor be denoted $F^{\mu\nu}$ (cf. Eq. 91).

Let us also gauge weak hypercharge, applying the gauging algorithm to each member of the doublet, and of course to the singlet. Denote the weak hypercharge coupling constant as g' and its Abelian gauge boson as B_μ and its Abelian Faraday tensor as $G^{\mu\nu}$.

We must invoke the covariant derivatives wherever derivatives appear, in both the fermion and Higgs sectors of the Lagrangian. Finally we must remember to include the gauge field Lagrangians that represent the energy brought into the system by the gauge fields themselves. Now our fully gauged Lagrangian becomes

$$\begin{aligned} L^{(\text{gauged})} = {} & \bar\psi_L\,\gamma^\mu(\partial_\mu - ig\,\tfrac{1}{2}\,\boldsymbol{\tau}\cdot\mathbf{A}_\mu - i\,g'B_\mu)\,\psi_L \\ & + \bar\psi_R\,\gamma^\mu(\partial_\mu - i\,g'B_\mu)\psi_R \\ & + \tfrac{1}{2}\,|\partial_\mu - ig\,\tfrac{1}{2}\,\boldsymbol{\tau}\cdot\mathbf{A}_\mu\phi - i\,g'B_\mu\,\phi|^2 \\ & - M^2\,\phi^\dagger\phi - h\,(\phi^\dagger\phi)^2 \\ & - G_e(\bar\psi_L\phi\,\psi_R + \bar\psi_R\phi^\dagger\,\psi_L) . \\ & - \tfrac{1}{4}F_{\mu\nu}\,F^{\mu\nu} - \tfrac{1}{4}G_{\mu\nu}\,G^{\mu\nu} . \end{aligned} \tag{111}$$

Now when we choose the vacuum expectation value for the Higgs field as done previously, the first two and last two groupings of terms remain the same as before. But the terms involving the Higgs boson become

(continued on page 17)

$-\frac{1}{2}\lambda^2 g^2 [(A^1_\mu)^2 + (A^2_\mu)^2]$ Term I

$-\frac{1}{2}\lambda^2 (gA^3_\mu + g'B_\mu)^2$ Term II

$-\lambda G_e(\bar{\psi}_e\psi_e)$ Term III

where ψ_e is the electron wave function. Had an electron mass m been included initially it would have been as a term

$$m(\bar{\psi}_L\psi_L + \bar{\psi}_R\psi_R) = m\,\bar{\psi}_e\psi_e \quad (112)$$

so that upon comparing this to Term III we see that the Higgs-induced electron mass is

$$m = \lambda G_e . \quad (113)$$

If we define new electrically charged gauge bosons $W_\mu^\pm$ in terms of the original ones,

$$W_\mu^\pm = \sqrt{\tfrac{1}{2}}\,(A^1_\mu + iA^2_\mu) , \quad (114)$$

then compare this to a Lagrangian for a vector field with mass introduced from the outset, we see from Term I that $\frac{1}{2}\lambda^2 g^2 = \frac{1}{2}M_W^2$ so that

$$M_W = \tfrac{1}{2}\lambda g . \quad (115)$$

These gauge bosons carries ±1 unit of electric charge because A^1_μ and A^2_μ are the coefficients of τ_1 and τ_2 which transform neutrinos and left-handed electrons into one another.

The gauge bosons A^3_μ and B_μ are uncharged because they mediate electron-to-electron, or neutrino-to-neutrino interactions. If we take the two linear combinations of them,

$$Z_\mu = (g^2 + g'^2) - \tfrac{1}{2}(gA^3_\mu + g'B_\mu) \quad (116)$$

$$A_\mu = (g^2 + g'^2) - \tfrac{1}{2}(-g'A^3_\mu + gB_\mu) \quad (117)$$

then writing Term II in terms of Z_μ and A_μ instead of A^3_μ and B_μ, we find that

$$M_Z = \tfrac{1}{2}\lambda (g^2 + g'^2)^{1/2} \quad (118)$$

and

$$M_A = 0 \quad (119)$$

which identifies A_μ with the photon. If the Lagrangian is rewritten in terms of the W, Z, and A, by comparing the couplings in this derived Lagrangian to one that would exist had one written everything in terms of these gauge bosons originally with the usual electric and weak coupling constants, one obtains the relations

$$e = gg'(g^2 + g'^2)^{-1/2} \quad (120)$$

and

$$\sqrt{\tfrac{1}{2}}\,G_{weak} = g^2/8M_W^2 = \tfrac{1}{2}\lambda^2 \quad (121)$$

Because e and G_{weak} are known from laboratory measurements, from Eq. (121) we have λ.

Then Eq. (120) tells us that both g and g' are larger than e, which means that $M_W > 40$ GeV/c^2, and $M_Z > 80$ GeV/c^2, at least 40 and 80 proton masses respectively for the W and Z bosons.

What has happened? Let's go back to the Higgs-fermion vertex. Initially we took the fermion masses to be zero and the Higgs field to be a *function* throughout spacetime. Then we invoked a non-zero vacuum expectation value for the Higgs field—both Higgs fields in the weak isospin doublet became *constants*—one of them nonzero, the other one zero. This turned the vertex *interaction* of the fermion with the Higgs into a *mass* term for the fermions. And because the vacuum expectation value of the Higgs field became constant, the derivatives of this field were zero (its kinetic energy vanishes); but in the *covariant* derivatives the gauge terms survived, and their couplings with the Higgs field likewise turned what would have been Higgs-gauge boson interactions into mass terms for some of the gauge bosons also.

Is this *real*? The $W^\pm$ bosons were produced in the laboratory in 1983, and its mass was about 80 GeV/c^2. The Z^0 boson was found soon after with a mass of about 90 GeV/c^2. And the photon has been known since 1905 to be massless. Although the Higgs field and its corresponding spin-0 bosons are *conjectured*, something very much like them seems to be operating in nature.

The transition matrix for the electroweak interaction can now be written:

	ν_e	e_L	e_R
ν_e	Z	W^+	0
e_L	W^-	Z, A	A, Z
e_R	0	A, Z	Z, A

The electromagnetic and weak interactions have been unified because any process that could go between charged fermions through photon exchange can also go through Z exchange, and the Z can therefore "talk" to neutrinos. Therefore such diagrams as Fig. 10 are possible-and the processes they describe really happen.

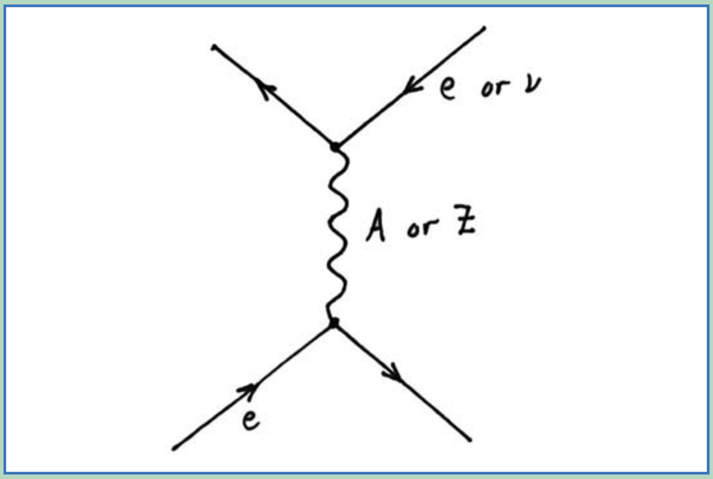


Fig. 10. An electron-positron pair can annihilate into a photon or a Z-boson. In the latter case, through the intermediate Z, we can have the overall effect of $e^+ + e^- \rightarrow \bar{\nu} + \nu$ thereby unifying the electromagnetic and weak interactions into the "electroweak" interaction.

This essay offers an informal introduction to some of the elements of the Standard Model of elementary particle physics of the electromagnetic, weak, and strong interactions. The Standard Model

(continued on page 18)

still has many unsolved problems. The number of parameters are too many, the constraints relating them too few. Everyone wonders: To what deeper theory is the Standard Model an approximation? What lies "Beyond the Standard Model"?[29] What, really, *is* the Higgs field? What are its cosmological consequences? More players are needed.

Exercise: Put three quarks and two leptons into a five-member multiplet. Invoke SU(5) and propose a transition matrix. See what *familiar* gauge bosons you can come up with, and what *new* gauge bosons (perhaps with fractional electric charge) are predicted that transform quarks into leptons and leptons into quarks. Such a model was worked out in 1974 by Howard Georgi and Sheldon Glashow.[30] Their model predicts proton decay with a half-life of $\sim 10^{32}$ years, or one decay annually in a sample of matter holding 10^{32} protons. How large a tank of water would be necessary to have this many protons? Unfortunately for the SU(5) model, despite valiant efforts, proton decay has not been observed.

Acknowledgments

With much appreciation, I thank Lee Sawyer for examining a draft of this article, offering clarifications and making many excellent suggestions.

References

[1] E. Fermi, *Nuovo Cimento* 11, 1 (1934); *Zeit. f. Physik* **88**, 161 (1934).
[2] See "Elegant Connections in Physics, Angular Momentum and Spin," *SPS Observer*, Summer 2006, www.spsnational.org.
[3] G. Sudarshan and I. Duck, *Pauli and the Spin-Statistics Theorem* (World Scientific, 1997).
[4] J. Bjorken and S. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, 1964), Ch. 2.
[5] Contemporary cosmology shows strong evidence that "ordinary" matter makes up about five percent of the energy density in the universe. About 25 percent is thought to be non-baryonic "dark matter", about 70 percent "dark energy." The term "lepton" comes from a Greek word that means "lightweight"; the term "quark" was encouraged by Murry Gell-Mann, from a line in James Joyce's *Finnigan's Wake*, "three quarks for Muster Mark...."
[6] "Who ordered these?" —Attributed to I. I. Rabi. No one knows.
[7] E. F. Taylor and J. A. Wheeler, *Spacetime Physics* (Freeman, 1966, 1992).
[8] If the charge disappeared from within a region then popped up instantly in some other region, the charge would be *globally* but not *locally* conserved.
[9] Bjorken and Drell, Ref. 4, Ch. 2.
[10] Bjorken and Drell, Ref. 4, Ch. 1.
[11] These sections on parity nonconservation in the weak interaction is adapted from "Elegant Connections in Physics: Radioactivity, Symmetry, and Parity," *Radiations*, Fall 2002, pp. 20-24.
[12] C. N. Yang and T. D. Lee, "Question of Parity Conservation in Weak Interactions," *Phys. Rev.* **101**, 254-258 (1956).
[13] C. S. Wu, et. al., "Experimental Test of Parity Conservation in Beta Decay," *Phys. Rev.* **105**, 1413-1414 (1957).
[14] From Ref. 11, p. 23.
[15] The "minus" in "*V–A*" becomes "plus" in other conventions on the gamma matrices; nevertheless, the jargon says "*V minus A*."
[16] Bjorken and Drell, Ref. 4, Ch. 2.
[17] R. Marshak and G.E. Sudarshan, *Phys. Rev.* **109**, 1860 (1958).
[18] R. P. Feynman and M. Gell-Mann, *Phys. Rev.* **109**, 193 (1958).
[19] The experimental uncertainties *on measuring* the photon's mass put $m_\gamma < 4\times10^{-48}$ gm , some 20 orders of magnitude less than the electron's mass. See J. D. Jackson, *Classical Electrodynamics* (Wiley, 1975), pp. 5-9.
[20] In Eq. (80) the term $\partial_\mu A^\mu$ can be made to vanish (the "Lorentz gauge") if it does not already with a given A^μ, by carrying out a gauge transformation and finding a χ that makes $\partial_\mu A'^\mu = 0$.
[21] The gluons are strictly massless, but quarks remain "confined" within hadrons...but that's another story.
[22] J. Schwinger, *Ann. Phys.* **2**, 407 (1957).
[23] S. A. Bludman, *Nuovo Cimento* **9**, 443 (1958).
[24] P. W. Higgs, *Phys. Rev. Lett.* **12**, 132 (1964); *Phys. Rev. Lett.* **13**, 508 (1964); *Phys. Rev.* **145**, 1156 (1966).
[25] S. Glashow, *Nucl. Phys.* **22**, 579 (1961).
[26] A. Salam and J. C. Ward, *Phys. Lett.* **13**, 168 (1964).
[27] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.* **19**, 1264-1266 (1967). The present description of the electroweak model follows Weinberg's 1967 paper.
[28] "Spontaneous symmetry breaking" should be contrasted to "dynamical symmetry breaking" where internal dynamics, with multiple-vertex diagrams describing the same fermion emitting and absorbing gauge bosons, generates mass. But that, too, forms another story.
[29] E.g., B. Bonner and H. Miettinen (Eds.), *Prodeedings of the Rice Meeting*, Vol. 2 (World Scientific, 1990); K. Milton, R. Kantowski, and M. Samuel (Eds.), *Beyond the Standard Model II* (World Scientific, 1991).
[30] H. Georgi and S. Glashow, "Unity of All Elementary-Particle Forces," *Phys. Rev. Lett.* **32**, 438-441 (1974).

◆