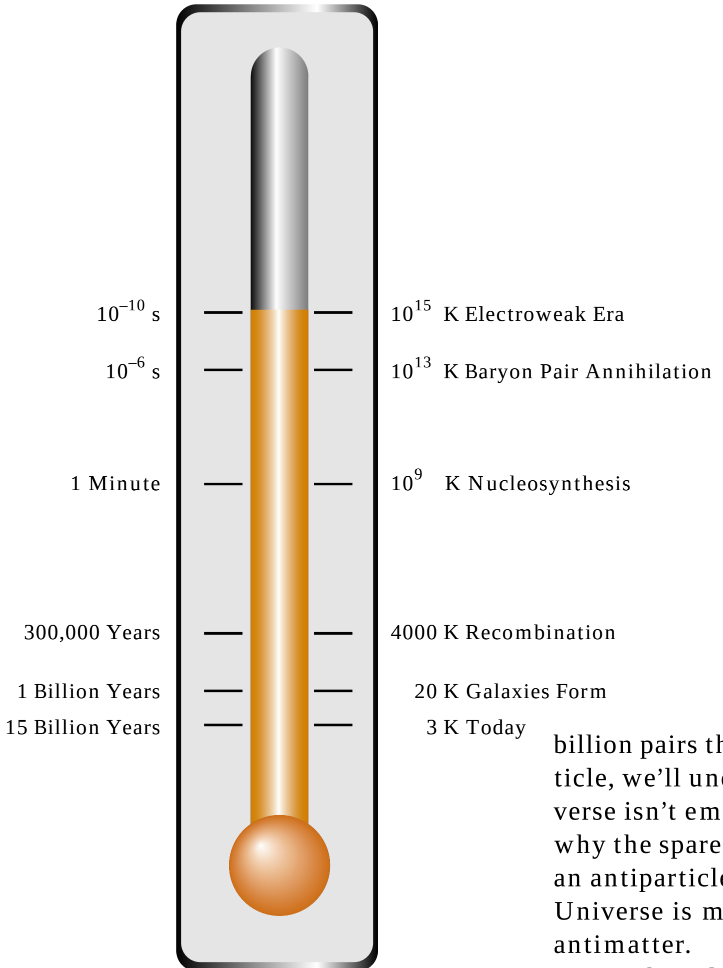
The Mystery of the Matter Asymmetry

by ERIC SATHER

Why is our Universe made of matter and not antimatter? The answer might be found in the laws that govern elementary particles.

WE'VE DISCOVERED a lot about our Universe by asking questions and then looking for answers. For example, we asked how stars are powered and found the answer in the transformations of atomic nuclei. But there are still simple questions that we can ask. And one is: Why is our Universe full of things like us and stars, and not empty?

It doesn't seem remarkable that there are things in our Universe. But if we look back toward the beginning of the Universe we can see that having it turn out not empty was a close thing. The Universe has cooled over its long history, but was extremely hot just after it was born in a Big Bang: so hot that there was lots of energy for making particles and antiparticles in pairs. As the Universe cooled, these particles and antiparticles annihilated in pairs. Had the amounts of matter and antimatter been equal, everything would have annihilated and the Universe would be empty. So when the Universe was hot there must have been more matter than antimatter, so that after it cooled we and stars would be left over.



Important events in the known history of the Universe (times and temperatures are approximate). The Universe has cooled since its formation in a hot Big Bang, so the earliest times correspond to the highest temperatures. As indicated by the level of the mercury, this article concerns the electroweak era. Subsequent events shown are baryon-antibaryon annihilation, which left the residual baryon asymmetry; the synthesis of light nuclei; recombination, when electrons and nuclei combined into neutral atoms, leaving the Universe transparent to light; galaxy formation; and today, when the Universe is filled with 3-degree-Kelvin microwave background radiation, which is the light released at the time of recombination redshifted by the subsequent expansion of the Universe.

If we work out what the Universe was like one billionth of a second after it began, it turns out that for every billion particle-antiparticle pairs there was just one extra particle. To that particle we and stars owe our existence. If we can explain why for every billion pairs there was one spare particle, we'll understand why the Universe isn't empty. And if we can say why the spare was a particle and not an antiparticle, we'll know why the Universe is made of matter and not antimatter.

So why, ultimately, is the world made of matter and not antimatter or nothing at all? Perhaps it is just how the Universe was composed at the instant of the Big Bang or an accident of subsequent history. But it could be the result of laws of nature which we can discover. While we're still looking for the answer, in recent years we've come to realize that we might find it if we can penetrate the next layer of microscopic physics.

BARYOGENESIS

Before looking for the origin of the excess of matter over antimatter, we first need to understand a little about the excess itself. The matter in our Universe is not static but is transformed in stars. Nuclear transformations such as the reaction *proton* → *neutron* + *positron* + *neutrino* change the populations of particle species. Here protons decrease in number while neutrons increase. However the population of the more general class of particles called baryons, which includes protons and neutrons, doesn't change. Baryons and antibaryons can be created and annihilated in pairs, but the excess of baryons over antibaryons, known as baryon number, is constant. In fact, baryon number is conserved in all reactions that have been observed. Hence the baryon number has remained constant for as far back into the history of the Universe as we can describe it using the physics we have observed and understand.

The above reaction also preserves a similar quantity called lepton number, because a neutrino is an example of a lepton, and a positron (antielectron) is an antilepton. And all observed reactions conserve lepton number. However, to determine the lepton number we'd have to count neutrinos, and neutrinos are hard to detect. Baryons meanwhile make up most of the mass of the things we can see. Therefore the matter excess that we can observe, that dates back to the first moments of the Universe, and that we need to explain is an excess of baryons.

High-energy experiments have revealed that each baryon, such as a proton or neutron, is actually a composite of three more fundamental objects, called quarks. To date six kinds, or flavors, of quarks have been discovered. Similarly each antibaryon consists of three antiquarks. The baryon number of the Universe is then one-third the quark number. The composite nature of baryons implies that the ultimate explanation of the matter-antimatter asymmetry must be framed in the language of quarks. Nevertheless, for historical reasons, the matter excess of the Universe is referred to as the baryon asymmetry. And the production of the matter excess is called baryogenesis.

SAKHAROV CONDITIONS

The birth of the field of baryogenesis and the idea that the matter excess could be explained by microscopic physics came in 1967. In that year Andrei Sakharov listed three conditions necessary for an explanation of the baryon asymmetry. In so doing, he laid the foundation for all future attempts to explain the matter excess of the Universe.

Sakharov pointed out that in order to produce a baryon excess where none existed before there first must be processes that change the baryon number. Such baryon-number-violating processes have not yet been observed. Second, the laws of nature must be biased so that a matter excess results and not an antimatter excess. Third, and less obvious, the baryon-number-violating processes must be out of thermal equilibrium. Otherwise, in equilibrium, these processes would even the amounts of baryons and antibaryons and nullify the baryon number. Providing these three ingredients—baryon-number violation, matter-biased laws, and thermal nonequilibrium—is the starting point for any attempt to explain the matter-antimatter asymmetry of the Universe.

The earliest ideas about baryogenesis centered on speculative theories that provide the desired baryon-number violation. Unfortunately, these theories describe physics at energies far beyond the current reach of experiment. Current experiments try to test the long-prevailing theory of elementary-particle physics, the very successful electroweak theory. After these early investigations of baryogenesis, it was discovered that the electroweak theory itself could provide the necessary baryon-number violation. With this realization, that the origin of the matter asymmetry might be found in the layer of physics now being revealed by experiment, the focus of the baryogenesis quest shifted.

If we work out what the Universe was like one billionth of a second after it began, it turns out that for every billion particle-antiparticle pairs there was just one extra particle. To that particle we and stars owe our existence.

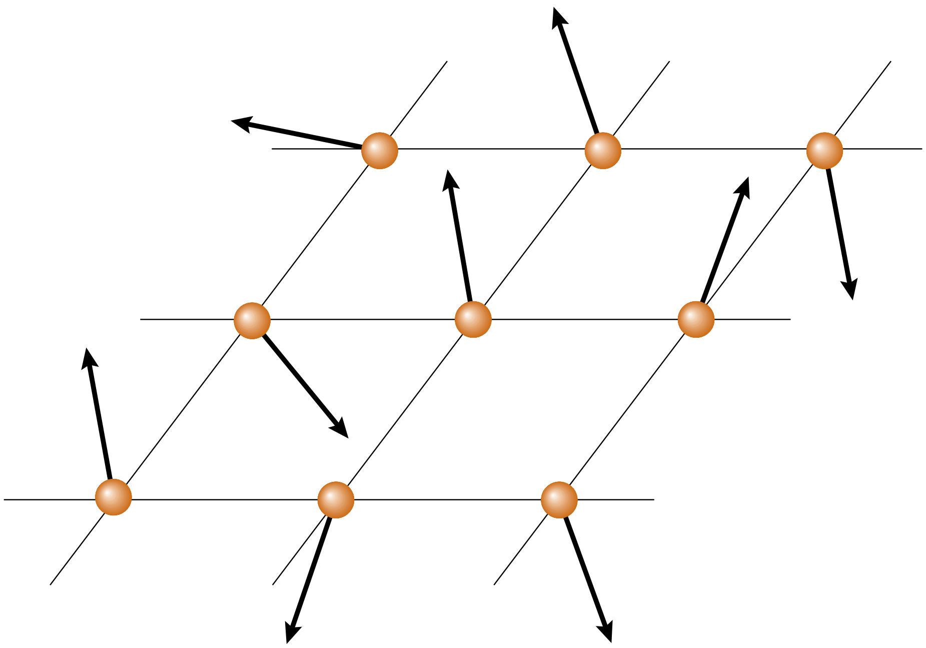
ELECTROWEAK BARYOGENESIS

To see how the baryon asymmetry could be produced by electroweak baryogenesis, we need to know some of the basics of electroweak physics. The electroweak theory summarizes our deepest insights into the ultimate laws of nature. It synthesizes the electromagnetic theory of charges and light and the weak theory of nuclear β-decay. In embracing these disparate theories, the electroweak theory predicts a wealth of new phenomena. Over the last twenty-five years, experiments have observed many of these phenomena and shown that the predictions of the theory hold to remarkable accuracy.

Our understanding of the electroweak interactions, and indeed all the physics of elementary particles, relies heavily on the ideas of symmetry and broken symmetry. To illustrate these ideas, consider the example of a ferromagnetic material like iron. In hot iron, the spins of the electrons point randomly, oriented in all directions with equal probability. There is no net magnetization, and the iron exhibits rotational symmetry, appearing the same from all directions (see the illustration on the next page). The energy of a ferromagnet is least when the electron spins all point in the same direction. So in cold iron the spins align, the iron is magnetized, and the overall rotational symmetry is broken. Some rotational symmetry persists, however. The iron still appears the same when rotated about the direction of magnetization (see the illustration on page 35).

The most important symmetries in nature are the so-called gauge symmetries, which give rise to the known forces. Gauge symmetry lies at the heart of the electroweak theory, producing the electromagnetic and weak forces. These forces are transmitted by messenger particles called gauge bosons: the photon transmits electromagnetism, while *W* and *Z* bosons transmit weak interactions.

We can understand the basics of electroweak gauge symmetry by analogy with a ferromagnet. Like the rotational symmetry of cold, magnetized iron, the electroweak



Schematic showing the electron spins in a portion of a ferromagnet. Above: At high temperature the spins point in random directions, and the ferromagnet is rotationally symmetric.

symmetry is broken, but not completely. The weak symmetry breaks but the electromagnetic symmetry survives. Because electromagnetic symmetry survives, the photon is massless, and electromagnetic forces carry over large distances. In contrast, because weak symmetry breaks, the *W* and *Z* are massive, and weak interactions act only over a very short range and thus appear weak.

We don't yet know what breaks the electroweak symmetry. The simplest explanation is that there is a field called the Higgs which, just like a ferromagnet, breaks symmetry when it falls into its state of lowest energy. All we really know is that there is some mechanism that breaks the electroweak symmetry and thereby gives mass to the *W*, *Z*, and all other massive particles. (See the previous article in this issue by Lance Dixon in which he illuminates the Higgs mechanism using an analogy with superconductivity.)

ELECTROWEAK PHASE TRANSITION

Returning to the example of a ferromagnet, hot iron occupies a state, or phase, of symmetry, while cold iron lies in a phase of broken symmetry. When hot iron is cooled below a critical temperature, magnetization develops and rotational symmetry breaks. At this Curie temperature, the iron suffers a phase transition from the symmetric phase to the broken phase.

In direct analogy with the ferromagnet, which loses its magnetization and exhibits maximum symmetry at high temperature, the electroweak symmetry was unbroken when the Universe was born in a hot Big Bang. The critical temperature for electroweak symmetry breaking is, however, enormously higher than the Curie temperature of iron. The Universe had cooled to this temperature and experienced an electroweak phase transition only one ten-billionth of a second after its birth.

During this transition, just as bubbles of steam form in boiling water, bubbles of broken phase formed. These bubbles expanded until they filled the Universe, leaving it in its current phase of broken symmetry. Throughout the transition, the Universe was out of equilibrium, thus satisfying one of Sakharov's conditions. Therefore, if the origin of the baryon asymmetry lies in electroweak physics, the asymmetry must have formed during the electroweak phase transition.

ELECTROWEAK BARYON NUMBER VIOLATION

What about the other Sakharov conditions, for instance baryon-number violation? At first glance, the electroweak theory appears to conserve baryon number; there are no explicit interactions that change it. However, due to quantum-mechanical subtleties, there are baryon-number violating processes.

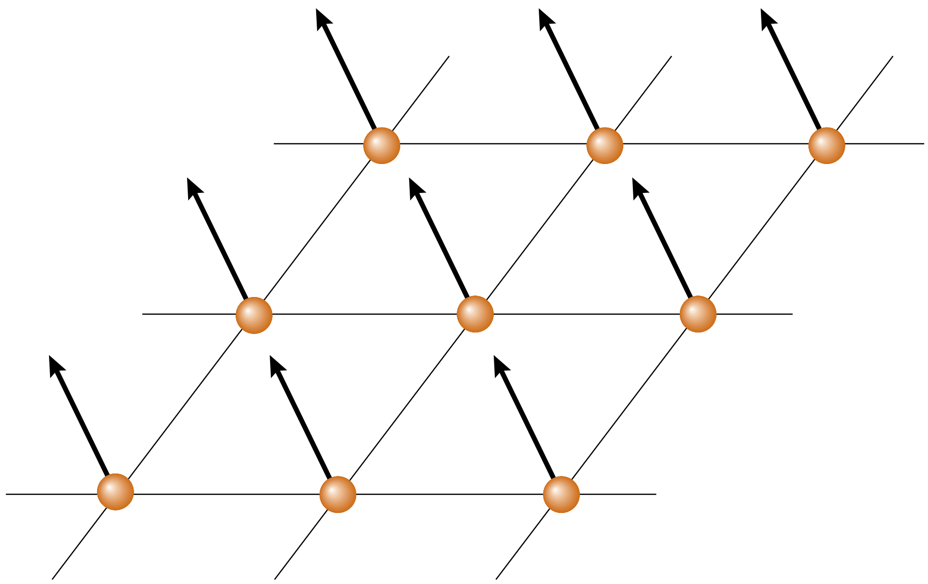
Then why has baryon-number violation escaped detection? Because today, in the broken phase, such violation requires quantum-mechanical tunneling through a large energy barrier and is in consequence suppressed. But in the symmetric phase, both before and during the electroweak phase transition, this barrier was absent and baryon number could fluctuate. This insight led to the study of electroweak baryogenesis.

C AND CP

So the electroweak theory can provide both baryon-number nonconservation and thermal nonequilibrium. What about the remaining Sakharov condition? Does the electroweak theory distinguish between matter and antimatter?

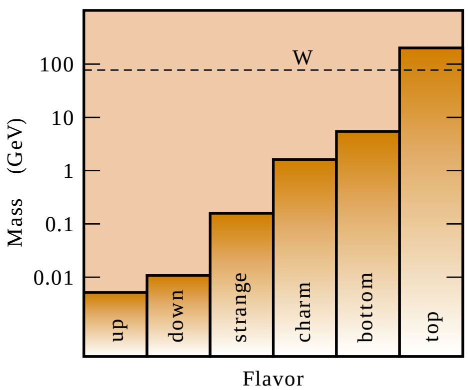
To answer this question, we need to consider two transformations which relate matter to antimatter. The first, called charge conjugation and denoted *C*, simply interchanges particles with antiparticles. The second, denoted *CP*, is a composite of *C* and the parity transformation *P*. Parity, like a mirror, reverses the direction of particle motion but preserves spins. Hence the combined operation *CP* turns a particle into an antiparticle with reversed momentum but identical spin. Since both *C* and *CP* relate particles to antiparticles, if either were a symmetry of the laws of nature, particle production would always be countered by equal antiparticle production, and no baryon asymmetry could result.

Are *C* and *CP* symmetries of nature? *C* and *P* are symmetries of the electromagnetic interactions and also of the strong interactions which bind quarks into protons and neutrons, and bind these, in turn, into nuclei. These transformations were assumed to be exact symmetries of all laws of nature. But in 1956 Lee and Yang realized that *C* and *P* are only approximate symmetries: weak interactions violate them. Soon after, parity violation was observed in nuclear β-decay. The composite operation *CP* still appeared to be an exact symmetry, but in 1964 a group led by Cronin and Fitch discovered *CP* violation in the weak decays of



At low temperature the spins align, the ferromagnet is magnetized, and the overall rotational symmetry is broken. The ferromagnet is still symmetric under rotations about the direction of magnetization.

		Generation		
		1	2	3
Charge	2/3	up	charm	top
	−1/3	down	strange	bottom



The six known quark flavors. Top: The known quarks and leptons occur in three sets called generations. Each generation consists of a pair of quarks, with charges 2/3 and –1/3, and a pair of leptons. This illustration shows how the six quark flavors pair to fit into generations. To a first approximation, charge-changing weak interactions only interconvert quarks within a generation. Bottom: Illustration showing quark masses, together with a dashed line indicating the mass of the W *boson, that characterize the mass scale of the weak interactions. All quark flavors other than the top are very light compared to the* W*. This suppresses baryogenesis that relies on Kobayashi-Maskawa* CP *violation.*

particles called kaons. The quark constituents of these kaons must therefore have *CP*-violating weak interactions. Hence the weak interactions violate both *C* and *CP*.

A RECIPE FOR BARYOGENESIS

So the electroweak theory contains all three ingredients for baryogenesis required by Sakharov. Now we just need a recipe for how to combine them to make a baryon asymmetry. Let us assemble the ingredients at the time of the electroweak phase transition: Initially, the Universe is filled with the symmetric phase, but bubbles of broken phase form and expand, supplanting the symmetric phase. Baryon-number-violating processes are rapid in the symmetric phase but shut off—out of equilibrium—inside the bubbles. Finally, a plasma of quarks and antiquarks with *C*- and *CP*-violating interactions permeates the Universe.

Now suppose that as the surfaces of expanding bubbles sweep through the plasma, antiquarks are less likely to enter the bubbles than quarks because of *C* and *CP* violation. The excess of antiquarks left outside the bubbles is simply erased by the baryon-number changing processes active in the symmetric phase. However, an opposite excess of quarks is deposited inside the bubbles, where baryon-number is conserved. This excess would survive to the present day as the baryon asymmetry of the Universe.

THE FAILURE OF STANDARD CP VIOLATION

Thus we have the ingredients and a recipe for producing a baryon excess. But can they reproduce the baryon asymmetry that we observe? Unfortunately, the standard electroweak theory fails. The reason lies in the origin of *CP* violation. In the standard electroweak theory, *CP* violation originates from charge-changing weak interactions that change the charge and flavor of quarks. The six known quark flavors divide evenly between two charge states. There are three "up-type" quarks with charge 2/3: up, charm, and top; and three "down-type" quarks with charge –1/3: down, strange, and bottom. The up-type and down-type flavors can be thought of as coming in pairs: up-down, charm-strange, and top-bottom (see top figure on the left). To a good approximation, the charge-changing interactions only operate within a pair, for example turning up into down or vice versa. But more precisely, these interactions turn an up quark into a quantum mixture of down-type quarks which is mostly down but is partly strange and bottom. The charm and top quarks turn into similar, but orthogonal, mixtures, which are mostly strange and bottom respectively.

By considering the most general mixing of this kind, Kobayashi and Maskawa discovered that these charge-changing interactions can violate *CP*. Because the mixing is observed to be small, the *CP* violation is a small effect. This Kobayashi-Maskawa *CP* violation vanishes if any two quark flavors with the same charge have the same mass. In reality, no two flavors have the same mass. However, setting aside the top, the other five flavors are very light compared to the typical mass scale of the weak interactions, for example the mass of the *W* (see bottom illustration on the previous page). Compared to the typical weak scale, these five flavors all have nearly the same mass, namely zero mass. Therefore, in any process characterized by the weak scale, the *CP* violation will be tiny because of these small quark masses and also the small quark mixing.

Electroweak baryogenesis, a *CP*-violating, weak-scale process, would thus have been ineffectual. Of course, as mentioned at the outset, the excess of matter over antimatter was only one part per billion at this early epoch. Nevertheless, baryogenesis using Kobayashi-Maskawa *CP* violation falls far short of even this tiny number.

OUTLOOK FOR ELECTROWEAK BARYOGENESIS

Is electroweak baryogenesis a failure then? Actually, the inadequacy of Kobayashi-Maskawa *CP* violation for baryogenesis was recognized immediately. From this we learn that before we can explain the baryon asymmetry, we must first improve our understanding of physical laws, either at the electroweak scale or else at an even deeper level.

From the beginning, work on electroweak baryogenesis has considered generalizations of the standard electroweak theory which include new, nonproblematic sources of *CP* violation. Usually the mechanism of symmetry breaking is modified, which is allowed since so little is known about this mechanism. Several variants of the electroweak theory appear capable of producing the observed baryon asymmetry.

A definitive answer to the mystery of the baryon asymmetry thus awaits the next generation of high-energy experiments, which hope to at last shed light on the far-reaching phenomenon of electroweak-symmetry breaking. The Fermilab Tevatron, the Large Electron Positron (LEP) collider and Large Hadron Collider (LHC) at CERN, and potentially a high-energy electron collider (NLC) will all attempt to probe the symmetry-breaking mechanism directly. Meanwhile *B*-meson factories at SLAC, KEK, and elsewhere will look for the origin of *CP* violation.

As these facilities begin to reveal the foundations of electroweak physics, we'll learn why weak forces are weak, what gives particles mass, and how nature distinguishes matter from antimatter. Our knowledge of history will then reach back a little further, to a time of baryon-number violation and symmetry breaking, when perhaps the baryon asymmetry was forged. At last we might understand why our Universe is made of matter and not antimatter. And we'd know why it isn't empty.

A definitive answer to the mystery of the baryon asymmetry thus awaits the next generation of high-energy experiments, which hope to shed light on the far-reaching phenomenon of electroweak-symmetry breaking.