Spontaneous symmetry breaking as a basis of particle mass

Chris Quigg

Theoretical Physics Department, Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA
and
Theory Group, Physics Department, CERN, CH-1211 Geneva 23, Switzerland

E-mail: quigg@fnal.gov

Received 30 March 2007
Published 8 June 2007
Online at stacks.iop.org/RoPP/70/1019

Abstract

Electroweak theory joins electromagnetism with the weak force in a single quantum field theory, ascribing the two fundamental interactions—so different in their manifestations—to a common symmetry principle. How the electroweak gauge symmetry is hidden is one of the most urgent and challenging questions facing particle physics. The provisional answer incorporated in the ‘standard model’ of particle physics was formulated in the 1960s by Higgs, by Brout and Englert and by Guralnik, Hagen, and Kibble: the agent of electroweak symmetry breaking is an elementary scalar field whose self-interactions select a vacuum state in which the full electroweak symmetry is hidden, leaving a residual phase symmetry of electromagnetism. By analogy with the Meissner effect of the superconducting phase transition, the Higgs mechanism, as it is commonly known, confers masses on the weak force carriers \( W^\pm \) and \( Z \). It also opens the door to masses for the quarks and leptons, and shapes the world around us. It is a good story—though an incomplete story—and we do not know how much of the story is true. Experiments that explore the Fermi scale (the energy regime around 1 TeV) during the next decade will put the electroweak theory to decisive test, and may uncover new elements needed to construct a more satisfying completion of the electroweak theory. The aim of this article is to set the stage by reporting what we know and what we need to know, and to set some ‘big questions’ that will guide our explorations.

(Some figures in this article are in colour only in the electronic version)
## Contents

1. Introduction to mass 1021
2. Sources of mass in the electroweak theory 1023
3. The significance of the Fermi scale 1029
4. A dynamical approach to electroweak symmetry breaking 1030
5. A world without the Higgs mechanism 1031
6. The problem of identity 1032
7. A closer look at fermion masses 1033
8. In search of the standard-model Higgs boson 1036
9. More new physics on the Fermi scale? 1044
10. Another look at dynamical symmetry breaking 1046
11. The vacuum energy puzzle 1047
12. Big questions 1048
   - Acknowledgments 1049
   - References 1049
1. Introduction to mass

In the opening lines of his *Principia*, [1, 2], Newton defines mass as ‘the quantity of matter ... arising from its density and bulk conjointly.’ That intuitive notion of mass as an intrinsic attribute of matter, sharpened by $F = ma$ and the law of universal gravitation, is a foundation of classical physics. Mass, for Newton, is at once a measure of inertia and a source of gravitational attraction. It follows directly that mass is conserved: the mass of an object is the sum of the masses of its parts, in agreement with everyday experience. The extension of the law of conservation of mass to the realm of chemical reactions by Lavoisier and Lomonosov was central to the development of chemistry as a quantitative science, leading—through the work of Dalton and others—to the empirical underpinnings of the modern atomic theory. But in the classical worldview, mass does not arise, it simply is.

Mass remained an essence—part of the nature of things—for more than two centuries, until Abraham (1903) and Lorentz (1904) sought to interpret the electron mass as electromagnetic self-energy\(^1\). Our modern conception of mass has its roots in Einstein’s pregnant question [5], ‘Does the inertia of a body depend upon its energy content?’ and his powerful conclusion, ‘The mass of a body is a measure of its energy content; if the energy changes by $L$, the mass changes in the same sense by $[L/c^2]$, where $c$ is the speed of light.’ \textit{Mass is rest energy}\(^2\). Among the virtues of identifying mass as $m = E_0/c^2$, where $E_0$ designates the body’s rest energy, is that mass, so understood, is a Lorentz-invariant quantity, given in any frame as $m = (1/c^2) \sqrt{E^2 - p^2 c^2}$. But not only is Einstein’s a precise definition of mass, it invites us to consider the origins of mass by coming to terms with a body’s rest energy.

For the first few steps down the quantum ladder, the difference between the Einsteinian conception of mass as rest energy and the Newtonian expectation that the mass of an object is the sum of the masses of its parts is subtle but telling. We understand the mass of an atom or molecule in terms of the masses of the atomic nuclei, the mass of the electron, and small corrections for binding energy that are given by quantum electrodynamics. ‘Small corrections’ is perhaps an understatement. The 13.6 eV binding energy of the 1S electron in the hydrogen atom is but $1.45 \times 10^{-8}$ of the atom’s mass. And the 13 MJ liberated in burning a cubic metre of hydrogen in the reaction $2\text{H}_2 + \text{O}_2 \rightarrow 2\text{H}_2\text{O}$ corresponds to a fractional mass difference between reactants and products of only $9 \times 10^{-11}$. The law of conservation of mass holds to an impressive degree; our fossil-fuel economy feeds on the tiny deviations.

In precise and practical—if not quite ‘first-principle’—terms, the masses of all the nuclei follow from the proton mass, the neutron mass, and our semi-empirical knowledge of nuclear forces. The deeply bound $\alpha$ particle ($^4\text{He}$) has a mass defect of only $\frac{3}{2} \%$, so even in the nuclear realm the notion that the mass of an object is the sum of the masses of its parts is an excellent first approximation. On a macroscopic scale, that small mass defect is a proxy for a prodigious store of energy.

Nucleon mass is an entirely different story, the very exemplar of $m = E_0/c^2$. Quantum chromodynamics (QCD), the gauge theory of the strong interactions, teaches that the dominant contribution to the nucleon mass is not the masses of the quarks that make up the nucleon, but the energy stored up in confining the quarks in a tiny volume [9]. The masses $m_u$ and $m_d$ of the up and down quarks are only a few MeV each [10]. The quarks contribute no more than 2% to the 939 MeV mass of an isoscalar nucleon (averaging proton and neutron properties),

\(^1\) For Lorentz’s résumé of his thinking, see [3]; a modern perspective, with extensive citations to other work, appears in [4].

\(^2\) For a vivid presentation of the context and impact of the 1905 papers, see [6], and for a brisk tour of the concept of mass, see [7]. To hear $E = mc^2$ pronounced in Einstein’s own words, consult [8].
because
\[ \frac{3m_u + m_d}{2} = 7.5 - 16.5 \text{ MeV}. \] (1)

Hadrons such as the proton and neutron thus represent \textit{matter of a novel kind}. In contrast to macroscopic matter, and to a degree far beyond what we observe in atoms, molecules and nuclei, the mass of a nucleon is not equal to the sum of its constituent masses (up to small corrections for binding energy); \textit{it is confinement energy} (up to small corrections for constituent masses)!

QCD formulated on a spacetime lattice brings a quantitative dimension to these statements. The CP-PACS Collaboration (centred in Tsukuba, Japan) has made a calculation omitting virtual quark–antiquark pairs that matches the observed light-hadron spectrum at the 10% level [11]. That discrepancy is larger than the statistical and systematic uncertainties, and so is interpreted as an artefact of the quenched (no dynamical fermions) approximation. New calculations that include virtual quark–antiquark pairs should show the full quantitative power of lattice QCD, and give us new insights into the successes and shortcomings of the simple quark model [12, 13].

The tiny u and d quark masses do account for an important detail of the nucleon spectrum that is essential to the world we know. The counterintuitive observation that the neutral neutron (udd) is 1.29 MeV more massive than the charged proton (uud) is explained by the fact that \( m_d \) exceeds \( m_u \) by enough to overcome the proton’s greater electromagnetic self-energy. Together with nuclear binding forces, the neutron–proton mass difference determines the pattern of radioactive decays and the roster of naturally occurring stable nuclei.

Let us be clear about the meaning of the successful calculation of the hadron spectrum using the methods of lattice QCD. In identifying the energy of quark confinement as the origin of the nucleon mass, \textit{quantum chromodynamics has explained nearly all the visible mass of the Universe}, since the luminous matter is essentially made of protons and neutrons in stars and clouds. The oft-repeated assertion that the Higgs boson is the source of all mass in the Universe is simply incorrect—even if we restrict our attention to the luminous stuff made of ordinary baryonic matter.

The Higgs boson and the mechanism that distinguishes electromagnetism from the weak interactions are nevertheless of capital importance in shaping our world, accounting for the masses of the weak-interaction force particles and—at least in the standard electroweak theory—giving masses to the quarks and leptons. Understanding the Higgs mechanism—or whatever stands in its stead—will give us new insight into why atoms exist, how atoms can form chemical bonds, and what makes possible stable structures. These are some of the deepest questions humans have ever pursued, and they are coming within the reach of particle physics. Over the next decade, experiments will carry out definitive explorations of the Fermi scale, at energies around 1 TeV for collisions among quarks and leptons. This \textit{nanonanoscale physics} probes distances smaller than \( 10^{-18} \text{ m} \), where we confidently expect to find the key to the mechanism that drives electroweak symmetry breaking. A pivotal step will be the search for the Higgs boson and the elaboration of its properties. In the same fresh terrain, we suspect that other new phenomena will give new insight into why the electroweak scale is so much smaller than the Planck scale. A class of weakly interacting dark-matter candidates could also populate the Fermi scale.

Resolving the conundrums of the Fermi scale should permit us to see more clearly the little-known territory at still shorter distances, where we may uncover new challenges to our understanding. We could well find new clues to the unification of forces or indications for a rational pattern of constituent masses, viewed at a high-energy scale. If the agent of electroweak symmetry breaking turns out to be an elementary scalar, as the standard model would have it, it
would be the first such particle known to experiment. Learning how it behaves could contribute important new intuition about the early-time dynamics of the inflationary Universe \[14, 15\] and the origins of dark energy \[16, 17\].

We shall begin our tour by reviewing the standard electroweak theory, paying particular attention to the mechanism that hides the electroweak symmetry and generates masses for the weak gauge bosons, the quarks and the leptons. The electroweak theory points to the energy scale around 1 TeV, or \(10^{12}\) eV, for crucial information. The Large Hadron Collider \[18\] soon to operate at CERN has been designed to empower experiments to carry out a thorough exploration of the Fermi scale. In the future, we expect the International Linear Collider to enrich our portrait of electroweak symmetry breaking and new phenomena on the 1 TeV scale \[19\]. Then we shall look briefly at a simple alternative mechanism in which the electroweak symmetry is dynamically broken by QCD. Those considerations will prepare us to uncover the broad significance of electroweak symmetry breaking by asking what the world would be like if there were no Higgs mechanism to hide the electroweak symmetry.

Next, we will consider the problem of identity—what distinguishes this quark from that lepton—and be led to take a closer look at fermion mass and mixings. We will describe signatures that will be important in the search for the Higgs boson. Then we will argue, independent of any specific mechanism for electroweak symmetry breaking, that (something like) the Higgs boson must exist. We shall find that additional considerations also single out the 1-TeV scale as fertile terrain for new physics. We then look again at the idea of dynamical symmetry breaking, transplanting the QCD strategy to the Fermi scale. We shall describe a great gap in our understanding: the vacuum energy problem, and its connection to the electroweak theory. We close with a catalogue of big questions for the decade of discovery ahead.

\section{Sources of mass in the electroweak theory}

We build the standard model of particle physics\(^3\) on a set of constituents that we regard provisionally as elementary: the quarks and leptons, as depicted in figure 1, plus a few

\(^3\) For general surveys of the standard model of particle physics, and a glimpse beyond, see \[20, 21\].
fundamental forces derived from gauge symmetries. The quarks are influenced by the strong interaction, and so carry colour, the strong-interaction charge, whereas the leptons do not feel the strong interaction and are colourless. We idealize the quarks and leptons as pointlike, because they show no evidence of internal structure at the current limit of our resolution ($r \approx 10^{-18}$ m). The charged–current weak interaction responsible for radioactive beta decay and other processes acts only on the left-handed fermions. Whether the observed parity violation reflects a fundamental asymmetry in the laws of Nature, or a left–right symmetry that is hidden by circumstance and might be restored at higher energies, we do not know.

The electroweak theory\(^4\) (like QCD) is a gauge theory, in which interactions follow from symmetries. Already in the 1930s, Fermi \([31]\) and Klein \([32]\) proposed descriptions of the weak interaction in analogy to the emerging theory of quantum electrodynamics (QED). The correct electroweak gauge symmetry, which melds the $SU(2)_L$ family (weak-isospin) symmetry suggested by the left-handed doublets of figure 1 with a $U(1)_Y$ weak-hypercharge phase symmetry, emerged through trial and error, guided by experiment. We characterize the $SU(2)_L \otimes U(1)_Y$ theory by the left-handed quarks

$$L^{(1)}_q = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad L^{(2)}_q = \begin{pmatrix} c \\ s \end{pmatrix}_L, \quad L^{(3)}_q = \begin{pmatrix} t \\ b \end{pmatrix}_L,$$

with weak isospin $I = \frac{1}{2}$ and weak hypercharge $Y(L_q) = \frac{1}{6}$; their right-handed weak-isoscalar counterparts

$$R^{(1,2,3)}_u = u_R, c_R, t_R \quad \text{and} \quad R^{(1,2,3)}_d = d_R, s_R, b_R,$$

with weak hypercharges $Y(R_u) = \frac{4}{3}$ and $Y(R_d) = -\frac{2}{3}$; the left-handed leptons

$$L_e = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad L_\mu = \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \quad L_\tau = \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L,$$

with weak isospin $I = \frac{1}{2}$ and weak hypercharge $Y(L_\ell) = -1$; and the right-handed weak-isoscalar charged leptons

$$R_{e,\mu,\tau} = e_R, \mu_R, \tau_R,$$

with weak hypercharge $Y(R_\ell) = -2$. (Weak isospin and weak hypercharge are related to electric charge through $Q = I_3 + \frac{1}{2}Y$.) Here we have idealized the neutrinos as massless; we will touch on possible sources of neutrino mass in section 7.

The $SU(2)_L \otimes U(1)_Y$ electroweak gauge group implies two sets of gauge fields: a weak isovector $b_\mu$, with coupling constant $g$, and a weak isoscalar $A_\mu$, with independent coupling constant $g'$. The gauge fields compensate for the variations induced by gauge transformations, provided that they obey the transformation laws $b_\mu \rightarrow b_\mu - \alpha \times b_\mu - (1/g)\partial_\mu \alpha$ under an infinitesimal weak-isospin rotation generated by $G = 1 + (i/2)\alpha \cdot \tau$ (where $\tau$ are the Pauli isospin matrices) and $A_\mu \rightarrow A_\mu - (1/g')\partial_\mu \alpha$ under an infinitesimal hypercharge phase rotation. Corresponding to these gauge fields are the field-strength tensors

$$F^\ell_{\mu\nu} = \partial_\mu b^\ell_\nu - \partial_\nu b^\ell_\mu + g\epsilon_{ijk} b^j_\mu b^k_\nu,$$

for the weak-isospin symmetry, and

$$f^\ell_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

for the weak-hypercharge symmetry.

We may summarize the interactions by the Lagrangian

$$L = L_{\text{gauge}} + L_{\text{leptons}} + L_{\text{quarks}},$$

\(^4\) Many textbooks develop the electroweak theory; see in particular \([22–25]\). For a look back at how the electroweak theory came to be, see the Nobel Lectures by some of its principal architects \([26–30]\).
Origins of mass

with

\[ \mathcal{L}_{\text{gauge}} = -\frac{1}{4} F_{\mu\nu} \cdot F^{\mu\nu} - \frac{1}{4} f_{\mu\nu} f^{\mu\nu}, \]  \hfill (9)

\[ \mathcal{L}_{\text{leptons}} = \overline{\mathcal{R}_\ell} i \gamma^\mu \left( \partial_\mu + i \frac{g'}{2} A_\mu Y \right) \mathcal{R}_\ell + \mathcal{L}_{\text{quarks}} \]  \hfill (10)

where \( \ell \) runs over e, \( \mu \), \( \tau \), and

\[ \mathcal{L}_{\text{quarks}} = \overline{\mathcal{R}^{(n)}_u} i \gamma^\mu \left( \partial_\mu + i \frac{g'}{2} A_\mu Y \right) \mathcal{R}^{(n)}_u + \overline{\mathcal{R}^{(n)}_d} i \gamma^\mu \left( \partial_\mu + i \frac{g'}{2} A_\mu Y \right) \mathcal{R}^{(n)}_d \]

\hfill (11)

where \( n \) runs over 1, 2, 3.

Although the weak and electromagnetic interactions share a common origin in the \( SU(2)_L \otimes U(1)_Y \) gauge symmetry, their manifestations are very different. Electromagnetism is a force of infinite range, while the influence of the charged–current weak interaction responsible for radioactive beta decay only spans distances shorter than about \( 10^{-15} \) cm. The phenomenology is thus at odds with the theory we have developed to this point. The gauge Lagrangian (9) contains four massless electroweak gauge bosons, namely \( A_\mu \), \( b_{1\mu} \), \( b_{2\mu} \), and \( b_{3\mu} \), because a mass term such as \( \frac{1}{2} m^2 A_\mu A^\mu \) is not invariant under a gauge transformation. Nature has but one: the photon. Moreover, the \( SU(2)_L \otimes U(1)_Y \) gauge symmetry forbids fermion mass terms \( \overline{\mathcal{R}} \mathcal{R} \) in (10) and (11), because the left-handed and right-handed fields transform differently.

To give masses to the gauge bosons and constituent fermions, we must hide the electroweak symmetry, recognizing that a symmetry of the laws of Nature does not imply that the same symmetry will be manifest in the outcomes of those laws. How the electroweak gauge symmetry is spontaneously broken—hidden—to the \( U(1)_\text{em} \) phase symmetry of electromagnetism is one of the most urgent and challenging questions before particle physics.

The superconducting phase transition offers an instructive model for hiding the electroweak gauge symmetry\(^5\). To give masses to the intermediate bosons of the weak interaction, we appeal to the Meissner effect—the exclusion of magnetic fields from a superconductor, which corresponds to the photon developing a nonzero mass within the superconducting medium. What has come to be called the Higgs mechanism \([35–40]\) is a relativistic generalization of the Ginzburg-Landau phenomenology \([41]\) of superconductivity\(^6\). The essential insight is that the Goldstone theorem \([49, 50]\) does not operate when a local gauge symmetry, as opposed to a continuous global symmetry, is broken. Instead, a miraculous interplay between the would-be Goldstone bosons and the normally massless gauge bosons endows gauge bosons with mass and removes the massless scalars from the spectrum.

Let us see how spontaneous symmetry breaking operates in the electroweak theory. We introduce a complex doublet of scalar fields

\[ \phi \equiv \left( \begin{array}{c} \phi^s \\ \phi^0 \end{array} \right) \]  \hfill (12)

\(^5\) See section 4.4 of \([33]\) and section 21.6 of \([34]\).

\(^6\) Early steps towards understanding spontaneously broken gauge symmetry are recalled in \([42–48]\).

\(^7\) If the Lagrangian of a local, manifestly Lorentz-invariant quantum field theory with positive-definite metric on the Hilbert space is invariant under a continuous symmetry, then either the vacuum state is also invariant under that symmetry or a massless spin-zero ‘Goldstone boson’ corresponds to each broken generator. Nonrelativistic evasions are exhibited in \([51–53]\).
with weak hypercharge $Y_\phi = +1$. Next, we add to the Lagrangian new (gauge-invariant) terms for the interaction and propagation of the scalars,
\[ \mathcal{L}_\text{scalar} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi_\dagger \phi), \] (13)
where the gauge-covariant derivative is
\[ D_\mu = \partial_\mu + i \frac{g}{2} A_\mu Y + i \frac{g}{2} \tau^a b_\mu, \] (14)
and (inspired by Ginzburg and Landau) the potential interaction has the form
\[ V(\phi_\dagger \phi) = \mu^2 (\phi^\dagger \phi) + |\lambda| (\phi^\dagger \phi)^2. \] (15)

We are also free to add gauge-invariant Yukawa interactions between the scalar fields and the leptons ($\ell$ runs over $e$, $\mu$, $\tau$ as before),
\[ \mathcal{L}_\text{Yukawa–\ell} = -\frac{\lambda_\ell}{\sqrt{2}} \left[ (L_\ell \phi_\dagger) R_\ell + \overline{R_\ell} (\phi \ell_\dagger) \right], \] (16)
and similar interactions with the quarks—about which we shall have more to say presently, in section 7.

We then arrange their self-interactions so that the vacuum state corresponds to a broken-symmetry solution. The electroweak symmetry is spontaneously broken if the parameter $\mu^2$ is taken to be negative. In that event, gauge invariance gives us the freedom to choose the state with weak hypercharge $Y$. The vacuum expectation value is taken to be negative. In that event, gauge invariance gives us the freedom to choose the state with weak hypercharge $Y$.

\[ \langle \phi \rangle = \left( \begin{array}{c} 0 \\ v/\sqrt{2} \end{array} \right), \] (17)

where $v = \sqrt{-\mu^2/|\lambda|}$.

Let us verify that the vacuum of (17) does break the gauge symmetry $SU(2)_L \otimes U(1)_Y \rightarrow U(1)_{\text{em}}$. The vacuum state $\langle \phi \rangle$ is invariant under a symmetry operation corresponding to the generator $G$ provided that $e^{-ig'G} \langle \phi \rangle = \langle \phi \rangle$, i.e. if $G \langle \phi \rangle = 0$. Direct calculation reveals that the original four generators are all broken, but electric charge is not. The photon remains massless, but the other three gauge bosons acquire masses, as auxiliary scalars assume the role of the third (longitudinal) degree of freedom.

Introducing the weak mixing angle $\theta_W$ and defining $g' = g \tan \theta_W$, we can express the photon as the linear combination $A = A \cos \theta_W + b_3 \sin \theta_W$. We identify the strength of its (pure vector) coupling to charged particles, $g g'/\sqrt{g^2 + g'^2}$, with the electric charge $e$. The mediator of the charged–current weak interaction, $W^\pm = (b_1 \mp i b_2)/\sqrt{2}$, acquires a mass $M_W = g v/2 = e v/2 \sin \theta_W$. The electroweak gauge theory reproduces the low-energy phenomenology of the $V–A$ theory of weak interactions, provided we set $v = (G_F \sqrt{2})^{-1/2} = 246 \text{ GeV}$, where $G_F = 1.16637(1) \times 10^{-5} \text{ GeV}^{-2}$ is Fermi’s weak-interaction coupling constant. It follows at once that $M_W \approx 37.3 \text{ GeV} / \sin \theta_W$. The combination of the $I_3$ and $Y$ gauge bosons orthogonal to the photon is the mediator of the neutral–current weak interaction, $Z = b_3 \cos \theta_W - A \sin \theta_W$, which acquires a mass $M_Z = M_W / \cos \theta_W$. The weak neutral–current interaction was not known before the electroweak theory. Its discovery in 1973 [54, 55] marked an important milestone, as did the observation a decade later [56] of the $W^\pm$ [57, 58] and $Z^0$ [59, 60] bosons.

Three decades of extensive studies of the weak neutral current culminated in experiments at the $e^+e^- \rightarrow Z$ factories. The ALEPH, DELPHI, L3, and OPAL detectors accumulated 17 million $Z$ decays at LEP, and the SLD detector recorded 600 thousand $Z$ decays using polarized beams at the Stanford Linear Collider [61]. A broad collection of experimental measurements and the supporting theoretical calculations have elevated the electroweak theory to a law of Nature, tested as a quantum field theory at the level of one part in a thousand [10, 62–64]. The mass of the neutral weak boson is known to impressive precision, $M_Z = 91.1876 \pm 0.0021 \text{ GeV}$,
while the world average $W$-boson mass is $M_W = 80.398 \pm 0.025\,\text{GeV}$ [65]. By themselves, these measurements imply an ‘on-shell’ weak mixing parameter $\sin^2 \theta_W \equiv 1 - M_W^2 / M_Z^2 = 0.22265 \pm 0.00052$, in fine agreement with the Particle Data Group 2006 grand average, $0.22306 \pm 0.00031$ [10]. The quantum (loop) corrections to many observables, including the ratio $M_W/M_Z$, are sensitive to the top-quark mass, and showed a preference for a heavy top before the discovery. Now the comparison of indirect inferences with the measured top-quark mass is one of many consistency checks for the electroweak theory. The 2006 inferred value, $m_t^{\text{indirect}} = 172.3^{+10.2}_{-7.6}\,\text{GeV}$ [10] matches the (more precise) Tevatron average of direct measurements, $m_t = 170.9 \pm 1.8\,\text{GeV}$ [66].

One noteworthy achievement is a clear test of the electroweak gauge symmetry in the reaction $e^+e^- \rightarrow W^+W^-$. Neglecting the electron mass, this reaction is described by three Feynman diagrams that correspond to s-channel photon and $Z^0$ exchange, and t-channel neutrino exchange, figures 2(a)–(c). For the production of longitudinally polarized $W$-bosons, each diagram leads to a $J = 1$ partial-wave amplitude that grows as the square of the c.m. energy, but the gauge symmetry enforces a pattern of cooperation. The contributions of the direct-channel $\gamma$- and $Z^0$-exchange diagrams of figures 2(a) and (b) cancel the leading divergence in the $J = 1$ partial-wave amplitude of the neutrino-exchange diagram in figure 2(c). The interplay is shown in figure 3. If the $Z$-exchange contribution is omitted (middle line) or if both the $\gamma$- and $Z$-exchange contributions are omitted (upper line), the calculated cross section grows unacceptably with energy. The measurements compiled by the LEP Electroweak Working Group [63] agree well with the benign high-energy behaviour predicted by the full electroweak theory, confirming the gauge cancellation in the $J = 1$ partial-wave amplitude.

Three of the four scalar degrees of freedom that we introduced to forge a vacuum state that does not manifest the gauge symmetry have become the longitudinal components of $W^+$,
Figure 3. Cross section for the reaction $e^+e^- \rightarrow W^+W^-$ measured by the four LEP experiments, together with the full electroweak-theory simulation and the cross sections that would result from $\nu$-exchange alone and from $(\nu + \gamma)$-exchange [63].

Table 1. Masses of the charged leptons and quarks [10].

| Particle | Mass (MeV) |
|----------|------------|
| e        | $0.510\,998\,92 \pm 0.000\,000\,04$ |
| $\mu$    | $105.658\,369 \pm 0.000\,009$ |
| $\tau$   | $1\,776.99^{+0.29}_{-0.26}$ |
| u        | $2.25 \pm 0.75$ |
| d        | $5.00 \pm 2.00$ |
| s        | $95 \pm 25$ |
| c        | $1\,250 \pm 90$ |
| b        | $4\,200 \pm 70$ |
| t        | $170\,900 \pm 1\,800$ [66] |

$W^-$ and $Z$. What of the fourth? It appears as a vestige of the spontaneous symmetry breaking, in the form of a massive spin-zero particle, called the Higgs boson, $H$. Its mass is given symbolically as $M_H^2 = -2\mu^2 > 0$, but we have no prediction for its value. What we take to be the work of the Higgs boson is all around us⁸, as we shall detail below, but the Higgs particle itself has not yet been observed!

The masses of the elementary fermions collected in table 1 are a more mysterious story. Each fermion mass involves a new Yukawa coupling $\zeta$ (cf (16)). When the electroweak symmetry is spontaneously broken, the electron mass emerges as $m_e = \zeta e v/\sqrt{2}$. The Yukawa couplings that reproduce the observed quark and lepton masses range over many orders of magnitude, as shown in figure 4. The origin of the Yukawa couplings is obscure: they do not follow from a known symmetry principle, for example. In that sense, therefore, all fermion masses involve physics beyond the standard model.

⁸ We shall see in section 8 that the Higgs-boson contribution of figure 2(d) ensures acceptable high-energy behaviour of the $J = 0$ partial-wave amplitude.
Let us summarize what we have learned about the sources of particle mass in the standard electroweak theory. Unless the electroweak gauge symmetry is hidden, the four gauge bosons and all the constituent fermions are massless. Spontaneous symmetry breaking, in the form of the Higgs mechanism, gives masses to the weak gauge bosons and creates the possibility for the fermions to acquire mass. Once the weak mixing parameter $\sin^2 \theta_W$ is fixed by the study of weak-neutral–current interactions, the theory makes successful quantitative predictions for the $W^\pm$- and $Z^-$-boson masses. Although the natural scale of fermion masses would seem to be set by the electroweak scale, the masses themselves are determined by Yukawa couplings of the fermions to the Higgs field. We do not know what fixes the size of the Yukawa couplings. Finally, the theory requires a scalar Higgs boson, but does not make an explicit prediction for its mass.

3. The significance of the Fermi scale

The electroweak theory does not make a testable prediction for the Higgs-boson mass because we do not know the values of the Higgs-potential parameters $|\lambda|$ and $\mu^2$ that enter the symbolic prediction $M_H^2 = -2\mu^2 = 2 |\lambda| v^2$. However, a thought experiment [67] leads to a conditional upper bound on the Higgs-boson mass that identifies a high-value experimental target. Imagine pairwise collisions among the $W^\pm$ bosons that mediate charge-changing weak interactions, the $Z$ boson responsible for weak neutral–current interactions and the Higgs boson $H$. Most channels ‘decouple,’ in the sense that partial-wave amplitudes are small at all energies (except very near the particle poles or at exponentially large energies), for any value of the Higgs boson mass $M_H$. Four neutral channels—$W_0^+ W_0^-$, $Z_0 Z_0$, $HH$ and $HZ_0$, where the subscript 0 denotes longitudinal polarization—are particularly interesting. Amplitudes calculated at lowest order in the electroweak theory make sense in the limit of high energies, in that the probability of a scattering event does not exceed unity, provided that the Higgs-boson mass is not too large. Specifically, the $J = 0$ partial-wave amplitudes involving these channels are all asymptotically constant (i.e. well-behaved) and proportional to Fermi’s constant $G_F$ times the square of the Higgs-boson mass. Requiring that the largest eigenvalue respect the partial-wave...
unitarity condition $|a_0| \leq 1$ yields

$$M_H \leq \left( \frac{8\pi \sqrt{2}}{3G_F} \right)^{1/2} \approx 1 \text{ TeV},$$

which characterizes the Fermi scale.

If $M_H$ respects the bound, weak interactions remain weak at all energies, and perturbation theory is everywhere reliable. If the Higgs boson were heavier than 1 TeV, the weak interactions among $W^\pm$, $Z$ and $H$ would become strong on the Fermi scale, and perturbation theory would break down. At TeV energies, we might then observe multiple production of weak bosons, $W^+W^-$ resonances and other phenomena evocative of pion–pion scattering at GeV energies [68]. One way or another, something new—a Higgs boson or strong scattering, if not some other new physics [69, 70]—is to be found in electroweak interactions at energies not much larger than 1 TeV [71, 67, 72].

4. A dynamical approach to electroweak symmetry breaking

The analogy between electroweak symmetry breaking and the superconducting phase transition led to the insight of the Higgs mechanism. The macroscopic order parameter of the Ginzburg–Landau phenomenology, which corresponds to the wave function of superconducting charge carriers, acquires a nonzero vacuum expectation value in the superconducting state. The microscopic Bardeen–Cooper–Schrieffer theory [73] interprets the dynamical origin of the order parameter through the formation of correlated states of elementary fermions, the Cooper pairs of electrons.

The elementary fermions—electrons—and gauge interactions—QED—needed to generate the correlated pairs are already present in the case of superconductivity. Could a scheme of similar economy account for the transition that hides the electroweak symmetry? Consider an $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ theory of massless up and down quarks. Because the strong interaction is strong and the electroweak interaction is feeble we may treat the $SU(2)_L \otimes U(1)_Y$ interaction as a perturbation. For vanishing quark masses, QCD displays an exact $SU(2)_L \otimes SU(2)_R$ chiral symmetry. At an energy scale $\sim \Lambda_{QCD}$, the strong interactions become strong, fermion condensates appear, and the chiral symmetry is spontaneously broken to the familiar flavour symmetry, isospin:

$$SU(2)_L \otimes SU(2)_R \rightarrow SU(2)_V.$$ (19)

Three Goldstone bosons appear, one for each broken generator of the original chiral invariance. These were identified by Nambu [74] as three massless pions.

The broken generators are three axial currents whose couplings to pions are measured by the pion decay constant $f_\pi$, which is measured by the charged-pion lifetime. When we turn on the $SU(2)_L \otimes U(1)_Y$ electroweak interaction, the electroweak gauge bosons couple to the axial currents and acquire masses of order $\sim g f_\pi$. The mass-squared matrix,

$$M^2 = \begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & gg' \\ 0 & gg' & gg' & g^2 \end{pmatrix} \frac{f_\pi^2}{4}$$

(20)

(where the rows and columns correspond to $W_1$, $W_2$, $W_3$, and $A$) has the same structure as the mass-squared matrix for gauge bosons in the standard electroweak theory. Diagonalizing the matrix (20), we find that the photon, corresponding as in the standard model to the combination $A = (gA + g'B)/\sqrt{g^2 + g'^2}$, emerges massless. Two charged gauge bosons,
W^\pm = (b_1 \mp ib_2)/\sqrt{2}, acquire mass-squared \( M_W^2 = g^2 f_\pi^2 / 4 \), and the neutral gauge boson Z = (-g' A + gb_3)/\sqrt{g^2 + g'^2} obtains \( M_Z^2 = (g^2 + g'^2) f_\pi^2 / 4 \). The ratio,

\[
\frac{M_Z^2}{M_W^2} = \frac{(g^2 + g'^2)}{g^2} = \frac{1}{\cos^2 \theta_W},
\]

reproduces the standard-model result.

The massless pions thus disappear from the physical spectrum, having become the longitudinal components of the weak gauge bosons. Despite the structural similarity to the standard model, the chiral symmetry breaking of QCD does not yield a satisfactory theory of the weak interactions. The masses acquired by the intermediate bosons are 2500 times smaller than required for a successful low-energy phenomenology; the W-boson mass is only \( M_W \approx 30 \text{ MeV} \), because its scale is set by \( f_\pi \).

5. A world without the Higgs mechanism

Exploring the Fermi scale will bring us a new appreciation of what lies behind the complexity and diversity of the everyday world. To see what we can hope to learn, let us consider how the world would be changed if we could not rely on something like the Higgs mechanism to hide electroweak symmetry.

What a different world it would be! First, the quarks and leptons would remain massless, because mass terms are not permitted if the electroweak symmetry remains manifest. Eliminating the Higgs mechanism does nothing to alter the strong interaction, so QCD would still confine the (now massless) colour-triplet quarks into colour-singlet hadrons, now with many ‘light hadrons’ because there are no heavy quarks.

If the electroweak symmetry were unbroken, the asymptotically free weak-isospin force would confine objects that carry weak isospin into weak-isospin singlets. But as we have just seen in section 4, the chiral condensate of QCD hides the electroweak symmetry, even in the absence of a Higgs mechanism. Because the weak bosons have acquired mass, the SU(2)_L interaction does not confine. The familiar light-hadron spectrum persists, but with a crucial difference. In the no-Higgs-boson scenario, with no u–d quark mass difference to tip the balance, the proton would outweigh the neutron. The pattern of radioactive beta decay would be turned on its head. In our world, a free neutron decays \((n \rightarrow p e^- \bar{\nu}_e)\) with a mean life of about fifteen minutes. If quark masses vanish and \( M_W \approx 30 \text{ MeV} \), a free proton decays in less than a millisecond: \( p \rightarrow n e^+ \nu_e \). There is no hydrogen atom, and the lightest ‘nucleus’ would be one neutron.

It seems likely that some light elements would be produced in the early no-Higgs universe [76–79]. But even if some nuclei are produced and survive, they would not form atoms we would recognize. A massless electron means that the Bohr radius of an atom—half a nanometre in our world—would be infinite. (Now, it is nearly inevitable that effects negligible in our world would, in the Higgsless world, produce fermion masses many orders of magnitude smaller than those we observe. The Bohr radius of a would-be atom would be macroscopic, sustaining the conclusion that matter would lose its integrity.)

A world without compact atoms would be a world without chemical valence bonds and without stable composite structures like our solids and liquids. All matter would be insubstantial—and we would not exist!
6. The problem of identity

Contemplating the variety of the quarks and leptons invites the tantalizing question, ‘What makes a top quark a top quark, an electron an electron, and a neutrino a neutrino?’ In more operational terms, we may ask, ‘What determines the masses and mixings of the quarks and leptons?’ It is not enough to answer, ‘The Higgs mechanism,’ because the fermion masses are a very enigmatic element of the electroweak theory. Once the electroweak symmetry is hidden, the electroweak theory permits—welcomes—fermion masses, but the values of the masses are set by the couplings of the Higgs boson to the fermions, which are of unknown provenance. Nothing in the electroweak theory is ever going to prescribe those couplings. It is not that the calculation is technically challenging; there is no calculation.

The exciting prospect, then, is that quark and lepton masses, mixing angles and the subtle differences in the behaviour of particles and their antiparticles manifested as CP violation put us in contact with physics beyond the standard model. The challenge is to construct what the big question really is. We may find new phenomena that suggest the origin of some or all of the quark and lepton masses. The extremely light (\(< 2 \text{ eV}\)) neutrinos—which might be their own antiparticles, because they are electrically neutral—may be special, acquiring some or all of their mass from a mechanism not open to the quarks and charged leptons. And it might just be that we have not recognized a latent pattern in the masses because we were not seeing the whole picture yet. Perhaps it will take discovering the spectrum of a new kind of matter—superpartners, or something entirely different—before it all begins to make sense.

Should we expect to find an ultimate resolution to the problem of identity? According to a recurring dream for generations of physicists, the theory of the world might prove to be so restrictive that things have to turn out the way we observe them. Is this really the way the world works, or not? Are the elements of our standard model—the quarks and leptons and gauge groups and coupling constants—inevitable, at least in a probabilistic sense, or did it just happen this way?

It may be instructive to call to mind Johannes Kepler’s quest to understand why the Sun should have exactly six planetary companions in the observed orbits [85]. Kepler sought a symmetry principle that would give order to the universe following the Platonic–Pythagorean tradition. Perhaps, he thought, the six orbits were determined by the five regular solids of geometry, or perhaps by musical harmonies. We now know that the Sun holds in its thrall more than six planets, not to mention the asteroids, periodic comets and planetini, nor all the moons around Kepler’s planets. But that is not why Kepler’s problem seems ill-conceived to us; we just do not believe that it should have a simple answer. Neither symmetry principles nor stability criteria make it inevitable that those six planets should orbit our Sun precisely as they do. This example holds two lessons for us: first, it is very hard to know in advance which aspects of the physical world will have simple, beautiful, informative explanations, and which we shall have to accept as ‘complicated.’ or environmental parameters. Second, and here Kepler is a particularly inspiring example, we may learn very great lessons indeed while pursuing challenging questions that—in the end—do not have illuminating answers.

9 Either family symmetries or unified theories of the strong, weak, and electromagnetic interactions, in which quarks and leptons are members of extended families, offer the prospect of simple relations among fermion masses at a high-energy scale. These relations, which reflect the symmetries and the pattern of symmetry breaking, are modulated by the running of the masses down to the low-energy scales on which we measure them [80–82]. Models that incorporate extra spacetime dimensions present new ways to think about the exponential range of Yukawa couplings. If the standard-model fields were constrained to a thick wall, the wave packets representing different fermion species might be fixed on different tracks within the extra dimension [83,84]. Yukawa couplings would measure the extra-dimensional overlap of the left-handed and right-handed fermion wave packets and the Higgs field, presumed pervasive. Small offsets in the new coordinate could yield exponentially large mass differences.
Are Nature’s Laws the same at all times and places? Yes, of course they are, to good approximation, in our experience. Otherwise science would have had to confront a universe that is in some manner capricious. But all times and all places is a very strong conclusion, for which we cannot have decisive evidence. Some cosmologists argue that our Universe is but one trifling pocket within a multiverse of mind-boggling proportions. Proponents of the string-theory landscape—a mathematical meta-space representing all the alternative environments that theory allows—see a grand panorama of self-consistent possibilities in which what we take to be the laws of nature apply only in our corner. Everywhere in the landscape, the home team plays by its own rules of physics, derived from the locally prevalent gauge symmetries, elementary-particle spectra and coupling constants [86–88].

An eventual explanation of masses and mixings and CP violation might come in the form of inevitability or probability or possibility. Exploring the Fermi scale, it seems to me, is highly likely to resolve the question of mechanism often attributed to Richard Feynman, ‘Why does the muon weigh?’ The follow-up question, ‘What does the muon weigh?’ may be trickier to settle.

7. A closer look at fermion masses

In the standard electroweak theory, the quarks and leptons are taken to be elementary particles. The masses and mixings of the quarks arise from Yukawa interactions with the Higgs condensate,

\[ \mathcal{L}_{\text{Yukawa}} = -\xi_d^{ij}(\bar{L}_i\phi)d_{Rj} - \xi_u^{ij}(\bar{L}_i\phi)u_{Rj} + \text{h.c.} , \]

where the Yukawa couplings \( \xi_{ud} \) are 3 \( \times \) 3 complex matrices, \( i \) and \( j \) are generation indices, \( L_i \) are left-handed quark doublets, \( u_{Rj} \) and \( d_{Rj} \) are right-handed quark singlets, and \( \phi = i\sigma_2\phi^* \).

Equation (22) yields quark mass terms when the Higgs field \( \phi \) acquires a vacuum expectation value (17), as we saw in the penultimate paragraph of section 2. The mass eigenstates are obtained by diagonalizing the Yukawa matrices \( \xi_f^{\text{diag}} = U_f L^\dagger_f \), where \( f = u, d \) refers to up-like or down-like quarks and \( U_{fL}^f \) are unitary matrices (cf figure 4). Accordingly, the charged-current interactions among the left-handed quarks \( u_L = (u_L, c_L, t_L) \) and \( d_L = (d_L, s_L, b_L) \) are specified by

\[ \mathcal{L}^{(q)}_{\text{CC}} = -\frac{g}{\sqrt{2}} \bar{u}_L \gamma^\mu V d_L W^\mu \text{ h.c.} , \]

where

\[ V \equiv U_{uL}^u U_{dL}^d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} . \]

The quark-mixing matrix (24) is called the Cabibbo–Kobayashi-Maskawa (CKM) matrix [89, 90]. We observe significant mixing amplitudes between the first and second generations, modest mixing between the second and third and small mixing between the first and third [10]:

\[ |V| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.974 & 0.227 & 0.004 \\ 0.227 & 0.973 & 0.042 \\ 0.008 & 0.042 & 0.999 \end{pmatrix} . \]

The Higgs scalar is the only element of the standard model that distinguishes among the generations. In Veltman’s phrase [91], it knows something that we do not know.
Figure 5. Favoured values for the light, medium and heavy neutrino masses $m_l$, $m_m$, $m_h$, as functions of the lightest neutrino mass in the three-neutrino oscillation scenario for the normal (left panel) and inverted hierarchy (right panel).

In similar fashion, the charged–current interactions among the left-handed leptonic mass eigenstates $\ell_L = (e_L, \mu_L, \tau_L)$ and $\nu = (\nu_1, \nu_2, \nu_3)$ are specified by

$$\mathcal{L}^{(q)}_{\text{CC}} = -\frac{g}{\sqrt{2}} \bar{\nu} \gamma^\mu V^{\dagger}_{\nu L} W^\mu_{\nu_L} + \text{h.c.},$$

(26)

where

$$\nu = \begin{pmatrix} \nu_{e1} & \nu_{e2} & \nu_{e3} \\ \nu_{\mu1} & \nu_{\mu2} & \nu_{\mu3} \\ \nu_{\tau1} & \nu_{\tau2} & \nu_{\tau3} \end{pmatrix}.$$

(27)

The standard form of the neutrino mixing matrix is owed to the greater familiarity of the flavour eigenstates. It is sometimes called the Pontecorvo—Maki–Nakagawa–Sakata (PMNS) matrix in tribute to neutrino-oscillation pioneers.

By convention, $\nu_1$ and $\nu_2$ are the solar pair with $m_1 < m_2$ and $m_2^2 - m_1^2 = \Delta m^2_{\odot}$. The third mass eigenstate, $\nu_3$, is separated from $\nu_1$ and $\nu_2$ by the atmospheric mass splitting $\Delta m^2_{\text{atm}}$, but current experiments do not tell us whether it lies above (normal hierarchy) or below (inverted hierarchy) the solar doublet. We do not know the absolute scale of neutrino masses, but the effective electron–neutrino mass is constrained by kinematic measurements to be $< 2\text{ eV}$ [10].

We display in figure 5 the normal and inverted spectra that would correspond to a plausible range of values for the lightest neutrino mass, taking the current central values, $\Delta m^2_{\odot} = 7.9 \times 10^{-5} \text{ eV}^2$ and $\Delta m^2_{\text{atm}} = 2.5 \times 10^{-3} \text{ eV}^2$ [96, 97]. If the neutrino masses are Dirac masses generated (in analogy to the quark and charged-lepton masses) by $H \bar{\nu} \nu$ interactions, the Yukawa couplings are $\zeta_\nu \approx 10^{-11}$.

The observed structure of the neutrino mixing matrix differs greatly from the pattern of the quark-mixing matrix. Data indicate that $\nu_3$ consists of nearly equal parts of $\nu_\mu$ and $\nu_\tau$, perhaps with a trace of $\nu_e$, while $\nu_2$ contains similar amounts of $\nu_e, \nu_\mu$ and $\nu_\tau$, and $\nu_1$ is rich in $\nu_e$, with approximately equal minority parts of $\nu_\mu$ and $\nu_\tau$. A recent global fit [98] yields the

10 A plausible limit from the analysis of astronomical data is $\sum m_\nu < 0.62 \text{ eV}$ at 95% CL [95].
following ranges for the magnitudes of the neutrino mixing matrix elements:

\[
|V| = \begin{pmatrix} 0.79 - 0.88 & 0.47 - 0.61 & < 0.20 \\ 0.19 - 0.52 & 0.42 - 0.73 & 0.58 - 0.82 \\ 0.20 - 0.53 & 0.44 - 0.74 & 0.56 - 0.81 \end{pmatrix}. \tag{28}
\]

What are the possible forms for neutrino mass terms? The familiar Dirac mass term that we have encountered for the quarks and charged leptons connects the left-handed and right-handed components of the same field,

\[
\mathcal{L}_D = -D \bar{\psi} \psi = -D (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L), \tag{29}
\]

where we have used the chiral decomposition of a Dirac spinor,

\[
\psi = \frac{1}{2} (1 - \gamma_5) \psi + \frac{i}{2} (1 + \gamma_5) \psi \equiv \psi_L + \psi_R. \tag{30}
\]

The mass eigenstate is thus \( \psi = \psi_L + \psi_R \).

We may seek to accommodate neutrino mass in the electroweak theory by adding to the spectrum a right-handed neutrino \( N_R \) and constructing the (gauge-invariant) Dirac mass term

\[
\mathcal{L}_D^{(\nu)} = -\zeta \left( [\bar{L}_1 \phi] N_R + \bar{N}_R (\phi^1 L) \right) \rightarrow -m_D \left[ \bar{\nu}_L N_R + \bar{N}_R \nu_L \right], \tag{31}
\]

where \( L_1 \) is the left-handed lepton doublet defined in (4), \( \phi \) is the conjugate Higgs field, \( m_D = \zeta \mu / \sqrt{2} \), and as usual \( v = (G_F \sqrt{2})^{-1/2} \). The right-handed neutrino \( N_R \) is sterile: it is an \( SU(2) \) singlet with weak hypercharge \( Y = 0 \), and has no weak interactions except those induced by mixing. A Dirac mass term conserves the additive lepton number \( L \) that takes on the value +1 for neutrinos and negatively charged leptons, and −1 for antineutrinos and positively charged leptons.

Because neutrinos carry neither colour nor electric charge, they might—alone among the standard-model fermions—be their own antiparticles, so-called Majorana fermions. The charge conjugate of a right-handed field is left-handed, \( \psi_R^c \equiv (\psi^c)_L = (\psi^c)_R \). Majorana mass terms connect the left-handed and right-handed components of conjugate fields,

\[
\mathcal{L}_{MA} = A (\bar{\nu}_L \nu_R + \bar{\nu}_R \nu_L) = A \bar{\chi} \chi
\]

\[
\mathcal{L}_{MB} = B (\bar{N}_R N_L + \bar{N}_L N_R) = B \omega \omega. \tag{32}
\]

The self-conjugate Majorana mass eigenstates are

\[
\chi \equiv \nu_L + \nu_R = \chi^c
\]

\[
\omega \equiv N_R + N_L = \omega^c. \tag{33}
\]

A Majorana fermion cannot carry any additive \([U(1)]\) quantum number. The mixing of particle and antiparticle fields means that the Majorana mass terms correspond to processes that violate lepton number by two units. Accordingly, the exchange of a Majorana neutrino can mediate neutrinoless double beta decay, \((Z, A) \rightarrow (Z + 2, A) + e^+ + e^-\). Detecting neutrinoless double beta decay [99] would offer decisive evidence for the Majorana nature of the neutrino.

The mass of the active \( \nu_L \) may be generated by a Higgs triplet that acquires a vacuum expectation value [100], or by an effective operator that involves two Higgs doublets combined to transform as a triplet [101].

It is interesting to consider both Dirac and Majorana terms, and specifically to examine the case in which Majorana masses corresponding to an active state \( \chi \) and a sterile state \( \omega \) arise from weak triplets and singlets, respectively, with masses \( M_3 \) and \( M_1 \). The neutrino mass matrix then has the form

\[
\begin{pmatrix} \bar{\nu}_L & \bar{N}_L^c \end{pmatrix} \begin{pmatrix} M_3 & m_D \\ m_D & M_1 \end{pmatrix} \begin{pmatrix} \nu_R^c \\ N_R \end{pmatrix}. \tag{34}
\]
In the highly popular seesaw limit \([102–106]\), with \(M_3 = 0\) and \(m_D \ll M_1\), diagonalizing the mass matrix (34) yields two Majorana neutrinos,

\[
n_{1L} \approx \nu_L - \frac{m_D}{M_1} N_{cL}^c \quad n_{2L} \approx N_{cL}^c + \frac{m_D}{M_1} \nu_L,
\]

with masses

\[
m_1 \approx \frac{m_D^2}{M_1} \ll m_D \quad m_2 \approx M_1.
\]

The seesaw produces one very heavy ‘neutrino’ and one neutrino much lighter than a typical quark or charged lepton. Many alternative explanations of the small neutrino masses have been explored in the literature \([107]\), including some in which collider experiments exploring the Fermi scale could reveal the origin of neutrino masses \([108]\).

The discovery of neutrino mass through the observation of neutrino flavour metamorphosis offers two possible paths to physics beyond the standard model. One is to accommodate right-handed sterile neutrinos in the fermion spectrum of the electroweak theory and to endow neutrinos with mass by the Higgs mechanism. To my mind, this would be a minor adjustment to the electroweak theory—correcting an oversight, we might say. Whether the small Yukawa couplings of neutrinos to the Higgs field \((\zeta \approx 10^{-11})\) are qualitatively more puzzling than the factor of \(3 \times 10^5\) that separates the electron and top-quark couplings is for now a question for intuition. Self-conjugate Majorana neutrinos would represent a much greater revision to the standard model because the mass of the right-handed Majorana neutrino evidently is not due to the usual Higgs mechanism. This Majorana mass might be extraordinarily large—perhaps offering a window on the unification scale for the strong, weak, and electromagnetic interactions.

8. In search of the standard-model Higgs boson

The evidence in hand suggests that the agent of electroweak symmetry breaking represents a novel fundamental interaction operating on the Fermi scale. We do not know what that force is.

A leading possibility is that the agent of electroweak symmetry breaking is an elementary scalar, the Higgs boson of the electroweak standard model. An essential step towards understanding the new force that shapes our world is, therefore, to search for the Higgs boson and to explore its properties by asking the following questions.

(i) Is it there? Is there only one?
(ii) What are its quantum numbers?
(iii) Does the Higgs boson generate mass both for the electroweak gauge bosons and for the quarks and leptons?
(iv) How does the Higgs boson interact with itself?

The search for the Higgs boson has been a principal goal of particle physics for many years, so search strategies have been explored in great detail\(^\text{11}\). A brief profile of Higgs-boson properties and production mechanisms is in order here.

Consider first the most prominent decay modes of the standard-model Higgs boson. Decays \(H \to f \bar{f}\) into fermion pairs, where \(f\) occurs in \(N_c\) colours, proceed at a rate

\[
\Gamma(H \to f \bar{f}) = \frac{G_F m_f^2 M_H}{4\pi \sqrt{2}} \cdot N_c \cdot \left(1 - \frac{4m_f^2}{M_H^2}\right)^{3/2},
\]

\(^\text{11}\) For a useful sampler, see \([67, 109–116]\).
which is proportional to $N_c m_f^2 M_H$ in the limit of large Higgs mass. The partial width for decay into a $W^+ W^-$ pair is

$$
\Gamma(H \rightarrow W^+ W^-) = \frac{G_F M_H^3}{32\pi \sqrt{2}} (1 - x)^{1/2} (4 - 4x + 3x^2), \quad (38)
$$

where $x \equiv 4 M_W^2 / M_H^2$. Similarly, the partial width for decay into a pair of $Z^0$ bosons is

$$
\Gamma(H \rightarrow Z^0 Z^0) = \frac{G_F M_H^3}{64\pi \sqrt{2}} (1 - x')^{1/2} (4 - 4x' + 3x'^2), \quad (39)
$$

where $x' \equiv 4 M_Z^2 / M_H^2$. The rates for decays into weak-boson pairs are asymptotically proportional to $M_H^3$ and $\frac{1}{2} M_H^3$, respectively. In the final factors of (38) and (39), $2x^2$ and $2x'^2$, respectively, arise from decays into transversely polarized gauge bosons. The dominant decays for large $M_H$ are into pairs of longitudinally polarized weak bosons.

Branching fractions for decay modes that may hold promise for the detection of a Higgs boson are displayed in figure 6. In addition to the $f \bar{f}^0$ and $VV$ modes that arise at tree level, the plot includes the $\gamma\gamma$, $Z\gamma$, and two-gluon modes that proceed through loop diagrams. The rare $\gamma\gamma$ channel offers an important target for LHC experiments, if the Higgs boson is light, because the relatively benign backgrounds may be overcome by fine resolution.

Below the $W^+ W^-$ threshold, the standard-model Higgs boson is rather narrow, with $\Gamma(H \rightarrow \text{all}) \approx 1 \text{ GeV}$. Far above the threshold for decay into gauge-boson pairs, the total width is proportional to $M_H^6$. As its mass increases towards 1 TeV, the Higgs boson becomes highly unstable, with a perturbative width approaching its mass. It would therefore be observed as an enhanced rate, rather than a distinct resonance. The Higgs-boson total width is plotted as a function of $M_H$ in figure 7.

The most telling searches for the Higgs boson have been carried out at CERN’s Large Electron Positron collider, LEP [117, 118]. Because the $He^+e^-$ coupling is very small, being proportional to the electron mass, the cross section for resonant Higgs-boson formation, $\sigma(e^+e^- \rightarrow H \rightarrow \text{all})$, is minute. The small cross section sets aside a traditional strength of electron–positron annihilations—the ability to tune the collider energy to excite a resonant
Figure 7. Total width of the standard-model Higgs boson versus mass, from [112].

Figure 8. Higgsstrahlung mechanism for the reaction \( e^+e^- \rightarrow HZ \)

The most promising process is the radiation of a \( Z \)-boson and a Higgs boson from a virtual \( Z \), depicted in figure 8, for which the cross section is

\[
\sigma(e^+e^- \rightarrow HZ) = \frac{\pi \alpha^2}{24\sqrt{s}} \frac{K(K^2 + 3M_Z^2)[1 + (1 - 4x_W)^2]}{(s - M_Z^2)^2} x_W^2 (1 - x_W)^2 ,
\]

(40)

where \( K \) is the c.m. momentum of the outgoing Higgs boson and \( x_W \equiv \sin^2 \theta_W \). The LEP experiments set a lower bound on the mass of the standard-model Higgs boson, \( M_H > 114.4 \) GeV, at 95\% confidence level [117, 118].

The search is now the province of the proton accelerators. The 2-TeV proton–antiproton Tevatron Collider is operating now, its integrated luminosity having surpassed 2 fb\(^{-1}\) [119], and the 14 TeV Large Hadron Collider at CERN [18] will provide high-luminosity proton–proton collisions beginning in 2008\(^1\). The largest cross section for Higgs production at these machines occurs in the reaction \( p\bar{p} \rightarrow H + \text{anything} \), which proceeds by gluon fusion through

\(^1\) For a prospectus on early running at the LHC, see [120].
heavy-quark loops, as shown in the left panel of figure 9. The cross section is given by

\[ \sigma(p^\pm p \rightarrow H + \text{anything}) = \frac{G_F^2 \alpha^2}{32\pi \sqrt{2}} |\eta(\epsilon)|^2 \times (\text{gg luminosity}), \tag{41} \]

where the structure of the loop diagram is captured in

\[ \eta(\epsilon) = \frac{\epsilon}{2} [1 + (\epsilon - 1)\varphi(\epsilon)], \tag{42} \]

with \( \epsilon = 4m^2_Q/M^2_H \).

\[ \varphi(\epsilon) = \begin{cases} -\arcsin^2(1/\sqrt{\epsilon}), & \epsilon > 1 \\ \frac{1}{4} \left[ \ln \left( \frac{\zeta_+}{\zeta_-} \right) + i\pi \right]^2, & \epsilon < 1 \end{cases}, \tag{43} \]

and \( \zeta_{\pm} = 1 \pm \sqrt{1 - \epsilon} \). The behaviour of \(|\eta(\epsilon)|^2\) is plotted in the right-hand panel of figure 9. Only the top-quark loop contributes significantly to gg \( \rightarrow H \), in the remaining search range for \( M_H \).

For small Higgs-boson masses, the dominant decay is into \( b\bar{b} \) pairs, but the reaction \( p^\pm p \rightarrow H + \text{anything followed by the decay } H \rightarrow b\bar{b} \) is swamped by QCD production of \( b\bar{b} \) pairs. Consequently, experiments must rely on rare decay modes (\( \tau^+\tau^- \) or \( \gamma\gamma \), for example) with lower backgrounds or resort to different production mechanisms for which specific reaction topologies reduce backgrounds. Cross sections for the principal reactions under study are shown in figure 10. (The peak of \(|\eta(\epsilon)|^2\) near \( \epsilon = 0.68 \) is reflected in the shoulder near \( M_H = 400 \text{ GeV} \) in the gg \( \rightarrow H \) cross section at \( \sqrt{s} = 14 \text{ TeV} \).)

The production of Higgs bosons in association with electroweak gauge bosons is receiving close scrutiny at the Tevatron. Data being recorded now should give the Tevatron experiments sensitivity to raise the lower bound on the standard-model Higgs-boson mass beyond the limit set by LEP \([121, 122]\). A twofold increase in sensitivity may begin to open the possibility of finding evidence for a light Higgs boson at the three-standard-deviation level\(^\text{13} \).

\(^\text{13} \) The status of the ongoing Tevatron searches may be tracked at \([123, 124]\).
At the LHC, the multipurpose detectors ATLAS [125] and CMS [126] will make a comprehensive exploration of the Fermi scale, with high sensitivity to the standard-model Higgs boson reaching to 1 TeV. Current projections suggest that a few tens of fb$^{-1}$ will suffice for a robust discovery [127].

Once the Higgs boson is found, it will be of great interest to map its decay pattern, in order to characterize the mechanism of electroweak symmetry breaking. It is by no means guaranteed that the same agent hides electroweak symmetry and generates fermion mass. We saw in section 5 that chiral symmetry breaking in QCD could hide the electroweak symmetry without generating fermion masses. If it should turn out that the quarks and leptons are composite, with masses set largely by confinement energy, then the Higgs boson might couple to the masses of the constituents, not to the quark and lepton masses, with unpredictable consequences for branching fractions and production rates. Though out of theoretical fashion, quark and lepton compositeness remains a logical possibility that history reminds us not to neglect [128]. Indeed, many extensions to the standard model significantly alter the decay pattern of the Higgs boson. In supersymmetric models, five Higgs bosons are expected, and the branching fractions of the lightest one may be very different from those presented in figure 6 [113].

A Higgs-boson discovery in gluon fusion (gg → H) or in the Ht̄t channel would argue for a nonzero coupling of the Higgs boson to top quarks, and should in time constrain the Ht̄t coupling. With the large data sets the LHC will provide, it is plausible that Higgs-boson couplings can eventually be measured at levels that test the standard model and provide interesting constraints on extensions to the electroweak theory [129]. Precise determinations of Higgs-boson couplings is one of the strengths of the projected International Linear Collider [130].

We have seen the Higgs boson arise as an artefact of the mechanism we chose to hide the electroweak symmetry in section 2. What assurance do we have that a Higgs boson, or something very like it, will be found? It is instructive to examine the role of the Higgs boson in the cancellation of high-energy divergences. The most severe divergences of the individual $\nu$-, $\gamma$-, and $Z$-exchange diagrams for the production of longitudinally polarized $W^+W^-$ pairs in electron–positron collisions are tamed by a cooperation among the three diagrams of figures 2(a)–(c) that follows from gauge symmetry. This is not quite the end of the high-energy story. Because the electrons are massive and may therefore be found in the ‘wrong’ helicity state, we must also consider a $J = 0$ partial-wave amplitude, which grows as one power the c.m. energy. This unacceptable high-energy behaviour is precisely cancelled by...
Figure 11. $\Delta \chi^2 = \chi^2 - \chi^2_{\text{min}}$ from a fit to a universe of electroweak measurements as a function of the standard-model Higgs-boson mass. The solid line is the result of the fit. The blue band represents an estimate of the theoretical uncertainty due to missing higher-order corrections. The vertical yellow band shows the 95% CL lower bound on $M_H$ from direct searches at LEP. The dashed curve shows the sensitivity to a change in the evaluation of $\alpha_{\text{em}}(M_Z^2)$. (From the LEP Electroweak Working Group [63].)

the Higgs-boson graph of figure 2(d). Something else would have to play this role if the Higgs boson did not exist. From the point of view of $S$-matrix analysis, the $He\bar{e}$ coupling must be proportional to the electron mass, because the strength of ‘wrong-helicity’ configurations is measured by the fermion’s mass.

If the gauge symmetry were unbroken, there would be no Higgs boson, no-longitudinal gauge bosons, and no extreme divergence difficulties, but we would not have a viable low-energy phenomenology of the weak interactions. The most severe divergences of individual diagrams are eliminated by the gauge structure of the couplings among gauge bosons and leptons. A lesser, but still potentially fatal, divergence arises because the electron has acquired mass—because of the Higgs mechanism. Spontaneous symmetry breaking provides its own cure by supplying a Higgs boson to remove the last divergence. A similar interplay and compensation must exist in any satisfactory theory.

The sensitivity of electroweak observables to the (long unknown) mass of the top quark gave early indications for a very massive top, and the consonance of indirect and direct determinations is one element of the experimental support for the electroweak theory. Now that the top-quark mass is known to be about 1.4% from observations at the Tevatron, it becomes profitable to look beyond the quantum corrections involving top to the next most important effects, which arise from the Higgs boson. The Higgs-boson contributions are typically smaller than those from the top quark, and exhibit a more subtle (logarithmic) dependence on $M_H$ than the $m_t^2$ dependence characteristic of the top-quark contributions.

Figure 11 shows how the goodness of the LEP Electroweak Working Group’s Winter 2007 global fit depends upon $M_H$\textsuperscript{14}. That the fit is improved by the inclusion of Higgs-boson

\textsuperscript{14} See [131] for an introduction to global analyses.
effects does not constitute proof that the standard-model Higgs boson exists, but does show that the consistency of a standard-model analysis favours a light Higgs boson. The precision electroweak measurements on their own argue for $M_H < 144$ GeV, a one-sided 95% confidence level limit derived from $\Delta \chi^2 = 2.7$ for the blue band in figure 11. Imposing the exclusion $M_H > 114.4$ GeV from the LEP searches leads to an upper bound of $M_H \lesssim 182$ GeV [63].

The mass favoured by the global fit, $M_H = 76^{+13}_{-24}$ GeV, lies in the region excluded by direct searches. Chanowitz [133, 134] has cautioned that the values of $M_H$ preferred by fits to different observables are not entirely consistent. In particular, the forward-backward asymmetry in $e^+e^- \rightarrow b\bar{b}$ on the $Z$ resonance is best reproduced with $M_H \approx 400$ GeV. This is the observable most discrepant ($\approx 2.9\sigma$) with the overall fit; omitting it would improve the global fit, but lead to a very small Higgs-boson mass that would coexist uncomfortably with the LEP exclusion. Whether this tension is a harbinger of new physics or merely a statistical fluctuation is a tantalizing question.

By demanding consistency of the electroweak theory as a quantum field theory, we can establish bounds on the Higgs boson mass, and uncover another reason to expect that discoveries will not end with the Higgs boson. Scalar field theories make sense on all energy scales only if they are noninteracting, or ‘trivial’ [135]. The vacuum of quantum field theory is a dielectric medium that screens charge. Accordingly, the effective charge is a function of the distance or, equivalently, of the energy scale. This is the famous phenomenon of the running coupling constant.

In $\lambda\phi^4$ theory (compare the interaction term in the Higgs potential), it is easy to calculate the variation of the coupling constant $\lambda$ in perturbation theory by summing quantum corrections given by bubble graphs. The coupling constant $\lambda(\kappa)$ on a physical scale $\kappa$ is related to the coupling constant on a higher scale $\Lambda$ by

$$\frac{1}{\lambda(\kappa)} = \frac{1}{\lambda(\Lambda)} + \frac{3}{2\pi^2} \log (\Lambda/\kappa) . \quad (44)$$

In order for the Higgs potential to be stable (i.e. for the energy of the vacuum state not to race off to $-\infty$), $\lambda(\Lambda)$ must not be negative. Applied to (44), this condition leads to an inequality,

$$\frac{1}{\lambda(\kappa)} \geq \frac{3}{2\pi^2} \log (\Lambda/\kappa) , \quad (45)$$

that implies an upper bound,

$$\lambda(\kappa) \leq 2\pi^2/3 \log (\Lambda/\kappa) , \quad (46)$$

on the coupling strength at the physical scale $\kappa$. If the theory is to make sense to arbitrarily high energies—or short distances—we must consider the limit $\Lambda \rightarrow \infty$ while holding $\kappa$ fixed at some reasonable physical scale; the bound (46) then forces $\lambda(\kappa)$ to zero. The scalar field theory has become free field theory; in theorist’s jargon, it is trivial.

Rearranging and exponentiating both sides of (46) gives the condition

$$\Lambda \leq \kappa \exp \left( \frac{2\pi^2}{3\lambda(\kappa)} \right) , \quad (47)$$

from which we can infer a limit on the Higgs-boson mass. Choosing the physical scale as $\kappa = M_H$, using the definition $M_H^2 = 2\lambda(M_H)v^2$, we find that

$$\Lambda \leq \Lambda^* \equiv M_H \exp \left( \frac{4\pi^2v^2}{3M_H} \right) . \quad (48)$$

15 See [132] for an estimate of how strong nonresonant $WW$ scattering would modify radiative corrections.
16 Lattice field theory allows us to treat the strong-coupling regime beyond the small-$\lambda$ realm in which perturbation theory can be trusted.
For any given Higgs-boson mass, we can identify a maximum energy scale $\Lambda^*$ at which the theory ceases to make sense. The description of the Higgs boson as an elementary scalar is at best an effective theory, valid over a finite range of energies.

A lower bound on $M_H$ is obtained by computing quantum corrections to the classical potential (15) and requiring that $\langle \phi \rangle = v/\sqrt{2}$ be an absolute minimum of the Higgs potential. The upper and lower bounds plotted in figure 12 are the results of full two-loop calculations [142, 143]. There I have also indicated the upper bound on $M_H$ derived from precision electroweak measurements in the framework of the standard electroweak theory, as well as the lower limit from direct searches for the Higgs boson. Evidently the theory can be self-consistent up to very high energies, provided that the Higgs boson is relatively light. For the electroweak theory to make sense all the way up to a unification scale $\Lambda^* = 10^{16}$ GeV, the Higgs-boson mass must lie in the interval $134 \text{ GeV} \lesssim M_H \lesssim 177 \text{ GeV}$. If $M_H$ is not within this chimney, the electroweak theory is incomplete; it is an effective theory that will be subsumed in a more comprehensive description.

This perturbative analysis leading to (48) breaks down when the Higgs-boson mass approaches 1 TeV and the interactions become strong. Lattice analyses [144] indicate that, for the theory to describe physics to an accuracy of a few percent up to a few TeV, the mass of the Higgs boson can be no more than about $710 \pm 60$ GeV. If the elementary Higgs boson takes on the largest mass allowed by perturbative unitarity arguments, the electroweak theory lives on the brink of instability.

The condition of absolute vacuum stability that leads to the lower bound on $M_H$ displayed in figure 12 is more stringent than is required by observational evidence. It would suffice to ensure that the presumed ground state of the electroweak theory has survived quantum fluctuations until now, so that the mean time to tunnel from our vacuum to a deeper vacuum at large values of $|\phi|$ exceeds the age of the Universe, $T_U \approx 13.7 \text{ Gy}$ [10]. Figure 13
Figure 13. Metastability region of the standard-model vacuum in the \((M_H, m_t)\) plane, from [145]. The solid curves are calculated for the choice \(\alpha_s(m_Z) = 0.118\). Dashed and dot-dashed curves show the effects of varying \(\alpha_s(m_Z)\) by \(\pm 0.002\). The horizontal green band indicates the measured top-quark mass, \(m_t = 170.9 \pm 1.8\) GeV. The yellow band at the left shows the region of Higgs-boson masses excluded by searches at LEP, \(M_H \geq 114.4\) GeV. (After [145].)

Figure 14. Examples of loop diagrams that enter in the computation of quantum corrections to the Higgs-boson mass. The dashed lines represent the Higgs boson, solid lines with arrows represent fermions and antifermions and wavy lines stand for gauge bosons.

shows the outcome of a renormalization-group-improved one-loop calculation of the tunnelling probability at zero temperature [145–147]. For \(M_H = 115\) GeV, the Higgs potential develops an instability below the Planck scale for values of the top-quark mass \(m_t \approx 166\) GeV, but the electroweak vacuum’s lifetime exceeds \(T_U\) so long as \(m_t \approx 175\) GeV. Should \((M_H, m_t)\) settle in the metastable zone, we would have a provocative hint for new physics below the Planck scale.

9. More new physics on the Fermi scale?

Contemplating quantum corrections to the Higgs-boson mass leads to the inference that more new phenomena may appear at energies near 1 TeV. Suppose that, in keeping with the unitarity argument reviewed in section 3, a Higgs boson is found with mass \(M_H \approx 1\) TeV. How is the scale of electroweak symmetry breaking maintained in the presence of quantum corrections?

Beyond the classical approximation, scalar mass parameters receive quantum corrections from loops that contain particles of spins \(J = 0, \frac{1}{2}\) and 1, as indicated schematically in figure 14.

The loop integrals that determine the running mass lead potentially to divergences, which we may characterize by

\[
M_H^2(p^2) = M_H^2(\Lambda^2) + C g^2 \int_{p^2}^{\Lambda^2} dk^2 + \cdots ,
\]

where \(\Lambda\) defines a reference scale at which the value of \(M_H^2\) is known, \(g\) is the coupling constant of the theory and the coefficient \(C\) is calculable in any particular theory. The loop integrals appear to be quadratically divergent, \(\propto \Lambda^2\). In order for the mass shifts induced by quantum
In the absence of new physics, the reference scale $\Lambda$ would naturally be large. If the fundamental interactions are described by quantum chromodynamics and the electroweak theory, then a natural reference scale is the Planck mass, $\Lambda \sim M_{\text{Planck}} = (\hbar c/G_{\text{Newton}})^{1/2} \approx 1.2 \times 10^{19}$ GeV. In a unified theory of the strong, weak and electromagnetic interactions, a natural scale is the unification scale, $\Lambda \sim M_{\text{U}} \approx 10^{15}$–$10^{16}$ GeV. Both estimates are very large compared with the electroweak scale, and so imply a very long range of integration. The challenge of preserving widely separated electroweak and reference scales in the presence of quantum corrections is known as the hierarchy problem. Unless we suppose that $M_{\text{H}}^2/\Lambda^2$ and the quantum corrections are finely tuned to yield $M_{\text{H}}^2(p^2) \approx (1 \text{ TeV})^2$, some new physics must intervene at an energy of approximately 1 TeV to bring the integral in (49) under control\(^\text{17}\).

A fine-tuning problem may be seen to arise even when the scale $\Lambda$ is not extremely large. What has been called the ‘LEP Paradox’ \([150, 151]\) refers to a tension within the precise measurements of electroweak observables carried out at LEP and elsewhere. On the one hand, the global fits summarized in figure 11 point to a light standard-model Higgs boson. On the other hand, a straightforward effective-operator analysis of possible beyond-the-standard-model contributions to the same observables gives no hint of any new physics—of the kind needed to resolve the hierarchy problem—below about 5 TeV. Figure 15 shows that even with a cutoff $\Lambda = 5$ TeV, a careful balancing act is required to maintain a small Higgs-boson mass in the face of quantum corrections. The chief culprit is the large contribution from the top-quark loop. We are left to ask what enforces the balance or how we might be misreading the data.

Let us review the argument for the hierarchy problem: the unitarity argument showed that new physics must be present on the 1-TeV scale, either in the form of a Higgs boson or other new phenomena. But a low-mass Higgs boson is imperiled by quantum corrections. New physics not far above the 1-TeV scale could bring the reference scale $\Lambda$ low enough to mitigate the threat. If the reference scale is indeed very large, then either various contributions to the Higgs-boson mass must be precariously balanced or new physics must control the contribution

\[^{17}\text{The analysis is more subtle if the reference scale } \Lambda \text{ does not represent a physically meaningful cutoff. In that event, evaluating the loop integral by dimensional regulation \([148]\) yields no } \Lambda^2 \text{ piece. However, the scale of } M_{\text{H}}^2 \text{ is naturally set by the heaviest particles to which the Higgs boson couples—directly, or even through gauge bosons, so a hierarchy problem persists. For a careful treatment, see section 1 of \([149]\).}\]
of the integral in (49). We do not have a proof that Nature is not fine tuned, but I think it highly likely that both a Higgs boson and other new phenomena are to be found near the 1-TeV scale.

A new symmetry, not present in the standard model, could resolve the hierarchy problem. Exploiting the fact that fermion loops contribute with an overall minus sign relative to boson loops (because of Fermi statistics), supersymmetry \[113,149,152\] balances the contributions of fermion and boson loops\(^1\). In unbroken supersymmetry, the masses of bosons are degenerate with those of their fermion counterparts, so the cancellation is exact. If supersymmetry is present in our world, it must be broken. The contribution of the integrals may still be acceptably small if the fermion–boson mass splittings \(\Delta M\) are not too large. The condition that \(g^2 \Delta M^2\) be ‘small enough’ leads to the requirement that superpartner masses be less than about 1 TeV. It is provocative to note that, with superpartners at \(\mathcal{O}(1\text{ TeV})\), the \(SU(3)_c \otimes SU(2)_L \otimes U(1)_Y\) coupling constants run to a common value at a unification scale of about \(10^{16}\) GeV \[154\].

Theories of dynamical symmetry breaking (cf section 10) offer a second solution to the problem of the enormous range of integration in (49). In technicolour models, the Higgs boson is composite, and its internal structure comes into play on the scale of its binding, \(\Lambda_{TC} \simeq \mathcal{O}(1\text{ TeV})\). The integrand is damped, the effective range of integration is cut off, and mass shifts are under control.

Dark matter offers one more independent indication that new phenomena should be present on the Fermi scale. An appealing interpretation of the evidence that dark matter makes up roughly one-quarter of the energy density of the Universe \[155\] is that dark matter consists of thermal relics of the big bang: stable—or exceedingly long-lived—neutral particles. If the particle has couplings of weak-interaction strength, then generically the observed dark-matter density results if the mass of the dark-matter particle lies between approximately 100 GeV and 1 TeV \[156\]. Whether based on extra dimensions, new strong dynamics, or supersymmetry, scenarios to extend the electroweak theory and resolve the hierarchy problem typically entail dark-matter candidates on the Fermi scale\(^1\).

10. Another look at dynamical symmetry breaking

Quantum chromodynamics offers a suggestive model for the idea of dynamical symmetry breaking, even if it does not account itself satisfactorily for the observed electroweak symmetry breaking. The minimal technicolour model \[162,163\] transcribes the same ideas from QCD to a new setting. The technicolour gauge group is taken to be \(SU(N)_{TC}\) (usually \(SU(4)_{TC}\)), so the gauge interactions of the theory are generated by

\[
SU(4)_{TC} \otimes SU(3)_c \otimes SU(2)_L \otimes U(1)_Y.
\]

The technifermions are a chiral doublet of massless colour singlets

\[
\begin{pmatrix}
U \\
D
\end{pmatrix}_L, \quad U_R, \quad D_R.
\]

With the electric charge assignments \(Q(U) = \frac{1}{2}\) and \(Q(D) = -\frac{1}{2}\), the theory is free of electroweak anomalies. The ordinary fermions are all technicolour singlets.

In analogy with our discussion of chiral symmetry breaking in QCD, we assume that the chiral symmetry of the massless technifermions is broken,

\[
SU(2)_L \otimes SU(2)_R \otimes U(1)_Y \rightarrow SU(2)_V \otimes U(1)_Y.
\]

\(^1\) ‘Little Higgs’ models \[70\] and ‘twin Higgs’ models \[153\] employ different conspiracies of contributions to defer the hierarchy problem to about 10 TeV.

\(^{19}\) Other dark-matter candidates, notably the axions \[157,158\] implicated in a possible solution \[159,160\] to the strong CP problem, do not select the 1-TeV scale. For recent reviews, see \[10,161\].
Three would-be Goldstone bosons, the technipions $\pi^+ T, \pi^0 T, \pi^- T$, emerge, for which we are free to choose the technipion decay constant as

$$F_\pi = \left( G_F \sqrt{2} \right)^{-1/2} = 246 \text{ GeV},$$

which amounts to choosing the scale on which technicolour becomes strong. When the electroweak interactions are turned on, the technipions become the longitudinal components of the intermediate bosons, which acquire canonical standard-model masses

$$M_{W \pi}^2 = \frac{g^2 F_\pi^2}{4} = \frac{\pi^2 \alpha}{G_F \sqrt{2} \sin^2 \theta_W},$$

$$M_{Z \pi}^2 = \frac{(g^2 + g'^2) F_\pi^2}{4} = M_{W \pi}^2 \cos^2 \theta_W,$$

thanks to our choice (53) of the technipion decay constant.

Technicolour shows how the generation of intermediate boson masses could arise without fundamental scalars or unnatural adjustments of parameters. It thus provides an elegant solution to the naturalness problem of the standard model. However, it offers no explanation for the origin of quark and lepton masses, because no Yukawa couplings are generated between Higgs fields and quarks or leptons. Consequently, technicolour serves as a reminder that particle physics confronts two problems of mass: explaining the masses of the gauge bosons, which demands an understanding of electroweak symmetry breaking; and accounting for the quark and lepton masses, which requires not only an understanding of electroweak symmetry breaking but also a theory of the Yukawa couplings that set the scale of fermion masses in the standard model.

We can be confident that the origin of gauge-boson masses will be understood on the Fermi scale. We do not know where we will decode the pattern of the Yukawa couplings. In extended technicolour models [165–167], separate gauge interactions hide the electroweak symmetry and communicate the broken symmetry to the quarks and leptons. Specific implementations of these ideas face phenomenological challenges pertaining to flavour-changing neutral currents, the large top-quark mass, and precision electroweak measurements, but the idea of dynamical symmetry breaking remains an important alternative to the standard elementary scalar [168, 169].

11. The vacuum energy puzzle

The cosmological constant problem—why empty space is so nearly massless—is one of the great mysteries of science [170, 171]. It is the reason why gravity has weighed on the minds of electroweak theorists [91, 172, 173], despite the utterly negligible role that gravity plays in particle reactions.

At the vacuum expectation value $\langle \phi \rangle$ of the Higgs field, the (position-independent) value of the Higgs potential is

$$V(\langle \phi \rangle) = \frac{\mu^2 v^2}{4} = -\frac{|\lambda| v^4}{4} < 0.$$ (55)

Identifying $M_H^2 = -2\mu^2$, we see that the Higgs potential contributes a field-independent constant term,

$$\varrho_H = \frac{M_H^2 v^2}{8}.$$ (56)

An alternative variation on the QCD theme posits that QCD-induced chiral symmetry breaking in exotic (colour 6, 8, 10, ...) quark sectors may drive electroweak symmetry breaking [164].
which plays the role of a vacuum energy density $\varrho_H$ in the Lagrangian. In the context of gravitation, this is equivalent to adding a cosmological constant term, $\Lambda = (8\pi G_{\text{Newton}}/c^4)\varrho_H$, to Einstein’s equation \[16\].

Recent observations of the accelerating expansion of the Universe [174, 175] raise the intriguing possibility that the cosmological constant may be different from zero, but the essential fact is that the observed vacuum energy density must be very small indeed [10],

$$\varrho_{\text{vac}} \approx 10^{-46} \text{GeV}^4 \approx (\text{a few meV})^4.$$ (57)

Therein lies the puzzle: if we take $v = (G_F\sqrt{2})^{-\frac{1}{2}} \approx 246 \text{GeV}$ and insert the current experimental lower bound [117] $M_H \approx 114.4 \text{GeV}$ into (56), we find that the Higgs field’s contribution to the vacuum energy density is

$$\varrho_H \approx 10^8 \text{GeV}^4,$$ (58)

some 54 orders of magnitude larger than the upper bound inferred from the cosmological constant. This mismatch has been a source of dull headaches for more than three decades.

The problem is still more serious in a unified theory of the strong, weak and electromagnetic interactions, in which other (heavy!) Higgs fields have nonzero vacuum expectation values that may give rise to still larger vacuum energies. At a fundamental level, we can therefore conclude that a spontaneously broken gauge theory of the strong, weak and electromagnetic interactions—or merely of the electroweak interactions—cannot be complete. The vacuum energy problem must be an important clue. But to what?

The tentative evidence for a nonzero cosmological constant recasts the problem in two important ways. First, instead of looking for a principle that would forbid a cosmological constant, perhaps a symmetry principle that would set it exactly to zero, we may be called upon to explain a tiny cosmological constant. Second, if the interpretation of the accelerating expansion in terms of dark energy is correct, we now have observational access to some new stuff whose equation of state and other properties we can try to measure. Maybe that will give us the clues that we need to solve this old problem and to understand how it relates to the electroweak theory.

12. Big questions

Opening the Fermi scale to exploration means entering a new world. The quest for the origins of electroweak symmetry breaking—testing the Higgs mechanism—heads what promises to be a very rich experimental agenda. What we learn from the new round of experimentation, and the interplay with theory, will transfigure particle physics and deepen our understanding of the everyday world. Moreover, we have the strong suspicion that many of the outstanding problems of particle physics and cosmology may be linked—and linked to the Fermi scale.

Here are some of the questions that will shape the explorations to come.

(i) What is the agent that hides the electroweak symmetry?
(ii) Is there a Higgs boson? Might there be several?
(iii) Does the Higgs boson give mass to fermions, or only to the weak bosons? What sets the masses and mixings of the quarks and leptons?
(iv) How does the Higgs boson interact with itself? What shapes the Higgs potential?
(v) Could we be living in a false (metastable) vacuum?
(vi) Is the Higgs boson elementary or composite?
(vii) Does the pattern of Higgs-boson decays imply new physics? Will unexpected or rare decays of the Higgs boson reveal new kinds of matter?
(viii) What stabilizes the Higgs-boson mass on the Fermi scale? Is Nature supersymmetric? Is electroweak symmetry breaking an emergent phenomenon connected with strong dynamics? Is electroweak symmetry breaking related to gravity through extra spacetime dimensions?

(ix) How can a light Higgs boson coexist with the absence of signals for new phenomena?

(x) What resolves the vacuum energy problem?

(xi) What lessons does electroweak symmetry breaking hold for unified theories of the strong, weak, and electromagnetic interactions? For the inflationary Universe? For dark energy?

It is an inspiring list. Within a decade, we can expect to have many answers—and even better questions!

Acknowledgments

Fermilab is operated by Fermi Research Alliance, LLC under Contract No DE-AC02-07CH11359 with the United States Department of Energy. It is a pleasure to thank Luis Álvarez-Gaumé and other members of the CERN Theory Group for warm hospitality in Geneva. I thank Abdelhak Djouadi and Olga Mena for providing figures, and Ian Aitchison, Michael Chanowitz, Mu-Chun Chen, JoAnne Hewett, J D Jackson, Andreas Kronfeld, Joe Lykken, Tom Rizzo, and Ruth Van de Water for helpful advice.

References

[1] Newton I 1999 *The Principia: Mathematical Principles of Natural Philosophy* (Berkeley, CA: University of California Press) Transl. I B Cohen and A Whitman
[2] Chandrasekhar S 1995 *Newton's Principia for the Common Reader* (Oxford: Oxford University Press)
[3] Lorentz H A 2003 *The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat* (Mineola: Dover) (reprint of the 1909 edition)
[4] Jackson J D 1998 *Classical Electrodynamics* 3rd edn (New York: Wiley) sections 16.3–6
[5] Einstein A 1905 *Ann. Phys.* 18 639 (available in English translation at URL www.fourmilab.ch/etexts/etexts/einstein/E=mc2/www/)
[6] Rigden J S 2005 *Einstein 1905: the Standard of Greatness* (Cambridge, MA: Harvard University Press)
[7] Okun L B 2006 *The concept of mass in the Einstein year* Preprint hep-ph/0602037
[8] Einstein A et al 2005 *Albert Einstein: Historic Recordings 1930-1947* (London: The British Library Board) Audio CD NSACD 23 produced by the British Library Sound Archive on the occasion of the World Year of Physics; track 1, ‘E = mc²’
[9] Wilczek F 1999 *Phys. Today* 52 11–13
[10] Yao W M et al (Particle Data Group) 2006 *J. Phys.* G 33 1–1232 http://pdg.lbl.gov
[11] Aoki S et al (CP-PACS Collaboration) 2003 *Phys. Rev.* D 67 034503 (Preprint hep-lat/0206009)
[12] Namekawa Y et al (CP-PACS Collaboration) 2004 *Phys. Rev.* D 70 074503 (Preprint hep-lat/0404014)
[13] Davies C 2005 *Lattice QCD—A guide for people who want results* Preprint hep-lat/0509046
[14] Guth A H 1997 *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins* (New York: Perseus)
[15] Peacock J A 1999 *Cosmological Physics* (Cambridge: Cambridge University Press) chapter 11
[16] Peebles P J E and Ratra B 2003 *Rev. Mod. Phys.* 75 559–606 (Preprint astro-ph/0207347)
[17] Feng J L 2005 ILC cosmology Proceedings of the 2005 Linear Collider Workshop ed Hewett J L eConf C050318 0013 (Preprint hep-ph/0509309)
[18] The Large Hadron Collider Project http://lc.web.cern.ch/lhc
[19] Dawson S and Oreglia M 2004 *Annu. Rev. Nucl. Part. Sci.* 54 269–314 (Preprint hep-ph/0403015)
[20] Gaillard M K, Grannis P D and Scullia F J 1999 *Rev. Mod. Phys.* 71 S96–S111 (Preprint hep-ph/9812285)
[21] Quigg C 2006 *Particles and the Standard Model The New Physics: for the 21st Century* ed G Fraser (Cambridge: Cambridge University Press) pp 86–118 chapter 4
[22] Aitchison I J R and Hey A J G 2003 *Gauge Theories in Particle Physics* vol 2 3rd edn (London: Taylor & Francis)
[65] The CDF Collaboration 2007 First Measurement of the W Boson Mass with CDF in Run II CDF Note 8665 (available at URL http://www.fnal.gov/physics/ewk/2007/wmass/wmass_conf.ps)
[66] Tevatron Electroweak Working Group 2007 Combination of CDF and DØ results on the mass of the top quark Preprint hep-ex/0703034
[67] Lee B W, Quigg C and Thacker H B 1977 Phys. Rev. D 16 1519–31
[68] Chanowitz M S 2005 Czech. J. Phys. 55 B45–58 (Preprint hep-ph/0412203)
[69] Csaki C, Hubisz J and Meade P 2006 Electroweak symmetry breaking from extra dimensions Physics in D ⩾ 4: TASI 2004 ed J Terning et al (Singapore: World Scientific) pp 703–76 (Preprint hep-ph/0510275)
[70] Schmaltz M and Tucker-Smith D 2005 Annu. Rev. Nucl. Part. Sci. 55 229–70 (Preprint hep-ph/0502182)
[71] Veltman M J G 1977 Acta Phys. Pol. B 8 475–92
[72] Dicus D A and Mathur V S 1973 Phys. Rev. D 7 3111–14
[73] Bardeen J, Cooper L N and Schrieffer J R 1957 Phys. Rev. 106 162–4
[74] Nambu Y 1960 Phys. Rev. Lett. 4 380–2
[75] Weinberg S 1973 Phys. Rev. D 8 2511–24
[76] Agrawal V, Barr S M, Donoghue J F and Seckel D 1998 Phys. Rev. Lett. 80 1822–5
[77] Chen M C, de Gouvêa A and Dobrescu B A 2006 Gauge trimming of neutrino masses Phys. Lett. B 647 1–6
[78] Horgan C J 2000 Rep. Prog. Phys. 63 1729–39
[79] Yoong J J and Scherrer R J 2003 Phys. Rev. D 68 052003 (Preprint hep-ph/0302155)
[80] Chen M C and Mahanthappa K T 2003 Phys. Scr. T 104 89–93 (Preprint hep-ph/0305088)
[81] Raby S 2004 Rep. Prog. Phys. 67 755–811 (Preprint hep-ph/0401155)
[82] Mohapatra R N and Smirnov A Y 2006 Annu. Rev. Nucl. Part. Sci. 56 569–628 (Preprint hep-ph/0603118)
[83] Arkani-Hamed N and Schmaltz M 2000 Phys. Rev. D 61 033005 (Preprint hep-ph/9903417)
[84] Mirabelli E A and Schmaltz M 2000 Phys. Rev. D 61 113011 (Preprint hep-ph/9912265)
[85] Wolfenstein L 2003 Preprint hep-ph/0202129
[86] Schwetz T 2006 Observables sensitive to absolute neutrino masses: a reappraisal after WMAP-3-year and first MINOS results Preprint hep-ph/0602215
[87] Lee B W and Shrock R E 1977 Phys. Lett. B 54 1–5 (Preprint hep-ph/0606060)
[88] Fogli G L et al 2006 Observables sensitive to absolute neutrino masses: a reappraisal after WMAP-3-year and first MINOS results Preprint hep-ph/0608060
[89] Gonzalez-Garcia M C 2005 Phys. Scr. T 115 1–5 (Preprint hep-ph/0503170)
[90] Eliott S R and Vogel P 2002 Annu. Rev. Nucl. Part. Sci. 52 115–51 (Preprint hep-ph/0202264)
[91] Gelmini G B and Roncadelli M 1981 Phys. Lett. B 99 411–15
[92] Weinberg S 1979 Phys. Rev. Lett. 43 1566–70
[93] Minkowski P 1977 Phys. Lett. B 67 421–28
[94] Yanagida T 1979 Horizontal gauge symmetry and masses of neutrinos Proc. Workshop on the Baryon Number of the Universe and Unified Theories ed O Sawada and A Sugamoto (Tsukuba: KEK) pp 95–98 (reprinted in SEESAW 25: Proc. Int. Conf. on the Seesaw Mechanism ed J Orloff et al (Singapore: World Scientific) pp 261–4)
[95] Gell-Mann M, Ramond P and Slansky R 1979 Complex spinors and unified theories Supergravity ed P van Nieuwenhuizen and D Z Freedman (Amsterdam: North-Holland) pp 95–101
[96] Mohapatra R N and Senjanovic G 1980 Phys. Rev. Lett. 44 912–15
[97] Schechter J and Valle J W F 1980 Phys. Rev. D 22 1227–35
[98] Smirnov A Y 2005 Alternatives to the seesaw mechanism SEESAW 25: Proc. Int. Conf. on the Seesaw Mechanism ed J Orloff et al (Singapore: World Scientific) pp 221–36 (Preprint hep-ph/0411194)
[99] Chen M C, de Gouvêa A and Dobrescu B A 2006 Gauge trimming of neutrino masses Preprint hep-ph/0612017
Origins of mass

[147] Einhorn M B and Jones D R T 2007 The effective potential, the renormalisation group and vacuum stability Preprint hep-ph/0702295
[148] ’t Hooft G and Veltman M J G 1972 Nucl. Phys. B 44 189–213
[149] Martin S P 2006 A Supersymmetry Primer Preprint hep-ph/0709356v4
[150] Barbiere R and Strumia A 2000 The LEP paradox Preprint hep-ph/0007265
[151] Burdman G, Chacko Z, Goh H S and Harnik R 2006 Folded supersymmetry and the LEP paradox Preprint hep-ph/0609152
[152] Pape L and Treille D 2006 Rep. Prog. Phys. 69 2843–3067
[153] Chacko Z, Goh H S and Harnik R 2006 Phys. Rev. Lett. 96 231802 (Preprint hep-ph/0506256)
[154] Amaldi U, de Boer W and Fürstenau H 1991 Phys. Lett. B 260 447–55
[155] Spergel D N et al Wilkinson Microwave Anisotropy Probe (WMAP) three year results: implications for cosmology Preprint astro-ph/0603449
[156] Bertone G, Hooper D and Silk J 2005 Phys. Rep. 405 279–390 (Preprint hep-ph/0404175)
[157] Weinberg S 1978 Phys. Rev. Lett. 40 223–6
[158] Wilczek F 1978 Phys. Rev. Lett. 40 279–82
[159] Pecci R D and Quinn H R 1977 Phys. Rev. Lett. 38 1440–3
[160] Pecci R D and Quinn H R 1977 Phys. Rev. D 16 1791–7
[161] Raffelt G G 2006 Axions: Motivation, limits and searches Preprint hep-ph/0611118
[162] Weinberg S 1976 Phys. Rev. D 13 974–96
Weinberg S 1979 Phys. Rev. D 19 1277 (addendum)
[163] Susskind L 1979 Phys. Rev. D 20 2619–25
[164] Marciano W J 1980 Phys. Rev. D 21 2425–28
[165] Dimopoulos S and Susskind L 1979 Nucl. Phys. B 155 237–52
[166] Eichten E and Lane K D 1980 Phys. Lett. B 90 125–30
[167] Appelquist T, Piai M and Shrock R 2004 Phys. Rev. D 69 015002 (Preprint hep-ph/0308061)
[168] Hill C T and Simmons E H 2003 Phys. Rep. 381 235–402
Hill C T and Simmons E H 2003 Phys. Rep. 390 553–4 (erratum) (Preprint hep-ph/0203079)
[169] Lane K 2002 Two lectures on technicolor Preprint hep-ph/0202255
[170] Weinberg S 1989 Rev. Mod. Phys. 61 1–23
[171] Abbott L 1988 Sci. Am. 258 N5 106–13
[172] Linde A D 1974 Pisma Zh. Eksp. Teor. Fiz. 19 320–22
Linde A D 1974 JETP Lett. 19 183 (Engl. Transl.)
[173] Veltman M J G 1975 Phys. Rev. Lett. 34 777
[174] Riess A G et al (Supernova Search Team) 1998 Astron. J. 116 1009–38 (Preprint astro-ph/9805201)
[175] Perlmutter S et al (Supernova Cosmology Project) 1999 Astrophys. J. 517 565–86 (Preprint astro-ph/9812133)