

STANFORD LINEAR ACCELERATOR CENTER

Spring/Summer 1996, Vol. 26, No. 1

# Beam Line



# Beam Line

A PERIODICAL OF PARTICLE PHYSICS

SPRING/SUMMER 1996 VOL. 26, NUMBER 1

## Editors

RENE DONALDSON, BILL KIRK

## Contributing Editor

MICHAEL RIORDAN

## Editorial Advisory Board

JAMES BJORKEN, GEORGE BROWN

ROBERT N. CAHN, DAVID HITLIN

JOEL PRIMACK, NATALIE ROE

ROBERT SIEMANN

## Illustrations

TERRY ANDERSON

## Distribution

CRYSTAL TILGHMAN

The *Beam Line* is published quarterly by the Stanford Linear Accelerator Center, PO Box 4349, Stanford, CA 94309.

Telephone: (415) 926-2585

INTERNET: [beamline@slac.stanford.edu](mailto:beamline@slac.stanford.edu)

FAX: (415) 926-4500

Issues of the *Beam Line* are accessible electronically on the World Wide Web at <http://www.slac.stanford.edu/pubs/beamline>

SLAC is operated by Stanford University under contract with the U.S. Department of Energy. The opinions of the authors do not necessarily reflect the policy of the Stanford Linear Accelerator Center.

*Cover: Four stages in the fabrication of the rf cavities for PEP-II at SLAC. The 26 cavities will supply the power to the circulating electron and positron beams. The bowl shapes are pressed out of copper plate and electron-beam welded together. Water channels cut into the outside are filled with blue wax and plated over to form cooling circuits. After the wax is removed, the remaining ports and flanges are added.*

*(Courtesy Lawrence Livermore National Laboratory)*

Printed on recycled paper



## FEATURES

### 2 WHY ARE WE BUILDING B FACTORIES?

*Physicists on three continents are building experimental facilities to search for CP violation in B meson decays.*

Natalie Roe & Michael Riordan

### 10

*The National Laboratory for Synchrotron Light*  
**THE BRAZIL EXPERIENCE**

*The Director of the Brazilian synchrotron light source relates his experiences in building a complex scientific instrument in a developing nation.*

Cylon Gonçalves da Silva

### 16

**GAMMA-GAMMA COLLIDERS**

*Photon beams can be made so energetic and so intense that when brought into collision with each other they can produce copious amounts of elementary particles.*

Kwang-Je Kim & Andrew Sessler



M. Belloni

# CONTENTS

- 23 FROM SUPERCONDUCTORS TO SUPERCOLLIDERS  
*Superconductivity sheds light on the mystery of the Higgs phenomenon in the weak interactions.*  
Lance Dixon

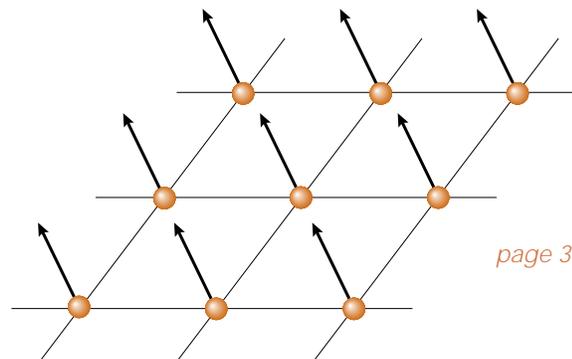
- 31 THE MYSTERY OF THE MATTER ASYMMETRY  
*Why is our Universe made of matter and not antimatter? The answer might be found in the laws that govern elementary particles.*  
Eric Sather



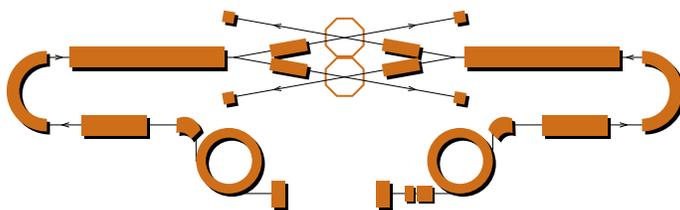
page 23

## DEPARTMENTS

- 38 THE UNIVERSE AT LARGE  
*The Department of Celestial Magnetism Part II*  
Virginia Trimble
- 49 CONTRIBUTORS
- 52 DATES TO REMEMBER



page 31



page 16



page 38

# Why Are We Building $B$ Factories?

by NATALIE ROE & MICHAEL RIORDAN

*Physicists on three continents are building experimental facilities to search for CP violation in  $B$  meson decays.*

**D**URING THE PAST TWO YEARS construction has begun on two particle colliders known as “asymmetric  $B$  factories”—one in California and the other in Japan. These new machines will collide electrons and positrons at unequal energies to produce copious pairs of  $B$  mesons in a clean, low-background environment. About one in every four interactions yields a pair of  $B$  mesons. And as no additional particles are produced when such an event occurs, every single track observed in the surrounding particle detector can be ascribed to one of the two  $B$  mesons.

At the same time there is strong and growing interest in studying  $B$  mesons that will be produced by the billions at proton colliders, once Fermilab’s Main Injector (Batavia, Illinois) and CERN’s Large Hadron Collider, the LHC (Geneva, Switzerland), begin operating. And at DESY’s electron-proton collider HERA (Hamburg, Germany) there are ambitious plans to study  $B$  mesons generated by inserting a stationary target into the proton beam’s halo. The raw number of  $B$ ’s produced at these machines is much greater than that generated by asymmetric  $B$  factories. But only a small fraction of interactions will contain  $B$  mesons, and there

will be lots of extraneous tracks in every event. Extracting a clean signal in the face of such difficult conditions is a challenging but not impossible task—one that promises great rewards for dedicated experimenters.

At the core of all the recent interest in  $B$  mesons is the prospect of encountering additional examples of  $CP$  violation. In 1963 Jim Cronin, Val Fitch, and their colleagues discovered this asymmetry between matter and antimatter in certain decays of neutral  $K$  mesons. But despite more than a quarter century of searching, no other instance of this intriguing phenomenon has ever been observed. It may be a natural consequence of the Standard Model of elementary particle physics, or perhaps our first glimpse of new physics *beyond* the Standard Model. Painstaking studies of  $K$  meson decays have yet to resolve this issue (see article by Jack Ritchie in the last issue of the *Beam Line*, Vol. 25, No. 4), in part due to the intrinsic theoretical uncertainties in these processes. Finding another example of  $CP$  violation—and measuring it in detail—is the crucial next step in understanding this phenomenon, which many cosmologists reckon to be the central element in explaining the matter-antimatter asymmetry of the Universe.  $B$  mesons promise to be an especially powerful tool in this search, having many decay modes in which the Standard Model predicts large particle-antiparticle asymmetries with little theoretical ambiguity.

And there are other important reasons for studying  $B$  mesons, too. One of their constituents, the bottom,

or  $b$ , quark is the second-heaviest quark in the Standard Model. The heaviest is the recently discovered top, or  $t$ , quark, whose observation by the CDF and  $DØ$  experiments at Fermilab put the capstone on the matter content of the Model. But the  $t$  quark is so top-heavy that it falls apart before it can form a bound state with other quarks; in a tiny fraction of an eyeblink, it disintegrates into a  $b$  quark and a  $W$  boson. By contrast, the much stabler  $b$  quark survives long enough to form bound states with other quarks, each exhibiting a rich variety of decay modes and offering opportunities to study the dynamics of heavy-quark interactions. This research is being pursued not only at Fermilab but also at Cornell, which operates a symmetric (equal electron and positron energies)  $B$  factory—as well as at the LEP and SLC colliders at CERN and Stanford Linear Accelerator Center (SLAC) respectively, where  $B$ 's produced in the decays of massive  $Z$  bosons have yielded a variety of interesting results.

**T**HE  $b$  QUARK and  $B$  mesons were discovered in the late 1970s and early 1980s, during the consolidation of the Standard Model as the dominant particle-physics theory. Physicists were beginning to recognize that the elementary building blocks of matter come in families with two quarks and two leptons apiece. Thus the 1976 discovery by Martin Perl and his colleagues (see article by Perl in last issue of the *Beam Line*, Vol. 25, No. 4) of a third charged lepton, the tau lepton, suggested the existence of another pair of quarks. An initially obscure—but now famous—paper

by two Japanese theorists, Makoto Kobayashi and Toshihide Maskawa, indicated  $CP$  violation could occur naturally in the Model if and only if a third such family existed.

The discovery of the bottom quark was not long in coming. In 1977 a group of physicists led by Leon Lederman reported the discovery of the massive Upsilon particle  $\Upsilon$  in collisions of high-energy protons with nuclei at Fermilab. It was widely believed to be a neutral meson composed of a charge  $-1/3$  quark plus its antiquark. Perhaps the  $b$ ? In the early 1980s, the CLEO experiment at Cornell's electron-positron collider CESR bore out this picture with the discovery of  $B$  mesons, composed of a  $b$  quark and a light quark, in decays of an excited state of the Upsilon, designated the  $\Upsilon(4S)$ .

The  $B$  meson lifetime was first measured at SLAC using the PEP collider, which ran at a higher energy than CESR. Although the rate of  $B$  meson production was much lower at PEP, the  $B$ 's came flying out with enough momentum to travel a measurable distance before they disintegrated—almost 1 mm, on the average. This was a big surprise, as it meant the  $B$  lifetime was much longer than expected.

In 1986–87 the high energy physics community became very excited by another surprise: the observation of “ $B^0\bar{B}^0$  mixing,” in which a neutral  $B$  meson spontaneously converts into its antiparticle. This phenomenon was discovered by the UA1 experiment at CERN's proton-antiproton collider and confirmed by the ARGUS experiment at DESY's electron-positron collider DORIS. Similar mixing behavior had been

observed decades earlier for neutral  $K$  mesons, prior to the discovery of  $CP$  violation. The high degree of  $B^0\bar{B}^0$  mixing, when combined with the long  $B$  meson lifetime, soon suggested to physicists that  $CP$  violation might be observed in decays of neutral  $B$  mesons. Suddenly,  $B$  physics became a very hot topic.

**M**OST OF THE DATA collected at electron-positron  $B$  factories is taken on the  $\Upsilon(4S)$  resonance. What makes this resonance so special is that, at a mass of 10.58 GeV, it has just enough energy to decay into two  $B$  mesons—and nothing else. This is an ideal experimental situation, because all the tracks in a given event can be assigned to one  $B$  meson or the other. Another advantage at this energy is that  $\Upsilon$  production accounts for one-fourth of all interactions, and  $B$ -meson decays are readily distinguished by their topology from events containing lighter quarks.

Physicists have exploited these favorable experimental conditions at CESR and DORIS, producing many exciting discoveries. In a collegial competition, these two groups have tried to outdo each other in devising clever analysis techniques in order to be the first with groundbreaking results. DORIS has now ceased operation, but CESR has been upgraded to higher luminosity and continues to set records for machine performance. It is the world's highest-luminosity electron-positron collider and will maintain such an enviable position at least through 1998. Physicists working there are reaping record numbers of  $B$  mesons, continually improving our understanding of their behavior.

But the study of  $CP$  violation in  $B$  meson decay will be extremely challenging at CESR because of a major, fundamental limitation. In order to extract  $CP$ -violating asymmetries from the data (see box on the right), it is crucial to know the *order* in which the two  $B$ 's decay. But  $B$  mesons at a symmetric  $\Upsilon(4S)$  machine like CESR are produced almost at rest, traveling only about 30 microns or so before decaying. Such a distance is much too short to be resolved using present detector technology—which therefore makes it impossible to establish the exact moment at which each meson expires. Without this crucial information, it is exceedingly difficult to observe  $CP$  violation at a  $\Upsilon(4S)$ -producing machine.

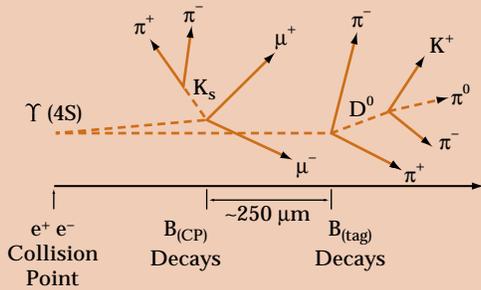
$CP$  violation appears when there is a difference between the decay rates of  $B^0$ 's and  $\bar{B}^0$ 's to the same final state. The Standard Model predicts that this phenomenon will occur for certain special states known as " $CP$  eigenstates," which are symmetric in their matter-antimatter content (see box). Having observed such a decay, however, the poor experimenter does not know whether it originated from a  $B^0$  or a  $\bar{B}^0$ . The solution to this conundrum is to observe the *other*  $B$ ; if it decays to a final state that can be clearly identified as matter or antimatter, it can be used as a "tag." An excess of either kind of tag accompanying a given  $CP$  eigenstate is evidence for  $CP$  violation. But we are not quite home yet; we still need the timing information mentioned above. This is the primary reason for building asymmetric  $B$  factories—to give the  $\Upsilon(4S)$  a boost in one direction. The  $B$  mesons that emerge from

### Observing $CP$ Violation at an Asymmetric $B$ Factory

IN THE STANDARD MODEL  $CP$  violation should occur in certain rare  $B$  decays to final states with equal matter and antimatter content called " $CP$  eigenstates."  $CP$  is a quantum-mechanical operator that simultaneously changes particles into their antiparticles and reverses the parity of a system; a  $CP$  eigenstate is a system of particles that is left unchanged by the action of the  $CP$  operator. When a  $B^0$  decays to  $\pi^+\pi^-$ , for example, it yields a  $CP$  eigenstate. The charge-conjugation operator  $C$  changes the  $\pi^+$  to a  $\pi^-$  and the  $\pi^-$  to a  $\pi^+$ , while the parity operator  $P$  has no effect at all because the pion has no spin, and the pair is produced in a state with zero angular momentum; we end up just as we started. Another  $CP$  eigenstate is  $J/\Psi K_S$ , because both the  $J/\Psi$  and the  $K_S$  are self-conjugate particles: they are their own antiparticles. Certain  $B$  decays to "near- $CP$  eigenstates" (such as  $B^0 \rightarrow \rho^+ \pi^-$  and  $B^0 \rightarrow J/\Psi K^*$ ), in which matter-antimatter symmetry occurs at the quark level, can also be used in studies of  $CP$  violation.

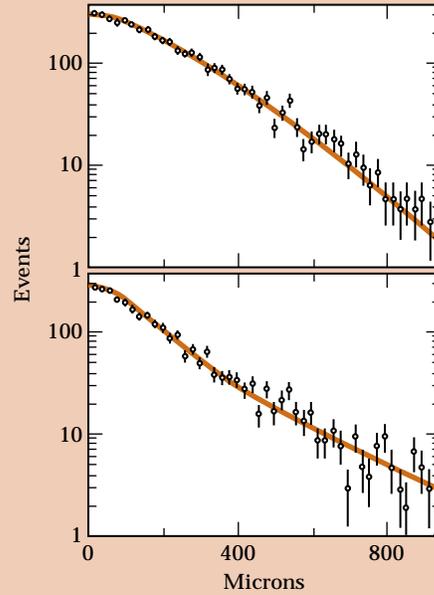
If there were perfect symmetry between matter and antimatter,  $B^0$  and  $\bar{B}^0$  mesons would decay to a  $CP$  eigenstate at exactly the same rate. The observation of a difference between these two decay rates would be evidence of  $CP$  violation. But there is a catch-22. Because both  $B^0$  and  $\bar{B}^0$  mesons can decay to a  $CP$  eigenstate, it is impossible to know which of the two processes occurred merely by observing the outcome. Fortunately,  $b$  quarks are always produced in quark-antiquark pairs; thus by establishing the identity of the other  $B$

meson, often called the “tagging  $B$ ,” we can deduce the identity of its partner. Although the tagging  $B$  can oscillate before decaying, no penalty is paid for this effect in measurements on the  $\Upsilon(4S)$ ; this is an advantage of resonant production. The disadvantage is that if we don’t observe the order in which the two  $B$ ’s decay, the  $CP$  asymmetry vanishes.



A typical “golden event” at an asymmetric  $B$  factory. The first  $B$  meson decays into a  $J/\psi$  (which immediately decays to  $\mu^+\mu^-$ ) and a  $K_S$ , while the second  $B$  yields a clear tag.

The above drawing illustrates a “golden event” at an asymmetric  $B$  factory. The typical separation between the two  $B$  meson decays is about 250 microns at PEP-II or KEKB energies. There are four possible event categories: (i) a  $\bar{B}^0$  tag followed by the decay to a  $CP$  eigenstate; (ii) a  $B^0$  tag followed by the decay to a  $CP$  eigenstate; (iii) the decay to a  $CP$  eigenstate followed by a  $\bar{B}^0$  tag; (iv) the decay to a  $CP$  eigenstate followed by a  $B^0$  tag. By combining events from categories (i) and (iv), and those from categories (ii) and (iii), and plotting these two resulting event classes versus the time interval (or, equivalently, the distance) between the two  $B$  decays, we obtain distributions such as those shown at right above. In both there is an obvious departure from pure exponential falloff, with the first distribution showing a slight excess of events over the second. The difference between these two distributions represents a clear indication of  $CP$  violation. If we add the two distributions, however, we blur out this difference, and an exponential falloff is all that we can observe. This illustrates why the order of the  $B$  decays is crucial information to obtain in making measurements at the  $\Upsilon(4S)$ .



Distributions of two categories of simulated golden events versus the separation between their two  $B$ -decay points. An experimentally observed difference between these two distributions would be firm evidence for  $CP$  violation. The top plot has events from categories i and iv while the bottom plot contains events from categories ii and iii. Consult the text for definitions of the four categories.

Another interesting feature of  $B$  meson decays is that there are three distinct  $CP$ -violating asymmetries that are in principle measurable. They are related to each other and can be expressed as three angles of a triangle (called the “unitarity triangle”; see article by David Hitlin and Sheldon Stone, *Beam Line*, Vol. 21, No. 4, Winter 1991). Given two asymmetry measurements, the third is uniquely determined within the Standard Model framework, and their sum cannot exceed 180 degrees. In addition, there are several decay modes which can be used to measure each asymmetry, providing a consistency check for each angle individually. Other studies of heavy-quark decays will provide additional constraints on the allowed values for each of these three asymmetries, as well as on the magnitude of the three sides of the triangle. The complete set of  $B$ -meson decay measurements will therefore give us a very stringent self-consistency test of the Standard Model.

it generally decay at measurably different points along that direction, allowing an accurate determination of which expired first.

The idea for an asymmetric  $B$  factory was first proposed in 1987 by Pier Oddone of Lawrence Berkeley National Laboratory (LBNL) in California. Elegant in its simplicity, his concept nevertheless met initially with considerable skepticism because of its challenges for accelerator builders. In conventional symmetric colliders the electron and positron bunches occupy the same beam pipe—traveling in opposite directions and guided by the same sets of focusing magnets. An asymmetric  $B$  factory requires two separate beam pipes, each with its own magnet system guiding electron or positron beams on independent journeys around the ring. It must also

have a complicated interaction region, with magnets that bring the two beams together briefly (so that they can collide) and then immediately separate them. Previous two-ring colliders (such as DORIS at DESY in Germany) had run into severe difficulties and never reached their design luminosities. But Oddone soon launched a feasibility study at LBNL to examine this problem, joining

physicists. Their bold conceptual design spawned a number of asymmetric  $B$  factory proposals all over the world, two of which are now under construction.

KEK in Japan and SLAC are both recycling existing electron-positron colliders into asymmetric  $B$  factories. SLAC is upgrading the PEP collider into PEP-II, while KEK is transforming its TRISTAN collider into KEKB. Both colliders were originally designed to operate at much higher energies but with lower currents, and with equal-energy beams counter-rotating inside a single beam pipe. In their new incarnations, they will be transformed into two-ring colliders, with completely new systems to power and control beams that have almost 100 times more circulating current.

There is an important difference in approach, however. In KEKB the electron and positron beams collide at a slight angle, which allows the beams to be separated easily after they cross one another, while in PEP-II they collide head-on. Although there is a risk that the KEKB scheme could result in lower luminosity, as happened on DORIS, the advantage is that no bending magnets are needed close to the interaction region to separate the beams. This allows more room for particle detectors and reduces the backgrounds from synchrotron radiation. At PEP-II the beams are brought together and then separated by means of permanent bending magnets positioned a mere 20 cm from the interaction point. The large synchrotron radiation background generated by bending the beams is absorbed in a series of water-cooled masks located inside the beampipe.

*Typical magnets for the low-energy ring (above) and high-energy ring (below) in PEP-II.*



This rather conservative approach to accelerator design forces experimenters to be very creative in devising means to squeeze their detectors into the limited space remaining around the machine components.

KEK and SLAC have also begun building ambitious new state-of-the-art particle detectors to record the tracks and energy deposits left by decaying  $B$  mesons. The detector at KEKB is known as Belle, while the one at PEP-II is called BaBar—after the elephant in Laurent DeBrunhoff's children's stories. These two detectors are fairly similar in most aspects, differing only in certain technical details. At the heart of both detectors is a precision vertex detector that can determine the  $B$ -decay vertices to better than 100 microns. In BaBar the vertex detector is mounted directly on the final bending magnets. Based on silicon-strip detectors (see article by John Jaros and Alan Litke, *Beam Line*, Vol. 20, No. 1, Spring 1990), this precision tracking device is crucial for the measurement of  $CP$ -violating asymmetries and must be located as close as possible to the interaction point.

In both Belle and BaBar, low-mass drift chambers will measure particle momenta while minimizing multiple scattering, and precision electromagnetic calorimetry based upon cesium-iodide crystals will determine photon energies. Pions, kaons and protons will be identified in new and different ways. In Belle a threshold Cherenkov counter made of an extremely light, diaphanous substance called aerogel (see photograph right) will be augmented by high-precision time-of-flight counters, while BaBar will use a novel device known

as DIRC, for Detection of Internally Reflected Cherenkov light.

An important feature of these detectors is their ability to record tens of millions of  $B$ -meson events per year. Massive computing power will be needed to sift through these enormous data samples offline to find the truly interesting events. "Factory-like" operation is required to produce such huge samples because experimenters expect to detect  $CP$  violation only in rare events that occur less than once in a thousand or so  $B$  decays.

For example, the decay process  $B^0 \rightarrow J/\psi K_s$  is a particularly obvious mode for observing this phenomenon because the  $J/\psi$  occasionally breaks up into two easily identified muons. Such events comprise only about 0.05 percent of all  $B^0$  decays, however, and only a small fraction of them is easily reconstructed. Thus only 1 in every 40,000  $B^0$  decays results in such an observable golden event. This yield is further depleted by the effects of imperfect detector acceptance and efficiency. And the tagging  $B$  must be accurately identified, resulting in a further 70 percent loss. The net result is that for every 10 million  $\Upsilon(4S)$  events produced at an asymmetric  $B$  factory and recorded by the detector, slightly more than 100 tagged golden events will be found that are useful in the search for  $CP$  violation.

When operating at design luminosity, PEP-II and KEKB will record about 30 million  $B^0\bar{B}^0$  events per year, which should be sufficient to observe at least two of the three  $CP$ -violating asymmetries expected in the Standard Model (see box on pages 4 and 5). In the clean environment of the  $B$



*KEK physicist peers through a sample of aerogel, to be used for the Cherenkov counter in the Belle detector.*

factory, we can reconstruct several different final states for each asymmetry measurement and combine them to enhance the statistical significance. If  $CP$  violation is observed as expected, these tests will tell us conclusively how the phenomenon originates. And if unexpected results occur, the capability of observing many different final states will be a powerful tool to scout for potential new physics.

**I**N ADDITION to the Babar and Belle experiments at SLAC and KEK, there are proposals to search for  $CP$  violation in  $B$  decays using Fermilab's Tevatron collider and the HERA collider at DESY. Like the  $B$  factories, these experiments are scheduled to begin taking data in 1998 or 1999. There are also experiments being planned for CERN's new LHC that have targeted  $CP$  violation; they promise to make measurements at much higher precision than the planned initial round of "discovery" experiments.

Although these experiments will each confront different challenges, they have several features in common. Because  $B$  mesons are produced nonresonantly in hadron collisions, it is enough to observe that one  $B$  decays to  $J/\psi K_S$  (for example) and to tag the other one, without measuring the timing of the two disintegrations. And at the higher energies of hadron colliders,  $B$  mesons travel visible distances before decaying—almost a full *centimeter* at HERA-B, for example—so lifetime information is readily available to help reject undesirable backgrounds and improve the experimental sensitivity.

The raw numbers of  $B$ 's produced will be orders of magnitude larger than at KEKB and PEP-II, but  $B$  events constitute only a small fraction of all events, and most ordinary  $B$  decays are difficult to distinguish from background processes. Selection of the most interesting  $B$  decays to be written onto tape must be done in a few microseconds using a sophisticated electronic triggering system. (In contrast, Belle and BaBar will be able to record essentially all  $B$  decays for later offline analysis.) Such an event trigger is required to recognize certain decays based on rapid reconstruction of important event characteristics. This feat is fairly easy with such an obvious final state as  $J/\psi K_S$ , but it is more difficult for other states such as  $\pi^+\pi^-$ . Several experiments are developing powerful event triggers based on quick detection of a displaced vertex; if successful, they could open the door to the study of additional  $CP$  eigenstates beyond the  $J/\psi K_S$ .

A different sort of asymmetric collision will take place at DESY, whose electron-proton collider HERA will be adapted for a novel experiment called HERA-B. The plan is to collide HERA's 820 GeV proton beam with stationary nuclei in a wire target inserted into the halo of the beam; that way the experiment can run without interfering with the normal HERA program.

The rate of  $B$ -meson production in proton-nucleus collisions is not known to better than a factor of 2 at this energy, but a conservative estimate is several hundred million  $B$ 's per year. Only about one in a million interactions will actually contain  $B$

mesons, which will be embedded in an extremely high total data rate—vastly higher than a particle detector can hope to record. In addition, an average of three background events will be superimposed on every  $B$  event. Experimenters must therefore extract the  $B$  decay products from up to 200 tracks in order to reconstruct a golden event and tag the other  $B$  without being confused by all the other particles. To achieve this goal, the detector must be highly segmented to resolve so many particles, and radiation-resistant to survive the intense environment close to the beam line.

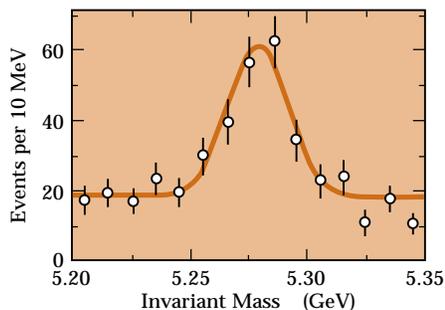
These are daunting challenges, but physicists working on HERA-B have one very important thing in their favor—a working accelerator. They have already performed tests with an internal wire target and have successfully obtained the necessary interaction rate without seriously degrading the operation of HERA's other experiments. As detector prototypes are constructed, they will be subjected to online testing in order to assess their performance and begin to get experience studying  $B$ 's in this unique environment. HERA-B is scheduled to have its first full-scale run in 1998, about a year before PEP-II and KEKB begin colliding beams.

The main goal of HERA-B is to observe  $CP$  violation using  $B^0 \rightarrow J/\psi K_S$  events, which are by far the easiest to identify in high backgrounds. Several hundred of these golden events are anticipated per year. Other final states, which could increase the statistical accuracy and provide important cross-checks, will be much harder to isolate.

**B**ACK AT FERMILAB, meanwhile, physicists have already demonstrated that the Tevatron  $\bar{p}p$  collider is competitive in  $B$  physics. They have ambitious plans to search for  $CP$  violation after the Main Injector upgrade is completed in 1999, increasing the luminosity by a factor of 10. Once this is achieved, the Tevatron will become the most prolific  $B$  factory in the world, producing about one hundred billion  $B$ 's per year. At 2 TeV about one in a thousand  $\bar{p}p$  interactions contain  $B$  mesons, and these events typically contain extraneous tracks produced in the underlying collision. This situation is much more favorable than at HERA-B, although more challenging than the absolutely clean environment of an asymmetric electron-positron collider.

Of the two experiments now running at the Tevatron, CDF is better optimized for  $B$  physics at present; the collaboration has published interesting measurements on  $B$ -meson lifetimes, mixing, production cross sections and masses. In addition, it has already reconstructed over 100  $B^0 \rightarrow J/\psi K_s$  events (see graph above) and over 600  $B^+ \rightarrow J/\psi K^+$  events that can be used to evaluate tagging techniques.

Both the CDF and DØ collaborations are planning extensive upgrades to improve their sensitivity to  $B$ -meson decays; they expect to be able to obtain several hundred tagged  $J/\psi K_s$  events per year. And CDF physicists are designing a sophisticated silicon vertex track processor that will allow them to trigger on displaced vertices. If successful, it would also give them access to other



*Preliminary CDF data on the production of  $J/\psi K_s$  events, plotted versus the invariant mass of the  $J/\psi K_s$  system. About 138 events are included under the peak.*

final states such as  $\pi^+\pi^-$ . In addition to seeking  $CP$  violation, both Tevatron experiments will continue to improve their lifetime and mixing measurements, search for rare  $B$ -decay modes and examine the formation of expected but as-yet-unobserved mesons and baryons. In many respects, the Tevatron program is complementary to those of the asymmetric  $B$  factories—offering greater breadth in  $B$  physics topics, although perhaps less depth in the variety of accessible  $CP$ -violating decay modes.

**T**HE PROJECTED time scales for completion of the asymmetric  $B$  factories at KEK and SLAC, the Main Injector upgrade of the Tevatron, and the HERA-B experiment are all roughly the same—with first results expected by the turn of the millenium. With so many physicists hot on its trail,  $CP$  violation will soon be examined more closely than ever before in its elusive thirty-year history. The race to be the

first to observe  $CP$  violation in the golden-decay mode  $B^0 \rightarrow J/\psi K_s$  will be intense, and should bring out the best in experimental effort and ingenuity.

But there is much more to this  $B$ -physics program than just a single asymmetry measurement. The Standard Model predicts a rich and complex pattern including three different  $CP$ -violating asymmetries—each potentially observable in a variety of  $B$ -decay modes. The measurements of different modes that probe the same asymmetry must agree with one another, and the three asymmetry angles must add up to 180 degrees, if the Standard Model is correct. An experimental program capable of measuring  $CP$  violation in a variety of different channels, including at least two of the three different asymmetries, will tightly constrain this theory and provide the ultimate precision test.

And if Nature is kind, if there is still something truly exciting lying in wait, this broad physics program may provide enough clues to lead us into unexplored territory beyond the Standard Model.



*The National Laboratory for Synchrotron Light*  
**THE BRAZIL EXPERIENCE**

by CYLON GONÇALVES DA SILVA

*The Director of the Brazilian synchrotron light source relates his experiences in building a complex scientific instrument in a developing nation.*

**W**HY IS BRAZIL BUILDING a synchrotron light source? Have synchrotron light sources become the technological status symbol of the 90s for developing countries, as nuclear reactors were in the past? Not very likely. However, how can one explain that so many of these countries—Taiwan, South Korea, India, China, Brazil, Thailand—have built, are building, or are talking about building their own synchrotron facilities? The lure of “Big Science”? The “keeping up with the Jones’” syndrome? In spite of many differences—historical, cultural, economic, and political—I believe that the role of science and technology in modern life is the prime reason behind all of these projects. Unfortunately, nothing is more difficult to pinpoint in a clear and immediately perceivable way than the elusive relationship between science and technology on one hand and economic development on the other. In the major industrial nations, this relationship—which was taken for granted at least since the last World War—is now the object of serious questioning. In a developing country, a project such as a synchrotron light source may spark debates that compare in acrimony with those surrounding the Superconducting Super Collider project in the United States.

Synchrotron light sources are assumed to be evidences of and contributors to a modern advanced economy. Industrial applications are an important selling point, even when in practice they still account for only a small—not to say, insignificant—fraction of their use. What counts for a developing country, at least initially, is not actual applications of synchrotron light, but building a complex scientific instrument. The technologies behind a storage ring are seen as “enabling” tools for further developments. If we can do this (the storage ring), then we can do that (modern production technologies) also. A country must break into the virtual circle of economic development on many fronts simultaneously. Building capability to do quality R&D is one of the most important social functions of large scientific projects in these countries. The process is rarely uniform or follows an efficient, logical path. It is instead history dependent owing to chance events, highly non-uniform, and messy (witness the somewhat empty experimental halls of many new synchrotron laboratories).

**C**ONCEPTUALLY, the Brazilian project was sold on what I have called—in homage to the high-energy physics of my student days—the three-fold way: a strategy combining engineering, science, and organization. The three-fold way is depicted schematically in the pie chart on the right. Engineering meant designing and building as much as possible of the storage ring and instrumentation in Brazil, with the help, whenever possible, of local

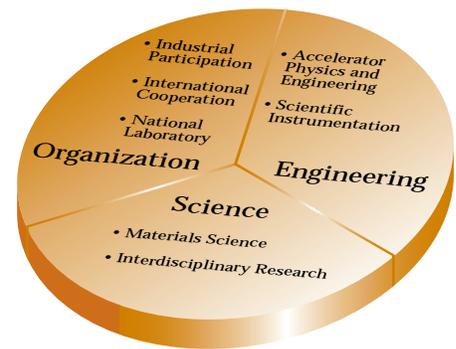
industry. The idea was to have accelerator technology without going into costly high-energy physics, in which we could not be competitive. This settled the choice for science: materials science done with photons from a storage ring. The third leg of the three-fold way was the concept of a national laboratory. A synchrotron light source would serve a broad community—practically all disciplines in exact, life, and earth sciences would benefit. The best devised strategy, however, still has to survive the tests of real life. Where did the Brazilian National Laboratory for Synchrotron Light (LNLS) stand back in the mid-1980s?

There was only one person in Brazil, Ricardo Rodrigues, a young physicist from the University of São Paulo, who was available, willing, and qualified to be the technical leader of the project. When he agreed to be Technical Director, I knew there were no challenges we could not meet. Rodrigues was given the task of running the construction of the accelerators (with very profitable side incursions into everything else!).

The engineering leg of our strategy depended on a huge bet that in a short time LNLS could train a minimal staff to design and build the accelerators. At this point a decision was made—we would bootstrap ourselves into the business by training the staff in-house as much as possible. (From three-fold way to bootstrap, we held firm, albeit tongue-in-cheek, to particle physics.) The argument made a lot of sense to Rodrigues and to myself—there was no time to send people abroad for extended training periods; whatever experience they gained would not be

*A view of the LNLS under construction.*

N. Chinaglia, March 1996



*The three-fold way: the basic strategy for setting up LNLS may be summarized as an engineering effort to build a synchrotron light source and its scientific instruments to be used for materials research in the institutional setting of a national laboratory. International scientific cooperation, especially in Latin America, but also with similar laboratories in other countries, in addition to the participation of local industry in the construction of the light source, form an integral part of this basic strategy.*

## Why Build a Synchrotron Light Source in a Developing Country?

**#1** A SYNCHROTRON LIGHT SOURCE (SLS) may be used to introduce a novel type of science organization—a national laboratory—in a developing



L. Camara, September 1993

country. National laboratories may initially be seen as spending money that would be better spent by spreading it throughout the scientific community. However, if they are properly managed as open facilities, with access based on the quality of the proposals and peer review, they are eventually accepted. The resources that can be amassed at one site are much larger than anything a university department can offer. The crucial point is open access.

**#2** SYNCHROTRON LIGHT SOURCES produce photons for materials research in the broadest description of the term—organic and inorganic—encompassing fields as varied as engineering, exact, life, and environmental sciences. They are not restricted to a small constituency of users. By serving a majority of the scientific community in a country, they can gain political support across disciplines and institutions.

**#3** ANOTHER IMPORTANT ASPECT of an SLS for the development of science is the dynamics of the experimental hall floor. All disciplines are represented; graduate students from different backgrounds and their advisors rub shoulders (sometimes literally, as the floor can get very crowded). This leads to many fruitful cross-disciplinary interchanges and helps to break down artificial barriers between domains of science that are normally departmentalized in universities.

**#4** DESIGNING AND BUILDING a SLS can be an opportunity for the scientific community to experience the process of technological development. It exposes the individuals involved to real-life situations where they can interact with the industrial and private sectors of the economy.

immediately applicable to the working environment in Brazil, and construction had to start immediately. It is not clear it made sense to anybody else. Fortunately, the LNLS Board of Directors bought our idea. This decision was complemented by two related ones: (i) send technical staff abroad for short periods to learn specific techniques or to solve clearly defined problems after they had tackled the difficulties by themselves for a while; (ii) from time to time have experts review the project (for a variety of reasons this actually happened only twice, in 1989 and again in 1991).

As to the Science leg of our strategy, LNLS had to start by building up a users' community. A community of users of synchrotron light is, first and foremost, a research community, the size and composition of which will vary from country to country, owing to local historical experiences. One comment about the recent Brazilian efforts to develop science and technology may be of interest. The National Council for Scientific and Technological Development, CNPq, the organization which sponsors LNLS, was created in the early 1950s. Influenced by the post-war American example and constrained by the lack of industrial demand for R&D, emphasis was given to basic research. CNPq was, and still is, an agency dedicated to the support of basic research. Thanks to its efforts, over the last four decades Brazil built up a small, but politically visible, scientific community. In the meantime, industrial development was geared to imported technological black boxes and turnkey installations, so that science has remained largely isolated from mainstream economic life. For most scientists, technology still smacks of lower quality, not a calling for higher talents and better brains. This led to the somewhat paradoxical situation in which it was easy to build rapidly a community of users but there was widespread initial opposition to the idea of building a synchrotron light source.

Late in 1986, Aldo Craievich accepted the position of Deputy Director of LNLS, responsible for the scientific program. Hence, in parallel with the effort to build the accelerators, LNLS began a series of workshops to “market” research with synchrotron light sources. These topical workshops, in addition to advertising LNLS and the potential of light sources as research tools, allowed the local community to establish useful links with users abroad. This was instrumental to increase the number



*Participants at the 5th Annual Users' Meeting, December 1994.*

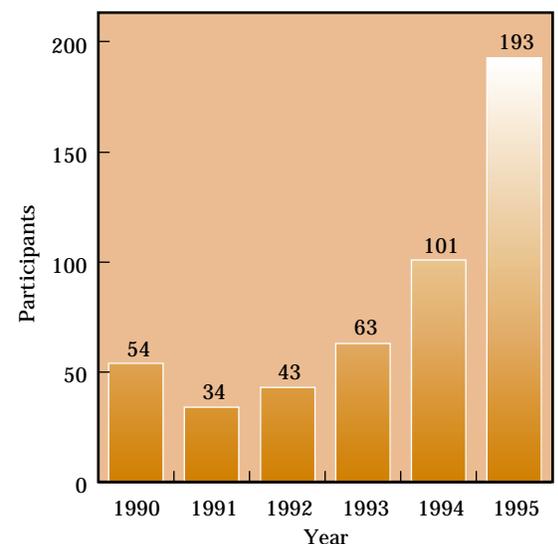
of trained users in Brazil. The bottom figure shows the evolution of the number of participants in the Annual Users' Meeting. What is not shown, but is perceptible to those who have followed these meetings, is the qualitative evolution in the profile of participants, thanks to the training obtained in foreign synchrotron light laboratories.

The development of scientific instrumentation for using synchrotron light has been one of the main concerns of the Scientific Department of LNLS over the years. The existence of a reasonably strong research basis in the country made it possible to rapidly form high quality groups for VUV and X-ray instrumentation. This also allowed a considerable reduction in the cost of beam lines—so much so that in spite of severe budgetary constraints LNLS has seven beam lines scheduled to come into operation soon after synchrotron light becomes available, and the design of its four-crystal high resolution X-ray monochromator is being copied by the European Synchrotron Radiation Facility. In 1992, thanks to Volker Saile's enthusiastic support, LNLS installed its first beam line at the Center for Advanced Microstructures and Devices of Louisiana State University in Baton Rouge. To my knowledge, this was the first time that a complex scientific instrument manufactured in Brazil crossed the equator (thereby reversing the usual flux).

The third slice of the strategic pie turned out to be, as expected, the most difficult. There was no previous experience with a national laboratory for physicists, chemists, or biologists. The prevailing culture was that of small science done in a

compartmentalized way. Laboratories in university departments were (and still are) very much self-contained. Hence, the reaction of the establishment against LNLS was fierce—it was seen as an unfair competitor for resources, dominated by a bunch of insolent youngsters. The idea that it could be something different—a laboratory operated on a professional basis, managed for efficiency and pooling of scarce resources, with allocation of time based on peer review of qualified projects—was entirely foreign to the majority of the scientific community. Even the Brazilian Physical Society publicly opposed LNLS. We quickly learned that technical problems are trivial compared with cultural ones. Fortunately, opposition got swamped by the growth of the scientific community. The younger generation without vested interests to defend supported LNLS. Influential scientists who initially opposed the project eventually changed their minds. We knew we had arrived when the president of the Brazilian Physical Society referred to LNLS as “our” light source.

The concept of a national laboratory concentrating resources but offering free access to the scientific and technological communities of a developing country may be the most important fringe benefit of a light source. National laboratories are a cost effective way to speed the



*Participants at the LNLS annual users' meetings.*

## A BRIEF HISTORY

### 1981–1986 The Early Years

During this period there were extensive discussions with the scientific community and the National Science Council (CNPq) about the possibility of building a synchrotron light source in Brazil. LNLS was formally created at the end of 1984, but nothing really happened for another two years.

### 1987–1989 From Words to Action

In these three years, LNLS is set up by CNPq in Campinas, state of São Paulo; the technical staff is assembled and work starts on the linac injector and on the conceptual design of the storage ring. Only half of the planned injector linac gets built owing to insufficient funds to house the 100-MeV linear accelerator. In December 1989 the first beam is obtained.

### 1990–1993 From Action to Inaction

These four years were the crossing of the desert for LNLS. The political winds changed in Brasília; a new President practically killed off Science. LNLS is forced to go slow, very slow. Even so, work proceeds on prototypes for various components of the storage ring and scientific instrumentation. The first beam line gets built and is installed in CAMD in Louisiana.

### 1994–Present Revival

Finally, funds begin to flow again. Construction of the experimental hall and storage ring begin. In December 1995 the linac is successfully operated. Storage ring construction proceeds at a healthy pace. Injection and first stored beam are expected for May 1996.

### Cost

The price tag of a large project is usually the first issue raised by friends and foes alike. In developing countries the cost of a synchrotron light source may represent a substantial fraction of the budget allocated to science and technology. Since the inception of the project, LNLS has spent approximately \$50 million, including salaries. To this should be added the cost of the land for the campus (approximately \$6 million), donated by the State of São Paulo. Overall not an impressive sum compared with the annual budget of CNPq, the Brazilian National Research Council and LNLS sponsor, which has oscillated between \$350 million and \$600 million from bad to good years. It is even less impressive compared with the Brazilian GNP of the order of \$500 billion. So, do not expect synchrotron light sources to have a large direct impact on the economy of a developing country. The impact is longer term and diffuse—upgrading of the technological basis of the country and a superb R&D and human resources training facility, with a useful lifetime to be measured in decades—that is, a large number of young people who were not even born when the installation was first discussed will benefit from its existence.



N. Chinaglia, March 1996

*Ricardo Rodrigues, part-time crane operator and full-time Technical Director, patiently explains to a largely unmoved Director the need to hire a new staff member, presumably to replace him at the crane in the storage ring hall.*

The Stanford Linear Accelerator Laboratory (SLAC) and the Stanford Synchrotron Radiation Laboratory (SSRL) played an important role in the early history of the Brazilian Synchrotron Light Source. In the early 1980s, Roberto Lobo, then Director of the Brazilian Center for Physical Research (CBPF), in Rio de Janeiro, and Roberto Salmeron, a Brazilian expatriate working at CERN and the Ecole Polytechnique in Paris, were thinking about ways to stimulate experimental research in Brazil. They hit upon the idea of a synchrotron light source. A call to Stanford produced a visit by Helmut Wiedemann to Rio, where, in 1982, he gave an introductory course on synchrotron light sources. In this way, SLAC and SSRL played a major role in initiating the discussions about light sources in Brazil. (Wiedemann also helped design the first storage ring for LNLS—one that never got built but was instrumental in training the future Technical Director of the project, Ricardo Rodrigues.)

An early incident of interest involving Fermilab and SLAC centered around getting local industry to develop the capacitors needed for the linac modulator. Greg Loew of SLAC let us know that SLAC had a set of spare capacitors that could be made available to us; however, red tape on both Brazilian and US sides made it a very difficult operation. In those days, Fermilab, through Leon Lederman and Roy Rubinstein, was responsible for an NSF grant to help science in Latin America. The solution found was that Fermilab would buy these capacitors from SLAC and ship them to LNLS. However, by the time the whole operation could be set up, we had found an industry in São Paulo with whom we jointly developed the components with the required specifications. This was one of our first successes in interacting with local industry and, at the same time, it showed the interest and willingness of the international community to help.

scientific and technological development, provided they are outward looking in their policies. In addition, the broad spectrum of disciplines that can be covered by synchrotron light research is a vital element for the decision to build such a facility.

Interaction with local industry was an important part of the strategy for setting up LNLS. However, it was not an easy task given the paucity of the budget, the irregularity of the cash flow, government regulations concerning procurement, and industry's lack of experience with "high" tech demands. LNLS could not pay a premium price for components and equipment that had to be custom built or developed specifically for the Laboratory. The total cost of the project, thus far, has been about \$50 million (in US dollars). Many times industry did not or could not respond to our requests for a reasonable price in a reasonable time. Curiously enough, given the substantial historical differences, LNLS had an experience similar to that of CERN, as related by Brian Southworth, ". . . when the CERN PS and even ISR were being built almost all technologically advanced design, prototype, and assembly work had to be done in house" (CERN *Courier*, June 1992, p.12).

Every synchrotron light laboratory around the world has its own experience to tell, strongly influenced by local history. The Brazilian experience is, perhaps, of interest to poorer developing countries for it shows that a light source can be affordable if the right strategy is chosen. In spite of many differences, the unifying principle of synchrotron light sources is the wide spectrum of



L. Camara, May 1995

the science that can be done with the photons they generate. Unfortunately, as a technology for producing photons, storage rings are pitifully inefficient machines, even if they are the best that we can produce right now. Bright and original ideas are urgently needed for new and more efficient ways to convert electrical power into high flux, brilliant, and tunable photon beams. The story is just beginning. ○

*Putting the finishing touches on a quadrupole magnet for the storage ring.*

### Progress Update

ON MAY 30, 1996, the first thousand turns of the injected beam at 111 MeV were observed in the LNLS storage ring. On July 30, the beam was successfully ramped to the design energy of 1.15 GeV. Commissioning is now under way to deliver the storage ring to the users' community according to schedule. The highest energy particle accelerator and storage ring in Latin America will then be a reality.

# GAMMA-GAMMA COLLIDERS

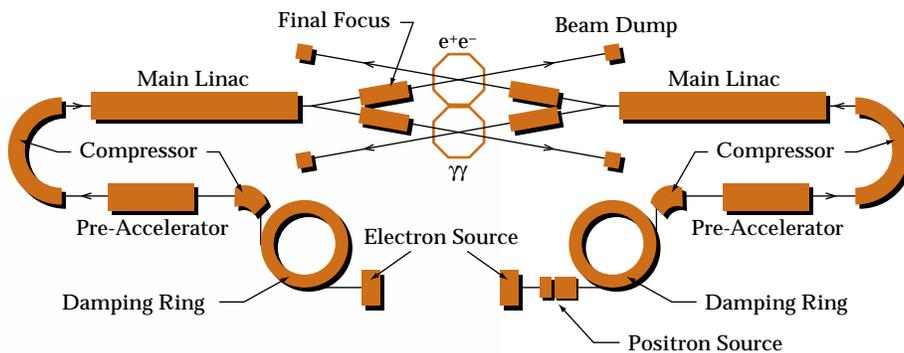
by KWANG-JE KIM & ANDREW SESSLER

*Photon beams can be made so energetic and so intense that when brought into collision with each other they can produce copious amounts of elementary particles.*

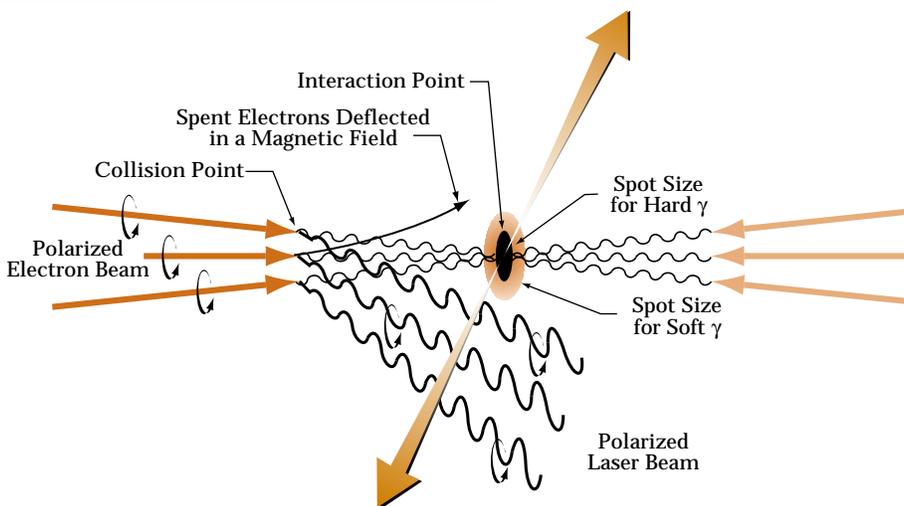
**A**LMOST EVERYONE KNOWS that light doesn't affect light. Suppose you stand near the corner of a darkened room with a flashlight in each hand and you shine the right-hand beam onto the left wall and the left-hand beam through the other beam onto the right-hand wall. Then if you move one flashlight up and down its spot on the wall will move up and down, but it will have no effect on the other beam. All this was described very elegantly in 1864 by James Maxwell, who made the electrodynamic equations linear.

However, as the flashlights are made brighter and brighter, one beam does in fact begin to affect the other. The effect is the scattering of light on light, and it is purely a quantum mechanical effect; it wouldn't happen classically. Now suppose the flashlight beam is made more energetic by increasing the frequency of the light so that it becomes more blue, then passes into the ultra-violet, becomes X rays, and even gamma rays. At first quantum electrodynamic effects occur, and pairs of electrons and positrons are produced; soon quantum chromodynamic effects occur, and strange particles are produced. Once again, common knowledge is not accurate knowledge.

*Schematic of a linear collider with a second interaction region devoted to  $\gamma\gamma$  collisions.*



Gamma-gamma colliders are all about what we just described; namely, making intense beams of gamma rays and having them collide so as to make elementary particles. We shall show in this article that constructing a gamma-gamma collider as an add-on to an electron-positron linear collider is possible with present technology and that it does not require much additional cost. Furthermore, we shall show that the resulting capability is very interesting from a particle physics point of view. An overview of such a linear collider with a second interaction region devoted to  $\gamma\gamma$  collisions is shown in the top illustration.



### PRODUCING THE GAMMA RAYS

Perhaps first of all we should describe how intense gamma rays can be made. The best manner is by Compton back-scattering of almost visible photons from an intense, high energy electron beam. A diagram of the creation of gamma rays in the so-called conversion region is shown in the bottom illustration on the right. The resulting gamma rays will have a spectrum that extends up to approximately 80 percent of the electron energy. The intensity can be adequately high if the incoming beam of photons is adequately intense. How intense must it be and how intense can it be?

Since the process is going to be an add-on to a linear collider, we have a measure of acceptability already at hand, namely, the luminosity of the  $e^+e^-$  collider. To match the  $e^+e^-$  luminosity in a  $\gamma\gamma$  collider requires that we produce one gamma ray from each electron. That is just correct

energetically, since in colliding with a photon the electron transfers essentially all of its energy to the photon and is therefore not available for the creation of a second gamma ray. Now we must ask how many photons we need to collide with the electron beam so as to make one gamma ray from each electron? The cross section for Compton back-scattering is approximately the Thomson cross section, namely  $(8/3)\pi r_0^2$ , where  $r_0$  is the classical electron radius. Combining this with the cross-sectional area of the laser pulse, we can deduce that about  $10^9$  photons need to be collided with each electron to make one photon back-scattered and thus to make one gamma ray. For a typical collider bunch of  $10^{10}$  electrons we thus need  $10^{19}$  photons. The mathematics to go with this discussion, and in particular a discussion of polarization phenomena, may be found in the box on the next page. If the incident photon intensity is too large, there will be undesirable non-

*A sketch showing the conversion region where laser photons are converted by Compton back-scattering into high energy gamma photons.*

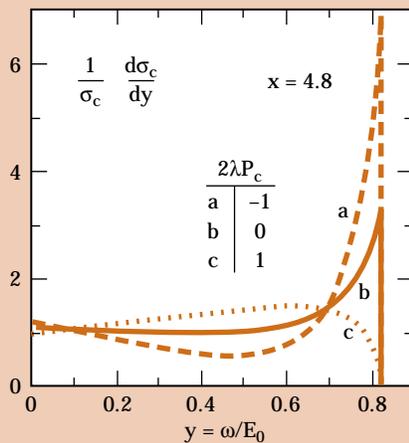
linear quantum electrodynamic effects. These can be simply avoided by making the conversion region longer.

Now the incident photon can not be too energetic or else the photons in the incident photon stream can interact with the back-scattered gamma rays and produce electron pairs. It is easy to see that the criterion for the absence of this process is that the wavelength of incident photons must be longer than  $3.93 E$  [TeV], where  $E$  is the energy of one (of the two) electron beams. Furthermore, one should be close to the limit. Thus for a  $250 \text{ GeV} \times 250 \text{ GeV}$  collider, the photon wavelength should be about 1 micron. Now a 1 micron photon has an energy of 1 eV and thus,

## Compton Scattering

**A** MEASURE OF THE ENERGY of the colliding electron-photon system in Compton scattering is the invariant quantity  $x = 12.3 E \text{ (TeV)} / \lambda \text{ (microns)}$ , where  $\lambda$  is the wavelength of the laser photon. The value of  $x$  should be smaller than 4.83; otherwise the high energy gamma photons produced by the Compton backscattering have a good chance to disappear by colliding with other laser photons to produce  $e^+e^-$  pairs (the Breit-Wheeler process). On the other hand, the maximum energy of the back-scattered gamma photon is  $[x/(1+x)]E$ . For a gamma-gamma collider,  $x$  is chosen to be about 4.8 in order to produce gamma photons with as high an energy as possible while avoiding the pair-production threshold. It then follows that the maximum photon energy is about 80 percent of the electron energy, and that the optimum laser wavelength for  $E = 250 \text{ GeV}$  is about 1 micron.

The cross section for the Compton process  $\sigma_c$  ranges from the classical Thomson cross section  $(8/3)\pi r_0^2$  for  $x$  much smaller than unity to about  $\pi r_0^2 = 2.5 \times 10^{-25} \text{ cm}^2$  at  $x = 4.8$ . The illustration below shows the differential Compton scattering as a function of  $y = \text{the energy of the back-scattered gamma-photon}/E$ . We notice the marked difference in the  $y$ -dependence on the polarization state. When the helicity of the electrons is opposite to that of the laser



*The spectrum of the Compton scattered photons for different combinations of helicities (see text for further explanation).*

photons, the differential cross section shown by curve a is sharply peaked in the narrow region around the maximum value of  $y$ . The peak is less prominent for the unpolarized case shown by curve b, and disappears altogether when the helicities are parallel as shown by curve c. For a gamma-gamma collider, it is clearly most advantageous to choose case a.

To produce one gamma photon per electron, the laser photons, each with an effective cross section  $\sigma_c$ , must cover the full transverse area of the laser pulse  $A$ ;  $\sigma_c N_{\text{laser}} = A$ , where  $N_{\text{laser}}$  is the total number of laser photons in the pulse. On the other hand,  $A = \lambda \ell$  because of diffraction, where  $\ell$  is the effective length of the laser pulse which in turn is about the same as the electron pulse length. Taking the pulse length to be about 1 ps, and using the laser wavelength and the Compton cross section discussed above, it follows that  $N_{\text{laser}} = 10^{19}$ .

although  $10^{19}$  photons is a lot of photons, it is only about 1 joule of energy. A laser able to produce an energy of 1 joule is not excessive at all; in fact such lasers are rather commonplace.

Note that the high luminosity of an  $e^+e^-$  collider is the result of the very small spot size at the collision point: typically tens of nanometers. Note also that the conversion point at which the Compton back-scattering takes place can't be too far from the  $e^+e^-$  collision point or else the natural spreading angle of the produced gamma rays, of order  $1/\gamma$ , will diffuse the collision point too much. On the other hand, the greater the distance between the conversion point and the collision point, the more monochromatic will be the gamma-ray spectrum. For typical colliders the conversion point will be less than 1 cm from the collision point.

Nevertheless, at the conversion point the electron beam is much wider than it is at the crossing point, perhaps even 100 times wider. Thus gamma rays produced at the conversion point would seem to have a large radial extent when they reach the  $e^+e^-$  crossing point. If that were the case the luminosity of a  $\gamma\gamma$  collider would be very much reduced from that of an  $e^+e^-$  collider. But that isn't the case! The gamma rays produced at the conversion point are "focused" to the spot the electrons would have occupied at the collision point. Everyone knows you can't focus gamma rays, but you can produce them to be focused, and that is precisely what is done. This happens automatically, for the electrons have much more momentum than the laser photon, and the produced gamma ray proceeds along the direction of the electron. Thus, to make a  $\gamma\gamma$  collider, it is only necessary to focus the electrons which is just what is normally done in  $e^+e^-$  colliders.

Further discussion of the parameters for a design of a gamma-gamma collider as a second interaction region of the Next Linear Collider (NLC) is presented in the box on the next page. This is in a sense a minimal design, because the electron beam parameters optimized for  $e^+e^-$  collisions before the final focus system are taken as they are. The only modifications are in the design of the final focus optics and an adequate laser for Compton conversion. With this approach, the luminosity for the gamma-gamma collisions, within 20 percent bandwidth, is about 10 percent of that in the  $e^+e^-$  collider mode. This is sufficient to carry out a number of interesting physics

experiments. In particular, the cross sections for some reactions are larger at a gamma collider than at a conventional electron collider. Clearly what is needed is a careful discussion of expected reaction rates and background rates, for it is after all the signal to noise that is the relevant thing. If the electron beam is modified, for example by changing the damping ring so as to make an electron beam more favorable for a gamma-gamma collider (for example with a higher bunch charge), then the luminosity may be increased to a value comparable with or, according to some speculations, even higher than that of an electron-positron collider, because beam-beam effects are less severe for gamma-gamma collisions.

### THE PARTICLE PHYSICS

A discussion of the laser that provides the photons for a gamma collider will be given later, coupled with information about the requisite optics in  $\gamma\gamma$  colliders. Here we comment briefly on some of the particle physics opportunities that such a device would provide. One of the most interesting physics programs is a measurement of the Higgs boson partial width into  $\gamma\gamma$ . This partial width is sensitive to physics beyond the Standard Model. A further discussion of the Higgs width is presented in the box on the next page.

Another interesting set of reactions is the decay of the Higgs into either  $b\bar{b}$  or  $ZZ$ . One could also study the  $CP$  eigenvalue of the Higgs by means of polarized gamma rays. It is important to note that in a  $\gamma\gamma$  collider one can use the full center-of-mass energy to produce Higgs bosons in

### Parameters for a Gamma-Gamma Collider as a Second Interaction Region for the Next Linear Collider

THE TABLE BELOW shows the main parameters of a gamma-gamma collider as the second interaction region of the NLC. The electron beam parameters up to the final focus system are taken to be the same as that for the first interaction region (for  $e^+e^-$  collisions). The horizontal and the vertical beam sizes in the case of the  $e^+e^-$  collisions are designed to be as small as possible within the constraints of the beam-beam effects. In the case of the  $\gamma\gamma$  collisions, the electron beam sizes at the interaction point (assuming the electrons are allowed to collide without being Compton scattered) are determined from different considerations. The vertical beam size should not be smaller than that determined by the  $1/\gamma$  spreading of the gamma rays between the conversion point and the interaction point. The horizontal beam size should be made as small as possible and consistent with various lattice design constraints of the final focus system (such as chromaticity control, beam blow-up caused by synchrotron radiation in the final quadrupoles known as the Oide effects, etc.) A reasonable choice appears to be that the beta functions in the horizontal and the vertical directions have the same values, 0.5 mm. This leads to a geometric electron-electron luminosity (again assuming the electrons are allowed to collide without being Compton scattered) of about  $1 \times 10^{34}/\text{cm}^2/\text{s}$ , which is about twice that of the  $e^+e^-$  collisions.

#### Electron beam parameters:

Electron energy	250 GeV
Repetition rate	90 bunches separated by 1.4 ns, 180 Hz
Particles per bunch	$N_e = 0.65 \times 10^{10}$
Normalized rms emittance	$\gamma\epsilon_x = 5 \times 10^{-6} \text{m} \cdot \text{r}$ , $\gamma\epsilon_y = 8 \times 10^{-8} \text{m} \cdot \text{r}$
Beta function at the IP	$\beta_x^* = \beta_y^* = 0.5 \text{ mm}$
Rms bunch length	$\sigma_z = 0.1 \text{ mm}$
Polarization	Fully polarized with helicity switching capability
CP-IP distance	$b = 5 \text{ mm}$

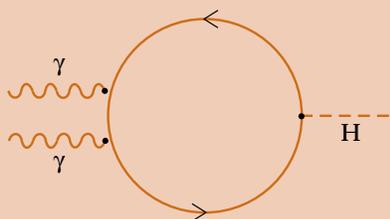
#### Laser beam parameters:

Wave length	$\lambda = 1.053 \mu\text{m}$
Micropulse energy	$A = 1 \text{ J}$
Rayleigh length	$Z_R = 0.1 \text{ mm}$
Rms micropulse length	$\sigma_{LZ} = 0.23 \text{ mm}$
Peak power	0.5 TW
Average power	16.2 kW
Polarization	Fully polarized with helicity switching capability

As noted before, the laser parameters are chosen to optimize the conversion of the electron beams into the high energy photon beams. With 1 joule per pulse focused to  $1 \times 10^{18} \text{ W}/\text{cm}^2$ , the probability of producing a high energy gamma photon per electron (the conversion efficiency) is about 65 percent. Therefore the luminosity of the  $\gamma\gamma$  collisions will be about  $(0.65)^2 = 0.43$  times the geometric electron-electron luminosity. However, the spectrum of the gamma-ray photons is fairly wide, with a peak near 83 percent of the electron energy. The luminosity near this peak, within a 20 percent bandwidth, is about 10 percent of the geometric luminosity, that is, about  $1 \times 10^{33}/\text{cm}^2/\text{s}$ .

## Higgs Two Gamma Width

**A** GAMMA-GAMMA collider is uniquely suited for a direct measurement of the partial decay width of a Higgs boson into two gamma quanta. The decay amplitude involves loops of any charged particles whose mass is derived from the Higgs mechanism, as shown in the illustration below.



*Feynman diagram for two photon decay of the Higgs boson. The loop can be of any charged elementary particles whose mass is generated via the Higgs mechanism.*

A measurement of the two-photon width would be a sensitive test of various models predicting higher mass particles without producing them directly in accelerators. There are several such models with two-photon couplings different from that of the Standard Model: supersymmetric models, technicolor models, and other extensions of the Standard Model.

the s-channel, whereas one would need to produce them in pairs at an  $e^+e^-$  collider. Thus the energy reach of a  $\gamma\gamma$  collider is much increased over that of an  $e^+e^-$  collider.

If supersymmetry exists, charged super particles can be produced at a  $\gamma\gamma$  collider with reasonable cross sections. The main source of background is  $WW$  pairs. Finally, if some of the particles in the Standard Model are composites, made of more fundamental particles, then they should either have excited states decaying into the ground state by emission of  $\gamma$  or  $Z$  particles, or anomalous interactions as low-energy limits of their form factors. If the  $W$ -boson is composite it may have an anomalous magnetic moment (or electric quadrupole moment), and either of these should be measurable given the copious production of  $W$  expected at a  $\gamma\gamma$  collider.

## THE LASER AND ITS OPTICS

Returning now to the technical problems associated with constructing a gamma collider, we need to discuss the laser optics in the conversion /interaction region. As the reader will appreciate, the region is very crowded. Furthermore, for detector reasons it is important that a minimum amount of mass is put into this region. The simplest configuration, and probably the one that will be used in practice, is to bring the opposing laser beam for each beam of electrons in from the opposite side, and after conversion dispose of (the essentially unaltered) laser beam.

In devising suitable optics for the intense laser beam we must consider a number of elements. First, trans-

missive optics are for the most part not feasible. Second, the overlap between the electron and photon beam must be good, that is, the two axes must be closely aligned, and the Rayleigh length must be adequately long. Because of the high peak and average power, the spent laser beam must be transported to an external beam dump. Because the two electron beams intersect at a small angle (so that the spent beams have separate channels to exit the detector), the used laser beam from one side will intersect the optics of the other side. Since two pulses will fall on the same mirror, it is necessary to locate mirrors at points where two pulses will not reach them at the same time. A possible mirror arrangement to bring the laser pulses to the conversion point and to dispose of them in the tight space limited by the vertex chamber, the mask, and quadrupoles is shown in the figure on page 22.

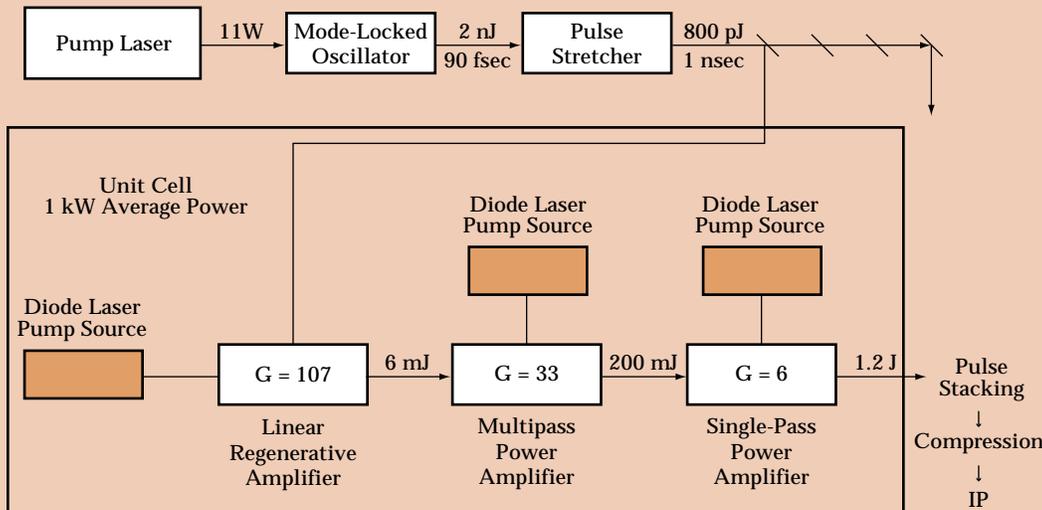
Finally, we must address the question of developing a laser for the purposes of a  $\gamma\gamma$  collider. The requirements listed in the box on page 19 are a pulse of about 1 joule, a pulse length of a few picoseconds (corresponding to a peak power of a few terawatts), an average power of 16 kilowatts (corresponding to about 100 bunches per pulse and a rep-rate of about 160 Hz), and variable polarization. Using the chirped pulse-amplification technique solid-state lasers already give peak powers that satisfy our requirements; they only fail our needs in that the average power is currently in the few watt range. Two critical technologies for constructing high average power lasers for a gamma-gamma collider

## Technologies for Constructing High Power Lasers

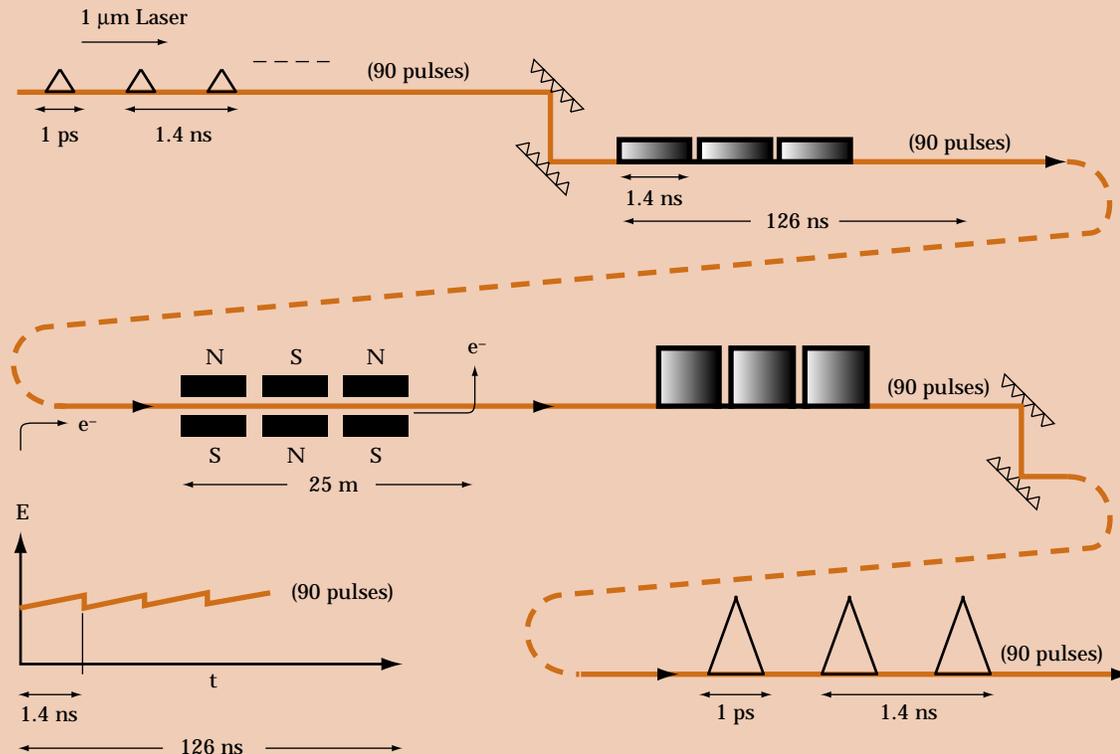
**HIGH POWER DIODE LASERS** for pumping and lasing materials that can handle high thermal loading are two major efforts on advancing diode laser technologies already under way as part of both military and civilian projects. Advanced solid state materials, either athermal glass hosts or new crystals specifically engineered for diode pumping, are also being actively developed. Based on these developments, the 16 kW laser needed for a gamma-gamma collider can be built out of 1 kW unit cells, schematically illustrated below. There needs to be two such lasers, one for each opposing electron beam. If the laser pulses could be reused, for example

by storing them in an optical cavity, the requirement on the average power could be significantly relaxed.

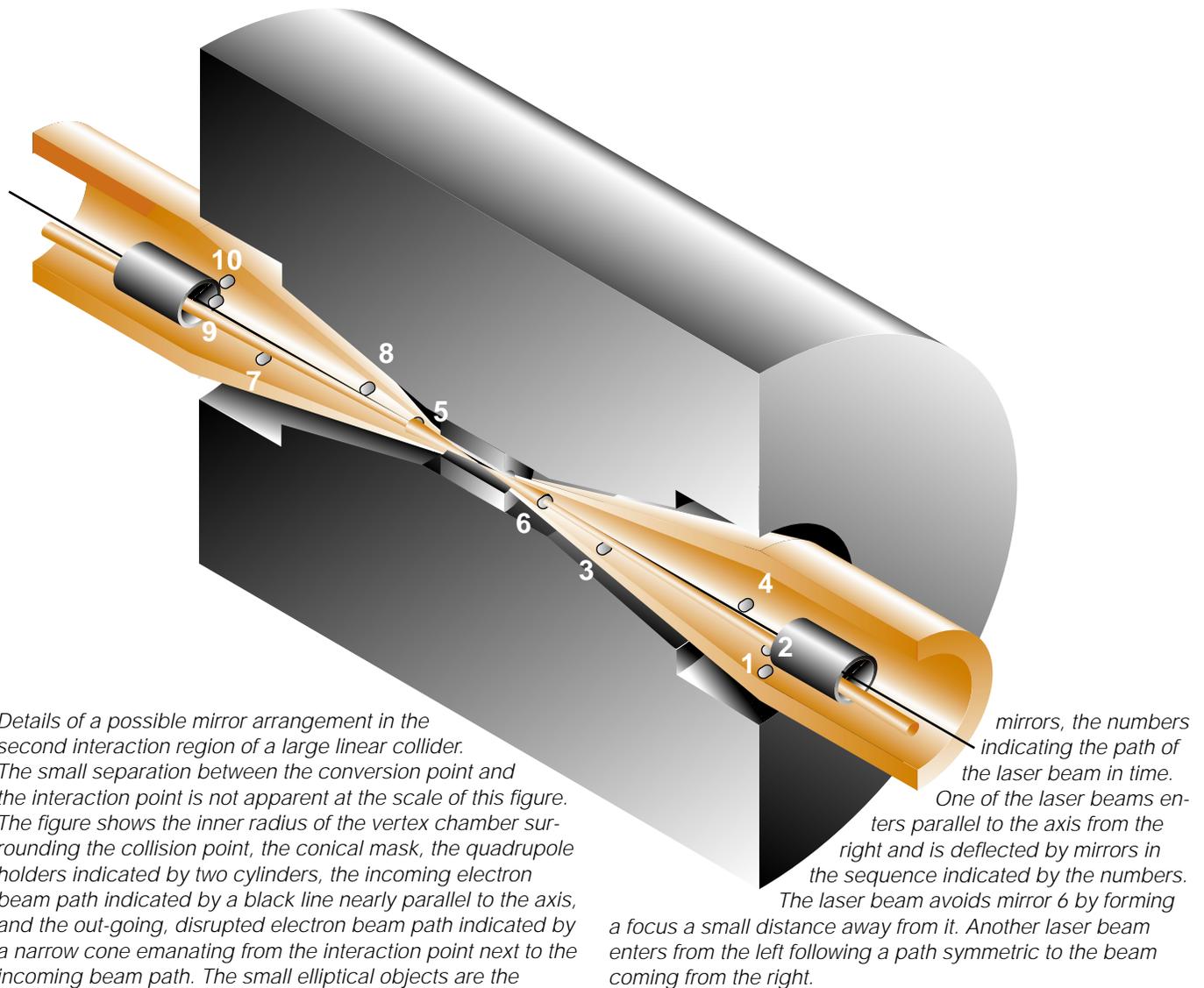
Another possibility is a free-electron laser (FEL). So far, FELs have generated neither the requisite peak power (gigawatt has been achieved so far) nor average power (a few watts is the present situation). On the other hand, they have considerable promise: The chirped pulse-amplification technique can be applied in an induction-linac-based free electron laser to produce optical pulses of the required characteristics for a gamma-gamma collider, as shown in the bottom figure.



Left: A solid-state laser concept for the gamma-gamma collider built of 1 kW unit cells. All of the cells are fed by a single phase-stabilized oscillator, ensuring synchronization of all pulses with the electron beam. Each of the unit cells consists of a series of diode-pumped, solid-state laser amplifiers. The pulses are subsequently compressed in a grating pair and stacked into a single pulse train.



Left: A free-electron laser that employs a pulse stretcher, an FEL amplifier, and a pulse compressor.



are underway and described in the box on the previous page.

The necessary research and development to make it real will consist primarily of work on laser development and on optical elements. Much of this work is “table top,” that is, it does not require a high energy beam of electrons. However, to have adequate confidence in the concept (so as to be able to build it into an NLC), it will be necessary to develop a  $\gamma\gamma$  device working in the tens of GeV range. This is necessary in order to study backgrounds, detector issues (such as compatibility with the required optical elements),

and lifetime of optical elements in a collision-point background.

### THE PROSPECTS

A very interesting possibility is the use of the SLAC Linear Collider (SLC), that is, extending the capability of the SLC (call it SLC II) to  $\gamma e^-$  collisions. Collisions of  $\gamma e^-$  have a greater energy reach and thus are much more interesting, at the SLC, for particle physics than  $\gamma\gamma$  collisions. Either, of course, is of interest for technological studies. One could study the spin asymmetry in the process  $\gamma e^- \rightarrow Ze^-$ , which tests the Standard

Model and probes anomalous  $\gamma ZZ$  couplings, while energy scans near the threshold of  $\gamma e^- \rightarrow W^- \nu$  may lead to an improved  $W$  mass measurement.

We have seen that a  $\gamma\gamma$  collider is technically possible and that it would open up important new possibilities for particle physics research. Such a project will require a significant investment in preparatory research and development. Our hope is that the high energy community will come to realize the outstanding promise of the gamma-gamma collider idea and will respond by offering its participation and support. ◻

# From Superconductors to Supercolliders

by LANCE DIXON

*Superconductivity sheds light on the  
mystery of the Higgs phenomenon  
in the weak interactions.*

**I**N THE NINETEENTH CENTURY, the concept of the “luminiferous ether” held sway in physics. The ether was imagined as a material substance, permeating all of space and supporting the propagation of light and other electromagnetic waves. Just as sound travels at a unique velocity through air, so should light travel at a unique velocity through the ether. An observer on Earth, moving through the ether, should then see different velocities for light moving parallel or perpendicular to the Earth’s motion. In 1887, Albert Michelson and Edward Morley found the velocity of light to be independent of direction, and the theory of the ether thus came crashing down, to be replaced by the theory of special relativity.

In the twentieth century, in the modern theory of the weak interactions, the ether has been resurrected in a relativistic and quantum-mechanical form. This new kind of ether—known as the Higgs field, after Peter Higgs,

*Nowadays W's and Z's are produced routinely at particle accelerators, yet the Higgs mechanism remains the most poorly tested part of the electroweak theory.*

one of the first to postulate it—is supposed to permeate all of space, and to be responsible for giving mass to the  $W$  and  $Z$  bosons, carriers of the weak force, and to the quarks and leptons that constitute all matter. What is this new ether made out of? Why should we believe it exists, and how does it give mass to elementary particles? How is it related to the Higgs particle, the main quarry of present and planned supercolliders? The answers to these questions, where known, are not that easy to describe. Fortunately, though, there is a fruitful and precise analogy between the Higgs mechanism for generating the  $W$  and  $Z$  masses and the phenomenon of superconductivity. In this analogy, the entire Universe is a superconductor, not for electromagnetism but for the weak interactions. The abstract quantum fields associated with a yet-undetected Higgs particle (or particles) are made more concrete in the analogy; they are represented by the concerted motions of many ordinary electrons in a solid. These parallels have been known for a long time; indeed they inspired many particle physicists in the late 1950s and 1960s, as the Higgs mechanism was being incorporated into models of the weak interaction. As we'll see, the analogy could operate at more than one level; it is certainly valid at a descriptive, phenomenological, or macroscopic level, but it might also work at a deeper one.

Before turning to superconductivity, let's take a glimpse at what we have learned about the weak interactions in the one hundred years since Henri Becquerel discovered them. The most obvious property of

the weak force is, naturally, its weakness. For example, a neutrino can easily pass through the entire Earth without interacting. The electroweak theory, initially developed by Sheldon Glashow, Steven Weinberg, and Abdus Salam in the 1960s, explains the weak interactions in terms of a triplet of particles called vector bosons, the electrically charged  $W^+$  and  $W^-$  and neutral  $Z$  bosons. These particles are analogous to the photon (the quantum of light) which carries electromagnetism. The photon is exactly massless because of a certain symmetry, called gauge invariance. The corresponding symmetry in the electroweak theory is said to be spontaneously broken by the Higgs mechanism, giving a mass to the  $W$  and  $Z$  bosons. We'll use the superconductor analogy to explain what that last sentence really means!

The weakness of the weak interactions is due to the large  $W$  and  $Z$  masses, almost one hundred times the proton mass. In the scattering of a neutrino off a nucleon, there is not enough energy to make a real  $W$  or  $Z$  particle. However, the uncertainty principle of quantum mechanics allows a "virtual"  $W$  or  $Z$  to be produced, but only for a short time, related to its large mass; this results in a very small probability for the scattering to occur. Another way of saying the same thing is that the

field of the  $W$  or  $Z$  only extends a tiny distance away from the neutrino (see below), thus making it hard for the neutrino to find a quark in the nucleon to interact with. Nowadays, real  $W$ 's and  $Z$ 's are produced routinely at particle accelerators, and their properties agree extremely well with theoretical predictions; yet the Higgs mechanism remains the most poorly tested part of the theory. The minimal version of the mechanism predicts a single new particle, the Higgs boson, but this version is probably just a surrogate for more complicated dynamics. One of the prime rationales for building new supercolliders (using superconducting magnets!) is to shed light on the mystery of the Higgs sector.

**SUPERCONDUCTIVITY** is itself a pretty remarkable phenomenon. Cool a chunk of lead or niobium down to a few degrees above absolute zero and its electrical resistance completely vanishes. Once an electrical current is established in a ring of superconducting material, it circulates essentially forever without an external power source. Magnets wound from superconducting wire have huge advantages over conventional magnets for many applications, ranging from magnetic resonance imaging to the bending magnets in modern proton storage rings such as the Fermilab Tevatron and the Large Hadron Collider to be built at CERN. Superconductors also expel external magnetic fields (the Meissner effect). The photograph on the next page shows a superconductor expelling the field of a permanent magnet and thereby levitating the magnet. This effect may eventually

find practical application in magnetically-levitated high-speed trains, for example.

A superconductor only works below some critical temperature  $T_c$ . Above  $T_c$  thermal effects disrupt the mechanism and the material becomes “normal.” Conventional metallic superconductors have critical temperatures ranging from a few to tens of degrees kelvin. In the past decade, however, a new class of ceramic materials has been discovered, which superconduct well above the boiling point of liquid nitrogen (77 degrees kelvin). Because liquid nitrogen is much cheaper to produce and store than liquid helium, there is considerable interest in developing commercial devices from these high- $T_c$  superconductors. The critical temperature for the weak interactions, viewed as a superconductor, is a quadrillion degrees kelvin ( $10^{15}$  K) making this truly the ultimate high- $T_c$  superconductor! Above this temperature, the masses of the  $W$  and  $Z$  bosons, and all the quarks and leptons, should vanish. Testing this conclusion in the laboratory would require heating a sample of the Universe to  $10^{15}$  K in order to make it go normal. One has to go back to the first instants of the Big Bang to achieve those conditions. In an accompanying article in this issue, Eric Sather describes how the predominance of matter over antimatter that we observe today may have originated as the Universe cooled through this electroweak phase transition. Since that moment, the Universe has been stuck firmly in the superconducting phase.

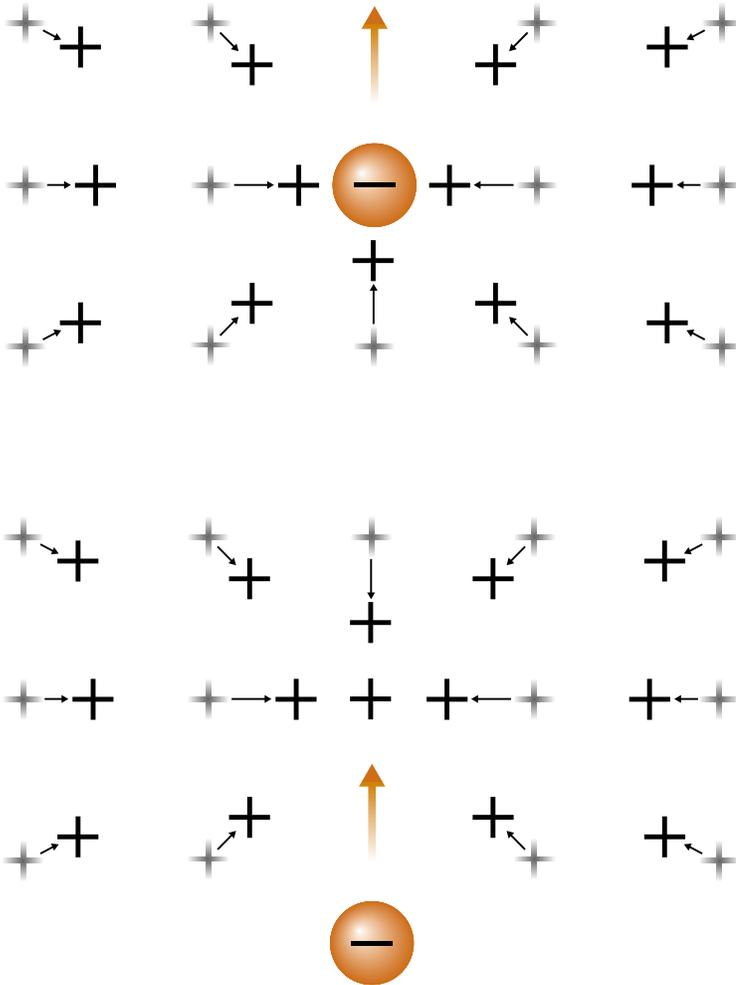
The mechanism of superconductivity in conventional supercon-

ductors has been well understood since the pioneering work of John Bardeen, Leon Cooper and J. Robert Schrieffer (BCS) in the mid 1950s. At room temperature these materials are metals—they have conduction electrons which are not localized on any particular atom but roam around the entire material, while the remaining positive ions stay fixed in a crystal lattice. A suggestive name for this collection of electrons is the Fermi sea. Like water molecules in the ocean, the electrons are free to move, but different electrons must occupy different locations, so the sea fills up to some level. Also like the ocean, the depths are calm, and all the action is at the surface.

As the temperature is lowered, an attractive interaction between conduction electrons near the surface of the Fermi sea binds them into “Cooper pairs.” It is somewhat surprising that a pair of electrons should attract, since the electrostatic (Coulomb) force between two like-sign charges is repulsive. In a metal this force is reduced, or screened, at longer distances: Other conduction electrons feel the charges of the two electrons in question, and move out of the way; this leaves a positive charge behind that makes the charges of each electron look less negative. Screening reduces the strength of the Coulomb force, but it remains



*A high- $T_c$  ceramic superconductor levitates a magnet. (Courtesy Science Kit & Boreal Laboratories)*



*Top: An electron moving through a solid polarizes the lattice of positive ions towards it.*

*Bottom: A second electron is attracted to the excess of positive charge.*

repulsive. The attractive interaction results instead from collective motions of the lattice of positive ions, whose modes of oscillation are called phonons. One electron polarizes the lattice by attracting the positive ions toward it; a second electron is then attracted to this build-up of positive-charge, as illustrated by the figure on the left. The net sum of the screened Coulomb repulsion and the lattice (phonon) mediated attraction can be attractive, and in that case it turns out that a Cooper pair will always form at sufficiently low temperature.

Superconductivity does not result merely from the formation of Cooper pairs, but rather from getting a significant fraction of them to occupy a single quantum state, a phenomenon known as Bose condensation. By themselves, electrons are forbidden from Bose condensing because they are fermions, not bosons—the Pauli exclusion principle states that no two of them (let alone a macroscopic number) can occupy the same state. In effect, the exclusion principle

holds up the Fermi sea. (In an atom, it similarly explains why the electrons don't all collapse onto the lowest energy atomic orbital.) On the other hand, a Cooper pair is a boson (see the box on the next page), and the Pauli exclusion principle does not apply to bosons. Instead, bosons prefer to occupy the same, lowest energy state, if the temperature is low enough. The term condensation refers to the sudden onset of the new phase as the temperature is lowered to the critical temperature  $T_c$ , reminiscent of the condensation of vapor into liquid. If you are trying to visualize what is happening to the Fermi sea, a better picture is the freezing of its surface. In a single quantum state, all the electrons must move in lockstep, like a sheet of ice floating on the surface. In reality, a complicated dance is going on, as individual electrons try to stay out of each other's way, yet as pairs they move together.

How does the macroscopic occupancy of a single quantum state endow a superconductor with its remarkable properties? In an ordinary metal, the electrical current is carried by a large number of electrons, with no quantum-mechanical relation to each other. Electrical resistance is generated when individual electrons (like individual water molecules) moving with the current scatter off impurities or phonons, dissipating energy into the Fermi sea. Energy can be lost in arbitrarily small units because there are empty quantum states nearby for the electrons to scatter into, and because no other electron “cares” what happens to a given electron. In contrast, in a superconductor, current is carried by

## Fermions, Bosons and Cooper Pairs

**E**LECTRONS ARE CALLED fermions because they carry  $1/2$  unit of the basic quantum of angular momentum (Planck's constant,  $\hbar$ ), and therefore they obey Fermi-Dirac statistics, which means that the quantum-mechanical wave function  $\psi$  for  $n$  electrons has to be antisymmetric under the exchange of each pair of electrons. For example, under the exchange of electrons 1 and 2,  $\psi$  picks up a minus sign,

$$\psi(r_1, s_1; r_2, s_2; \dots; r_n, s_n) = -\psi(r_2, s_2; r_1, s_1; \dots; r_n, s_n),$$

where  $r_i$  and  $s_i$  label the positions and spins of the electrons.

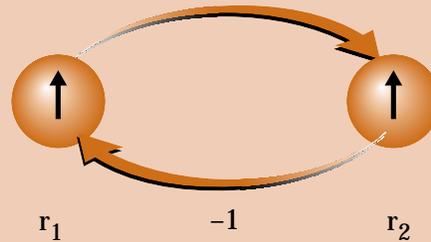
A particular consequence of these rules is the Pauli exclusion principle, that no two electrons can occupy exactly the same state. This means that they cannot condense directly. However, a pair of electrons, in any quantum state, must have an integer unit of angular momentum—the two spin  $1/2$ 's can add to make either 0 or 1, and the orbital angular momentum is always an integer. Thus the pair of electrons is a boson, obeying Bose-Einstein statistics, which requires the wave function to be symmetric under the exchange of any pair of bosons. The figures below illustrate how a system of two electron pairs can be symmetric

the Bose condensate of doubly-electrically-charged Cooper pairs, flowing in quantum-coherent unison, like the rigid ice sheet described earlier. Local effects are simply unable to stop its progress, and so the electrical resistance is precisely zero.

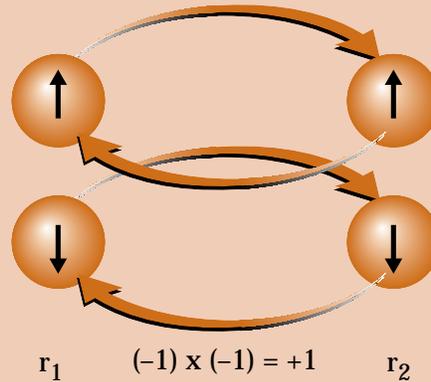
As the temperature rises, more and more of the electrons are kicked into excited states instead of the Bose condensate ground state (the ice sheet melts). However, the condensate can carry all the current with perfect conductivity, and so its depletion does not matter much until it disappears altogether at the critical temperature.

The Meissner effect—expulsion of a static magnetic field—requires this perfect conductivity. In response to an externally applied magnetic field, perpetual eddy currents circulate in the superconductor, producing an internal magnetic field that exactly cancels the applied one. An ordinary metal also generates eddy currents in response to an applied magnetic field, but these die out quickly owing to electrical resistance, and then the magnetic field penetrates. (One can observe the effects of these currents by placing a slab of aluminum vertically in a strong vertical magnetic field—it can take seconds for the slab to fall, as it tries to keep the magnetic field lines from penetrating!)

Before we can return to the weak interactions, we need to know that there are two important length scales in a superconductor. The first measures how efficiently the condensate expels a magnetic field. In fact, the expulsion is not quite complete (see the figure on the next page). There is a thin layer of depth  $\lambda$ , called the London penetration depth, over

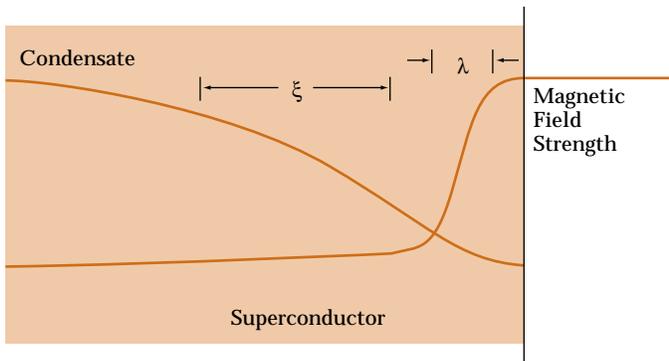


*The wave function for two electrons with the same spin (denoted by  $\uparrow$ ) at different locations  $r_1$  and  $r_2$ , must pick up a minus sign when  $r_1$  and  $r_2$  are exchanged.*



*The wave function for two pairs of electrons, at  $r_1$  and  $r_2$ , picks up two minus signs when the pairs are exchanged. (Opposite spins within each pair are also required by Fermi-Dirac statistics.)*

under exchange of the pairs, yet antisymmetric under exchange of individual electrons. In a Cooper pair—the lowest quantum state for a pair of electrons with a phonon-mediated attraction—the electrons have zero orbital angular momentum, and are antisymmetric in spin (spin 0). Thus a Cooper pair has zero total angular momentum. This did not have to be the case, and indeed there is mounting evidence that for high- $T_c$  superconductors the pairs that condense have spin 2, not 0.



*Behavior of the magnetic field and the superconducting condensate near the surface of a superconductor to which an external magnetic field has been applied.*

which the magnetic field drops exponentially to zero. The value for  $\lambda$  depends on the material, but a tenth of a micron is typical. In this same region, the magnetic field perturbs and reduces the condensate. The second length scale, called the coherence length  $\xi$ , governs how fast the condensate snaps back to its bulk value once the magnetic field has gone to zero. (Here you can imagine the stiffness of that ice sheet in response to some deformation, if you like.) Depending on the material, the coherence length can be either longer or shorter than the magnetic field penetration length; these two classes of superconductors (known as type I and type II) turn out to have quite different magnetic properties. In any case, it is this response of the condensate to a local disturbance that is the superconductor analog of the Higgs particle.

Finally, we have to accept that in quantum field theory, distance scales and energy-momentum scales are related through the uncertainty principle, and particles and fields become one and the same. A particle of mass  $M$  has associated with it a Compton wavelength given by  $\hbar/Mc$ , where  $\hbar$  is Planck's constant and  $c$  is the speed of light. The Compton wavelength tells how fast the field of a particle falls off with distance. For example, the exponentially falling penetration of the magnetic field into a superconductor means that the photon has acquired a mass  $M_\gamma = \hbar/\lambda c$ .

(The expert reader will object that this is not a true photon mass, because the condensate is nonrelativistic, and the electric field behaves differently from the magnetic field. But it's close enough for our purposes.)

**A**T LAST we are ready to draw the analogy between a superconductor and the Higgs mechanism. We imagine that there is a Bose condensate permeating all of space, the occupied quantum state of some boson that carries a charge under the weak interactions, just as the Cooper pairs carry an electric charge ( $2e$ ). A quark or lepton carries both weak and electromagnetic charges, so it produces fields of  $W$ 's,  $Z$ 's and photons around it. But the condensate responds to these  $W$  and  $Z$  fields by producing currents that conspire to cancel out the fields—expel them from the vacuum—except for a tiny region near the quark or lepton. That region is of order the  $W$  Compton wavelength, which we know to be 0.0025 fermi. By comparison, a proton is roughly a fermi ( $10^{-15}$  meters) across. The photon field, on the other hand, is not screened and extends to infinite distances (unless the quark or lepton happens to be inside a real superconductor!). Using the inverse of the penetration length as a measure of efficiency, the Higgs mechanism is about a billion times more efficient at screening the weak interactions than an ordinary superconductor is at screening electromagnetism. The screening is so efficient that it is hard to imagine the weak analogs of electrical currents, arising from the motion of a bunch of quarks and

*If we are to try  
to extend the Higgs  
superconductor analogy  
to a deeper, microscopic  
or mechanistic level,  
we must first ask:  
Is the Higgs field  $\phi$   
of the Standard Model  
really a fundamental  
scalar field?*

leptons. The  $W$  and  $Z$  fields fall to zero even before the next particle in the bunch is reached, for any bunch we can conceive of making in the laboratory.

In principle, a quark or lepton should also be able to “kick” the condensate, producing the exponentially decaying field of the Higgs particle itself, whose decay length (the coherence length) is also tiny. (Because we don’t know the mass of the Higgs particle, or particles, we don’t know this precise length. Whether it is large or small should control the qualitative nature of the Higgs sector—“weakly-coupled vs. strongly-coupled.” The analogous ratio of the coherence length to the penetration length similarly controls many properties of superconductors.) If the condensate were kicked hard enough, in a violent collision, a real Higgs particle could pop out. In practice, it is believed that the light quarks and leptons (which are all that we can easily fashion into particle beams), couple only very weakly to the Higgs, and most proposals for making the Higgs particle involve perturbing the condensate indirectly, by means of vector boson fields (the  $W$ ,  $Z$ , photon, or even gluon), much as the magnetic field perturbs the condensate in a superconductor.

We have seen that the phenomenon of superconductivity closely parallels the Higgs mechanism for  $W$  and  $Z$  mass generation, and that the two important length scales in a superconductor,  $\lambda$  and  $\xi$ , when reinterpreted in the electroweak context, become the Compton wavelengths, or inverse masses, of the  $W$  boson ( $\lambda$ ) and of the Higgs boson ( $\xi$ ). However, the analogy so far has been at the

macroscopic or phenomenological level. We did not ask what the Higgs condensate is made of. Likewise, we made no reference to the details of the BCS electron-pairing mechanism, beyond the fact that it generates a condensate. Indeed, in 1950, seven years in advance of the BCS theory, Vitaly L. Ginzburg and Lev D. Landau were able to accurately describe many phenomena of superconductivity without recourse to a microscopic theory. Their key advance was to account for the quantum-mechanical nature of the Bose condensate, and its ability to be deformed, by introducing a complex wave function  $\psi(x)$ , whose magnitude-squared  $|\psi(x)|^2$  they interpreted as the local density of the condensate at a point  $x$ .

The minimal Higgs mechanism in the electroweak theory can be obtained from the Ginzburg-Landau description of superconductivity by replacing their  $\psi$  by the Higgs field, usually called  $\phi$ . Just two additional modifications have to be made:

First, the theory must be relativistically invariant—or else Michelson and Morley would be very unhappy with us! A superconductor is not Lorentz invariant, in the sense that (as with any ordinary material) there is a preferred frame

where it is at rest. Technically, the necessary alteration to the theory is easily carried out; we just declare the Higgs to be a Lorentz-invariant (scalar or spin-zero) field. Conceptually, though, this change makes it much harder to visualize what kind of “material” makes up the Higgs field.

Second, a single complex value (at each point in space) will not suffice for  $\phi$ ; at least two turn out to be required in order to give mass to the triplet of weak vector bosons,  $W^+$ ,  $W^-$  and  $Z$ . Of course this vastly oversimplifies the way in which the electroweak theory was really developed.

**I**F WE ARE TO TRY to extend the Higgs-superconductor analogy to a deeper, microscopic or mechanistic level, we must first ask: Is the Higgs field  $\phi$  of the Standard Model really a fundamental scalar field? Or is it instead a cooperative effect of some more elementary objects—perhaps fermions—as in the BCS theory? Proponents of supersymmetry, a symmetry that transforms fermions into bosons and vice-versa, argue that it is just as natural for there to be fundamental scalars as fermions, and that one of them could and should be the Higgs. Proponents of technicolor argue that no fundamental scalar particle has been detected to date, and that the Higgs is more likely to be a composite particle, a pair of fermions bound together by a new strong gauge interaction. This new force is imagined as a super-strong version of the color force that binds quarks into nuclei, hence the name technicolor.

The jury is still out on which alternative is correct (if either!). Yet

it is striking how closely technicolor parallels the BCS theory at the microscopic level. The technicolor interaction is postulated to strongly bind together some new fermions, technifermions, into a state where they are moving at nearly the speed of light. Technifermions would also carry weak charges, so that a Bose condensate of pairs of them would give mass to the  $W$  and  $Z$  bosons, just as the superconducting electron-pair condensate gives mass to the photon. The BCS theory treats highly non-relativistic electrons that are weakly bound by phonons, so the two mechanisms still sound pretty different. However, remember that the important electrons for the BCS mechanism are those at surface of the Fermi sea. These electrons behave very much like relativistic particles—if they have a little extra momentum, their extra energy is proportional to that, like the relation for a particle moving at the speed of light,  $E = pc$ , except that the speed of light  $c$  gets replaced by something called the Fermi velocity  $v_F$ . The technicolor vector bosons and the phonons have this same massless, linear relation as well. The coupling between electrons and phonons may look weak, but like all couplings in relativistic field theories, it depends on the distance-scale. Indeed, both the technicolor and the electron-phonon couplings can be shown to change slowly, from a weak value at “very short” distances (the Planck scale for technicolor, the Debye frequency for BCS) to an arbitrarily strong value at “long” distances (the weak scale for technicolor, the coherence or penetration length for

BCS). So they are really very similar phenomena, something that was recognized, even before technicolor was fully developed, by Yoichiro Nambu and Giovanni Jona-Lasinio. Of course, just because the deeper theoretical analogy holds doesn't mean that technicolor is right.

While conventional superconductors are quite well understood today, via the BCS mechanism and many subsequent developments, the recent class of high- $T_c$  ceramic superconductors is much more of a mystery.

In some sense they are at the stage that the weak interactions were at in the 1960s and early 1970s, when the basic phenomenology, the full pattern of symmetry breaking, was not yet clear. Was there a  $Z$  boson or not? How were the  $W$  and  $Z$  masses related? In the case of the new superconductors, the analogous questions are about the charge of the condensate and its angular momentum, or spin. The charge appears to be  $2e$ , as for conventional superconductors, which means that an electron-pairing mechanism of some kind should be the culprit, but the BCS mechanism is probably too weak to do the job. The spin appears to be two, which means that the new condensate is not rotationally invariant, unlike the BCS condensate and Higgs condensate. As with the Higgs mechanism of the Standard Model, there are many strong opinions about the microscopic physics underlying high- $T_c$  superconductors, but no complete consensus. Only time and experiment will tell. ○

# The Mystery of the Matter Asymmetry

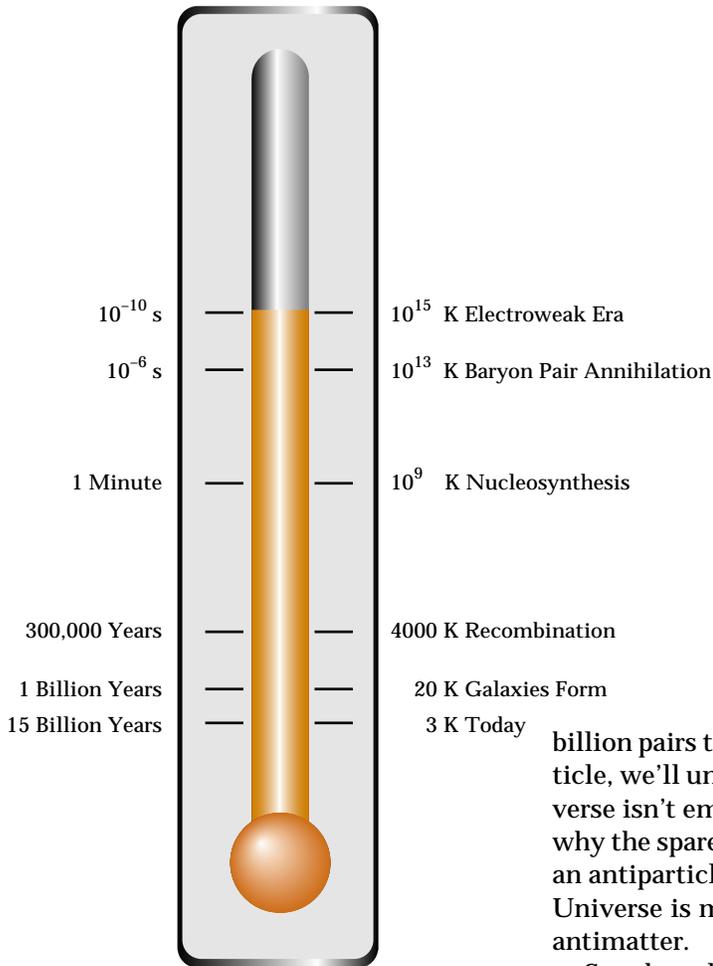
by ERIC SATHER

*Why is our Universe made  
of matter and not antimatter?*

*The answer might be found  
in the laws that govern  
elementary particles.*

**W**E'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*  $\rightarrow$  *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

*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.*

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.

### ELECTROWEAK BARYOGENESIS

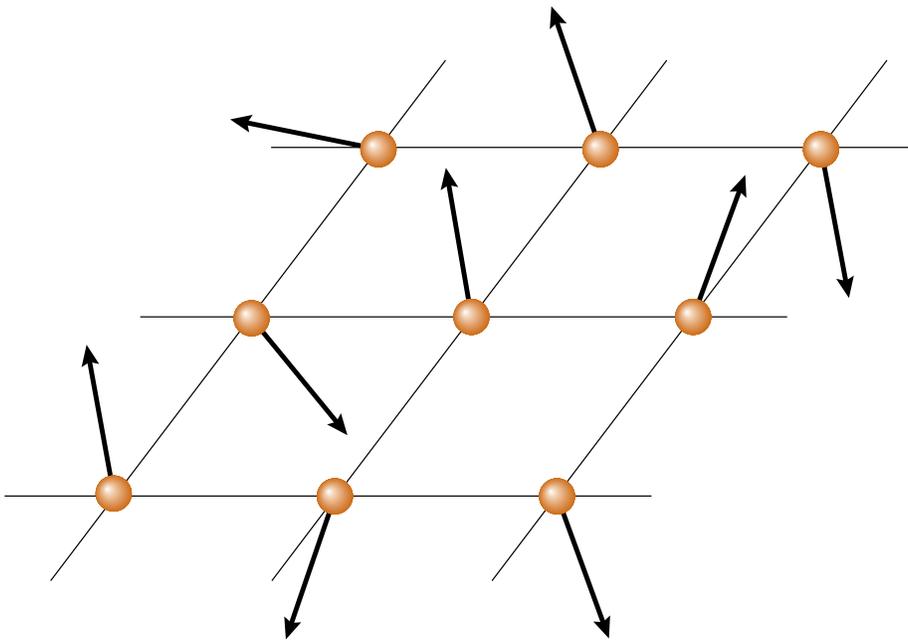
To see how the baryon asymmetry could be produced by electroweak baryogenesis, we need to know some of the basics of electroweak physics. The electroweak theory summarizes our deepest insights into the ultimate laws of nature. It synthesizes the electromagnetic theory of charges and light and the weak theory of nuclear  $\beta$ -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

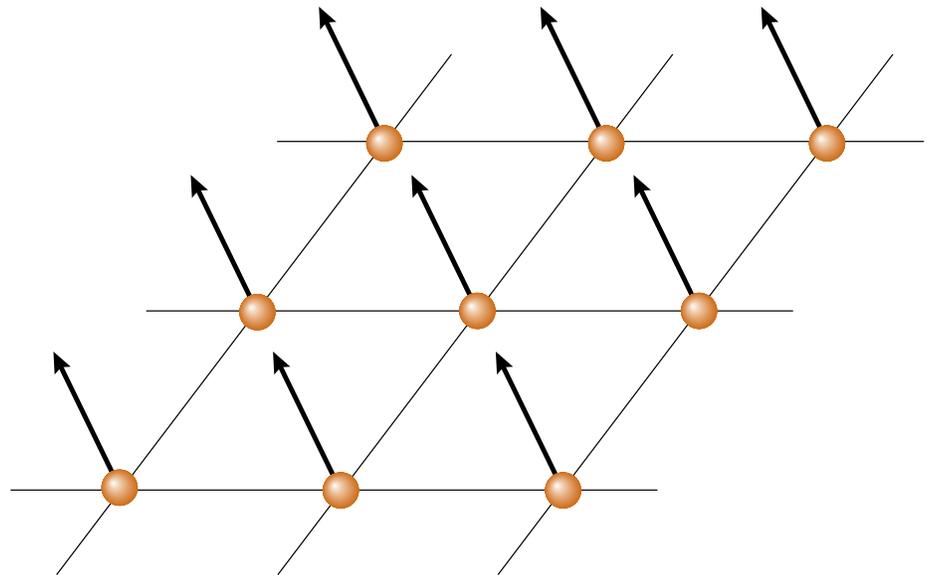
What about the other Sakharov conditions, for instance baryon-number violation? At first glance, the electroweak theory appears to conserve baryon number; there are no explicit interactions that change it. However, due to quantum-mechanical subtleties, there are baryon-number violating processes.

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

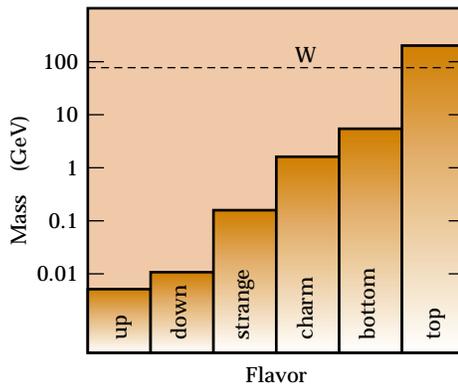


operation  $CP$  turns a particle into an antiparticle with reversed momentum but identical spin. Since both  $C$  and  $CP$  relate particles to antiparticles, if either were a symmetry of the laws of nature, particle production would always be countered by equal antiparticle production, and no baryon 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  $\beta$ -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	$\frac{2}{3}$	up	charm	top
	$-\frac{1}{3}$	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

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.

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

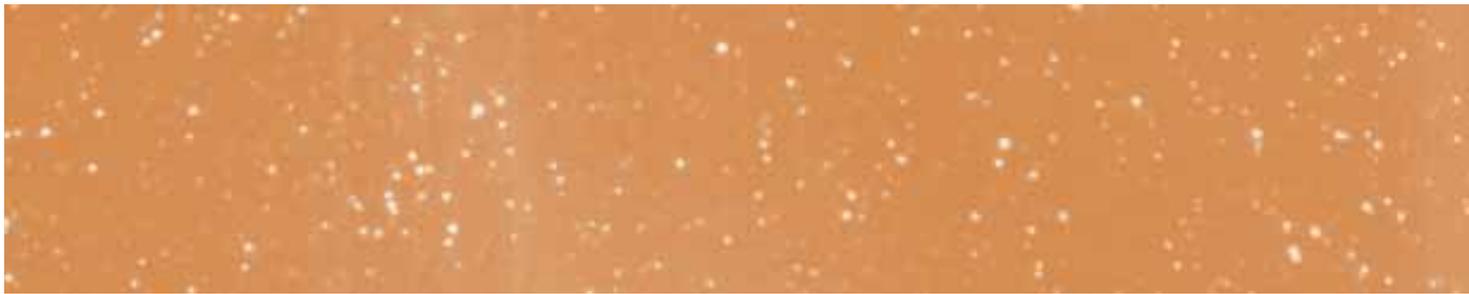
our understanding of physical laws, either at the electroweak scale or else at an even deeper level.

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. 



# Part II The Department of CELESTIAL MAGNETISM

by VIRGINIA TRIMBLE

*The larger the magnetic field, the stronger our ignorance.*

—Virginia Trimble (30 years after Nordwijk)

**P**ART I EXPLORED PLANETS AND STARS, whose magnetic fields are moderately well understood, but not energetically important. In Part II, we\* move on to interstellar, galactic, intergalactic, and cosmic magnetism, sites where the fields are energetically and dynamically much more important, but a good deal less well understood.

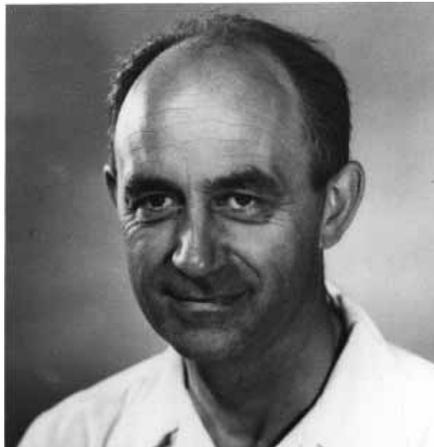
## BEGINNINGS

Among the good things to come out of World War II was an enormous boost to radio astronomy, in the form of left-over radar dishes and people with the vision to point them up rather than sideways. This included John Hey, Martin Ryle, and Bernard Lovell in England and Bernard Mills and John Bolton in Australia. They had a meager foundation to build on—the recognition by Karl Jansky in about 1935 that the disk and center of our galaxy were a major source of noise interfering with trans-Atlantic radio

*\*Or in any case I; but please do come along if you don't have anything else that absolutely has to be done in the next 20 minutes or so.*



*Hannes Alfvén. Early enthusiast for a large-scale galactic magnetic field and galactic cosmic rays. Alfvén waves (magnetohydrodynamic or MHD) constitute an important solution of Maxwell's equations not discovered by Maxwell.*



*Enrico Fermi. Perhaps the last person of whom one could say simply "physicist," meaning both theory and experiment on many topics. His mechanism for accelerating and confining cosmic rays requires a galactic magnetic field.*



*Jesse Greenstein. Among the first optical astronomers to take radio astronomy seriously and co-proposer of a mechanism for aligning interstellar grains so that they will polarize passing starlight. He was my first co-author (on white dwarfs, but not magnetic ones).*

telephony (he worked, of course, for Bell Labs) and by Grote Reber in 1944 that the sun is the brightest thing around at any wavelength (he was the quintessential backyard astronomer, having built his own radio reflector with which he also found radio emission in the direction of Cygnus).

The post-war heroes, struggling with high noise and low angular resolution, succeeded in demonstrating the existence of compact sources, including the already notable Crab Nebula (remnant of a supernova seen in 1054), but also others that didn't seem to be any particular interesting place at all, like Reber's Cygnus A.

The main stream (optical) astronomical community reacted with profound apathy, even when Reber, and soon after Jesse Greenstein, showed conclusively that known emission processes failed by orders of magnitude to account for what was being seen. At one point, the editor of the *Astrophysical Journal* thought it necessary to print an editorial assuring potential authors that papers in radio astronomy were not automatically rejected. He was not entirely believed.

Meanwhile, things were stirring on two other fronts. First, Hannes Alfvén and Enrico Fermi had been worrying about how to confine cosmic rays within the Milky Way and how to shove them up to relativistic energies in the first place. They concurred in 1949 papers that a large-scale galactic magnetic field of about  $10\mu\text{G}$  would be useful in both contexts. It is conceivable that this is the last (and first?) time that Alfvén agreed with anyone, but further research is needed on the point.

Second, back at the traditional optical ranch, John Hall and William Hiltner had gone out to look for polarization of starlight that, according to Chandrasekhar, should result from electron scattering of radiation in the atmospheres of hot (blue) stars. In 1949 they reported success of a sort, in separate papers (the collaboration by then having fallen apart, never to be reassembled—magnetic fields seem to have that effect on people). The light of many blue stars was indeed linearly polarized at the 1–3 percent level. But a map of polarization angles over the sky showed large-scale structure, mostly parallel to the plane of the Milky Way, with some loops and



*John Hall. Co-discoverer of interstellar polarization. He and Hiltner died within about a month of each other in fall 1991.*



*William Hiltner. Co-discoverer of interstellar polarization (in his days at the University of Michigan).*



*Josef Shklovsky (right) and Harold Zirin. Shklovsky was an early proponent of the synchrotron mechanism to account for continuum radiation from supernova remnants and active galaxies. Zirin, an expert observer of solar magnetic activity, translated and edited Shklovsky's memoirs with his wife Mary.*

wiggles, which made no sense as an effect of processes in the stars themselves.

The three fronts began to converge in the early 1950s. First, Alfvén and Herlofson said that synchrotron radiation (emitted by relativistic electrons spiraling in magnetic fields) was the key to understanding radio waves from the discrete sources. In the very next, also 1950, volume of *Physical Review*, Karl Kiepenheuer proposed the same mechanism for diffuse galactic emission. Ludwig Biermann imagined that such a field might be in equilibrium with galactic turbulent gas motions, and so have a strength of about  $10 \mu\text{G}$ .

Next, Leveritt Davis, Jr. and Jesse Leonard Greenstein proposed that the Hall-Hiltner polarization was light scattered by dust grains that had been lined up in a large-scale galactic magnetic field. The mechanism they proposed may or may not be the dominant one—it depends on the shapes of the grains and what they are made of. But their explanation of the polarization of starlight is now always advertised as the discovery of the galactic field. Jesse has said that he had magnetism on the brain at the time because of having heard Fermi talk about the

cosmic-ray problem at Chicago, thus tying directly to the third, theoretical, front.

Josef Samuelevich Shklovsky (his spelling, but authority has preferred Iosef, Shklovskii, and Shklovskij) pushed synchrotron emission upward to visible wavelengths for the Crab Nebula, predicting optical polarization of lots more than 1–3 percent, which his fellow countryman V. A. Dombrovsky (similar choice of spellings) found in the next year, 1954. The polarization of the Crab radio radiation turned up in 1959 in data collected by D. Kuz'min and V. A. Udal'tov (though many casual histories mention only a later American paper). And pretty soon the sky was filled with galactic and extragalactic polarized sources of synchrotron radiation.

Polarization of the diffuse galactic radio emission was firmly established (after much hard looking) in 1962 work by Gart Westerhout and Richard Wielebinski (each with several colleagues, named, of course, al.). The modern era can reasonably be said to begin with the 1968 discovery by Gerrit Verschuur of Zeeman broadening of the 21 cm emission from neutral gaseous hydrogen. Astronomers could then measure magnetic field strengths

and their place-to-place variations with some confidence. And it is doubly the beginning of the modern era for me and everybody else who received a PhD in 1968 (just as the line between history and “current events” is the year you started reading newspapers for yourself).

### LOCAL INTERSTELLAR FIELDS AND THE PROBLEM OF STAR FORMATION

Stars form when already-dense clumps of interstellar gas collapse further. They don't find it easy. The overall rotation of our galaxy means that any region containing one solar mass of gas has far too much angular momentum to become a one solar mass star rotating at less than break-up speed. Magnetic fields would seem to make things worse. Here, after all, is yet another source of pressure for gravity to overcome, in addition to  $P = nkT$ , cosmic rays, random turbulent motions, and the excess angular momentum.

It greatly reduces the effort of remembering how big all these things are that the energy density (or pressure) in magnetic field, cosmic rays, turbulent motion, and thermal kinetic energy are all about the same through much of the interstellar medium—about 1 eV per cubic centimeter. This is arguably not a coincidence, but rather the result of cosmic rays tugging on field lines tugging on clouds which collide and heat each other and tug on field lines which confine cosmic rays which . . . I have never been sure whether it is a coincidence that the energy density of starlight near us is also about 1 eV/cm<sup>3</sup>. The present 2.7K cosmic microwave background radiation also contributes a bit less than 1 eV/cm<sup>3</sup>, everywhere, which is surely a coincidence. Isn't it??

In any case, Zeeman measurements confirm that condensing clouds have taken some of the average galactic field with them and have  $B = 10\text{--}30 \mu\text{G}$ , vs.  $1\text{--}3 \mu\text{G}$  in more diffuse regions. And you might reasonably think this would make the star formation problem that much worse.

It doesn't. In fact, angular momentum and magnetic flux help to shove each other out of star formation regions, protostellar clouds, and very young stars. To

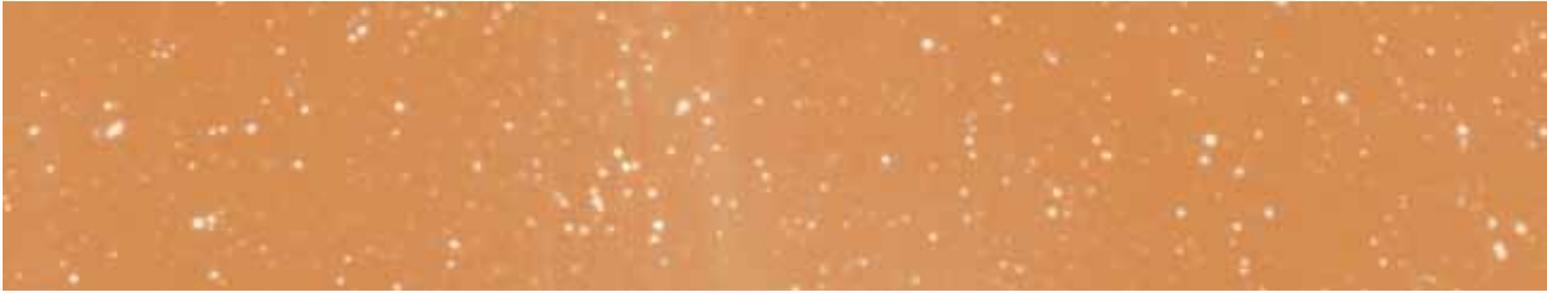


*Leon Mestel. One of the major contributors to the theory of cosmic magnetic fields and their influence on star formation and other processes. He also proposed, with Malvin Ruderman, the correct solution to a problem in cooling of white dwarfs—the carbon and oxygen nuclei eventually form a crystal lattice!*

attempt to apportion credit for the ideas would be to invite electronic, paper, and perhaps real over-ripe tomatoes to be aimed at Irvine, College Park, and even Stanford. I will merely say that I have heard the problems and the solutions most persuasively expressed by Leon Mestel (retired, but very much on active duty at the University of Sussex) and Telemachos Ch. Mouschovias of the University of Illinois.\* Now you can quarrel with my taste, but not with my perception of history!

The relevant process is simple, really, or, rather simplified (though I hope not beyond recognition). On the one hand, the field imposes co-rotation on the outer regions of cloud, protostar, or young stellar object until bits spin off carrying away more than their fair share of angular momentum per unit mass (the solar wind really does this now). And, on the other hand, the rotation first amplifies trapped field and, in due course, contributes toward dynamo generation of more. Thus the entities with strongest field and fastest rotation spin down and excrete magnetic flux most efficiently, leaving new-born stars that are faster rotators with more high-field star spots than their older cousins, but not unreasonably so. I have forgotten, though you must not,

*\*Prof. Mestel's undergraduate lectures in electromagnetism in his Cambridge days were reputed to feature two jokes, in alternate years. One concerned the disappearing vector, and the other did not. Prof. Mouschovias is not widely known to be associated with any jokes at all.*



ambipolar diffusion, which allows field to cheat its way out.

Observations tell you immediately where I have oversimplified. We definitely see gas flowing away from protostars and young stellar objects, but it looks more like jets and streams than like equatorial disks. Thus you have to think of the field twisting and turning in disks and collimating jets, which then do, nevertheless, carry off both angular momentum and flux.

The fields considered here belong to the galaxy as a whole, and we are not required to account for their origins until the next section.

### LARGE SCALE FIELDS IN THE MILKY WAY AND OTHER GALAXIES

Hall and Hiltner told us the direction of the average interstellar magnetic field and Verschuur something about its strength in particular clouds of neutral hydrogen. Synchrotron radiation provides information about average direction (from polarization angle), about total strength (from the brightness of the radiation, provided you know the density of relativistic electrons), and about the relative importance of ordered and chaotic field components (from the amount of polarization). Thus we\* came to the end of the 1960s reasonably sure that the galactic field had roughly equal strength in ordered and random components and that the ordered field lines were parallel to the plane of the galactic disk. Some spurs and loops out of the plane, which show both in polarized starlight and in synchrotron maps, were blamed on nearly shell-shaped old supernova remnants, of which more shortly.

Starting in the mid 1960s, the best estimate of coherent field strength gradually shrank from 10–30  $\mu\text{G}$  to 1–3  $\mu\text{G}$ . The larger value would have made the field dominant over cosmic rays, turbulence, and such, so that it might have protected spiral arms from winding up

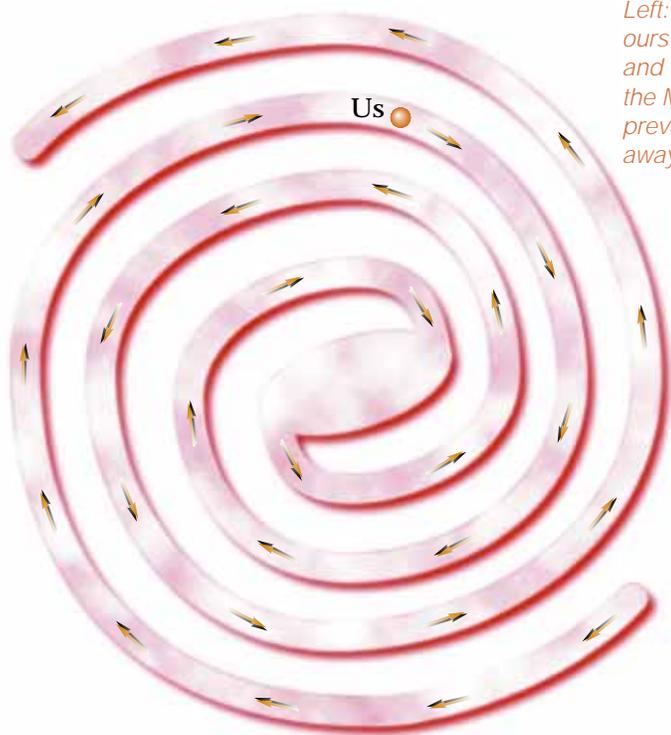
\*“We” in this context means “a majority of the astronomical community.” You will probably have noticed that “we” in these pieces more often means “author plus reader,” or, occasionally, “author and Queen Victoria.”

in the differential rotation of our own and other spiral galaxies. This was what Woltjer had in mind at the Nordwijk symposium when he associated strong magnetic fields with ignorance. But, at the same time the field was shrinking, C. C. Lin of MIT and his students, especially Frank Shu, were looking at instabilities in differentially rotating disks and concluding that the arms are probably soliton-like waves. Current opinion endorses this sort of picture.

The discovery of pulsars in 1968 opened a whole new window on galactic magnetism, glazed with the effects called dispersion measure and rotation measure. Dispersion measure (DM) means that longer wavelengths take longer to get here through intervening plasma. The time delay is proportional to the integral of electron density,  $n_e$ , along the line of sight. Rotation measure (RM) means that the intrinsic plane of linear polarization is rotated through an angle, larger for longer wavelengths, that is proportional to the integral of  $n_e B_{\parallel}$  along the line of sight. Divide RM by DM, and you have the average field along the line of sight out to the distance of that particular pulsar. In addition, the 180 degree ambiguity in field direction as found from optical and synchrotron polarization disappears. Faraday rotation of synchrotron radiation from quasars and such outside the Milky Way provides  $\int n_e B_{\parallel} dl$  for the entire galaxy in different directions.

With more information, naturally the field pattern looks more complicated. The coherent part is still mostly in the galactic plane, with some concentration toward the spiral arms. The best bet on directions is that the field lines follow the arms rather than circles. Perhaps most interesting, the field direction reverses from being clockwise where we are to being counterclockwise both inside and outside the solar distance from the galactic center, and reverses back again at least once still closer to the center. (I haven't said whether you are looking from above or below the plane, and won't).

The simplest explanation is a two-armed spiral with field lines coming out of the galactic center along one arm and going in on the other, in what is called a bisymmetric spiral pattern (see figure on the next page). The



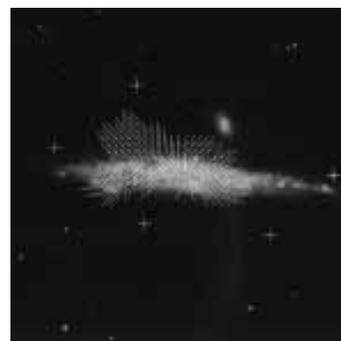
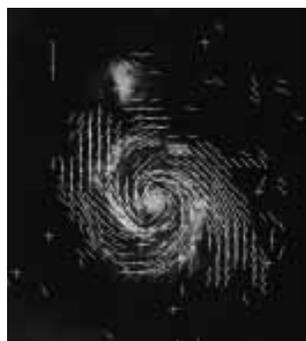
*Left: Greatly idealized version of a slightly barred spiral galaxy (which ours is), with ordered, Grand Design spirals arms (which ours has not), and a bisymmetric magnetic field. This last may or may not apply to the Milky Way, but would at least account for the several reversals of prevailing field direction (colored arrows) seen as we look toward and away from the galactic center.*

Other spiral galaxy fields are probably more or less like ours, although what you can measure varies. For face-on spirals, synchrotron polarization indicates field directions along arms and bisymmetry or axisymmetry perhaps correlated with whether the arms are tidy or messy (the technical terms are Grand Design and flocculent spirals). M51, shown below, is a classic Grand Design, two-armed spiral, with field directions along its arms. For edge-on ones, polarization necessarily tells us about fields perpendicular to the disk. My favorite is NGC 4631, shown below, with ought to be called the hedgehog galaxy. NGC 891, sometimes spoken of as a twin to the Milky Way, is rather similar.

alternative is called an axisymmetric spiral, and has field either going out or coming in on both arms.

What about the galactic halo? Information is sparse, because there is little dust to be aligned, few stars to have their light polarized, little synchrotron emission (how much was once bitterly fought over as “the existence of the radio halo”), and few pulsars to be Faraday rotated. But the rising loops, chimneys, fountains, and champagne bubbles of hot, supernova-driven gas that penetrate into the halo must carry some field with them, and it ought to be roughly perpendicular to the disk.

The galactic center is a region of stronger magnetism, milli- rather than micro-Gauss, structured roughly like a dipole perpendicular to the plane, but with twists, filaments, and snakes that either trace or drive gas flows and distributions.



*Left: Magnetic field vectors of the face-on spiral galaxy M51, from a 2.8 cm image obtained by Nikolaus Neininger. Clearly the field lines track the spiral arms at least approximately. (Copyright Max Planck Institute fuer Radioastronomie, Bonn, Germany; provided by Rainer Beck.) Right: Magnetic field vectors of the edge-on spiral galaxy NGC 4631, 6 cm image obtained by Goetz Golla and Ko Hummel. At least some spiral galaxies clearly also have a component of magnetic field extending more or less perpendicularly out of the disk. (Copyright Max Planck Institute fuer Radioastronomie, Bonn, Germany; provided by Rainer Beck.)*

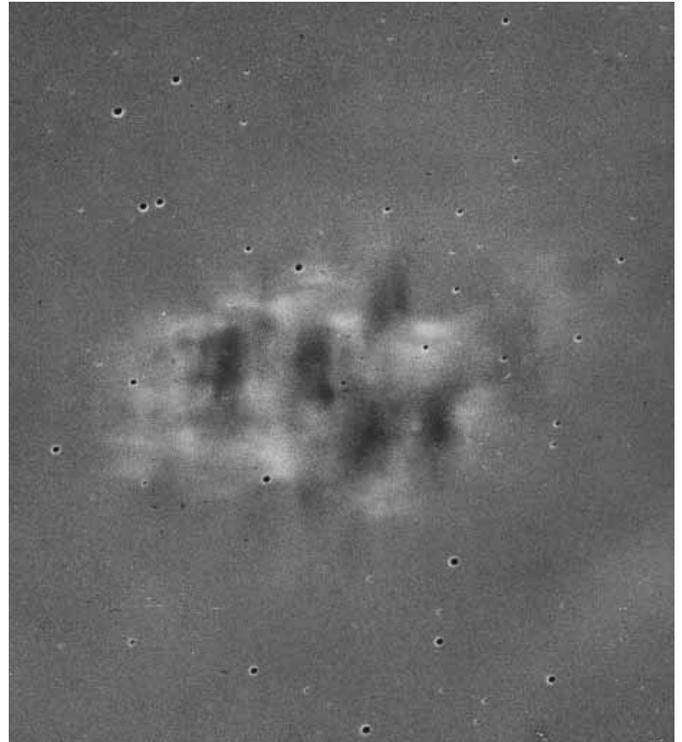


Information on magnetic fields in other types of galaxies is still sparser. The Large Magellanic Cloud, a sort of irregular galaxy and our nearest neighbor, has some coherent field, seemingly associated with star formation regions. As for ellipticals, a recent four-day conference on them doesn't even have magnetic fields as an index item in its proceedings. But a subset of giant ellipticals are the hosts of quasars and other strong radio sources, for which all models invoke significant fields, right on up to the (non-conventional) Strong Magnetic Field Model.

Galactic, at least spiral, magnetic fields have long been attributed to dynamos that derive their energy both from differential rotation and from gas turbulence (in turn driven by expanding supernova remnants and stellar wind bubbles). One set of theorists says that they can produce both axi- and bisymmetric spirals this way (though sometimes with another, unwanted, reversal of field direction across the galactic midplane). Another set periodically says that you get the wrong geometry, all the field energy cascades down to small length scales too fast, and that the saturation field is less than the real one . . . followed by rebuttals from the first set. Relevant observations are thin on the ground, but it does seem to be true that many galaxies at redshifts near two (when the universe was nine times its present density) already have fields as large as ours. This might be taken as a mild argument against dynamos, which take a while to get going. Alternatives to a dynamo are (1) "primordial fields," meaning you shove the problem back to a time whose physics is not easily probed in the laboratory or (2) smoothing out of localized fields originally contributed by supernovae.

### SUPERNOVA REMNANTS

Yes, Virginia says that the radio, visible, and X-ray photons from the Crab Nebula are all synchrotron emission. They are polarized, have the right (power law) spectra, and whatever else you want. There is a patch of coherent field near the pulsar of nearly a milliGauss, and regions of 10<sup>-4</sup>-5 G further out, with field lines vaguely



*Polarization structure of the optical synchrotron emission from the Crab Nebula. Images of this sort were first created by Fritz Zwicky. This particular one consists of two photographs, taken at about the same time, through Polaroid filters. One photo had the Polaroid oriented in position angle 45 degrees (vertical in the reproduction), the other in PA 135 degrees (horizontal in the reproduction), corresponding roughly to the major and minor axes of the elliptical silhouette of the nebula on the plane of the sky. A positive transparency was then created from one of the glass negatives, and prints made of light simultaneously shining through the positive and the other negative. Thus polarized regions look lighter or darker than the sky average (and stars do not perfectly cancel out). Polarized regions are patchy and tend to have substructure roughly parallel to the polarization angle. This image is the inverse of one Zwicky published in 1956 (PASP 68, 121). Curiously, it is also a mirror image, representing the sky as seen from outside. A number of Zwicky's images of the Crab Nebula came to me at his death in 1974, and one can reasonably suppose that this one is reproduced with his permission.*



aligned with filaments of densish gas and perpendicular to the remnant edge. This centrally-condensed field is surely a gift of the pulsar, and some other, less famous, remnants look similar.

Synchrotron radio (but not optical or X-ray) emission is common to all known supernova remnants—not remarkable, since this is how we distinguish them from other kinds of hot interstellar gas clouds that emit radio bremsstrahlung. In most SNRs, however, the field lines run around the perimeter, not radially through it. We think this means that the field is mostly just swept up interstellar stuff.

One way or another, supernovae and their remnants are important to the galactic magnetic field. They may be the source of the “alpha” (turbulence) half of an alpha-omega dynamo (omega is the rotational part). They surely carry field into the halo. Perhaps they are even the ultimate source. This last is attractive largely because stellar dynamos are easier to model than galactic ones—not quite the same as their being closer to reality, of course. So, why not make fields in stars, concentrate them in pulsars, blow them out in pulsar winds that feed Crab-like remnants, and let the remnants expand and merge. Eventually there will be a good deal of random field out there that can be twirled around into the right pattern by galactic rotation and concentrated into the nucleus and into dense clouds by gas flows. This idea has gone in and out of fashion several times since 1968, and I am not sure what the current phase is.

Planetary nebulae, the expelled gaseous envelopes of lower mass stars than the ones that make supernovae, do not have detectable fields, though of course someone once (1962) said that they should.

## RADIO GALAXIES AND QUASARS

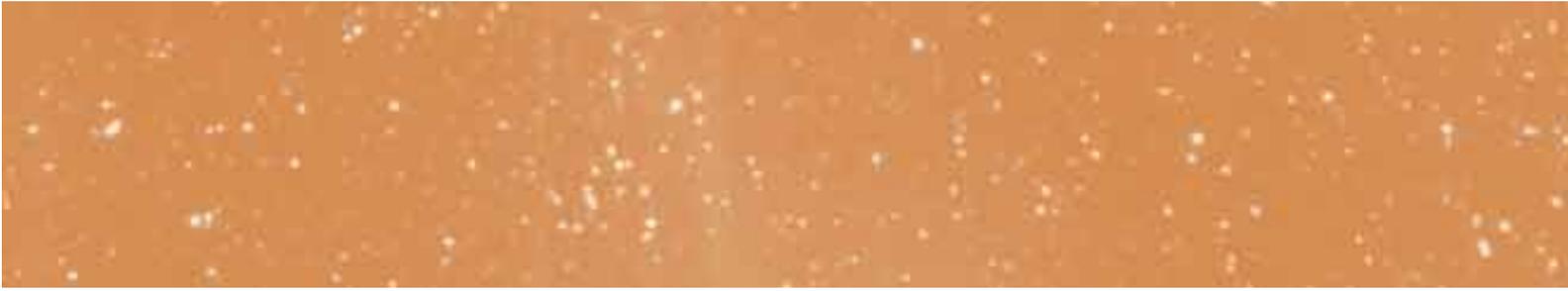
The third extragalactic radio source identification was that of Cygnus A, meaning the brightest radio source in the direction of that constellation. Walter Baade and Rudolph Minkowski were able to get the critical picture in 1954 because radio astronomy had provided an accurate position, and the information that the radio

photons were coming from two separate roundish bits of sky. The optical fuzz was in between the radio lobes and looked to Baade and Minkowski like two, flattened elliptical galaxies, colliding face to face. The modern interpretation is a single giant elliptical galaxy with a dust lane down its middle. Statistical evidence says, however, that interactions and mergers between galaxies do promote development of radio emission and other nuclear activity.

The early 1960s saw the discovery of polarization in radio galaxies and the discovery of the first quasars. These are, probably, also elliptical galaxies, but with nuclei so bright in visible light that the rest of the galaxy is nearly lost in the glare. Their radio emission is also linearly polarized, power-law-spectrum synchrotron. A much larger class of quasi-stellar objects (and also of less overwhelming active nuclei living in spiral galaxies and named for Carl Seyfert) are radio quiet, but bright in visible light and X-rays. Evidence for strong magnetic fields in them is less direct, and I will ignore them henceforth.

We have an official paradigm\* for active galaxies. A central black hole of  $10^{6-9}$  solar masses is accreting material from its surroundings, probably via a disk. The disk collimates jets of relativistic, magnetized plasma, which squirt out at high speed perpendicular to the disk, radiating as they go, and feeding energy out into large double lobes like those of Cygnus A. The jets and blobs are all more or less lined up from scales of a parsec near the black hole (resolved with very long baseline radio interferometry) to hundreds of kiloparsecs (a minute of arc on the sky, about the resolution of your eyes). One jet is bound to be pointing more or less toward us, the other away. Bulk relativistic motions thus make the approaching one look much brighter and result in structure changes observable from year to year.

*\*Within living memory, practicing scientists (won't we ever learn how?) used “paradigm” the way Kuhn had meant it, to mean an experiment that set an example for the way things ought to be done. Its current usage comes closer to the “best buy model” of Consumer Reports.*



Traditional calculations presume that the energy in relativistic electrons plus magnetic field ought to be the minimum to produce a given flux of synchrotron radiation. You get almost exactly the same numbers by assuming equal energy density in particles and field—yet another case where one mumbles to oneself “coincidence/obvious/causal/remarkable . . .” Extended radio blobs require a microGauss or two, compact bits more like a milliGauss. The equipartition field just outside the black hole horizon is about a Tesla, the only context I know in which this is a useful unit. Polarization studies show that lobe fields are fairly chaotic, while the core and jet ones are often sharply aligned with the jet axis.

Anyone bringing a fresh and vacant mind to the contemplation of quasars and their ilk will have some pretty pointed questions. For instance, if the extended lobes consist entirely of relativistic fluids—high speed electrons and magnetic field—why aren’t they expanding at the speed of light perpendicular to the source axis? This is called the “radio source confinement problem.” Various attempts to tie the field lines to some large, non-relativistic mass are fairly unpersuasive. The winner seems to be ram pressure, that is, the confining effect of a tenuous medium through which jets and blobs propagate. The confining pressure is then  $\rho_e v^2$ , where  $\rho_e$  is the external gas density. It has to balance the outward,  $B^2/8\pi$ , pressure. A reasonable estimate of circumsource gas density leads to minimum speed of at least 10–30 percent of  $c$ . Bigger than  $c$  is unlikely. The range is the same as what we deduce from rapid central structure changes, and implies a fairly narrow range of possible lifetimes for the extended sources, near  $10^8$  years. This is also about the time it takes for a central black hole to double its mass given the rate stuff must fall in to keep up all the activity, and so is probably about right. It is left as an exercise for the reader to consider water skiing as a problem in ram pressure support and to calculate the minimum speed the tow boat must move to keep you from sinking.

Next question: what and where are the currents that maintain these magnetic fields, or at least generate them initially? You can read a lot of quasar review articles

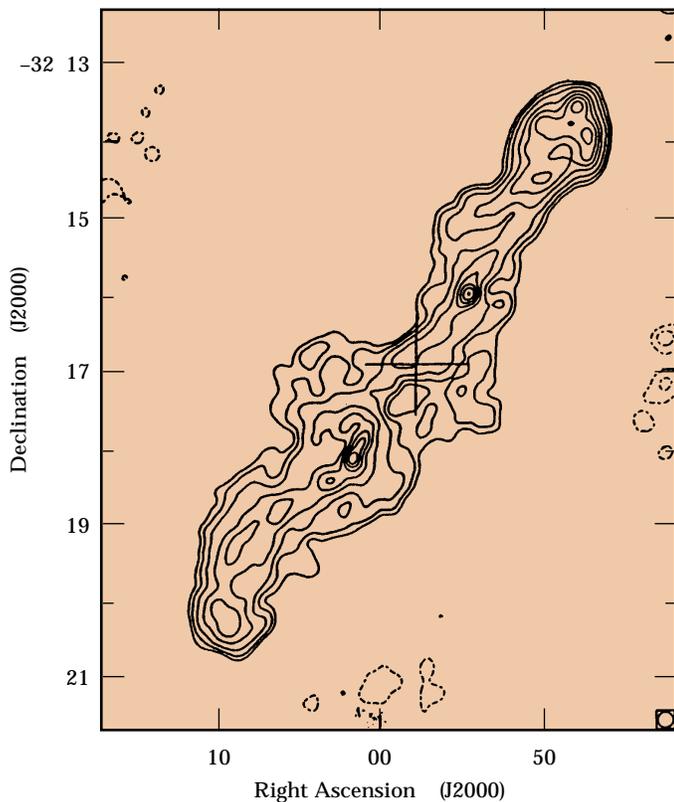
without finding many sentences devoted to this issue. The fields must have originated outside the black holes, and the commonest assumption is that inflowing gas has always carried its field with it, to be amplified in a differentially rotating accretion disk and then partly dumped inward and partly shot back out in the jets, helping to collimate them in the process. Most recent models lean heavily on the “twin exhaust” ideas put forward in 1974 by Roger Blandford and Martin Rees, then both at Cambridge.

The twin exhaust and related models also provide some answer to what the jets are made of—probably relativistic electrons and positrons (in equal number to within  $\pm 10^{-17}$  or so) plus field. An alternative idea, coming slightly earlier, also from Rees, had the jets starting out as pure low frequency electromagnetic waves from the central magnetized black hole (acting like a pulsar). Ambient, thermal electrons could ride the  $E$  vector waves, soon reaching relativistic speeds, and then perceive the  $B$  vector part as a nearly static magnetic field in which to spiral and emit synchrotron photons. This scenario falls afoul of detailed polarization data, but still seems to me to be very elegant.

#### JETS AND DISKS WHEREVER YOU LOOK

You have already been asked to believe in magnetized accretion disks and the jets they collimate as explanations for angular momentum removal and outflowing gas from young stars and for much of the phenomenology of active galaxies. Similar things happen other places as well, including the surroundings of neutron stars and black holes in X-ray emission binary star systems and in the binaries with a white dwarf that give rise to nova explosions. Two or three galactic X-ray sources have associated radio jets that move outward, like the quasar ones, at 25–75 percent of  $c$ , so that we see rapid structure changes and/or greatly redshifted and blueshifted optical emission lines.

Perhaps you won’t be surprised, since you know that the underlying physics is much the same, that all these disk/jet objects look quite similar, whether the central



*Contour diagram of a random, powerful radio source in the southern sky. Experienced radio astronomers claim that such diagrams are really more informative than ordinary halftone photographs, but they take a bit of practice to appreciate (I mean the pictures, not the radio astronomers.). Closely spaced contour lines mean that the surface brightness is changing rapidly with position there. The brightest bits are the ones most contour lines away from the edges. Some sources have their brightest bits at the tops of the lobes; others do not. Sources also vary considerably in their degree of asymmetry between the two sides (thought to be largely a relativistic beaming effect). Very similar configurations, on very different size scales, are associated with cores of some galaxies, X-ray emitting binary stars, and young stellar objects. (Courtesy Richard Hunstead, University of Sydney)*

core is a young star, a stellar mass black hole, or a supermassive black hole. In fact, in plots of radio contours like the illustration on above, even a fairly experienced astronomer will have to look at the angular scale of the image (or the name of the observer!) to be sure of whether he is contemplating a large double radio source, a compact core source, or a Herbig-Haro (young star) source. And for all these configurations, magnetic fields are dynamically and energetically important in a way they are not for planets and stars.

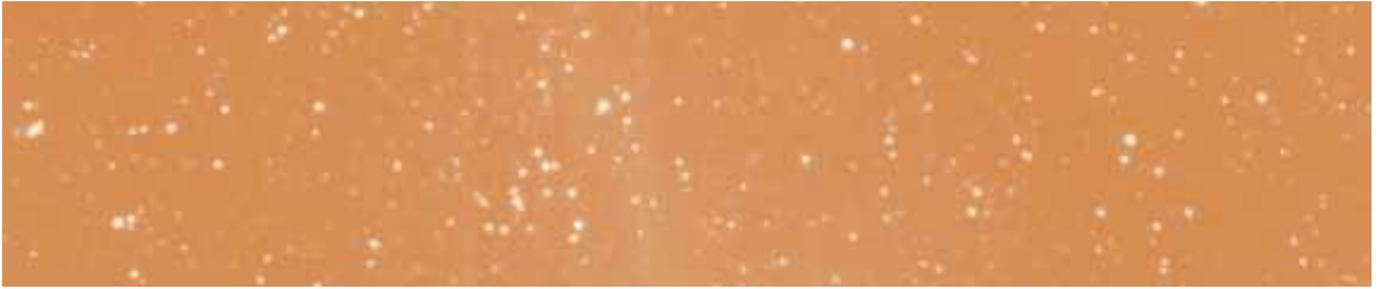
### LONG AGO AND FAR AWAY

Magnetic fields pervade at least some clusters of galaxies. The evidence includes diffuse radio emission and Faraday rotation and depolarization of radio lobes belonging to galaxies in the clusters and sometimes also

of radio sources behind the clusters. Field values are surprisingly high, not much smaller than galactic values. That is, one finds 1–10  $\mu\text{G}$  near large, central galaxies and tenths of microGauss further out in the clusters.

I can think of only three possible classes of explanation, and all have been defended within the present decade. First, these are fields originally belonging to normal and active galaxies, shot out with winds and jets. Second they are generated in situ by dynamos in hot intracluster gas (which is certainly there—we see the X-ray emission). Third, it is primordial magnetic field, brought into dark matter halos by the gas that flowed into form the cluster, later amplified by gas turbulence.

“Primordial” in this context essentially means “attributable to processes before galaxies formed.” As far as I am concerned, it also means “attributable to processes I don’t understand.” Recent suggestions have included vorticity in the early universe and the inflationary epoch invoked for other reasons, and, perforce, I stop here, finding, as Ehrenfest said, that it is quite difficult to explain something even if you understand it, and almost impossible if you don’t. There is a well-known Wittgenstein quote that would be equally apposite here, but it never seems to come out very well in English (perhaps because there is no way to put the “daruf” at the end?), and so you must supply it or some equivalent for yourself. ◻



## STILL MORE WORDS

THE CLASSIC TEXT is Ya. B. Zeldovich,

A. A. Ruzmaikin, and D. D. Sokoloff, **Magnetic Fields in Astrophysics**, Gordon & Breach (New York) 1983.

Two relevant symposia of the International Astronomical Union are published as: R. Beck, P. Kronberg, and R. Wielebinski, Eds., **Galactic and Inter-Galactic Magnetic Fields**, Kluwer, 1990.

F. Krause, K.-H. Raedler, and G. Ruediger, Eds, **The Cosmic Dynamo**, Kluwer, 1993.

The era of magnetically dominated spirals and the galactic radio halo is represented in articles by L. Woltjer (p. 531) and J.L. Pawsey (p. 219) in A. Blaauw & M. Schmidt, Eds., **Galactic Structure**, U. Chicago Press, 1965.

Readers who cultivate the truly arcane might be interested in: *The Strong Magnetic Field Model* which appears most recently as H. D. Greyber, p. 298, of V. Trimble & A. Reisenegger, Eds., **Clusters, Lensing and the Future of the Universe**, Astronomical Society of the Pacific Conference Series No. 88, 1996 (and references therein).

The very first appearance of “a new model for extragalactic radio sources” in the form of a meeting abstract, M. J. Rees & V. Trimble, *BAAS* **3**, 25 (1971).

This is probably the most important paper on which my name has ever appeared, and, sadly, I was merely the introducer, required by Society rules for papers by non-members.

A sharp-eyed colleague (Joe Tenn of Sonoma State University) found the missing Soul of Lodestone, author Alfred Still (1869–?), released in a late, 1946, edition by Murray Hill Books (New York and Toronto).

R. Kulsrud et al., 1996, *Physics Reports* (in press) provide an up-to-the-minute overview of the pros and cons of large-scale dynamos and primordial magnetic fields.

---

## CONTRIBUTORS

---



**MICHAEL RIORDAN** has been associated with the *Beam Line*, first as Editor and now as Contributing Editor, since he came to SLAC in 1988. He returned to his position here as Assistant to the Director this year after a leave of absence to complete his latest book, **Crystal Fire**, a history of the transistor (co-authored with Lillian Hoddeson) to be published in 1997 by W. W. Norton. He is also author of **The Hunting of the Quark** and co-author of **The Shadows of Creation**.

He did his PhD research at SLAC on the famous MIT-SLAC deep inelastic electron scattering experiments and continued at MIT as a postdoctoral research associate on these experiments.

He currently holds an additional appointment as a Research Physicist at the Santa Cruz Institute for Particle Physics, studying the history of the Superconducting Super Collider.



*Beam Line* Editorial Advisory Board member, **NATALIE ROE**, got her start in high energy physics as a student of Carlo Rubbia's while she was still an undergraduate at Harvard University working on the Home-stake Mine proton decay experiment. After a year at CERN working on the UA1 experiment, she went to Stanford for her PhD, where she searched for anomalous single photons on the ASP experiment with David Burke.

She started at Lawrence Berkeley National Laboratory (LBNL) as a postdoc on the D0 experiment at Fermilab, building calorimeters and studying the production of *W* and *Z* bosons. More recently she has been appointed to the scientific staff at LBNL and is co-manager of the silicon vertex tracker for the BaBar experiment. She is also co-manager of two young children, a project she finds equally challenging and rewarding.



**CYLON GONÇALVES DA SILVA** is Director of the Laboratório Nacional de Luz Síncrotron in Campinas, Brazil, and a Professor of Physics at the Instituto de Física “Gleb Wataghin.” He received his PhD in Physics from the University of California at Berkeley. Prior to dedicating himself to the light source project, he held a visiting research fellow position at IBM T. J. Watson Research Center and visiting professorships at the Ecole Normale Supérieure in Paris and the Université de Lausanne in Switzerland. He is also an associate member of the International Centre for Theoretical Physics in Trieste, Italy.

The author of over five books and many articles in theoretical condensed matter physics, Gonçalves da Silva is a member of the Brazilian Academy of Sciences and associate editor of *Solid State Communications*.



**KWANG-JE KIM** started out as an elementary particle theorist, worked at SLAC and Max-Planck -Institute for Physics and Astrophysics, and turned into an accelerator physicist when he joined Lawrence Berkeley National Laboratory in 1978, where he is the Deputy Head of the Center for Beam Physics.

He has always been interested in radiation sources with capabilities beyond what is currently available and in associated accelerator technology to bring that about, be it high brightness synchrotron radiation sources, variable polarization devices, free electron lasers, femtosecond X rays, rf photocathode guns, or the production of intense gamma rays. He is currently leading a multi-institution, international collaboration to design a second interaction region for the gamma-gamma collisions for the Next Linear Collider.



**ANDREW M. SESSLER** enjoys the outdoors. However, he has spent enough time indoors to publish 250 papers, win the Lawrence Award in 1970, the Particle Accelerator School Prize in 1988, and the Nicholson Medal for Humanitarian Service in 1994. He is a member of the National Academy of Sciences and Vice-President of the American Physical Society. He is former director (1973–1980) of the former Lawrence Berkeley Laboratory and can be reached at [TBALBL@LBL.GOV](mailto:TBALBL@LBL.GOV).



**LANCE DIXON** has oscillated between Princeton and SLAC for his entire professional career and has been at SLAC continuously since 1989, as Panofsky fellow and now associate professor. As you can see, he also moonlights as a bartender, but he would be ill-advised to give up his day job.



**ERIC SATHER** is a research associate in the Theory group at SLAC. He received an AB in Physics from the University of Chicago and a PhD in Physics from Massachusetts Institute of Technology. In addition to his studies of baryogenesis he has investigated the structure of hadrons as well as the consequences of physics beyond the Standard Model. In his spare time he enjoys brewing beer (as shown here with lauter tun).



**VIRGINIA TRIMBLE** graduated from Toluca Lake Grammar School in North Hollywood, California, and has since earned degrees from several other places. The picture was taken, depending on your point of view, (a) the year Dewey defeated Truman, (b) the year between the discovery of strong stellar magnet fields by Horace Babcock and the discovery of interstellar polarization by John Hall and William Hiltner, or (c) the year she began kindergarten at Toluca Lake. Bar magnets and iron filings were among her earliest toys, and her lack of heed to the advice to keep a piece of paper between the two has characterized her involvement with magnetic fields ever since.

---

## *DATES TO REMEMBER*

---

- Oct 3–6            Workshop on Heavy Quarks at Fixed Target (HQ 96), St. Goar, Germany (Kerstin Kern, HQ-96, Institut für Physik, Johannes Gutenberg-Universität Mainz, D-55099 Mainz, Germany (hq96@dipmza.physik.uni-mainz-de)).
- Oct 7–9            International Workshop on PCs and Particle Accelerator Controls, Hamburg, Germany (Ingrid Nikodem, DESY, Notkestrasse 85, D-22603, Hamburg, Germany (pcapac@desy.de)).
- Oct 7–11           Trieste Conference on Quarks and Leptons: Masses and Mixings, Trieste, Italy (ICTP, PO Box 586, I-34100 Trieste, Italy (smr938@ictp.trieste.it)).
- Oct 13–19          7th Workshop on Advanced Accelerator Concepts, Lake Tahoe, California (S. Chattopadhyay, Center for Beam Physics, Lawrence Berkeley National Laboratory, MS 71-259, 1 Cyclotron Road, Berkeley, CA 94720 or cbp@lbl.gov).
- Oct 14–18          43rd National Symposium of the American Vacuum Society, Philadelphia, Pennsylvania (American Vacuum Society, 120 Wall Street, 32nd floor, New York, NY 10005 or avsnyc@vacuum.org).
- Oct 15–20          Workshop on  $\mu^+\mu^-$  Colliders, Montauk, New York (Check URL for updates: <http://epswww.epfl.ch/conf/all.html>).
- Oct 21–24          9th International Symposium on Superconductivity (ISS 96), Sapporo, Hokkaido, Japan (ISTEC, Eishin Kaihatsu Bldg., 6F: 34-3, Shimbashi 5-chome, Minato-ku, Tokyo 105, Japan (chyoki@istec.mxa.meshnet.or.jp)).
- Oct 21–25          ITP Conference on Future High-Energy Colliders, Santa Barbara, California (Prof. James B. Hartle, Director, Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106-4030).
- Oct 21–Nov. 1      CERN Accelerator School Course on an Introduction to Accelerator Physics, Cascais, Portugal (Mrs. S. von Wartburg, CERN Accelerator School, AC Division, 1211 Geneva 23, Switzerland or cas.estoril@cern.ch)
- Oct 23              Workshop on Approaches to Modern and Advanced Analysis of XAS Data, Stanford, California (Suzanne Barrett, SSRL MS 99, Box 4349, Stanford, CA 94309 or barrett@slac.stanford.edu).
- Oct 23              Mini-Symposium on Structural Molecular Biology and Beamline 9 Dedication, Stanford, California (Suzanne Barrett, SSRL MS 99, Box 4349, Stanford, CA 94309 or barrett@slac.stanford.edu).
- Oct 24–25          SSRL 23rd Annual Users' Meeting, Stanford, California (Suzanne Barrett, SSRL MS 99, Box 4349, Stanford, CA 94309 or barrett@slac.stanford.edu).

- 
- 
- Oct 24–25 ICEC Symposium Cryogenics in Science and Industry, Geneva, Switzerland (Christine Petit-Jean-Genaz, CERN-AC, 1211 Geneva 23, Switzerland or [christin@cernvm.cern.ch](mailto:christin@cernvm.cern.ch))
- Nov 2–10 Nuclear Science Symposium and Medical Imaging Conference, Anaheim, California (Margit Sperakos, University of California, Department of Radiological Sciences, Irvine, CA 92717 or [mesperak@uci.edu](mailto:mesperak@uci.edu)).
- Nov 6–9 14th International Conference on the Application of Accelerators in Research and Industry, Denton, Texas (American Institute of Physics, 1 Ellipse Circle, College Park, MD 20740-3844).
- Dec 3–5 ITP Conference on Particle Beam Stability and Nonlinear Dynamics, Santa Barbara, California (Prof. James B. Hartle, Director, Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106-4030).
- Jan 20–31, 1997 US Particle Accelerator School Quality Education in Beams and Associated Accelerator Technology since 1980, Berkeley, California (USPAS Office, Fermilab, MS 125, Box 500 Batavia, IL 60510 or [uspas@fnal.gov](mailto:uspas@fnal.gov) and <http://fnalpubs.fnal.gov/uspas>).