Of Symmetries and Factories, Matter and Antimatter

CP violation may lie at the root of why there's more matter than antimatter in the universe.

by Alan J. Weinstein

Nestled in the lush countryside near Palo Alto, the Stanford Linear Accelerator Center, or SLAC, has been a rich source of data for high-energy physicists. The large building in the foreground houses the linear collider's particle detector, while the smaller building to its right will accommodate the proposed B factory's detector. The I-280 crosses the linear accelerator, which slingshots electrons and positrons to more than 99.99999 percent of the speed of light.

In 1991, the Department of Energy and the National Science Foundation commissioned a panel to identify the most important goals in high-energy physics and to recommend funding priorities. The committee concluded that one of the few top priorities for the next decade is the "comprehensive study of CP violation in B meson decays at… a B factory." What is CP and what is its violation? What are B mesons? What is a B factory, and why is it so important to particle physics? The observation of CP violation in B meson decay has become a focal point for high-energy physics in general. Even more significantly, CP violation may lie at the root of why there's more matter than antimatter in the universe.

The physics of elementary particles is a vast field whose ambitious goal is a fundamental description of the nature of matter and energy. The field's language and body of theory is the most accurate known to modern science. That language is called quantum field theory, a forced marriage between field theory and quantum mechanics. Field theory was developed in the mid-1800s by Faraday, Maxwell, and others to describe electricity and magnetism, and is now used to describe the distribution of all matter and energy in space and time. Quantum mechanics is a product of the early 20th century—the distilled wisdom of the great physicists of the era—and describes the laws that hold sway at atomic distances, where the classical physics of Newton and his successors breaks down and traditional notions of objective reality are challenged.

The marriage is forced in that both quantum mechanics and field theory stand alone quite happily until we consider high energies, where we're compelled to use both theories. Furthermore, the mathematics seems forced—it's difficult, it's devoid of the beautiful simplicity of theories such as general relativity, and it's plagued with unphysical results called infinities that must be eliminated through an ugly prescription called renormalization. Mind you, quantum field theory and renormalization are profound and important ideas; nonetheless, the marriage that produced them seems forced. What is important here is *inevitability*—physicists are happiest when their theories couldn't possibly turn out any other way. Quantum field theory is inevitable in this sense, but I wouldn't be surprised if, one day, someone writes down a much more satisfying, simple, and elegant mathematical formulation of it.

Quantum field theory also makes use of the theory of special relativity—Einstein's discovery in 1905 of the relationships between space and time, and between matter and energy. And, finally, the symmetry principles around which much of the rest of this article revolves are derived from group theory. The branch of mathematics dealing with the relationships of objects to one another, group theory was founded by the brilliant French mathematician Évariste Galois, who led a colorful life before being killed in a duel in 1832 at age 20.

Quantum field theory is the language of the so-called Standard Model. The Standard Model attempts to provide a full description of the behavior of all the particles (which the model

LEPTONS				QUARKS			
PARTICLE NAME	SYMBOL	REST MASS (MeV)	CHARGE	PARTICLE NAME	SYMBOL	REST MASS (MeV)	CHARGE
Electron neutrino	ν_e	~ 0	0	Up	u	~300	$+{}^2/_3$
Electron	e^-	0.511	−1	Down	d	~300	$-{}^1/_3$
Muon neutrino	ν_μ	~ 0	0	Charm	c	~1,500	$+{}^2/_3$
Muon	μ^-	106	−1	Strange	s	~500	$-{}^1/_3$
Tau neutrino	ν_τ	< 31	0	Top or Truth	t	> 50,000 (not yet discovered)	$+{}^2/_3$
Tau	τ^-	1,777	−1	Bottom or Beauty	b	~ 5,000	$-{}^1/_3$

FORCE	RANGE	RELATIVE STRENGTH	CARRIER	REST MASS (GeV)	CHARGE	STATUS
Gravity	Infinite	10^{-36}	Graviton	0	0	Not yet discovered
Electro-magnetism	Infinite	10^{-2}	Photon	0	0	Observed directly
Weak	$< 10^{-16}$ cm	10^{-13}	W^+	80	+1	Observed directly
			W^-	80	−1	Observed directly
			Z^0	91	0	Observed directly
Strong	$< 10^{-13}$ cm	1	Gluon	0	0	Confined

The standard model in a nutshell. Top: The six known leptons (particles that are not made up of quarks, and are thus "fundamental" in their own right) are mirrored by the six known or hypothesized quarks. Quarks and leptons come in three pairs, or "generations," based on their masses. MeV stands for million electron-volts— 1 electron-volt is 2×10^{-33} grams. Middle: The four fundamental forces of the universe and the particles that carry them. GeV stands for billion electron-volts. Bottom: A bestiary of popular particles.

PARTICLE NAME	QUARK CONTENT	CHARGE	MASS (MeV)
Proton	uud	+1	938
Neutron	udd	0	940
Λ	uds	0	1116
Σ^+	uus	+1	1189
Σ^0	uds	0	1193
Σ^-	dds	−1	1197
Ξ^0	uss	0	1315
Ξ^-	dss	−1	1321
Ω^-	sss	−1	1642
Λ_c^+	udc	+1	2285
π^+, π^-	$u\bar{d}, d\bar{u}$	+1, −1	140
π^0	$(u\bar{u} - d\bar{d})/\sqrt{2}$	0	135
K^+, K^-	$u\bar{s}, s\bar{u}$	+1, −1	494
$K^0, \bar{K}^0$	$d\bar{s}, s\bar{d}$	0, 0	498
D^+, D^-	$c\bar{d}, d\bar{c}$	+1, −1	1869
$D^0, \bar{D}^0$	$c\bar{u}, u\bar{c}$	0, 0	1865
B^+, B^-	$u\bar{b}, b\bar{u}$	+1, −1	5280
$B^0, \bar{B}^0$	$d\bar{b}, b\bar{d}$	0, 0	5280
J/Ψ	$c\bar{c}$	0	3097

describes as fields, in order to provide the mathematical tools needed to create and destroy particles via $E = mc^2$) and forces (also known as interactions) in the universe. The Standard Model doesn't quite do all that, but it does describe, in detail, three of the four known fundamental forces in terms of the particles (or fields) that carry them. The most familiar of these forces is the electromagnetic one, which pervades our daily existence in the form of light waves, TV signals, and so forth, and which is carried by the photon. Less well known is the strong nuclear force, which holds quarks together to form protons and neutrons (and protons and neutrons together to form atomic nuclei), and which is carried by gluons. Even more obscure is the weak nuclear force, which is responsible for certain forms of radioactivity and for the decay of heavy quarks into lighter ones, and which is carried by the W^+, W^-, and Z^0 bosons. Gravity, carried by the hypothetical graviton, is excluded, because it's difficult to quantize. The force and particle content of the Standard Model is summarized in the tables at left.

The Standard Model has been tremendously successful at describing atomic structure, semiconductor behavior, solar physics, radioactive decay, and virtually every other phenomenon to which it has been applied. In some cases, agreement between experiment and theory has been demonstrated to one part in a trillion. In the last 30 years, this theory has even been applied to the Big Bang—the birth of the universe itself. In the early universe, the relationship between matter-energy and space-time was particularly intimate, as predicted by Einstein's general theory of relativity. Cosmologists use the Standard Model to understand the creation and behavior of matter and energy in that earliest epoch. The very existence of matter may depend on CP—charge-parity—violation (see box opposite).

The Standard Model is extremely successful, but it is incomplete. Nagging questions remain about its mathematical consistency, and about the many parameters whose values cannot be derived from the theory but must be taken as they are found in the real world. For example, the theory does not predict what the mass of the electron ought to be, nor the amount of the charge on it. Some fundamental questions remain as well. What is the nature of mass? (We believe it has something to do with a particle called the Higgs boson, the discovery of which was the goal of the Superconducting Supercollider, or SSC, an ambitious project that Congress killed this past summer.) What lies beyond the Standard Model? Those nagging inconsistencies

The names of the quarks, and the term flavor, are arbitrary; they reflect the fact that most physicists don't know Latin.

hint at some more complete theory. Studying the origin and nature of CP violation will help complete the model.

So what does the Standard Model say? First off, it contains a classification scheme for particles. Just as the hundred-odd elements of the periodic table reduce to aggregates composed of only three building blocks (the protons and neutrons in the atomic nucleus, orbited by electrons) the hundred-odd denizens of the particle zoo derive from combinations of a handful of quarks. There are two important exceptions to this rule, however—the electron and its heavier brethren, the muon and the tau, are indivisible, quarkless particles, as are the three kinds of neutrinos associated with them. Quarks come in six known "flavors," called down, up, strange, charm, bottom, and top in order of increasing mass. The "bottom" and "top" quarks are also known as "beauty" and "truth," respectively. The names of the quarks, and the term flavor, are arbitrary; they reflect the fact that most physicists don't know Latin. However, no one knows why there are six different quarks (if in fact that's all there are), and no one knows why they have the masses they do.

But the key to the Standard Model is the mathematics of symmetries, which wrestle the model's welter of complex details into a semblance of order. Thus, for example, the strong nuclear interaction, which binds quarks together into particles, is flavor-symmetric. In other words, the strong interaction doesn't care what flavor the quark happens to be—it treats all six flavors alike.

CP Violation in the Early Universe

It's difficult, without the aid of quantum mechanics, to describe how K^0-$\bar{K}^0$ decays violate CP symmetry, but the consequences are nonetheless dramatic. For if the combination of charge, parity, and time is conserved in all interactions, and if some kind of interaction causes the combined symmetry with respect to charge and parity to be violated, then there must be an equal but opposite violation of time symmetry. Now, just as in the macroscopic world, the film no longer looks the same when run backward. This goes beyond the *statistical* time asymmetry implied by the second law of thermodynamics; it is a *fundamental* property of particle interactions at the microscopic level. Reactions that violate CP symmetry introduce a fundamental arrow of time into quantum physical systems.

Nowhere is this more important than in the early universe. If all the matter and energy in the universe was indeed created out of the gravitational potential energy of the Big Bang, then symmetry with respect to charge conjugation would have ensured that equal amounts of matter and antimatter would have been produced. Clearly, this is no longer the case—the universe we observe is composed virtually entirely of matter. It turns out that such an asymmetry cannot be produced by charge-conjugation symmetry violation alone; one needs CP, or equivalently, time symmetry violation, to evolve such an asymmetry out of an initially symmetric cosmos.

Andrey Sakharov established in 1967 that three ingredients are required in order for the universe to evolve a net surplus of protons versus antiprotons: CP violation; a stage in the evolution of the universe that was far from equilibrium; and proton decay. Experiments on neutral kaons have, in fact, established that at some small level, CP is indeed violated. The rapid expansion of the universe in its first moments postulated by current cosmological theories kept it far from equilibrium. And modern grand unified theories all predict that protons do indeed decay—diamonds are not forever! In fact, proton decay has not yet been observed, and sensitive experiments put the proton's lifetime in excess of 10^{32} years. (By contrast, the universe is roughly 10^{10} years old.) But even longer lifetimes, if finite, suffice to produce the matter-antimatter asymmetry observed in the universe. Cosmology suggests that this asymmetry was established some 10^{-35} seconds after the Big Bang.

As any lefty knows, the human world is far from being parity-symmetric; a parity-inverted pair of scissors is a welcome possession.

Likewise, symmetries inform the Standard Model's description of the fundamental forces. For example, electrically charged particles participate in the electromagnetic interaction. Without going into detail, electromagnetic charge can be thought of as a pointer arrow buried within each particle. No matter how an interaction rotates, or transforms, the pointer, the symmetry laws ensure that the particles' behavior remains unchanged as far as the outside world is concerned. This symmetry constraint is sufficient to derive all the properties of the electromagnetic force, including Maxwell's equations. The two other forces described by the Standard Model, the strong and weak nuclear forces, can be similarly described. It may even be possible to unite all three forces into a single grand unified theory at energies corresponding to 10^{20} K, where higher symmetries (again, with respect to abstract "pointers" within these subatomic particles) would become manifest, and the three forces would become indistinguishable from one another. Nothing in the universe has been that hot since 10^{-34} seconds after the Big Bang, when the observable universe was about a meter in diameter.

These symmetries are continuous because the physical quantities they describe vary smoothly—in shades of gray, as it were. For example, the strength of an electromagnetic field can be plotted as a smooth, continuous curve. And for every continuous symmetry, there is a conserved quantity—one whose net value doesn't change, although its distribution among the particles within the system may vary. Thus, the symmetry that gives rise to electromagnetism is associated with conservation of electric charge. The strong nuclear force also has a conserved "charge"—three of them, actually—known fancifully as color. And the weak nuclear force conserves something called hypercharge. A particle carrying a charge associated with one of these forces interacts with other such particles via that force.

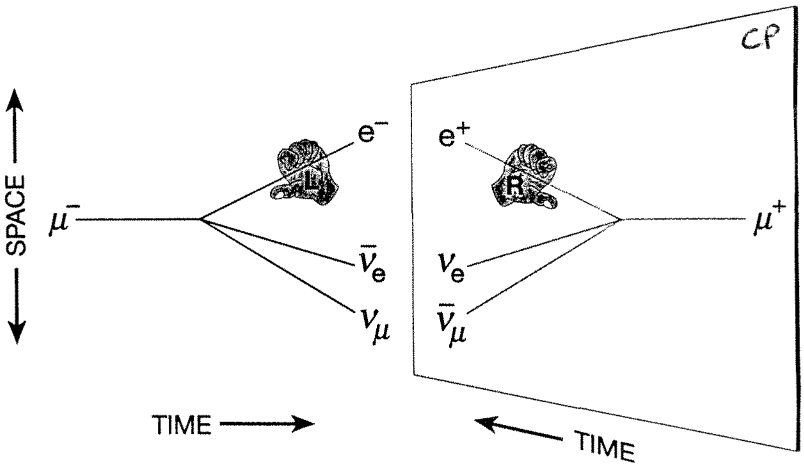
There are other, equally important discrete symmetries that describe discontinuous, black-or-white quantities—an electric charge is either positive or negative, for example. Chief among the discrete symmetries are the following three: Parity inversion is the inversion of all three spatial dimensions through some arbitrary point as if by a mirror, making right-handed systems appear left-handed and vice versa. Time reversal is the inversion of the temporal dimension, as when one runs a film backward. And charge conjugation is a change in the sign of the charges of all the particles in the system so that each particle becomes its antiparticle—for example, electrons become anti-electrons, called positrons, and vice versa.

These symmetries are respected at the subatomic level. Biological systems aside, the rates of chemical reactions rarely depend on whether a molecule is left-handed or right-handed. (Almost all of the molecules essential to life are left-handed. This is an example of spontaneous symmetry breaking—once one form of matter gains an advantage over another, even if by accident or happenchance, that dominance is maintained thereafter.) And a movie of two atoms colliding to form a molecule, run backward, is indistinguishable from a movie of that same molecule splitting up into the two atoms. Furthermore, when subatomic collisions convert energy to matter (via Einstein's $E = mc^2$), they produce exactly as many positrons as electrons.

However, these symmetries aren't perfect. The different quark flavors have different masses and electric charges, and thus they behave differently—in the real world, the flavor symmetry is broken. The electromagnetic, strong, and weak interactions are vastly dissimilar in our experience. Furthermore, as any lefty knows, the human world is far from being parity-symmetric; a parity-inverted pair of scissors is a welcome possession. And as anyone who has ever spilled milk knows, the world looks funny when the film runs backward, so time-reversal symmetry is broken. It's equally obvious that charge conjugation is a poor symmetry in nature, or positrons would be as abundant as electrons, anticarbon as common as carbon.

The lack of symmetry we see with respect to parity, time, and charge is the result of spontaneous symmetry breaking and the second law of thermodynamics. But at the fundamental level, these symmetries are conserved. In fact, in any universe in which one simultaneously performs the charge, parity, and time transformations, the laws of physics should be unchanged. This is known as the CPT theorem, and it is built into the mathematics of quantum field theory. It is difficult to even conceive of a universe in which the combination of charge, parity, and time transformations would yield different physical laws.

However, it *is* possible to imagine a universe in which one or more of these three symmetries is violated, as long as an equal but opposite violation occurs in one or both of the other two. Indeed, this is the case in our own universe. It was discovered in 1957 that the weak nuclear interaction, in which the ghostly particle known as the neutrino plays a crucial role, holds commerce only with left-handed neutrinos and right-handed antineutrinos. (The neutrino comes in both left-

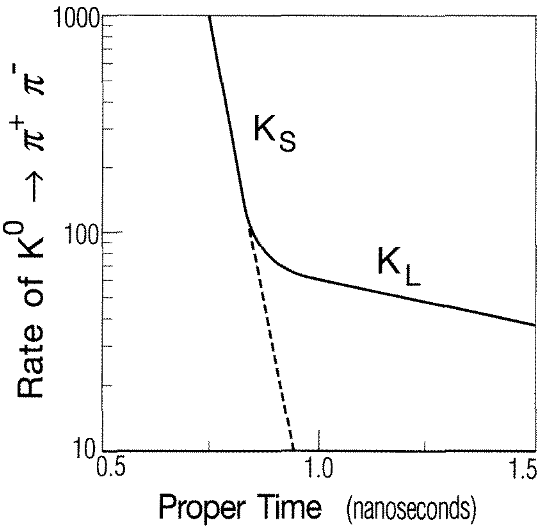


In this example of charge-parity conservation, a negatively charged muon (μ^-) decays into a left-handed electron (e^-), an electron antineutrino ($\bar{\nu}_e$), and a muon neutrino (ν_μ). The reaction looks just the same in the CP mirror, except that the particles' signs are reversed, and the positron (e^+), or antielectron, is now right-handed. An electron's handedness can be measured by bouncing it off a target of known spin polarization, and observing the scattering angle. (It's hard to measure the handedness of neutrinos, as they don't interact much with matter.)

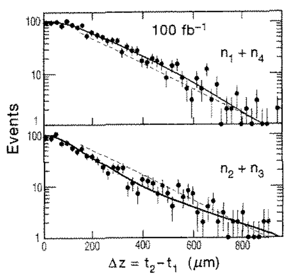
and right-handed, as well as particle and antiparticle, incarnations.) It appears that the weak interaction, unique among the known fundamental forces, violates parity conservation and charge conservation; but it was believed that, at least, the *product* of parity and charge was conserved. Thus, reactions would proceed in an identical manner to their mirror reflections, if all the reacting particles were replaced by their antiparticles. This is illustrated in the drawing (above) of muon decay, in which the handedness of the electron so produced is measured.

A universe in which both parity and charge have opposite senses with respect to ours would be indistinguishable from ours. Or would it? The weak interaction violates charge and parity individually, while conserving the combination—called CP for short—to a very high degree. But in 1964, while Christensen, Cronin, Fitch, and Turlay were studying the decays of the K meson, they observed that CP was violated in one decay in 500. If CP were strictly conserved, the decay rate for kaons, as K mesons are also called, as a function of time would be exactly equal to that for antikaons. In fact, there was a slight asymmetry in the decay rates. Not only was this very subtle phenomenon difficult to observe experimentally, it was also difficult to interpret theoretically. The explanation that eventually emerged runs as follows. The particles studied by Christensen et al. were neutral kaons—the K^0, which is composed of a down quark and a strange antiquark; and the K^0's antiparticle, the $\bar{K}^0$, which is composed of a down antiquark and a strange quark. But, says the theory, the particles coursing through the experiment's beam pipes and detectors were somehow neither and both: they were quantum-mechanical superpositions of K^0 and $\bar{K}^0$ mesons. This is a classic example of mixing, a bizarre phenomenon unique to quantum mechanics. As a final complication, the experiments revealed that the superposition itself exists in two states, called K_S and K_L because the former's lifetime is about 100 times shorter than the latter. Relatively speaking, of course—K_L lives to a ripe old age of about 50 billionths of a second.

If one applies the CP transformation to the K_S and K_L states—changing the quarks to antiquarks and vice versa (charge) and swapping the two quarks' positions within the particle, so that, as it were, the quark on the left is now on the right (parity)—something strange happens. The mathematical representation of the K_S, which is CP-even, retains its original sign—while the K_L, which is CP-odd, changes its sign. But things get stranger still. Experiments have revealed that the mixing pattern is only approximate. At some small level, the K_L has a CP-even piece, and the K_S has an equal but opposite CP-odd piece. One way in which this shows up is that the K_S, being primarily CP-even, decays into an even number of particles—in this case, a positively charged pion and a negatively charged antipion. The CP-odd K_L decays into an odd number of particles, with the third one being an uncharged pion, which is its own antiparticle. In the absence of CP violation, a plot of pion-antipion pairs produced versus time should go to zero very quickly as the short-lived K_S's decay. Seeing a K_L



Right: A plot of K^0s decaying to $\pi^+\pi^-$ pairs over time. If all the $\pi^+\pi^-$ pairs were produced by K_S decays, the rate would quickly fall to zero, as indicated by the dashed line. But instead, a small percentage of K_L mesons decay the "wrong" way, contributing a trickle of $\pi^+\pi^-$ pairs even after the K_S mesons are all gone. ("Proper time" is time corrected for the velocity at which the particles are traveling. At the near-lightspeed clip these guys are going, time slows down significantly.)



Below: How the data would actually appear in a particle detector. If there were no CP violation, the number of detection events versus difference in decay lengths (Δz) in both plots would be identical, as shown by the dashed lines. Instead, there is a slight excess of one decay, matching a slight deficit in the other.

go into only two pions indicates that it has a CP-even piece to it. Instead of dropping all the way to zero, a trickle of pion-antipion pairs remains, produced by K_L's that violate CP symmetry and decay the wrong way. This is what Christensen and his colleagues observed. Which brings us back to the question of a CP-inverted universe...

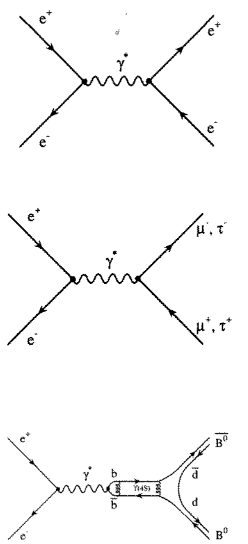
Suppose there existed, in some remote part of the universe, a realm peopled with intelligent beings made of antimatter—high-energy antiphysicists. If we could communicate with such beings, how could we determine whether they were indeed composed of antistuff? Touching them would suffice; if our matter and their antimatter annihilated ourselves and them, we could be sure that they were of a different breed. Short of that, we could imagine communicating via the exchange of light signals, i.e., photons. But photons are their own antiparticle, and would behave the same in an antiworld as they do in ours. And how could we agree, unambiguously, on what "left-handed" and "right-handed" mean? We could tell them that the electrons emitted from the radioactive decay of, say, cobalt-60 were left-handed; but this would be insufficient information unless we could also uniquely define which is "matter" and which is "antimatter" or, equivalently, a "negative" or "positive" charge. CP violation provides an unambiguous answer. Antiphysicists could perform experiments with K_L and K_S mesons. A negative charge is then defined as that of the electron associated with the slightly less abundant decay mode of the K_L. The spin of those electrons is defined as left-handed, and matter is composed of nuclei with a positive charge, while antimatter has negatively charged nuclei. Which are you?

If matter-antimatter symmetry were respected in the universe, sooner or later the matter and antimatter would find and annihilate each other. Stable chunks such as stars, planets, and people would not be around for long. In some sense, then, we have CP violation to thank for our enduring presence, and it behooves us to ask where it comes from. Except for the weak nuclear force, the fundamental forces of nature appear to respect CP symmetry. Is there something inherent in the weak interaction that permits such violations?

There does exist a mechanism in the Standard Model that can produce CP violations; it was proposed by Kobayashi and Maskawa in 1973. At the time, only three quarks (up, down, and strange) were known. Kobayashi and Maskawa anticipated the discovery of three more quarks, grouped by their masses into three pairs, or generations. The up and down quarks form the first generation; the then-undiscovered charm quark pairs with the strange quark in the second generation; and the still-undiscovered top quark rounds off the third generation, currently occupied only by the bottom quark. The quarks in the second and third generations have proved to be unstable. The second and third generation's heavier quarks (charm and top) decay primarily to the lighter particle in that same generation, via the emission of a W^+ boson (W^- for antiquarks). The subsequent decay of the lighter quark to a lower-generation particle occurs because the down-type quarks (down, strange, and bottom) quantum-mechanically mix with one another. This happens at a much slower rate than the decay from the heavy quark to the light quark within a generation, which is an unmixed reaction. There is a wave function that describes the mixed reaction, and another one that describes the unmixed reaction, and it is the complex phase difference between the two wave functions that produces the interference pattern characteristic of CP violation. Kobayashi and Maskawa knew that the down and strange quarks mixed, and boldly suggested that a third generation of quarks—and thus a third down-type quark that could mix with the down and the strange quark—existed, three years before there was any experimental evidence for that third family. (The mathematics of the Standard Model won't produce CP violation with only two generations of quarks—it takes three.)

To really test this idea, we have to perform experiments on the third generation of quarks. And it is the bottom quark, being the one that quantum-mechanically mixes in order to decay,

Feynman diagrams of three things that can happen when an electron (e^-) and a positron (e^+) collide. Top: They annihilate each other, forming a virtual photon (γ^*) that relapses into an electron-positron pair. Middle: The photon transmutes into either a muon-antimuon ($\mu^\pm$) or tau-antitau ($\tau^\pm$) pair. Bottom: The photon becomes a bottom quark (b)-bottom antiquark ($\bar{b}$) pair, bound together into "bottomonium" by the exchange of gluons—rendered here by coiled squiggles resembling telephone cords. The bottomonium acquires a paired down quark (d) and antiquark ($\bar{d}$) from the vacuum, forming a pair of B mesons (B^0 and $\bar{B}^0$).



that is the key to the puzzle. This is fortunate, since the top quark is so very massive that it hasn't yet been discovered! ($E = mc^2$; and the biggest particle-accelerator yet built—the Tevatron, at Fermilab near Chicago—is just at the threshold of generating enough E to make the top quark's m.) But a bottom quark can't be studied in isolation, because the strong nuclear force, which binds quarks into particles, is *so* strong that only combinations of quarks having no net color can be isolated. So we have to study B mesons, which are the commonest colorless combinations containing bottom quarks, instead. A bottom quark combined with an up antiquark makes a B^- meson. Other possibilities include its antiparticle, the B^+, formed from a bottom antiquark and an up quark. And a bottom antiquark and a down quark form a B^0. Its antiparticle is the $\bar{B}^0$, formed from a bottom quark and a down antiquark. B mesons can be produced in a high-energy collision, but they're rare because they are so massive. (Like car crashes, particle collisions tend to produce a very few massive chunks and a lot of small, lightweight debris.) The B mesons' large mass also makes it possible for them to decay to lighter particles in many different ways, and with a very short lifetime—about one trillionth of a second. The experimental study of B decays is no easy matter.

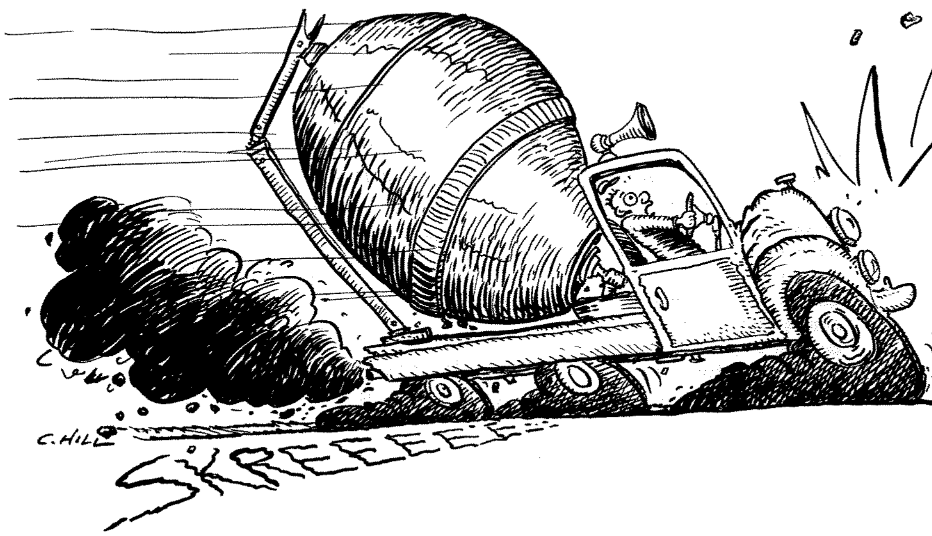
Most of what we know about B mesons comes from electron-positron colliders. The Cornell Electron Storage Ring (CESR) at Cornell University in New York, and the DORIS ring at the DESY Laboratory in Hamburg, Germany, are circular rings of magnets that store counter-rotating beams of electrons and positrons moving at 99.99999953 percent of the speed of light, and bring them into collision in the center of large and complex particle detectors. Each beam runs at about 5 billion electron volts, or 5 GeV, of energy. (An electron volt is the amount of energy it takes to move an electron across a one-volt electrical potential. One electron volt is not an awful lot of juice—by comparison, it takes 13.6 eV to ionize a hydrogen atom.)

Lots of different things can happen when electrons and positrons collide. Most often, they just bounce off each other. Less frequently, the electron and positron can annihilate into a virtual photon (the carrier of the electromagnetic force), or a Z^0 (the electrically neutral carrier of the weak nuclear force). (A virtual particle is one that fleetingly materializes from the quantum-mechanical vacuum of space. Virtual particles live on energy "borrowed" from the vacuum, not unlike the junk-bond kings of the 1980s, who lived on borrowed money from paper profits that weren't really there either.) Almost immediately thereafter, the photon or Z^0 will transmute into a particle-antiparticle pair. Any charged particle-antiparticle pair can be produced through a virtual photon, as long as the beam energies are sufficient to produce the mass of the two particles. And any particle-antiparticle pair that participates in the weak nuclear force—and all known particles do—can be produced through a Z^0, even the electrically neutral neutrinos. The Feynman diagrams associated with some of these reactions are shown at left. The cross section, or probability, for annihilation into quarks depends on the center-of-mass energy, which, in a symmetric collider, is twice the beam energy.

Virtual particles live on energy "borrowed" from the vacuum, not unlike the junk-bond kings of the 1980s, who lived on borrowed money from paper profits that weren't really there either.

If the energies of each beam are tuned to precisely 5.29 GeV, then the production of bottom quark–bottom antiquark pairs can just barely proceed, with almost no energy left over to kick the members of the pair thus produced in opposite directions. At this beam energy, the pair binds together into "bottomonium." This state decays, again almost immediately, into a pair of B mesons: B^+-B^- or B^0-$\bar{B}^0$. The B mesons lumber off slowly, at around 6 percent of the speed of light. They don't last long enough to make it to the detector themselves, but they decay into characteristic showers of less massive particles that do. By tracing the trajectories of these secondary particles back to their sources, the identities and flight paths of the original B mesons can be deduced. These lighter particles that actually register in the detector include photons, as well as electrons, muons, pions, kaons, protons, neutrons, and their antiparticles.

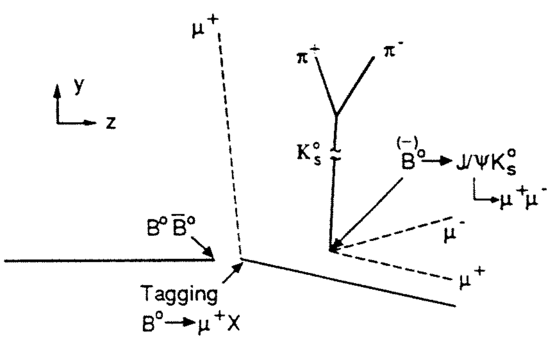
The J/Ψ particle has a double name because it was simultaneously discovered by two groups, neither of which was willing to relinquish the right to christen it.



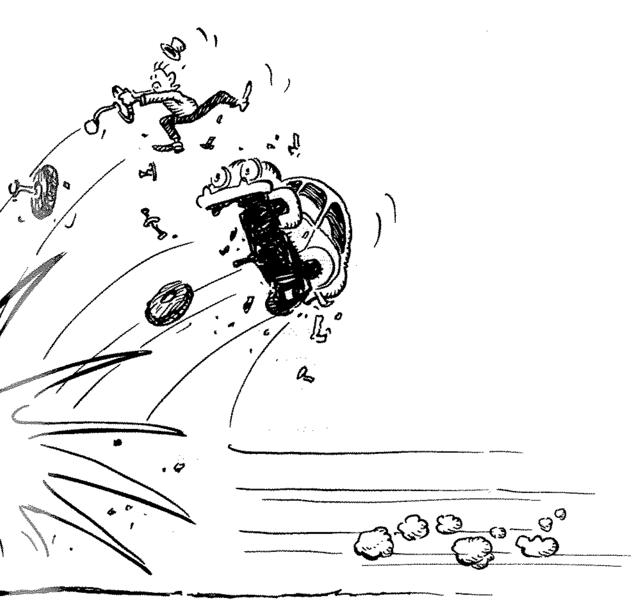
Neutrinos and antineutrinos typically sail through the detector unseen.

What makes the observation of CP violation possible here is the fact that B^0's and $\bar{B}^0$'s mix, just as K^0's and $\bar{K}^0$'s do. (This phenomenon was first observed by the ARGUS experiment at DORIS.) As the B mesons fly apart from each other, each one oscillates between its B^0 and $\bar{B}^0$ identities. A given B meson has a certain probability of mixing before it decays and a certain probability of decaying within a finite time. Specifically, a B meson has a mean life of 1.5 trillionths of a second, or 1.5 picoseconds. In that time, 63 percent of all B mesons will decay. If a B meson decays in exactly 1.5 picoseconds, it has a roughly 70 percent chance of having mixed first. In other words, if you have a sample of 100 B mesons, all of which decayed in exactly 1.5 picoseconds, about 70 of them will have mixed first. As in the kaon system, the interference between the two wave functions describing these two outcomes produces the observable CP-violating effect.

But the effect is very small. To look for such a small effect, one can try to measure a tiny asymmetry in common decays. However, a better approach is to look for a large asymmetry in certain very rare decays. Most promising are decays into systems of particles that are CP-eigenstates, that is, in which the CP product is conserved—either the particles formed by the decay are their own antiparticles, or the decay forms particle-antiparticle pairs. Either way, the final state is accessible to both B^0 and $\bar{B}^0$. One particularly promising decay is the formation of



How to tag a B^0: The bottomonium cruises along the solid line at lower left, decaying to a B^0-$\bar{B}^0$ pair at the arrowed end of the line. Shortly thereafter, the B^0 (the tag) shows up in the detector as a muon (dashed line labeled μ^+) and some other particles, shown here as a solid line. The $\bar{B}^0$ flies a bit farther before decaying into a K_S, which reaches the detector as a pion-antipion pair (π^+ and π^-), and a J/Ψ, which is detected as a muon-antimuon pair (μ^-, μ^+, dashed lines).



If a cement truck and a Volkswagen go head-to-head, the mangled metal will fly in the direction the cement truck was going.

a J/Ψ-K_S pair. (The J/Ψ particle has a double name because it was simultaneously discovered by two groups, neither of which was willing to relinquish the right to christen it.) The J/Ψ is a charm-anticharm bound state, and thus its own antiparticle; it shows up in the detector through its decay to an electron-positron or muon-antimuon pair. The K_S is not its own antiparticle, but it is a CP eigenstate; it is detected in its decay to a pion-antipion pair. Another promising decay, that of a B^0 or a $\bar{B}^0$ to a pion-antipion pair, can be observed directly. Under certain conditions, the CP asymmetry in these decays can be as large as many tens of percent. Unfortunately, only about one in 10,000 B mesons decays in the J/Ψ-K_S or pion-antipion mode.

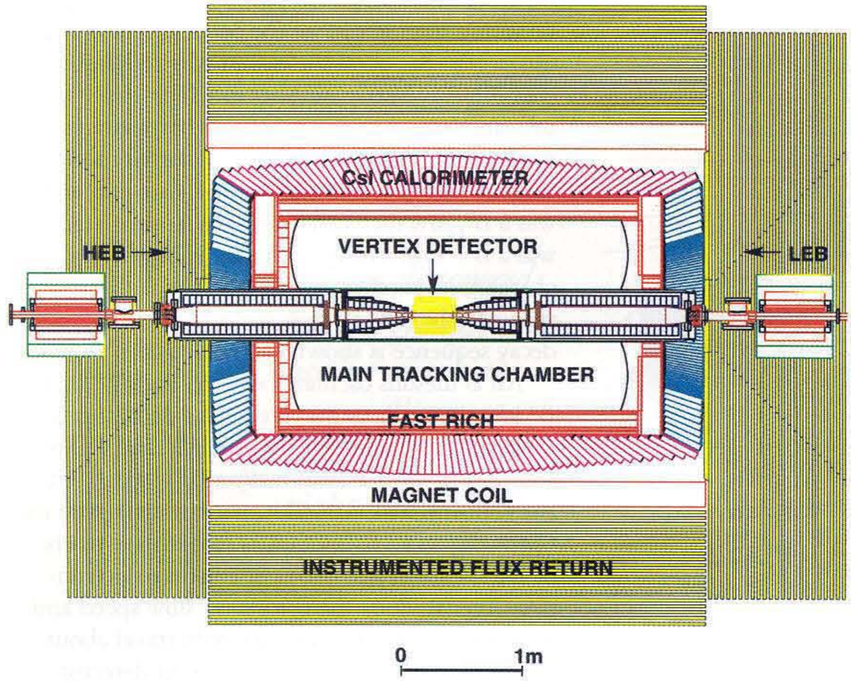
Therefore, one needs to produce many, many pairs of B mesons to see these decays. The highest luminosity—a measure of collision frequency—electron-positron collider in the world is the CESR machine at Cornell. On its best days, it produces one pair of Bs every five seconds. Taking into account the fact that not every pair produced is observed, the CLEO detector at CESR has still recorded nearly one million pairs over the last two years. But even with this unprecedentedly large amount of data, 10 to 100 times more data would be required to perform statistical analyses on these rare decays. Therefore, a B *factory* is needed.

Granted sufficient luminosity, observing CP violation remains a tricky matter. The simplest method is to observe the ratio of, say, B^0 to J/Ψ-K_S decays and $\bar{B}^0$ to J/Ψ-K_S decays. The first step in this process is to determine whether the *other* member of the B meson pair is a B^0 or a $\bar{B}^0$, thus revealing the identity of the particle that decayed into a J/Ψ-K_S. And the easiest way to identify that other B is to use events in which it decays into an electron or muon (plus other, perhaps unobserved, particles). This electron or muon, called a tag, is positively charged if its parent B was a B^0, in which case the B we're interested in was a $\bar{B}^0$. Conversely, if the tag is negatively charged, its parent was a $\bar{B}^0$ and the B that spawned the J/Ψ-K_S pair was a B^0. Such a decay sequence is shown on the opposite page.

All B mesons oscillate between their B^0 and $\bar{B}^0$ identity at the same rate before decaying, so in order to tell whether the particle that we are seeing decay as a B^0 was *really* a B^0 when it left the collision that formed it, we must measure its flight distance. (This applies to both members of the B meson pair.) But at existing electron-positron colliders, the B mesons' slow speed and short lifetime mean that they only travel about 25 microns before decaying. Present detector technology is simply not precise enough to measure such short decay lengths. And running the collider at higher beam energies, so that the outgoing B's are moving faster, is impractical because the production rate falls rapidly. Measuring both decay lengths is essential, however, so a new idea was clearly needed.

And a new idea was born, inspired by the asymmetry that was the goal of the experiment. An electron-positron collider with beams of equal energy will produce the bound quark pair called bottomonium at rest with respect to the detector. But in 1987, Pierre Oddone of Lawrence Berkeley Laboratory proposed an *asymmetric* B factory. If two loaded cement trucks run into each other head-on, the wreckage will remain where they collided. But if a cement truck and a Volkswagen go head-to-head, the mangled metal will fly in the direction the cement truck was going. Similarly, Oddone realized that by giving one of the collider's beams more energy than the other while keeping the total collisional energy equivalent to bottomonium's mass, the collision could produce a bottomonium pair moving in the direction of the higher-energy beam. When the bottomonium decays into a B meson pair, the B's will be moving, too, and at a significant fraction of the speed of light. Thus a B flies a lot farther in its fleeting lifetime. (Its lifetime gets longer too, due to the relativistic effect of time dilation as a particle's speed approaches that of light!) Sufficiently asymmetric collisions will produce decay lengths of 200 microns or more, lengths measurable with today's detector technology.

A particle detector works a bit like an eyeball.

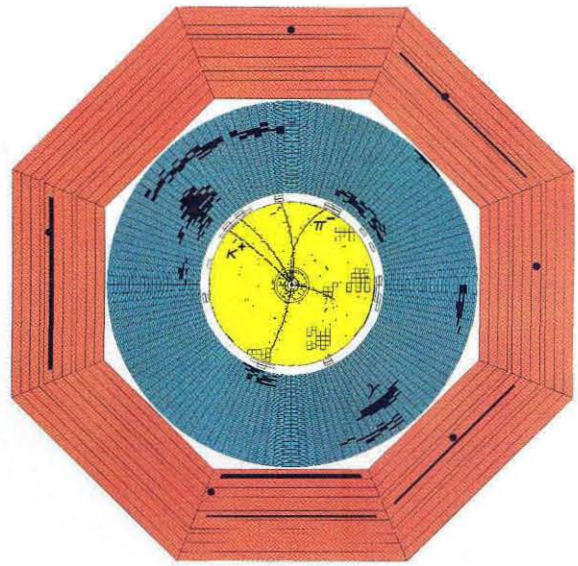


A cross section of an asymmetric B factory's particle detector. The high-energy beam, (HEB) enters from the left; the low-energy beam (LEB) from the right. The detector is asymmetric, too—its center is to the LEB's side of the collision point, as the debris will fly in that direction. The vertex detector is the small yellow square in the center. The nozzles to either side are magnets that cause the beams to collide. Drift chambers fill the "Main Tracking Chamber." The red rectangle called "Fast RICH" is a Ring Imaging Cherenkov counter, a particle-velocity detector. The cesium iodide crystals in the calorimeter are drawn as thin red and blue slabs—the blue ones form the cylindrical detector's circular end caps. The tracking system is embedded in a superconducting magnet. The "Instrumented Flux Return" contains iron sheets (yellow) and muon-detecting drift chambers (white).

We see something by measuring the direction, color, and intensity of light from a scene, from which our brain constructs an image. Analogously, a particle detector "sees" a collision by measuring the angles, energies, and number of particles coming from it, from which a computer constructs an image. However, a particle detector is considerably more sensitive than an eyeball. The detector can sense a single particle, whereas only the best-trained eyeballs can detect a single photon. And the detector measures a great range of energy carried by particles of all sorts, not just photons of visible light. Most importantly, the detector can distinguish different kinds of particles, instead of merely registering photons. Of course all this sensitivity and precision comes at a price—a particle detector is considerably larger and more massive than an eyeball. A detector suitable for a B factory would be about the size of a two-story house, would weigh some 500 tons, and would look something like the illustration at left. The detector must be capable of recording data at a tremendous rate; it might have 100,000 channels of electronic readouts, and be sensitive to collisions occurring every few nanoseconds.

The detector is actually a concentric set of cylindrical detectors, each of which measures different things, nested around the electron-positron collision point like the bun around a hot dog. The innermost layer is the vertex detector—in this case an array of silicon-wafer diodes that register the passage of charged particles. The diodes record the point at which the charged particle passes through them with an accuracy of better than 10 microns, or 10 millionths of a meter. The diodes are placed just outside the two-centimeter-radius beam pipe at the collision point, because their job is to distinguish between the trajectories of particles resulting directly from the collision and the trajectories of secondary particles whose progenitors—including B mesons—may have traveled a quarter of a millimeter before decaying. The vertex detector is surrounded by many layers of drift chambers, which are chambers filled with inert gas and containing high-voltage wires strung in arrays parallel to the cylinder's axis. A charged particle passing through a chamber will interact with a number of gas atoms en route, and will knock an electron loose from some of them. These electrons drift to the wires, registering as a pulse of electrical charge. The time it takes the electron to drift to the wire is proportional to the distance from the wire to the point where the particle hit the atom. By measuring the drift time, that point can be located to an accuracy of 150 microns or better. By collecting many such measurements, the

Right: An end-on view of how a B^0 decay registered at the CLEO detector at Cornell. The B^0 decayed into a K^{*0} and a photon. The K^{*0} then decayed into a K^+ and a pion (π), as labeled. (The unlabeled tracks belong to the decay products of the other B^0.) The drift chamber is tinted yellow; dots along each particle's path show where a wire was triggered. The squares in the same region show calorimeter crystals in the end cap that lit up. The blue region displays the crystals in the calorimeter barrel, but here the lit crystals are black. The flat rectangles in the intervening white ring are lit up elements of the time-of-flight detector. The photon (γ) hit the large black region at five o'clock. The red rectangles are muon detectors; all the black registrations, however, are spurious.



Below: Heisenberg Fellow Gerald Eigen holds a crystal of thallium-doped cesium iodide, scintillating here under an ultraviolet light. This crystal, the size of the ones to be used in the B factory calorimeter, is about 34 centimeters long and weighs some five kilograms.

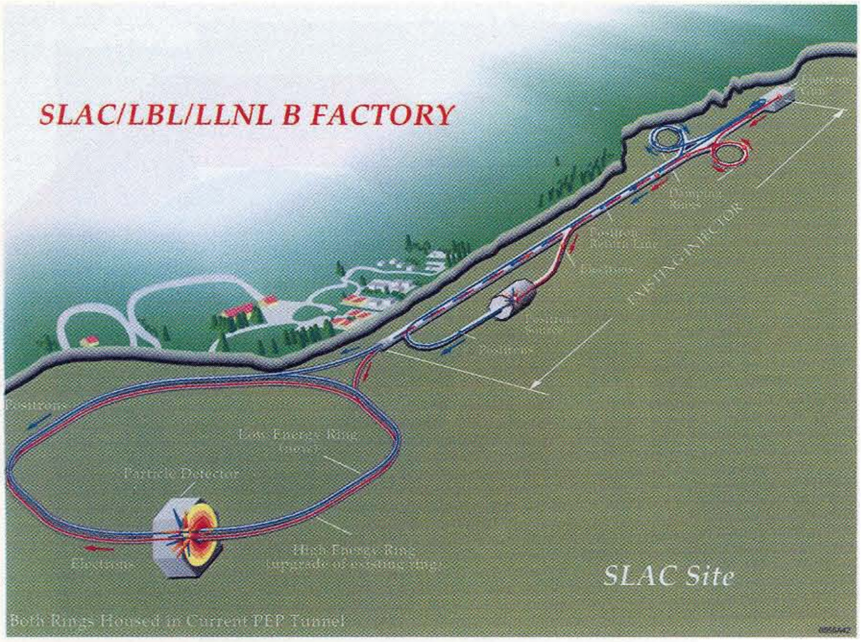
particle's trajectory can be reconstructed.

The next layer out is some kind of charged-particle identification system to distinguish pions from kaons from protons by measuring their velocities. (The tracking system has already measured their momentums; dividing a particle's momentum by its velocity gives its mass, which identifies the particle.) Then comes a precision crystal electromagnetic calorimeter to measure the energies and angles of photons and electrons. The calorimeter is an array of clear crystals, probably of cesium iodide, segmented into towers aimed at the collision point. An electron, positron, or photon entering a tower interacts with the crystal's massive atoms, creating an "electromagnetic shower" of countless electrons, positrons, and photons. The shower causes the crystal to scintillate—to absorb the energy and reemit it as light. A photodiode at the crystal's far end converts the scintillation light into an electrical pulse proportional to the energy of the incident particle. Energy measurement is like temperature measurement, so the device is called a calorimeter. These calorimeters aren't cheap; the crystals can cost as much as $20 million.

The entire apparatus is surrounded by a cylindrical magnet that generates a roughly one-tesla magnetic field. (Earth's magnetic field at the planet's surface is about 0.00005 tesla.) The field is solenoidal, which is to say that within the magnet it is everywhere parallel to the cylinder's axis. As the charged particles emanate from the collision point, the field bends their trajectories into helical arcs, and it is these arcs that the detectors see. (See drawing above.) The arcs contain a wealth of information about the particles tracing them. High-momentum particles leave nearly straight tracks, while low-momentum particle tracks curl up on themselves. Positively charged particles curve in one direction, and negatively charged ones veer in the other. Uncharged particles, of course, fly straight on through. And finally, outside the magnet, some 100 tons of iron return the magnetic flux and absorb any hadrons—particles composed of quarks—that have made it out this far. The iron allows muons, which contain no quarks, to escape into another set of drift chambers that identifies them.

Thus the key ingredients of the Asymmetric B Factory are an extremely high-luminosity electron-positron collider tuned to the production of bottomonium, asymmetric energy beams, and a high-precision detector. The B factory should produce three B^0-$\bar{B}^0$ pairs per second, or, assuming a normal operating schedule, 30 million pairs per year. Nothing like it has ever been built before, but teams of physicists around the world have spent the last five years convincing themselves that it can be done. They have now convinced the Clinton administration, too—the Department of Energy announced on October 5, 1993, that a B factory will be built at the Stanford Linear Accelerator Center (SLAC). The SLAC machine is illustrated on the next page. Similar proposals are being considered in Europe, Russia, and Japan.

A Caltech group led by Barry Barish, Linde Professor of Physics, and me has been exploring the physics of B mesons with Cornell's CLEO



A cutaway view of the B factory. A new ring for low-energy positrons will be added to the existing electron ring, which will be souped up. The electron beam comes from an electron gun—basically a bigger version of the cathode-ray gun in your TV set—at the drawing's upper right. A damping ring then squeezes the beam until it is thinner than a pencil lead, making it easier to inject. Next, the beam rockets down the linear accelerator to reach injection speed. Most of the beam gets injected clockwise into the high-energy electron storage ring. However, some electrons get shunted into a side tunnel, where they slam into a tungsten plate. The tungsten atoms emit positrons, which get piped back to the beginning of the linear accelerator. There they hang a tight U-turn, come back through a damping ring of their own, and shoot down the accelerator toward a counterclockwise injection in the low-energy ring.

detector since 1990. And Professor of Physics David Hitlin and Associate Professor of Physics Frank Porter have led the effort to build an asymmetric B factory at SLAC. Hitlin heads the detector-design team, and Porter is helping design the accelerator itself and the computing systems needed to analyze the large volume of data.

(It may also be possible to study CP violation at proton-proton colliders such as the Tevatron. There, B mesons are produced far more copiously, but it's considerably more difficult to reconstruct them from their decay products, because proton-proton colliders produce more of everything else, too. It's easy for the B decay products to literally get lost in the background.)

At the B factory, physicists would measure the B meson's decay rates into various CP eigenstates as a function of the decay length. If, indeed, Kobayashi and Maskawa's suggested mechanism is responsible for CP violation in the K and B systems, and if there are exactly three generations of quarks, then the measured asymmetries will satisfy certain mathematical relations. If all the factory's data fall in the expected range, then we will have performed a crucial test of the self-consistency of the Standard Model. If they do not, this would represent the first failure of the Standard Model, and we will have a "smoking gun" for physics beyond it. The observed deviations from the Standard Model's predictions will provide valuable clues as to the nature of the new physics. This would be even more noteworthy than observing CP violation itself.

If one ignores or fails to measure the decay flight distances, as in the case of a collider with beams of equal energy, the CP-violating asymmetry vanishes—you don't know which particle you started with, and therefore which set of decay products to expect. This is why you need an asymmetric collider to study CP violations. However, the asymmetry doesn't vanish, even at a symmetric collider, if charge, parity, and time are *all* violated. As mentioned above, CPT invariance is built into quantum field theories, and it is difficult to even conceive of a theory in which CPT is violated. Nevertheless, it should be possible to detect CPT violations in B meson decays at the percent level, should they exist, at a B factory. Such a discovery would revolutionize particle physics!

The B factory is ambitious and costly, but doable. It may be the best, perhaps even the only chance we have of observing CP violation at the fundamental level, outside of the kaon system. Either it will confirm Kobayashi and Maskawa's mechanism for producing CP violation—which is far from answering the question of why CP violation exists—or it will provide crucial clues to physics beyond the Standard Model. Either way, it is destined to be a landmark experiment. Such a machine will also permit physicists to perform very precise studies of the decays of other heavy particles as a spin-off from the ultimate goal of observing CP violation.

CP violation in the B system is one of the high frontiers of particle physics. Exploring this far-removed-from-everyday-life phenomenon has become a focus for hundreds of particle physicists around the world, who are eager to devote a large part of their careers and research dollars to it. This obscure effect is something of a keystone in the elaborate structure of particle physics, an edifice that encompasses everything we think we know about the fundamental nature of all forms of matter and energy. B uried in this phenomenon somewhere is a profound insight into the nature of space-time and matter-energy, and into the very basis of our existence. □

Assistant Professor of Physics Alan J. Weinstein earned his AB from Harvard in 1978 and his PhD in 1983, both in physics. He then joined the Institute for Particle Physics at the University of California, Santa Cruz, and spent most of his time at SLAC, where he worked on the linear collider, among other things. Weinstein came to Caltech in 1988.