In search of symmetry lost

Frank Wilczek

The Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA (e-mail: wilczek@mit.edu)

Powerful symmetry principles have guided physicists in their quest for nature's fundamental laws. The successful gauge theory of electroweak interactions postulates a more extensive symmetry for its equations than are manifest in the world. The discrepancy is ascribed to a pervasive symmetry-breaking field, which fills all space uniformly, rendering the Universe a sort of exotic superconductor. So far, the evidence for these bold ideas is indirect. But soon the theory will undergo a critical test depending on whether the quanta of this symmetry-breaking field, the so-called Higgs particles, are produced at the Large Hadron Collider (due to begin operation in 2007).

It has been almost four decades since our current, wonderfully successful theory of the electroweak interaction was formulated[1–4]. Central to that theory is the concept of spontaneously broken gauge symmetry[5,6]. According to this concept, the fundamental equations of physics have more symmetry than the actual physical world does. Although its specific use in electroweak theory involves exotic hypothetical substances and some sophisticated mathematics, the underlying theme of broken symmetry is quite old. It goes back at least to the dawn of modern physics, when Newton postulated that the basic laws of mechanics exhibit full symmetry in three dimensions of space — despite the fact that everyday experience clearly distinguishes 'up and down' from 'sideways' directions in our local environment. Newton, of course, traced this asymmetry to the influence of Earth's gravity. In the framework of electroweak theory, modern physicists similarly postulate that the physical world is described by a solution wherein all space, throughout the currently observed Universe, is permeated by one or more (quantum) fields that spoil the full symmetry of the primary equations. Thus, modern physicists hypothesize that what we perceive as empty space is actually a highly structured medium. In fact, as I will elaborate below, we vehemently suspect that the world is a multilayered, multicoloured, cosmic superconductor.

Fortunately this hypothesis, which might at first hearing sound quite extravagant, has testable implications. The symmetry-breaking fields, when suitably excited, must bring forth characteristic particles: their quanta. Using the most economical implementation of the required symmetry breaking, one predicts the existence of a remarkable new particle, the so-called Higgs particle. More ambitious speculations suggest that there should be not just a single Higgs particle, but rather a complex of related particles. The very popular and attractive idea of low-energy supersymmetry[7,8], to be discussed further below, requires at least five 'Higgs particles'.

The primary goal of fundamental physics is to discover profound concepts that illuminate our understanding of nature. Discovering new particles, as such, is secondary. In recent times, however, physicists have often found that their most profound concepts, when implemented with rigorous logic, are reflected in the existence of new particles. This happens because both quantum mechanics and special relativity are important in the regime of short distances and high energies, where high-energy physics explores fundamental laws. It is difficult to combine quantum mechanics and special relativity in a consistent way. The only way we know how to do it is by using quantum field theory, and the basic objects of quantum field theory are space-filling entities (quantum fields) whose excitations are what we perceive, concretely, as particles[9,10]. So, when our concepts are made consistent with quantum mechanics and relativity, they tend to be reflected rather directly in predictions about particles.

The W and Z bosons, carriers of the weak nuclear force, and gluons, carriers of the strong nuclear force, are outstanding examples of ideas embodied as particles. These so-called gauge particles are physical embodiments of the symmetry of physical law (gauge invariance)[11–13] — not merely metaphorically but in a very precise sense. Indeed, as a fact of history, the existence of these particles and their detailed behaviour was predicted before their experimental observation, starting from the concept of gauge symmetry. Harmony between mind and matter, in the form of mathematical abstractions conjuring up sensuous reality, has long figured in the dreams of mystics and the inspiration of visionaries — the 'music of the spheres'. Here it is realized in a form that is genuine, reproducible and precise.

Now we are faced with the opportunity for another synthesis. Ironically, the concept whose embodiment we now seek is a special, structured sort of symmetry breaking. This concept is a necessary complement to what has come before; for our symmetry-based understanding of the W and Z bosons — that is, of the electroweak interaction — relies on postulating symmetries that are broken in a very specific way. They are supposed to be spoiled by a form of cosmic superconductivity, with newly hypothesized fields having the role performed by electrons in ordinary superconductors. It is these new quantum fields that are the progenitors of Higgs particles.

So far, no Higgs particle has been observed. As yet, this failure does not represent a crisis. Detection of Higgs particles that are sufficiently heavy — specifically, those whose mass exceeds 114 GeV, which is the current lower bound[14] — will have to await more powerful accelerators than are now available. But theory tells us that this evasion cannot be maintained indefinitely. If the Higgs particle, or an appropriate complex of Higgs particles, does not turn up at the Large Hadron Collider (LHC), a major revision of our thinking will be required. The LHC is now under construction at CERN (European Centre for Particle Physics) near Geneva. It is due to begin operation in late 2007.

There are already indirect but significant indications that at least one Higgs particle with a mass below 250 GeV does exist[15]. If there is such a particle, it will certainly be observed at the LHC. That observation, if and when it occurs, will bring a glorious chapter in physics to a glorious conclusion. It will also provide a key to unlock new volumes that are currently sealed; for the circle of ideas around symmetry breaking and the Higgs particle includes, quite close to its elegant central core, some of the darkest and most forbidding zones of ignorance in the existing landscape of

fundamental physics. That is exciting because it means we will have an opportunity to learn. Big ideas and speculations about the unification of forces, and the cosmology of the early Universe, as well as supersymmetry, are very much in play. Thus the Higgs sector is not only a destination, but also a portal.

There is a vast technical literature on most of the topics discussed here[16,17] and quite a few popular and semipopular presentations. My goal here is to present a brief but substantial and critical review of the main concepts and prospects that is accessible to scientifically sophisticated non-experts, yet reflects the essence of present-day thinking.

The three standard systems

It has become conventional to say that our knowledge of fundamental physical law is summarized in a 'standard model'. But this convention lumps together three quite different conceptual structures. Instead, it is more accurate and informative to say that our current, working description of fundamental physics is made from three distinct parts: three 'standard systems'. These are the gauge system, the gravity system and the Higgs system.

Each of these systems concerns the interactions of a specific kind of particle: gauge bosons, gravitons and Higgs particles, respectively. It is remarkable that everything we know, or reliably infer, about the fundamental laws of nature can be interpreted as a statement about how one or another of these particles interacts with other forms of matter. To be precise, every known departure from trivial 'free' propagation — every nonlinear coupling of the quantum fields that describe matter in all its forms — involves interaction with a gauge particle, a graviton or a Higgs particle. Two of these three systems, the gauge and gravity systems, are governed by principled theories founded on deep, powerful concepts. Because of this they are tight, both conceptually and algorithmically.

The gauge system is constructed as the embodiment of extensive symmetries involving transformations among different kinds of 'colour' degrees of freedom. Colour used in this sense has nothing to do with optical phenomena; rather, it is a generalization of the concept of electromagnetic charge. Quantum chromodynamics (QCD) — our theory of the strong interaction[18,19] — works with three colour charges. The weak interaction invokes two other colour charges and ordinary electromagnetism introduces yet another charge. The precise symmetry is expressed in the language of group theory, as $SU(3) \times SU(2) \times U(1)$.

Gauge symmetries, when combined with the principles of quantum mechanics and special relativity, are extremely powerful. Gauge symmetries require the existence of appropriate gauge bosons, and vice versa. Through this connection between mathematics and physics — concept and reality — we arrive at a beautiful and tightly integrated theory of gauge bosons and their interactions with other forms of matter. A profound reflection of this is that the physics of the gauge system is almost fully determined from considerations of inner consistency. In other words, it contains few freely adjustable parameters. There are basically just three such parameters, one each for $SU(3)$, $SU(2)$ and $U(1)$. Given these three numbers, no further 'fudge factors' remain available. The gauge system provides precise predictions for many phenomena; predictions that are in excellent agreement with numerous accurate experiments. In fact no significant deviation has been found so far. Each term in the foundational equations underlying a complete description of strong, weak and electromagnetic interactions has been checked out repeatedly in precise experiments.

The gravity system is essentially Einstein's general theory of relativity. It is sometimes claimed that general relativity is impossible to reconcile with quantum mechanics, that there is a terrible crisis in quantum gravity, and so on. On hearing this, one might be puzzled about how physicists and astrophysicists manage to get on with their work. The truth is that there is a very concrete and precise working theory of gravity that fully conforms to the principles of quantum mechanics and which so far has proved adequate to describe all physical and astrophysical observations. It is simply the quantum field theory version of general relativity (for experts: the Einstein–Hilbert action for gravity itself, extended to matter using the minimal coupling procedure). This theory fails to give predictions for processes involving ultra-high-energy particles, but the energies at which it fails are much larger than those that are accessible to observation.

Like the gauge system, the gravity system is constructed as the embodiment of a powerful symmetry principle, in this case Einstein's general covariance. General covariance both requires the existence of the graviton, and tightly constrains its properties. It thereby generates a unique, principled theory of gravity. General relativity is fully specified in terms of just two freely adjustable parameters: one is Newton's gravitational constant G_N; the other is the so-called cosmological term, which parameterizes the energy density of empty space. Until very recently there were only upper bounds on the value of the cosmological term, but observations of the acceleration of distant galaxies, together with indirect inference from measured cosmic microwave background anisotropies, seem to require a positive value[20–22]. All other gravitational phenomena are predicted using only G_N. General relativity has scored many triumphs, both qualitative ones, such as providing foundations for Big Bang cosmology and black hole physics, and quantitative ones, such as accurately predicting the precession of Mercury and the time variation of binary pulsar frequencies.

The gauge system and the gravity system can both be written in appealing geometric forms. More precisely, quantum fields that describe different forms of curvature produce both gravitons and gauge bosons. For gravitons, the relevant curvature is that of space-time; for gauge bosons, it is curvature in so-called internal spaces, which are defined by the variables that describe configurations of colour charges. The coupling of these particles to other fields is also defined geometrically, technically, by the promotion of ordinary derivatives into covariant ones.

The third system, where the Higgs particle and its couplings to other forms of matter reside, is another story entirely. We know of no deep principle, comparable to gauge symmetry or general covariance, which constrains the values of these couplings at all tightly. In the Higgs system, the number of freely adjustable parameters mushrooms into dozens. Nor do the values of the couplings we now infer from the masses and mixing of quarks and leptons conform to any easily discernible pattern (see the section '...and a nest of kluges' below).

Clearly the Higgs system of fundamental physics is its least satisfactory part. Whether measured in terms of the large number of independent parameters it requires, or in terms of the small number of powerful ideas it contains, our theory of this sector falls far short of the high level we have achieved elsewhere. In particular, despite the phrase's connotation, no 'theory of everything' hitherto proposed has in practice materially improved our theory of the Higgs system. Having placed the Higgs system in context, let us now scrutinize it more closely.

Symmetry breaking in superconductivity

The most fundamental phenomenon of superconductivity is the Meissner effect, according to which magnetic fields are expelled from the bulk of a superconductor. The Meissner effect implies the possibility of persistent currents[23]. Indeed, if a superconducting sample is subjected to an external magnetic field, currents of this sort must arise near the surface of a sample to generate a cancelling field.

An unusual but valid way of speaking about the phenomenon of superconductivity is to say that within a superconductor the photon acquires a mass. The Meissner effect follows from this. Indeed, to say that the photon acquires a mass is to say that the electromagnetic field becomes a massive field. Because the energetic cost of supporting massive fields over an extended volume is prohibitive, a superconducting material finds ways to expel magnetic fields.

Bardeen, Cooper and Schrieffer (BCS) developed a satisfactory microscopic theory of superconductivity in metals[24]. BCS theory traces superconductivity to the existence of a special sort of long-range correlation among electrons. This effect is purely quantum-
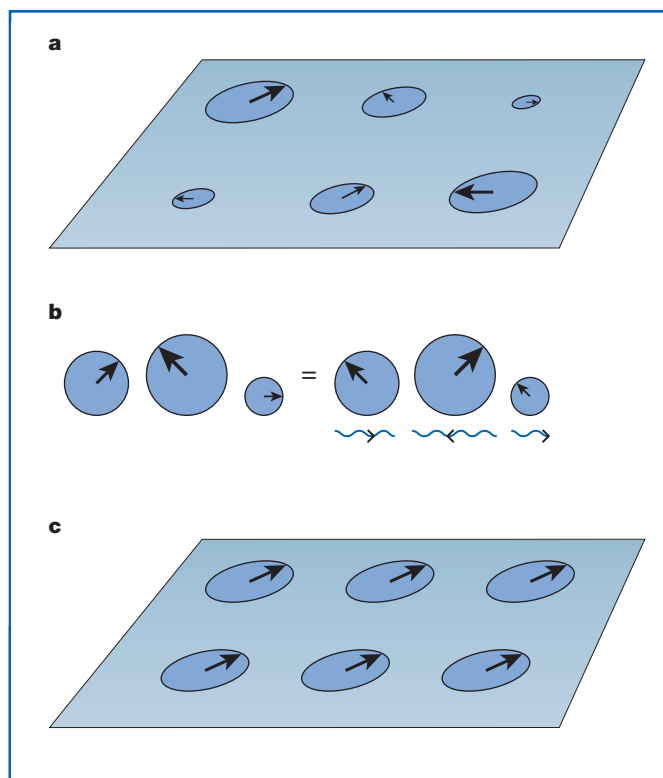
Figure 1 Visual metaphors for gauge symmetry and superconductivity. **a**, The quantum-mechanical wavefunction of an electrically charged field is a position-dependent complex number, which can be depicted as a collection of arrows with different sizes and orientations. **b**, Gauge symmetry states that the same physical situation can be described using different orientations of the arrows (phases of the wavefunction) with a compensating gauge field (electromagnetic potential, represented by wavy arrows). This is somewhat analogous to Einstein's equivalence principle, by which relatively accelerated reference frames become equivalent when a compensating gravity field is included. **c**, In the superconducting state, the charged-particle wavefunction is correlated over long distances and times — whereas in the non-superconducting state (**a** could be regarded as a typical, disordered slice of this state), the wavefunctions fluctuate randomly in space and time. The ordered structure of **c** is disrupted by gauge transformations, so now the electromagnetic potentials have an energetic cost — the photon has acquired a mass.

A fundamental principle of electrodynamics is its gauge symmetry. A gauge transformation is a mathematical transformation of electromagnetic potentials and the wavefunctions of charged particles. When we say electrodynamics obeys gauge symmetry, we mean that while the gauge transformation changes the variables that appear in the equations of electrodynamics, and consequently rearranges those equations, it nevertheless leaves the overall physical content of the equations unchanged (Fig. 1b). Weyl and London discovered the gauge symmetry of quantum electrodynamics in the 1920s (refs 11, 12). Yang and Mills proposed a more general concept of gauge symmetry[13] that supports a much wider class of transformations (non-abelian gauge theory). At first this generalization seemed to be a mathematical curiosity, but as physics developed it has come to be more and more central. We have come to recognize both that gauge symmetry is necessary for the consistency of the theory, and conversely, that the equations of electrodynamics can be derived from gauge symmetry, assuming the validity of quantum mechanics and special relativity. One aspect of this is that gauge symmetry enforces zero mass for the photon.

For present purposes, what is crucial is that gauge transformations rotate the wavefunction in the extra dimension, through an angle that can vary depending on location in space-time. Superconductivity occurs when the phases of the Cooper pairs all align in the same direction; that is, when they all have the same position within their extradimensional circles (Fig. 1c). Of course, gauge transformations that act differently at different space-time points will spoil this alignment. Thus, although the basic equations of electrodynamics are unchanged by gauge transformations, the state of a superconductor does change. To describe this situation, we say that in a superconductor gauge symmetry is spontaneously broken.

The phase alignment of the Cooper pairs gives them a form of rigidity. Electromagnetic fields, which would tend to disturb this alignment, are rejected. This is the microscopic explanation of the Meissner effect, or in other words, the mass of photons in superconductors.

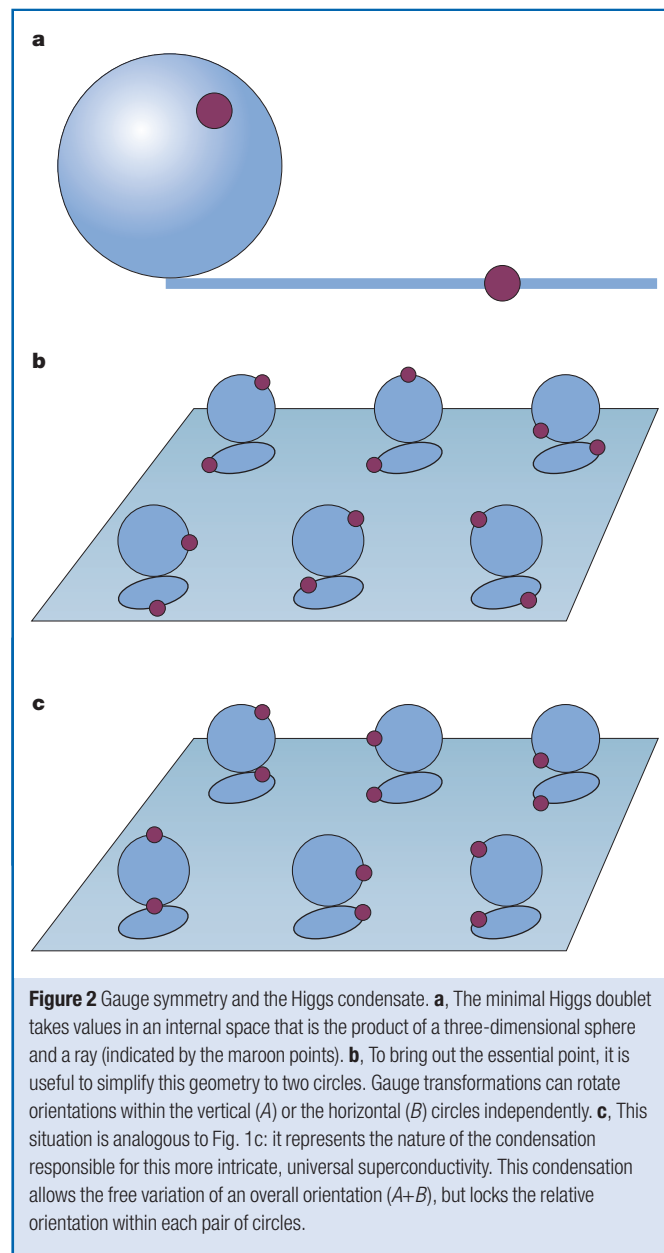## Electroweak symmetry breaking: some elegant details…
Several basic properties of the $W^{\pm}$ and Z particles, which are responsible for mediating the weak interaction, are quite similar to properties of photons. They are all spin-1 particles that couple with universal strength to appropriate charges and currents. This resemblance, together with consistency requirements of the sort mentioned above, suggests that the equations of the weak interactions must have an appropriate gauge symmetry, distinct from, but of the same general nature as, the gauge symmetry of electromagnetism.

Elaboration of this line of thought leads to our modern theory of the electroweak interactions, which is firmly based on the postulate of gauge symmetry. The gauge symmetry involved is more elaborate than that of electromagnetism. It involves a more intricate internal space (Fig. 2a). Instead of just a circle (orientation) and a ray (magnitude), we have a three-dimensional sphere (orientation) and a ray (magnitude) over each point of space-time, so four 'internal dimensions'. Gauge transformations rotate the spheres, although not all types of rotation are allowed. Mathematically, the allowed transformations define the group $SU(2) \times U(1)_Y$.

There is an essential supplement to this generalization, which provides our primary motivation for postulating the existence of the Higgs system. Unlike the photon, which acquires mass only inside superconductors, $W^{\pm}$ and Z are massive particles even in empty space. The equations of gauge symmetry, in their pristine form, predict the existence of these particles, but require that their masses vanish. To account for the masses of $W^{\pm}$ and Z, we must suppose that what we perceive as empty space is, in reality, a new form of superconductor, not for electromagnetism, but for its near-relation gauge interactions.

With this interpretation, what we perceive as empty space is not so empty. But what is it that has the role — for this new universal superconductivity — of the Cooper pairs whose alignment is responsible for conventional superconductivity? No form of matter identified so

mechanical. A classical phenomenon that is only very roughly analogous, but much simpler to visualize, is the occurrence of ferromagnetism owing to long-range correlations among electron spins (that is, their mutual alignment in a single direction). The sort of correlations responsible for superconductivity are of a much less familiar sort, as they involve not the spins of the electrons, but rather the phases of their quantum-mechanical wavefunctions. One cannot do full justice to this concept without some elaborate mathematical preparation. But as it is the leading idea guiding our construction of the Higgs system, I think it is appropriate to sketch an intermediate picture that is more accurate than the magnet analogy and suggestive of the generalization required in the Higgs system.

First we imagine that the electrons organize themselves into pairs, the so-called Cooper pairs. The wavefunction of a Cooper pair is a complex-valued function; it has both an amplitude and a phase. If we have a uniform density of Cooper pairs, then the amplitude is constant, but the phase can vary in space and time. We can represent the different possible values of the phase by points on a circle. So in representing the quantum dynamics of the Cooper pairs at each point of space-time, we have an overlying circular 'internal space' — an extra dimension if you like — and the position of the wavefunction in this extra dimension must be specified (Fig. 1a).

Figure 2 Gauge symmetry and the Higgs condensate. **a**, The minimal Higgs doublet takes values in an internal space that is the product of a three-dimensional sphere and a ray (indicated by the maroon points). **b**, To bring out the essential point, it is useful to simplify this geometry to two circles. Gauge transformations can rotate orientations within the vertical (A) or the horizontal (B) circles independently. **c**, This situation is analogous to Fig. 1c: it represents the nature of the condensation responsible for this more intricate, universal superconductivity. This condensation allows the free variation of an overall orientation (A+B), but locks the relative orientation within each pair of circles.

far provides a suitable candidate. We are therefore led to postulate that there is a new form of matter doing the job. Accordingly, what we perceive as empty space is in fact filled with an exotic, suitably aligned substrate: the Higgs condensate.

'Observing' this condensate only through its effect on W and Z bosons gives us very limited insight into its nature. The minimal hypothesis, within the conventional framework of contemporary physics, is to postulate the existence of a quantum field with just enough structure to support the desired pattern of symmetry breaking, imparting mass to some, but not all, of the gauge fields.

Let me spell this out with reference to the slightly simplified model of Fig. 2b. The pattern we want in reality is $SU(2) \times U(1)_Y \rightarrow U(1)_\gamma$; the simplified model exhibits $U(1)_A \times U(1)_B \rightarrow U(1)_{A+B}$, but the essential points are the same.

We begin with two independent rotation symmetries, but only a combination survives, which will represent ordinary electromagnetism. To break the symmetry in the appropriate way, we require the condensate to have a component in both circles (Fig. 2c). Furthermore, its orientations within the circles must be equal to one another, although their common angle can vary. Similarly to what was discussed for the Cooper-pair condensate, the Higgs condensate has a

form of rigidity that imposes an energetic price on fields that attempt to disrupt this favourable pattern. But the field that rotates both circles by the same amount preserves the pattern, so it remains massless. It can be identified with the electromagnetic field, which produces the photon γ. Note that the $A+B$ gauge symmetry of electromagnetism ($U(1)_\gamma$) is not the same as either 'pure' rotation A or B of the primary circles.

In the realistic $SU(2) \times U(1)_Y \rightarrow U(1)_\gamma$ case, a similar entwining of weak and electromagnetic gauge symmetry is mandatory, simply because the $W^\pm$ bosons are electrically charged. Thus, these interactions must be treated together, and we speak of 'electroweak' theory. It is often said that the $SU(2) \times U(1)_Y \rightarrow U(1)_\gamma$ theory of broken gauge symmetry is a unified theory of weak and electromagnetic interactions, but that seems to me an overstatement. The theory still involves two quite separate primary symmetries, which are related only by the fact that the same Higgs condensate spoils them both. The simplest way of implementing the symmetry-breaking pattern $SU(2) \times U(1)_Y \rightarrow U(1)_\gamma$ makes specific predictions for the masses and couplings of the W and Z bosons, which seem to be quite accurate. More complicated implementations are conceivable, but they seem neither desirable nor necessary.

Together with its orientations in the internal sphere, the quantum field associated with the Higgs condensate must have another degree of freedom, representing its overall magnitude. The degrees of freedom corresponding to variations in orientation are associated with the (broken-symmetry) gauge transformations. They get absorbed into the massive gauge fields; their quanta are longitudinally polarized W and Z bosons. The degree of freedom associated with changes in overall magnitude, however, has independent meaning. Its elementary excitation, or quantum, is what is usually referred to as the Higgs particle. The inescapable minimal consequence of our postulate of a Higgs condensate is a new, electrically neutral spin-0 particle.

One embellishment of the minimal scheme is especially well motivated, because it arises in connection with the important concept of low-energy supersymmetry[25]. The generalization simply involves imagining that two independent fields of the same general character (both representing positions on internal spheres and circles, and undergoing the same sorts of gauge transformations) contribute to the Higgs condensate. A striking consequence is that we then expect to have not one but five new Higgs particles, because we have introduced a field with four additional degrees of freedom (specified by its location within its four-dimensional internal space). Of these, three are electrically neutral; the other two are positively and negatively charged, each being the other's antiparticle.

...and a nest of kluges

It is not only the masses of W and Z bosons that are inconsistent with pristine $SU(2) \times U(1)$ gauge symmetry and that get tied up with the Higgs condensate. The masses of quarks and leptons present a similar difficulty. In the end, these masses too can be ascribed to interaction with the Higgs condensate, but the details are quite different and frankly (in my view) seem quite ugly and clearly provisional.

The proximate source of the difficulty is the observed parity violation of the weak interaction[26]. In its most basic form, the observed phenomenon is that when fast-moving quarks or leptons are emitted in weak decays, their spin tends to be aligned opposite to their direction of motion; they are therefore said to be left-handed. To accommodate this in the framework of gauge symmetry, we would like to say that the symmetry acts only on left-handed particles. That formulation, however, is inconsistent with special relativity. Indeed, a sufficiently fast-moving observer could overtake the particle, and to such an observer, its direction of motion would appear reversed, and it would be right-handed. This difficulty would be avoided if the quarks and leptons had zero mass, for then they would move at the speed of light and could not be overtaken. So we can implement the interactions of W bosons with

massless left-handed quarks and leptons in a consistent way using gauge symmetry; but non-zero masses for quarks and leptons must be tied up with gauge symmetry breaking.

Unlike in the case of W and Z bosons, there are no nice geometrical pictures associated with the masses of quarks and leptons, nor connections with profound concepts (gauge invariance, superconductivity). We are reduced to simply writing down the equations. For non-expert readers, the details are not vital, but I do want to convey a sense of what leads me to call them 'ugly' and 'provisional'.

At present there is no compelling theory that predicts the values of any quark or lepton masses. Yet there are striking facts to be explained that play a crucial part in the structure of the world as we know it. The mass of the electron, and of the up and down quarks are nearly a million times smaller than the 'natural' scale (250 GeV) set by the magnitude of the Higgs condensate, but the mass of the top quark is close to that scale; the CKM matrix (that describes the mixing of quark species) is measured to be almost, but not quite, the unit matrix. All this presumably indicates that principles — or perhaps accidents or even conspiracies — are lurking within the Higgs sector, which ensure that some, but not all, of the possible couplings are small. It is a great challenge for the future to discover these principles, or alternatively to understand why accident and conspiracy run rampant. So far, we have only had access to the masses and the CKM matrix, which encode highly processed versions of the basic couplings. The discovery of Higgs particles and measurements of their interactions will give us better access to fundamentals.

On the origin(s) of mass

As we have just seen, the masses of W and Z bosons, and of quarks and leptons, arise from the interaction of these particles with the pervasive Higgs condensate. This has inspired references to the Higgs particle as 'the origin of mass', or even 'the God particle'[27]. The real situation is interesting but rather different from what this hyperbole suggests. A few critical comments seem in order.

First, most of the mass of ordinary matter has an entirely different origin. This mass is contained in atomic nuclei, which are built up from nucleons (protons and neutrons), which in turn are built up from quarks (mainly up and down quarks) and colour gluons. Colour gluons are strictly massless, and the up and down quarks have tiny masses, compared with the mass of nucleons. Instead, most of the mass of nucleons (more than 90%) arises from the energy associated with the motion of the quarks and gluons that compose them[28] — according to the original form of Einstein's famous equation, $m = E/c^2$. This circle of ideas provides an extraordinarily beautiful, overwhelmingly positive answer to the question Einstein posed in the title of his original paper[29]: "Does the inertia of a body depend on its energy content?" And it has nothing to do with Higgs particles!

Second, as we have just seen, for quarks and leptons the Higgs mechanism appears more as an accommodation of mass than an explanation of its origin. We map observed values of masses and mixings through some distorting lenses into corresponding values of Higgs-field couplings, but only for the W and Z bosons do we have reliable, independent insight into what these values ought to be. And third, the Higgs field in no sense explains the origin of its own mass. A parameter directly equivalent to that mass must be introduced into the equations explicitly.

Finally, there is no necessary connection between mass and interaction with any particular Higgs field. As an important example, much of the mass of the Universe is observed to be in some exotic, 'dark' form. The dark matter has only been observed indirectly, through the influence of its gravity on surrounding ordinary matter. It has evaded more conventional observation using optical, radio or other telescopes, so evidently it has only a feeble coupling to photons. We do not yet know what this dark stuff is precisely, but the leading theoretical candidates (axions and weakly interacting massive particles, WIMPs) are massive particles that do not acquire their mass by interacting with the electroweak Higgs condensate. The general

point is that many kinds of hypothetical particle can have masses that, unlike the masses of W and Z bosons and quarks and leptons, do not violate the electroweak $SU(2) \times U(1)$ symmetry, and such masses need have no relation to that symmetry's spoiler, the electroweak Higgs condensate. The Higgs particle's own mass is also of this kind.

The genuine meaning of the Higgs field — that it embodies the concept of symmetry breaking and makes tangible the vision of a universal cosmic superconductor — is deep, strange, glorious, experimentally consequential and very probably true. This meaning has no need for embellishment, and can only be diminished by dubious oversimplification. So blush for shame, ye purveyors of hyperbole! End of sermon.

Searching for Higgs

From the observed masses of the W and Z bosons we infer the magnitude of the Higgs condensate to be 250 GeV. This sets the scale for structure in the Higgs system. The mass of the lightest excitation of the condensate could only be significantly larger than this if a large parameter appeared in the theory to amplify the scale. More specifically, what is relevant to this question is the coefficient of the nonlinear self-coupling of the Higgs field. But large self-coupling seems unlikely on both theoretical and experimental grounds. Theoretically, it takes us into a regime where the Higgs field undergoes violent quantum fluctuations, and appears to develop inconsistencies[30]. On the experimental side, precision measurements of various quantities can be compared to theoretical predictions at a level that is sensitive to the contribution of 'virtual particles'; that is, quantum fluctuations, including contributions that arise from coupling to the Higgs system. All indications so far are that these contributions are small, consistent with weak coupling[15]. Thus, on very general grounds, we expect to find new particles of some kind, recognizable as quanta of the Higgs condensate, with mass most probably smaller than 250 GeV, and in any case not much larger.

To go further and be quantitative, we must consider more specific models. The minimal implementation of symmetry breaking, which
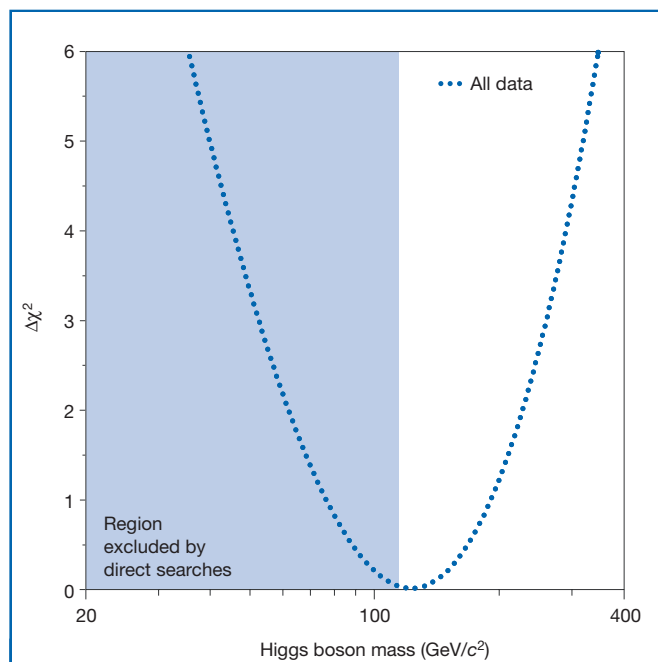
Figure 3 Constraints on the mass of the Higgs boson from experimental data. The dotted line shows, as a function of the Higgs mass, the χ^2 for a global fit to electroweak data (including such quantities as the mass of the W boson and the Z boson total width). Also included in the dataset is the current world-average value for the mass of the top quark[48]. A relatively light value for the Higgs mass is favoured, less than 200 GeV. Shading indicates the region excluded by direct searches for the Higgs boson[14].

a

t

g

h

g

b

e+

Z

e−

Z

h

Z

h Higgs particle
t Top quark
g Gluon

e Electron/positron
Z Z boson
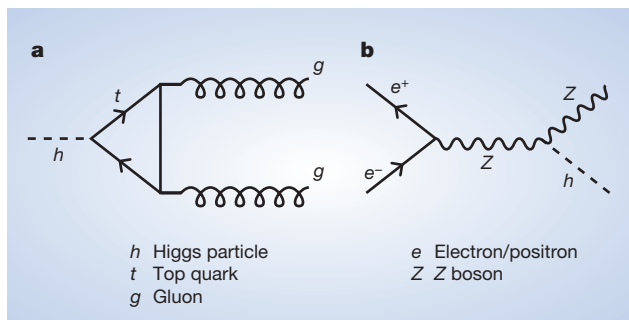
Figure 4 Higgs couplings. **a**, The dominant coupling of the Higgs particle to ordinary matter arises indirectly, through a virtual top/anti-top quark pair that annihilates into gluons. Direct couplings of the Higgs to up or down quarks, electrons, gluons or photons are very small, due to the tiny masses of these particles. **b**, Electrons can, however, give access to the Higgs through their coupling to the Z boson. The *Bremsstrahlung* process shown produces a distinctive *Zh* final state. Similar processes involving quark/antiquark annihilation produce either *Zh* or *Wh* final states.

assumes that only a single $SU(2) \times U(1)$ doublet exists (and thereby, as discussed above, predicts the existence of exactly one physical Higgs particle), is of course the canonical choice. Within this framework, which I shall adopt for the next few paragraphs, everything of interest is determined in terms of just one unknown parameter, conveniently taken to be the value of the Higgs particle's mass.

Now, with everything else pinned down, the aforementioned contributions (due to virtual particles, and often called radiative corrections) to a variety of measured quantities can be calculated precisely, as a function of the Higgs particle's mass. When this is done, one finds that the central value of the Higgs mass preferred by the experimental results is about 100 GeV, with a dispersion of about 50 GeV (Fig. 3).

The strength with which the Higgs particle couples to other fundamental particles is proportional to the mass of the particle involved (with a most important exception that I will come to in a moment). Thus, the dominant decay mode of the Higgs particle will be into the heaviest available particles. For a wide range of masses, including the most favoured values, the dominant decay is therefore into a heavy quark pair — bottom/anti-bottom ($b\bar{b}$); channels with heavier particles are energetically forbidden. The rates of decay into other quark or lepton pairs — charm/anti-charm and tau/anti-tau — are lower but not insignificant.

The exceptional case is gluons[31]. The direct coupling of the Higgs particle to gluons, inferred from the classical form of the theory, vanishes — as one might expect from the vanishing mass of the gluons. There is a purely quantum-mechanical process, however, in which the Higgs particle fluctuates into a virtual top/anti-top quark pair, which then materializes into a real gluon pair (Fig. 4a). As a fraction of all possible decays, this branch amounts to about 10%. A coupling to photons is similarly achieved by fluctuations through virtual *W* and *Z* bosons, as well as top quarks; it is smaller, owing to the weaker photon coupling.

Search strategies for the Higgs particle at electron–positron accelerators (the Large Electron Positron collider, LEP, or the planned International Linear Collider, ILC) and at proton–proton accelerators (the Tevatron at Fermilab, United States, or CERN's LHC) are quite different. The coupling of electrons and positrons to the Higgs particle is extremely small. A favoured strategy exploits processes in which the electron and positron annihilate into a *Z* boson, which then radiates the Higgs particle[32] (Fig. 4b). This has the further advantage that the final-state *Z* boson is very well characterized and can be reliably identified. Indeed, a general problem in Higgs-hunting is that one must distinguish the process of interest

from possible 'background' processes that can lead, from the point of view of what is observable in practice, to effectively the same final state. This channel (with a final state of $Zb\bar{b}$) was the subject of very intense scrutiny at LEP, especially during the final stages of its operation in 2000. The result, however, was inconclusive. A lower bound of 114.1 GeV, at the 95% confidence level, was put on the mass of the Higgs particle (assuming the minimal implementation). On the other hand, statistically weak (1.7σ) hints of excess events were observed at the upper limit of the machine's energy, consistent with what would be induced by a Higgs particle of mass 116 GeV.

At a hadron machine of sufficiently large energy, such as the LHC, Higgs particles are produced copiously through the gluon-fusion process, which exploits the Higgs' indirect coupling to gluon pairs[33]. The difficulty here is not production, but the signal-to-noise ratio in detection. Final states at hadron machines typically contain much extraneous debris, because the initial projectiles — protons — are complex objects. Indeed, our gluon-fusion production process relies on the fact that protons, in addition to quarks, contain a substantial fraction of gluons. The dominant final result of Higgs decay, a $b\bar{b}$ pair, is also very easily produced by ordinary particle interactions that in no way involve the Higgs particle. These form, from an experimental point of view, an almost impenetrable background. For this reason, it might be best to focus on the rarer decay of the Higgs into a pair of photons. Another strategy is to look for 'tagged' events that contain a *W* or *Z* in the final state, accompanying the Higgs particle (as at electron–positron machines). These could result from the radiation of Higgs particles from the *W* or *Z* bosons produced by quark/anti-quark annihilation.

Once we depart from the minimal implementation, many more possibilities for probing the Higgs sector open up, some much less demanding. For example, the additional Higgs particles predicted in the minimal supersymmetric model include electrically charged ones. These have very characteristic decay modes and they could be copiously produced at electron–positron machines once the energetic threshold for pair production is passed.

At this point it is appropriate to mention that whereas in the minimal model the predicted value of the Higgs mass is only weakly constrained from above, the minimal supersymmetric model generically predicts the existence of at least one Higgs particle with mass below 140 GeV, whose properties resemble those of the (absolutely) minimal Higgs particle with the same mass[17,34]. The above-mentioned indications from radiative corrections, and perhaps from LEP, might be interpreted as encouraging for the supersymmetric alternative.

This account of the practical strategy of Higgs searches is no more than a sketch of some of the simplest considerations involved. A vast, multi-faceted and growing literature is devoted to the subject. There is a general consensus that discovery of one or more Higgs particles cannot elude experimenters at the LHC, but that an exploration of the Higgs system worthy of the opportunities it affords will require the clean environment and high energies available in electron–positron collisions at the ILC.
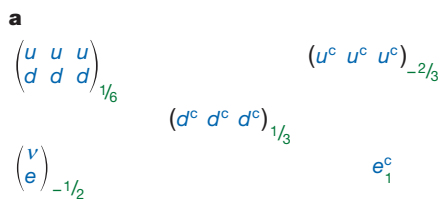
Extensions: unification, supersymmetry

If we merely count particles, then the minimal implementation of gauge symmetry breaking — using one doublet field and leading to one physical Higgs particle — recommends itself on grounds of economy. But considerations of logical coherence and structural integrity seem to lead us in another direction.

The structure of the gauge system gives powerful suggestions for its further fruitful development[35]. The product structure $SU(3) \times SU(2) \times U(1)$, the reducibility of the fermion representation (that is, the fact that the symmetry does not make connections linking all the fermions), and the peculiar values of the quantum number hypercharge assigned to the known particles all suggest a larger symmetry. The devil is in the details and it is not at all automatic that the observed, complex pattern of matter will fit neatly into a simple mathematical structure. But, to a remarkable extent, it does (Box 1).

Box 1

Unification

Within the standard gauge system the strong interactions are described by quantum chromodynamics (QCD), a theory based on the gauge symmetry group $SU(3)$ defined by transformations among

a

$$\begin{pmatrix} u & u & u \\ d & d & d \end{pmatrix}_{1/6} \qquad (u^c \; u^c \; u^c)_{-2/3}$$

$$(d^c \; d^c \; d^c)_{1/3}$$

$$\begin{pmatrix} \nu \\ e \end{pmatrix}_{-1/2} \qquad e^c_1$$

b

Quantum number Particle	R	W	B	G	P
u	+	−	−	+	−
u	−	+	−	+	−
u	−	−	+	+	−
d	+	−	−	−	+
d	−	+	−	−	+
d	−	−	+	−	+
u^c	−	+	+	−	−
u^c	+	−	+	−	−
u^c	+	+	−	−	−
d^c	−	+	+	+	+
d^c	+	−	+	+	+
d^c	+	+	−	+	+
ν	+	+	+	+	−
e	+	+	+	−	+
e^c	−	−	−	+	+
N	−	−	−	−	−

Hypercharge $Y = -\frac{1}{6}(R + W + B) + \frac{1}{4}(G + P)$

three colour charges; the weak interactions are described by an independent but mathematically similar gauge symmetry $SU(2)$ using two colour charges, and finally an independent 'hypercharge' symmetry $U(1)$ based on a single type of charge. Ordinary electromagnetism involves a combination of $SU(2)$ and $U(1)$ symmetry that remains valid after these separate symmetries are broken by the Higgs condensate.

In part **a** of the figure, the relationships of the quarks and leptons under these transformations are shown. The strong gauge symmetries act horizontally, the weak gauge symmetries act vertically and the values of the hypercharges are indicated by subscripts. The 15 quarks and leptons shown here fall into five unrelated clans. (There is a three-fold repetition of this entire structure, accommodating 45 quarks and leptons altogether.)

Theories of unified gauge symmetry propose that there is a more extensive symmetry that involves transformations among all these colour charges. That symmetry must be spontaneously broken, to explain why we observe the consequences of different types of charge to be quite different (strong interactions really are strong, and weak interactions weak). Nevertheless, this idea is not without consequences: the quarks and leptons must furnish the material for building unified structures that remain coherent under the extended symmetry.

One particular unified symmetry passes this test with flying colours (and is shown in part **b** of the figure). Although the smallest simple group into which $SU(3) \times SU(2) \times U(1)$ could possibly fit is $SU(5)$ (ref. 49) — it fits all the fermions of a single family into two representations ($\mathbf{10} + \bar{\mathbf{5}}$) and the hypercharges click into place — a larger symmetry group, $SO(10)$ (ref. 50), fits these and one additional $SU(3) \times SU(2) \times U(1)$ singlet particle into a single representation (the spinor **16**). All 15 quarks and leptons appear on the same footing, and the additional particle, which has the quantum numbers of a right-handed neutrino, is quite welcome: it plays a crucial role in the attractive 'seesaw' model for neutrino masses[43,44].

Where before we had a piecemeal accommodation of the observed particles, now we have a marvellous correspondence between reality and a unique, ideal mathematical object.

This unification of quantum numbers, although attractive, remains purely formal until it is embedded in a physical model. This requires realizing the enhanced symmetry in a local gauge theory. The enhanced symmetry must be broken. The Higgs system of electroweak theory supplies a precedent for that. What we need is another condensate, with a vastly larger density. We are proposing that the world is a multi-coloured, multi-layered superconductor.

There is an apparent difficulty with these ideas, which on closer scrutiny turns out to represent their greatest success. Non-abelian gauge symmetry requires that the relative strengths of the different couplings must be equal (universality), which is not what is observed. Fortunately, there is a compelling way to save the situation[36]. The higher symmetry is broken at a very large energy scale (equivalently, a small distance scale), but we observe interactions at much smaller energies (larger distances). The strength we observe differs from the intrinsic strength, because it is affected by the physics of vacuum polarization. A cloud of virtual particles surrounds the charge and can enhance or dilute its power. The resulting 'running' of couplings is an effect that can be calculated quite precisely, given a definite hypothesis about the particle spectrum. In this way we can test quantitatively the idea that the observed couplings originate from a single unified value at small distances.

The results of these calculations are quite remarkable and encouraging (Fig. 5). If we include vacuum polarization from the particles we know about in the minimal standard model, we find approximate

unification. If we include vacuum polarization from the particles needed in expanding the standard model to include supersymmetry[37,38], softly broken at the TeV (1,000 GeV) scale, we find accurate unification[39,40]. The unification occurs at a very large energy scale, of order 10^{16} GeV. This success is robust against small changes in the supersymmetry-breaking scale, and is not adversely affected by incorporation of additional particle multiplets, as long as they form complete representations of $SU(5)$ (ref. 37).

Low-energy supersymmetry is desirable on several other grounds as well. The most important has to do with the Higgs condensate. In the absence of supersymmetry, radiative corrections to the magnitude of this condensate diverge ('radiative corrections' is standard jargon for the effect of virtual particles or, alternatively, quantum fluctuations). One must fix the condensate's value, which sets the scale for electroweak symmetry breaking, by hand, as part of the definition of the theory. This 'renormalization' procedure leaves it utterly mysterious why the empirical value is so much smaller than unification scales. Moreover, enhanced unification symmetry requires that the Higgs doublet should come together with additional fields, to fill out a complete representation. Its partners, however, have the quantum numbers to mediate proton decay, so if they exist at all, their masses must be very large, at about the unification scale of 10^{16} GeV. This reinforces the idea that such a large mass is what is 'natural' for a scalar field[41]. The relatively small mass (of the order of 100 GeV) of the Higgs field that we need in our electroweak theory
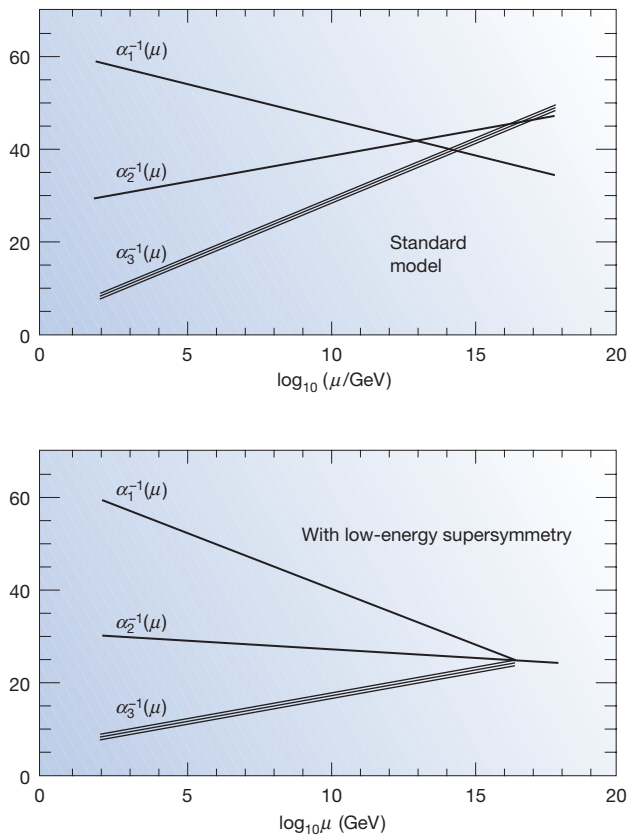
Figure 5 Unification of the forces. The strengths of the couplings of the weak, electromagnetic and strong forces are hugely disparate (represented here as α_1^{-1}, α_2^{-1} and α_3^{-1}). But their perceived strength changes with the energy scale of the process (μ), through corrections due to virtual particles. Assuming there are only the particles known to us in the standard model and extrapolating beyond the reach of experiment to very high energies, the couplings move towards each other but do not converge at a single point (top). If, however, the extra particles needed to implement low-energy supersymmetry are included in the calculation, the couplings meet neatly at an energy of about 10^{16} GeV (lower plot). Note that the energy scale is logarithmic (and the existence of other unknown particles is overlooked), so this calculation is a bold — perhaps reckless — extrapolation of the laws we know to apply to energies vastly larger than those at which these laws have been tested.

seems unnatural and requires some special justification. Supersymmetry, if it is not too badly broken, largely solves this problem, for it ensures that these unsavoury radiative corrections are small[42].

The fact that our unification calculations point to an enormous new mass scale for unification is profound. This enormous mass scale is inferred entirely from data taken at much lower energies. The disparity of scales arises from the slow (logarithmic) running of inverse couplings, which implies that modest differences in observed couplings must be made up by a long interval of running. The appearance of a very large mass scale is welcome on several grounds. I will mention three of the most important.

First, right-handed neutrinos, which as we have seen can enhance the symmetry of unification, naturally acquire masses of the order of the unification scale. Masses of that magnitude remove these particles from direct experimental accessibility, but they can have a most important indirect effect[43,44]. This is because, in second-order perturbation theory, the ordinary left-handed neutrinos make virtual transitions to their right-handed relatives and back. This exotic process generates non-zero masses for the ordinary neutrinos, but these are much smaller than the masses of other leptons and quarks.

The magnitudes that arise in this way are broadly consistent with the tiny observed masses of neutrinos. No more than order-of-magnitude success can be claimed because many relevant details of the models are poorly determined.

Second, unification tends to obliterate the distinction between quarks and leptons, and hence to open up the possibility that protons decay (their building-block quarks turn into electrons or muons). Heroic experiments to observe this process have so far come up empty-handed, with limits on partial lifetimes approaching 10^{34} years for some channels. It is very difficult to ensure that these processes are sufficiently suppressed, unless the unification scale is very large. Even the high scale indicated by the running of couplings and neutrino masses is barely adequate. Spinning it positively, experiments to search for proton decay remain a most important and promising probe into the physics of unification. Similarly, it is difficult to avoid the idea that unification brings in new connections among the different families. There are significant experimental constraints on flavour-changing neutral currents, lepton number violation and other exotic processes that must be suppressed, and this makes a high mass scale for the virtual particles that mediate them most welcome.

Third, with the appearance of this large scale, unification of the strong and electroweak interactions with gravity becomes much more plausible. Newton's constant has dimensions of mass, so it runs even classically. Or, to put it another way, gravity responds to energy/momentum, so it gets stronger at large energy scales. Nevertheless, because gravity starts out extremely feeble compared to other interactions on laboratory scales, it becomes roughly equipotent with them only at enormously high scales, comparable to the Planck energy of 10^{18} GeV. By inverting this thought, we gain a deep insight into one of the main riddles about gravity: if gravity is a primary feature of nature, reflecting the basic structure of space-time, why does it ordinarily appear so feeble? Elsewhere[45], I have traced the answer down to the fact that, at the unification (Planck) scale, the strong coupling is about 1/2!

In view of all this, our accounting of the 'economy of ideas' is altered. For it seems that with five Higgs particles you can buy a lot more than with one.

Cosmological implications

In the very early Universe, when temperatures were much higher, the Higgs condensate that now fills all space could not have maintained its alignment over extended distances. In a word, it melted[46]. Just as a superconductor heated beyond its critical temperature goes normal, or a magnet heated above its Curie temperature loses its magnetization, the Universe would then have been in a different, more symmetric phase. In this phase, W and Z bosons — like photons, colour gluons and gravitons — had zero mass, as did quarks and leptons. (Ironically, the Higgs particles themselves retained a finite mass.)

Thus, during the early evolution of the Universe there was a dramatic change in the properties of matter. The detailed physical nature of this change is at present unknown. It may have been a sharp phase transition in the thermodynamic sense, or a smooth crossover. Such a cosmic phase transition might have been accompanied by unusual or violent physical events that left lasting consequences. One possibility is that the current imbalance between the abundance of matter and antimatter might have been generated when the Higgs condensate froze in. Another is that the Higgs freeze-in catalysed an epoch of extremely rapid cosmic acceleration, akin to or even identical to the inflationary epoch, whose occurrence is widely conjectured in modern cosmology[47] but whose physical nature is highly uncertain. It is only by studying the Higgs system in detail that we can begin to assess these possibilities reliably.

The existence of a Higgs system with properties of the general sort I have discussed, notably including one or more accessible, recognizable Higgs particle, appears to be a compelling consequence of quantum field theory and the standard model of fundamental physics

— a complex of ideas that has been tremendously successful. This would be a beautiful thing to observe, and extremely instructive.

And yet, our standard system of gravity — general relativity — incorporates the principle that all forms of energy exert gravitational influence. As a special case, the postulated Higgs condensate, which fills all space, should weigh something. In fact, it should weigh a lot: estimating the energy density of this condensate using straightforward dimensional reasoning gives a value much larger than is allowed by observations. It would show up as such a large contribution to the cosmological term that the size of the Universe would double every 10^{-38} seconds! Thus, either the Higgs condensate does not exist, or its energy density is cancelled out by some other, still more exotic, contribution, or there is a profound lacuna in our understanding of gravity. The second and third alternatives seem sufficiently improbable as to suggest that just maybe we will be dragged to the first. If so, then our search for Higgs particles as the quanta of gauge symmetry breaking might instead turn up something quite different. Theoretical physicists, roused from their dogmatic slumbers, would be forced back to the drawing board. Wouldn't that be fun? □

1. Glashow, S. Partial symmetries of weak interactions. *Nucl. Phys.* **22,** 579–588 (1961).
2. Weinberg, S. A model of leptons. *Phys. Rev. Lett.* **19,** 1264–1266 (1967).
3. Salam, A. in *Elementary Particle Physics, Nobel Symp.* (ed. Svarthom, N.) No. 8, 367–377 (Almqvist & Wiksell, Stokholm, 1968).
4. 't Hooft, G. & Veltman, M. Regularization and renormalization of gauge fields. *Nucl. Phys. B* **44,** 189–213 (1972).
5. Englert, F. & Brout, R. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.* **13,** 321–323 (1964).
6. Higgs, P. Broken symmetries and the masses of gauge bosons *Phys. Rev. Lett.* **13,** 508–509 (1964).
7. Wess, J. & Zumino, B. A lagrangian model invariant under supergauge transformations. *Phys. Lett. B* **49,** 52–54 (1974).
8. Ferrara, S. (ed.) *Supersymmetry* (North Holland/World Scientific, Singapore, 1987).
9. Weinberg, S. *The Quantum Theory of Fields 1: Foundations* (Cambridge Univ. Press, Cambridge, 1995).
10. Wilczek, F. Quantum field theory. *Rev. Mod. Phys.* **71,** S85–S95 (1999).
11. Weyl, H. *Z.* Elektron and gravitation I. *Z. Phys.* **56,** 330–352 (1929).
12 London, F. Quantenmechanische deutung der theorie von Weyl. *Z. Phys.* **42,** 375–389 (1927).
13. Yang, C. N. & Mills, R. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.* **96,** 191–195 (1954).
14. ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration & The LEP Working Group for Higgs Boson. Search for the standard model Higgs boson at LEP. *Phys. Lett. B* **565,** 61–75 (2003).
15. LEP Electroweak Working Group [online] <http://lepewwg.web.cern.ch/LEPEWWG> (2004).
16. Einhorn, M. (ed) *The Standard Model Higgs Boson* (North Holland, Amsterdam, 1991).
17. Gunion, J., Haber, H., Kane, G., & Dawson, S. *The Higgs Hunter's Guide* (Addison Wesley, New York, 1990).
18. Gross, D. & Wilczek, F. Ultraviolet behavior of non-Abelian gauge theories. *Phys. Rev. Lett.* **30,** 1343–1346 (1973).
19. Politzer, H. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.* **30,** 1346–1349 (1973).
20. Riess, A. *et al.*, Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.* **116,** 1009–1038 (1998).
21 Permutter, S. *et al.* Measurements of Ω and Λ from 42 high-redshift supernovae. *Astrophys. J.* **517,** 565–586 (1999).
22. Spergel, D. *et al.*, Determination of cosmological parameters. *Astrophys. J.* **148** (suppl.), 175–194 (2003).
23. London, F. *Superfluids 1: Macroscopic Theory of Superconductivity* (Wiley, New York, 1950).
24. Bardeen, J., Cooper, L. & Schrieffer, R. Theory of superconductivity. *Phys. Rev.* **108,** 1175–1204 (1957).
25. Weinberg, S. *Quantum Theory of Fields, Vol.* 3*: Supersymmetry* (Cambridge Univ. Press, Cambridge 2000).
26. Lee, T. D. & Yang, C. N. Question of parity conservation in weak interactions. *Phys. Rev.* **104,** 254–258 (1956).
27. Lederman, L. & Teresi, D. *The God Particle* (Bantam, New York, 1993).
28. Wilczek, F. The origin of mass. *Ann. Phys. (MIT)* **16,** 24–35 (2003).
29. Einstein, A. Does the inertia of a body depend upon its energy content? [in German] *Ann. Phys.* **18,** 639–641 (1905).
30. Kuti, J., Lin, L. & Shen, Y. Upper bound on the Higgs-boson mass in the standard model *Phys. Rev. Lett.* **61,** 678–681 (1988).
31. Wilczek, F. Decays of heavy vector mesons into Higgs particles. *Phys. Rev. Lett.* **39,** 1304–1306 (1977).
32. Bjorken, J. in *Proc. SLAC Summer School* (SLAC publication 198, 1976).
33. Georgi, H., Glashow, S., Machacek, M. & Nanopoulos, D. Higgs bosons from two-gluon annihilation in proton–proton collisions. *Phys. Rev. Lett.* **40,** 692–694 (1978).
34. CMS collaboration. The compact muon solenoid [online] <http://www.lip.pt/~outreach/docs/cms2> (2004)
35. Pati, & Salam, A. Unified lepton-hadron symmetry and a gauge theory of the basic interactions. *Phys. Rev. D* **8,** 1240–1251 (1973).
36. Georgi, H., Quinn, H. & Weinberg, S. Hierarchy of interactions in unified gauge theories. *Phys. Rev. Lett.* **33,** 451–454 (1974).
37. Dimopoulos, S., Raby, S. & Wilczek, F. Supersymmetry and the scale of unification. *Phys. Rev. D* **24,** 1681–1683 (1981).
38. Dimopoulos, S. & Georgi, H. Softly broken supersymmetry and SU(5). *Nucl. Phys.* **193,** 150–162 (1981).
39. Amaldi, U., de Boer, W. & Fürstenau, H. Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP. *Phys. Lett. B* **260,** 447–455 (1991).
40. Ellis, J., Kelly, S. & Nanopoulos, D. Precision LEP data, supersymmetric GUTs and string unification. *Phys. Lett. B* **249,** 441–448 (1990).
41. Gildener, E. & Weinberg, S. Symmetry breaking and scalar bosons. *Phys. Rev. D* **13,** 3333–3341 (1976).
42. Witten, E. Dynamical breaking of supersymmetry. *Nucl. Phys. B* **188,** 513–554 (1981).
43. Gell-Mann, M., Ramond, P. & Slansky, R. in *Supergravity: Proceedings of the Supergravity Workshop at Stony Brook, September 27-29, 1979* (eds van Nieuwenhuizen, P. & Freedman, D.) 315–321 (North Holland, Amsterdam, 1979).
44. Yanagida, T. in *Workshop on Unified Theory and Baryon Number in the Universe* (eds Sawada, O. & Sugamoto, A.) 95–98 (KEK, Tsukuba, 1979).
45. Wilczek, F. Scaling mount Planck 3: is that all there is? *Phys. Today* **55N8,** 10–11 (2002).
46. Kirzhnits, D. & Linde, A. Symmetry behaviour in guage theories. *Ann. Phys.* **101,** 195–238 (1976).
47. Guth, A. Inflationary universe: a possible solution to the horizon and flatness problems. *Phys. Rev. D* **23,** 347–356 (1981).
48. A precision measurement of the mass of the top quark. *Nature* **429,** 638–642 (2004).
49. Georgi, H. & Glashow, S. Unity of all elementary-particle forces. *Phys. Rev. Lett.* **32,** 438–441 (1974).
50. Georgi, H. in *Particles and Fields 1974* (ed. Carlson, C.) 575–584 (AIP, New York, 1975).