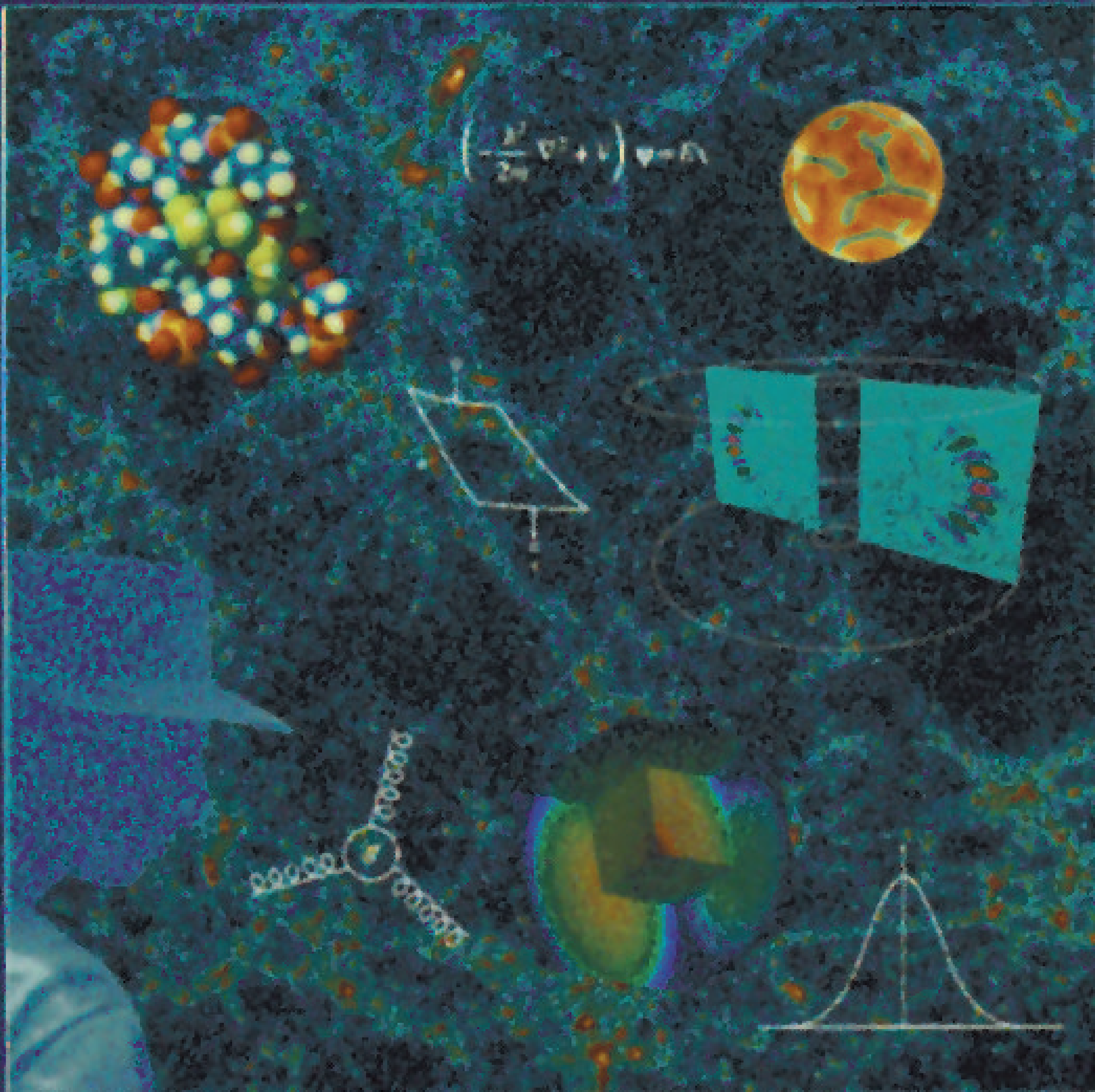
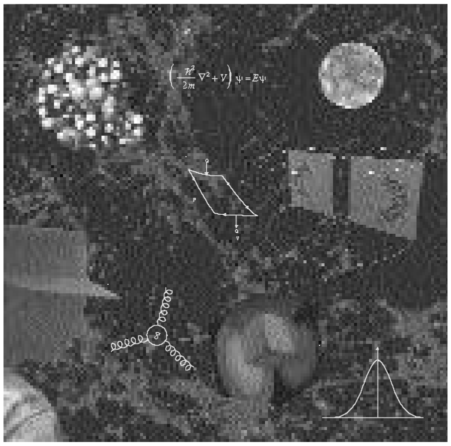
Los Alamos Science

LOS ALAMOS NATIONAL LABORATORY

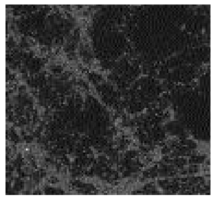


NUMBER 21 1993

THE LABORATORY'S 50th ANNIVERSARY

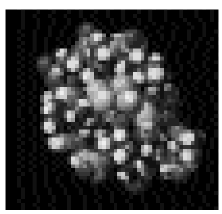


For fifty years, basic research in physics, mathematics, computers, chemistry, and life sciences, combined with ingenious engineering, have been the strength of Los Alamos National Laboratory. Shown on the cover are a few symbols of fundamental concepts in physics and mathematics—the Schrödinger equation, the Carnot cycle, the three-gluon vertex of QCD, and the Gaussian distribution. Also shown are state-of-the-art computer simulations performed at Los Alamos on problems of fundamental interest to science and the nation. High-power computing is opening up more and more complex systems to basic understanding and practical application.

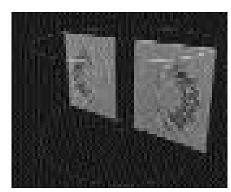


New computer hardware and software tools have enabled scientists, for the first time in history, to incorporate enough physics with sufficient accuracy to simulate the evolution of the universe and the formation of galaxies. The figure shows the distribution of dark matter in the universe from a computer simulation by M. S. Warren of the Laboratory's Theoretical Astrophysics Group and J. K. Salmon of the California Institute of Technology. Their ingenious "treecode" algorithm was able to keep track of the forces among 17 million particles. This algorithm has been picked up by others and adapted to more applied problems.

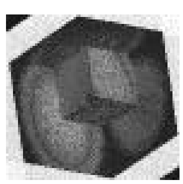
Global modeling of the earth's atmosphere, oceans, and interior is now being undertaken on Los Alamos supercomputers. Shown here is a simulation of convection in the earth's mantle developed by G. Glatzmaier in the Laboratory's Geoanalysis Group. The three-dimensional flow, dominated by viscosity, contains linear regions of downward flow (blue) reminiscent of the subduction zones found on the earth's surface at tectonic-plate boundaries.



The Laboratory's initiative in structural biology is exploiting supercomputers to understand the fundamental interplay between structure and function of biological macromolecules. Depicted here is one of many possible "hairpin" structures of a short DNA molecule in solution as calculated from nuclear-magnetic-resonance data with a new method developed by members of the Theoretical Biology and Biophysics Group. This method involves molecular dynamics and simulated annealing; it was implemented at the Advanced Computing Laboratory.



Research in magnetic fusion, begun in the early 1950s at the Laboratory, is currently an international effort. Progress requires an understanding of turbulence and instabilities in very high-temperature tokamak plasmas. The figure represents a cross section such a plasma showing the instantaneous electrostatic potential in the simulated development of an instability. Numerical simulations of the turbulent plasma provide information that is very difficult to obtain from fusion experiments because the plasma reaches temperatures on the order of 100 million kelvins. The calculation was performed at the Advanced Computing Laboratory at Los Alamos by S. E. Parker and W. W. Lee of the Plasma Physics Laboratory at Princeton University.



Even particle physicists studying the basic forces of nature are employing supercomputers to make predictions with their theories. Shown here is a wave function of a moving pion simulated with lattice quantum chromodynamics (a discrete version of the theory of strong interactions) by R. Brickner of the Computer Research and Applications Group and R. Gupta of the Elementary Particles and Field Theory Group. The contours of the wavefunction would be circles if the pion were at rest, but are distorted by the Lorentz contraction. This research is helping other scientists learn how to exploit the full power of parallel supercomputers at the Advanced Computing Laboratory.

Editor
Necia Grant Cooper

Managing Editor
Nadine Shea

Science Writers
Gerald Friedman
Richard Reichelt
Nancy K. Shera

Art Director
Gloria Sharp

Technical Illustration
Andrea Kron

Original Art
David R. Delano

Production
Nadine Shea

Photography
John Flower
Henry Ortega

Other Contributors
AnnMarie Dyson
Cheri Tiedman-Isham
Chad Kieffer
IS-9 Photography

Printing
Guadalupe D. Archuleta
Brenda R. Valdez

Address Mail to:
Nadine Shea
Los Alamos Science
Mail Stop M708
Los Alamos National Laboratory
Los Alamos, NM 87545

When the *Los Alamos Science* staff began this volume last September, we perceived it as a great challenge. What aspects would best represent the Laboratory on its fiftieth birthday? The accomplishment of the Manhattan Project and the great scientists who made it succeed inspire all those who come to Los Alamos. That work, those times, that spirit have been superbly documented in Richard Rhodes's *Making of the Atomic Bomb.* The equally extraordinary efforts to build thermonuclear devices, the early efforts to harness the energy of fission and fusion for peaceful uses, and the evolution of the Laboratory into a multipurpose scientific institution were the primary focus of the 1983 issue of *Los Alamos Science* on the occasion of the Laboratory's fortieth anniversary. At that time, many of the great scientists who took part in the Manhattan Project were alive and able to return to Los Alamos to recall for us the birth of the nuclear age.

The fiftieth anniversary is quite different; it marks the end of superpower struggle. The end of the Cold War is a great victory for freedom and for the West, and this volume could have been used to examine the many contributions of Los Alamos to that victory and to other challenges of the past. Given the great changes around the world and at home, however, it seemed more appropriate to just touch on the past and celebrate this anniversary by exploring the directions we can take in the future.

Here at Los Alamos we are experiencing something like an identity crisis, a time of doubt and self-evaluation. What are our missions now that the Cold War is over? What are the new opportunities? What are our strengths, our weaknesses? This volume begins the way we at the magazine began, with a discussion among Los Alamos scientists of those very questions. The round table, "Taking on the Future," suggests the growing pains and internal tensions of the Laboratory as it struggles to adjust to the changes of the last few years.

On one side are the nuclear-weapons scientists, who have always carried out their mission in the silence imposed by the nature of their work. Here they speak candidly about their new challenges and concerns. They face a comprehensive test ban, declining support for their mission, and, simultaneously, the tremendous responsibility of maintaining and managing the nation's nuclear-defense capability in the face of growing threats of nuclear proliferation. On the other side are the scientists working on nondefense projects and basic research. Some of them came from the nuclear-weapons program, but most have no experience with weapons work, nor, for that matter, do they know much about that work. They offered perspectives on new missions and new opportunities.

Though the separation between the two cultures is evident in the round table, their futures are inextricably tied together. As the opportunity to design and test weapons disappears, the weapons program will depend more and more on the basic-research side of the Laboratory for ideas, skills, and new talent. On the other hand, the basic-research effort will remain dependent on the strength of the weapons program. In the past, basic research has been supported primarily by the major applied programs at the Lab, of which the nuclear-weapons program has been by far the largest and best funded and will, in all likelihood, continue to be so for many years. The synergism between the two sides must grow, and together they must find new missions. Those efforts will be aided by the continuous interplay between basic and applied work and the ongoing interaction among the wide spectrum of disciplines that has always characterized the Laboratory.

Most refreshing in the round table as well as in all the articles that follow is the great display of diversity. As Edward Teller points out in "The Laboratory of the Atomic Age," the Los Alamos tradition of tolerance for widely different views began with Robert Oppenheimer and continues to this day. It was essential then and is essential now in this community of intensely creative individuals with diverse views, talents, and styles. Diversity and creativity are not easy to manage, but they are our hope and our strength in meeting the growing complexity of science, of our missions, and of the world.

"The Stewardship of Nuclear Weapons" emphasizes another essential hallmark of the Laboratory, its ongoing commitment to service. The men and women who work on nuclear weapons were drawn to that field not just by their fascination with the complexities of the physics and the challenges of designing and testing but also by their patriotism. The nation is asking them to change how they work, to go from the production of the new to the maintenance of the old, a tall order for creative people. Their response, outlined here, is no less than a redefinition of the entire scope and nature of their activities to meet the changing contingencies without, they hope, compromising the nation's nuclear capabilities. A particularly welcome result of their new situation is the opportunity to meet and collaborate with their Russian counterparts on unclassified projects of mutual interest. We also had the pleasure of meeting Russian scientists from Arzamas-16 on their visit to Los Alamos and of interviewing Alexander Pavlovskii, a long-time colleague of Andrei Sakharov. Pavlovskii's comments on the experience of the Russian weapons scientists, which close this section, quietly portray the drama of the Cold War in the Soviet Union and the difficult period that has followed.

"Science and Innovation," a sampling of the dazzling array of research at Los Alamos, exhibits the enormous energy, enthusiasm, and creativity of our scientists. The uncertainties of the future notwithstanding, they all put their hearts into their writing and were inspired on this special occasion to explain not only their latest advances but also the philosophy behind their work, its roots in their beloved disciplines, and the exciting opportunities they see in the future.

This issue closes with two pieces on science policy. The first places the Laboratory's work in historical perspective. The second, by our director, lays out the major directions for the Laboratory's future and many vehicles for our contributing to new national priorities in the civilian and commercial sectors.

It was a great pleasure to work with everyone who contributed to this issue. The volume does, I believe, clearly demonstrate why Los Alamos will continue to be what it has been for fifty years—a national treasure.

Los Alamos Science

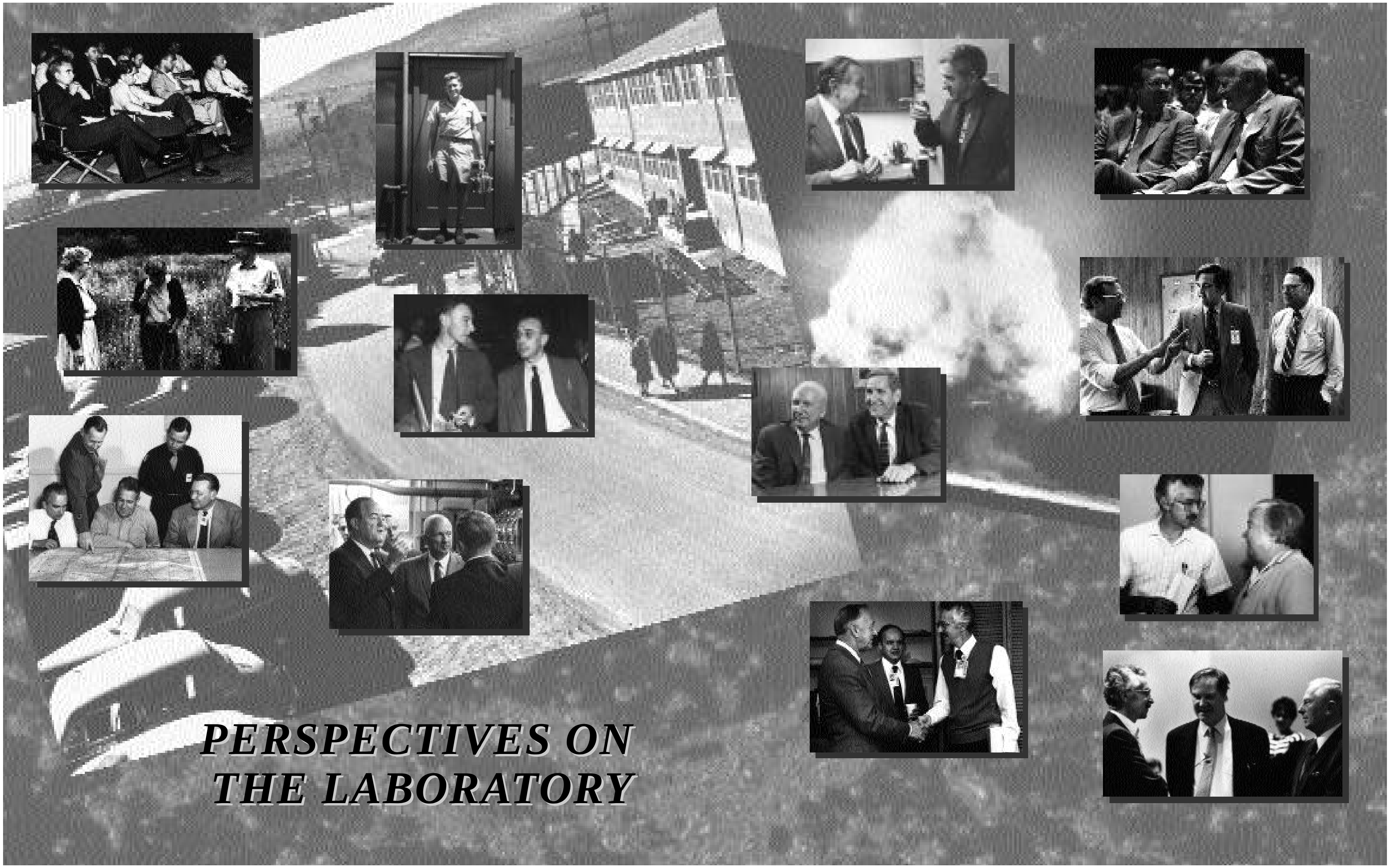
NUMBER 21 1993

Perspectives on the Laboratory 1

The Stewardship of Nuclear Weapons 40

Science and Innovation at Los Alamos 90

PERSPECTIVES ON THE LABORATORY



Milestones in the History of Los Alamos National Laboratory

J.Robert Oppenheimer, 1943–1945

1943 The Los Alamos laboratory, under the direction of J. Robert Oppenheimer, begins operation as Project Y of the Manhattan Project. The Bethe-Feynman formula, a simple method for calculating the yield of a fission bomb, is derived.

1944 The world's third nuclear reactor (a uranium-solution-fueled "Water Boiler" named LOPO) achieves criticality.

1945 The world's first nuclear bombs (Little Boy, a gun-type uranium bomb, and Fat Man, an implosion-type plutonium bomb) are proved successful. Norris E. Bradbury is named second director of the Laboratory.

1946 The world's first plutonium-fueled nuclear reactor (Clementine) first achieves criticality.

1947 The Monte Carlo technique for particle-transport computations is formulated.

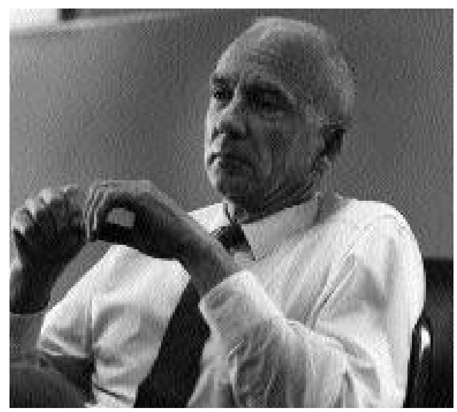
1948 Helium-3 is first liquefied.

1950 A new cyclotron-focusing method ("thomas" focusing) is developed that makes variable-energy machines possible.

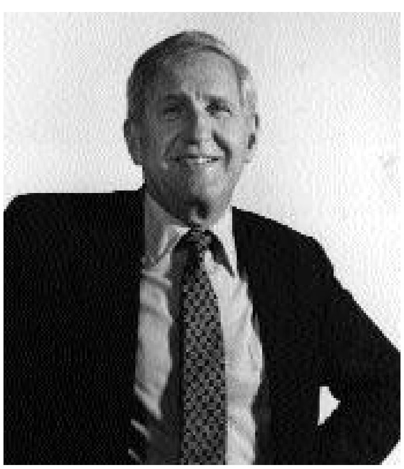
1951 First thermonuclear reaction is demonstrated in the George shot of the Greenhouse test series.

1952 The MANIAC computer becomes operational. The first thermonuclear explosion is achieved in the Mike shot of the Ivy test series. The first facility for handling liquid hydrogen on a large scale becomes operational. Plutonium-244, plutonium-246, americium-246, einsteinim-253, and fermium-256 are discovered in the debris of the Mike shot.

1953 The Lady Godiva critical assembly first achieves prompt criticality. The S_n, or discrete ordinates, method for solving neutron-transport problems is formulated.



Norris E. Bradbury, 1945–1970



Harold M. Agnew, 1970–1979

1954 The first thermonuclear bomb containing solid fusion fuel is demonstrated in the Bravo shot of the Castle test series.

1955 The Rover Project to investigate the use of nuclear reactors to power rockets is initiated.

1956 The neutrino is detected with the help of a recently developed liquid scintillator. The MANIAC II computer and the Omega West nuclear reactor become operational.

1957 The particle-in-cell (PIC) method for numerical fluid dynamics is invented.

1958 A helium-3 refrigerator providing temperatures below 0.45 kelvin is developed.

1959 Plutonium-238 is used as a power source in space.

1960 The KIWI nuclear reactor for the Rover Project is operated at full power.

1961 The Stretch computer is developed in collaboration with IBM.

1963 Satellite-borne sensors to verify adherence to the Limited Test Ban Treaty are developed. PHERMEX, the world's highest-intensity x-ray facility, is constructed.

_1964_ The world's highest-voltage Van de Graaff accelerator is completed.

_1965_ The Phoebus I-A Rover reactor is tested at full power.

_1967_ The side-coupled cavity is developed for the LAMPF linear accelerator.

_1968_ Funding for construction of LAMPF is approved by Congress and President Johnson.

_1969_ The ultra-high-temperature nuclear reactor (UHTREX) begins operation at 2400°F.

_1970_ Harold M. Agnew is named third director of the Laboratory.

_1971_ Naturally occurring plutonium-244 is isolated.

_1972_ LAMPF accelerates protons to design energy. Isotopes of uranium are separated by selective laser excitation of UF_6.

_1973_ Around this time insensitive high explosives for use in nuclear weapons are developed.

Donald M. Kerr, 1979–1985

_1974_ The Laboratory is named a national resource for stable isotopes.

_1976_ A portion of the Laboratory site is designated as a national environmental research park.

_1977_ Fusion neutrons are detected in a plasma confined by radiation from a carbon-dioxide laser.

_1978_ The Hot Dry Rock Program is initiated.

_1979_ Donald M. Kerr is named fourth director of the Laboratory. Universality of the approach to chaos in deterministic systems is discovered.

_1980_ The University of California establishes a branch of the Institute of Geophysics and Planetary Physics at the Laboratory. The Center for Nonlinear Studies is established.

_1981_ The Center for Materials Science is established.

_1982_ The Laboratory is designated as a national resource for flow cytometry. GenBank, the national database for nucleic-acid sequences, begins operation. A heavy-fermion superconductor is discovered.

_1983_ Congress approves long-term visits at LAMPF for citizens of the People's Republic of China.

_1984_ The radio-frequency quadrupole cavity is developed for a neutral-particle accelerator.

_1985_ Siegfried S. Hecker is named fifth director of the Laboratory. A new technique (CORRTEX) is developed to verify yields of underground nuclear explosions.

_1986_ The world's first high-temperature hot-dry-rock system is successfully tested.

_1987_ The first edition of nucleotide-sequence data for HIV samples is published.

_1988_ The Laboratory is designated as one of three national centers for human-genome studies. A new type of chemical bond is discovered in the binding of molecular hydrogen to the central metal atom in certain metal complexes.

Siegfried S. Hecker, 1985–

_1989_ A beam of energetic neutral particles is created in space.

_1990_ Superconducting tapes and thin films are fabricated.

_1991_ The Laboratory is designated as one of two centers for research on high-performance computing.

_1993_ Cross section for the scattering of electron neutrinos by electrons is determined experimentally.

TAKING ON THE FUTURE

Harold Agnew and Los Alamos scientists discuss the potential of the Laboratory

Harold Agnew and Paul White

The occasion of the fiftieth anniversary of Los Alamos National Laboratory offers an opportunity to celebrate the past and explore questions about the future. What is the Laboratory's mission now that the Cold War is over? What responsibility does the Laboratory have in maintaining the nation's nuclear capability? What kinds of activities are necessary to avoid surprises from foreign military technology? Why has basic research always been an essential part of the Laboratory? What is special about the Los Alamos culture? What new opportunities have been brought about by the end of the Cold War? How can the Laboratory contribute to the economic security of the nation?

Last October Los Alamos Science *invited Laboratory physicists, mathematicians, biologists, chemists, and computer scientists to discuss these questions in an open forum. To add historical perspective and a little more spice, we invited Harold Agnew to join us. Harold began his career at Los Alamos during the Manhattan Project and was director of the Laboratory from 1970 to 1979. He became president of General Atomic Company after leaving Los Alamos and is now retired. He is known for his candor, his enthusiasm for nuclear energy, his pride in Los Alamos, and his strength as a leader. Harold very effectively encouraged the participants to express their diverse interests and opinions.*

Here we present a condensed version of the day-long discussion. We thank everyone who participated and hope our readers will appreciate the individuality, talent, creativity, and passionate commitment to science and the nation that characterize these scientists and, in fact, the entire staff of the Laboratory.

Harold Agnew: I was asked to begin this discussion of the Laboratory and its future by commenting first on the past impact of the nuclear-weapons laboratories. Following World War II, the Lab's first big impact on world politics was in 1948 when NATO was first being formed. In a program called Backbreaker, Los Alamos and Sandia had the job of producing fifty Mark V fission weapons and then shipping them over to England on a newly initiated Air Force B-47 squadron. The Soviet Union had been gobbling up the Eastern European countries, but as soon as NATO was formed, backed up by the deployment of the Mark Vs overseas, Soviet expansion came to an abrupt halt. NATO's success in stopping Soviet expansion certainly enabled, in the long run, what transpired in the last two years; that is, the disintegration of the Soviet Union. Most people don't appreciate what transpired in the late forties and early fifties and the role the weapons labs played. We actually made all the weapons at Los Alamos and worked with Sandia on packaging those weapons in appropriate aerodynamic shapes. People also seem to forget that Sandia was part of Los Alamos in the early days. Now there's some talk of consolidating Sandia and Los Alamos under a single University of California contract as part of the plan to scale down the weapons program. It will be interesting to see what happens. I've given just one brief example of the importance of the United States weapons capabilities, hoping to illustrate how important our weapons labs are to the stability of both this nation and the world. Clearly the United States needs to maintain a credible nuclear weapons deterrent capability, and I think Los Alamos is in the best position to help do that job, not only because of its facilities but also because most of the weapons in the stockpile were designed at Los Alamos. I am somewhat concerned about proposals to make Los Alamos the sole nuclear-weapons laboratory, but if there is to be only one, I believe it should be at Los Alamos. On the other hand, I don't like the idea that Los Alamos may perhaps work on nothing but nuclear weapons. The Laboratory employs a tremendous group of technical people who can contribute to a broad range of national needs. I think your director, Sig Hecker, has done a tremendous job of fostering collaboration with General Motors, and with industry in general. That is the sort of thing the Labs are going to have to do. Now I'd like to hear your major concerns and ideas about the future of the Laboratory.

Greg Canavan: As I see it, the entire future of the nuclear-weapons program in this country is very uncertain. On the one hand, the military forces are essentially walking away from nuclear weapons as fast as they can. On the other hand, a lot of senior people in the defense establishment understand that nuclear weapons are a class unto themselves and that we must maintain nuclear competence regardless of whether these weapons are currently popular relative to smart conventional weapons. Those senior people will be around for the next decade or two and will ensure that the nuclear-weapons laboratories maintain some level of competence. But it is doubtful whether the laboratories will be asked to develop new nuclear technology. If you listen to the debates within the nuclear establishment, you hear some say it would be useful to develop an earth-penetrator weapon in case we have to fight a Gulf-type war again, but that's about the only new need that's mentioned. So it's clear that the size and scope of the nuclear establishment are going to contract pretty sharply over the next one or two decades. The question is: What does that mean for Los Alamos as a whole?

Los Alamos has always been a unique institution. Over the last five decades both the importance of nuclear weapons and the difficulty of advancing nuclear technology have required that Los Alamos be a fairly broad-spectrum laboratory at the forefront of a whole range of technologies, including materials, explosives, nuclear physics, atomic physics, radiation transport, and so on. But nuclear weapons are not going to be a catalyst for technology development in the next one to two decades.


Greg Canavan

It's clear that the size and scope of the nuclear establishment are going to contract pretty sharply over the next one or two decades. The question is: What does that mean for Los Alamos as a whole?

David Sharp: As I see it the need for national security is going to transcend the end of the cold war. The world is going to remain an uncertain place, and nuclear-weapons technology is not going to disappear. We will have to live with uncertainty on the political landscape as well as with the virtual certainty of new technological challenges. So we will need to maintain a nuclear-weapons capability. One of the keys to that capability is people. How are we going to retain a group of people who are smart, motivated, and knowledgeable about nuclear weapons? The only way is to give these people things to do that are interesting, challenging, and important. I have a couple of thoughts on what those things might be.

First, in the present political climate we need to develop the capability to design simple, robust nuclear weapons, the reliability of which can be guaranteed in the absence of full-scale testing. We have to learn to design weapons on the basis of better computations, better modeling, and the testing of components. Second, this Lab is rich in dual-use technologies, technologies that apply to both defense and non-defense problems, and we need to turn those strengths in new directions.

Ray Juzaitis: These are good suggestions, but the political process may have overtaken our technological preparedness. The constraints on nuclear testing imposed by the Hatfield provisions will prevent the type of deliberate and careful transition that you are describing. A few years ago Congress legislated a Test-Ban Readiness program, but at the same time, our budgets were cut, so we could not address all the technical issues and requirements that "test-ban readiness" implied. Now we are caught. We would like to put on the shelf more robust and safer weapon design as well as execute some good "bridging experiments" in anticipation of a zero-test environment, but the political climate may now prevent us from doing so.

Gene McCall: Just yesterday, during a private meeting, a fairly high political appointee in the defense department said to me, "If you think we're going to develop a new nuclear weapon, you're whistling up a drain pipe." I think that the Lab is going to have to adjust very quickly to a new and uncertain world. We will probably have to maintain a drastically reduced stockpile with a very small group of people. I suspect the number of people in the Los Alamos weapons program will be reduced to half of what it was a decade ago, which was 30 percent more than it is now, and that half of our weapons work will be nuclear and half conventional. That's probably a pretty good mix, and you might even find people who can transfer from one area to the other because the design codes and the experiments for both are similar—at least for the parts of nuclear weapons that you can test above ground.

As far as maintaining a nuclear-response capability, remember, Harold, you old-timers developed the first deliverable nuclear weapon in about three years. And you started out not knowing whether it would work and not having all the basic measurements needed for design. But we've got all that information now. If we wanted to respond to a national crisis, we'd be able to come up with a smaller, more efficient weapon—and do it faster—even if we started from scratch again. So, what do we want the program to be capable of doing now? We need to maintain the stockpile, but we need to specify what that entails. We must also be able to respond to future belligerent governments.

Unless we have a program that exercises our nuclear capabilities to at least the intermediate level of contained underground testing, there will come a time when no one will be able to certify a weapon and the whole image of deterrence with safe and reliable nuclear weapons will fall apart. Then I think we'll see a major political readjustment with respect to what it takes to maintain our nuclear capability.

Stirling Colgate: The key issue is the political perception of deterrence. Right now there's nothing to deter, but as soon as some government poses a threat to us, we'll have to face what we mean by an effective deterrence. My guess is that we'll be asked by Congress to certify nuclear weapons at the same high level of confidence as the airlines certify 747s. But unless we have a program that exercises our nuclear capabilities to at least the intermediate level of contained underground testing, there will come a time when no one will be able to certify a weapon and the whole image of deterrence with safe and reliable nuclear weapons will fall apart. Then I think we'll see a major political readjustment with respect to what it takes to maintain our nuclear capability. We will always be worrying about a dictator like Saddam Hussein building up a nuclear capability or a change in politics among the states in the former Soviet Union. So this Lab should be anticipating what it takes to maintain a responsible nuclear-weapons program that will generate an effective, acceptable deterrence for the future. We should be taking the high ground on this issue because we are the ones who will be counted on to provide the capability when the need arises.

Harold Agnew: And that capability can't be maintained without testing. Many people argue, incorrectly, that we need testing primarily to maintain reliability. But reliability is just one part of the whole enchilada.

Over the years the stockpile has been extremely reliable and quite safe. We've had an extensive surveillance program, and there have been very few glitches. Part of that confidence comes from continually designing, testing, and deploying many different weapons systems. When we start demilitarizing and cutting back, I worry that the diversity of systems is going to decrease and we will be in danger of having all our eggs, so to speak, in one basket. That could be disastrous.

Nobody is quite sure just what the right number is, but everybody starts getting uncomfortable when we start talking about reducing the stockpile below a certain number.

There is no question that the world ahead is going to be one of proliferation. More and more little nations will acquire some nuclear-weapons capability—maybe for terrorism, or to make themselves feel good, or to blow up a city or something. We're going to have to cope with that fact. But, the military's attitude is: "Look, we'll never get permission to use nuclear weapons. The rules for their maintenance and surveillance are a pain in the neck. And furthermore, they occupy a lot of ammunition-storage space. I don't want any part of them." Carson, you must remember General Kerwin's experience with the first group of Davy Crocketts that were sent overseas. These were small-yield nuclear weapons that had a range of a couple of thousand meters. It was a nightmare for him to maintain command and control of those things on board ship. He essentially had to put them under his bunk because they were short-range, tactical weapons, and you had to have them forward. If you had them forward then you were worried about command and control. They had to be authorized for use early, yet the probability of ever receiving such authorization wasn't realistic. Eventually the Army gave up on the whole concept.

Greg Canavan: In the strategic situation that persisted up to the collapse of the Soviet Union, we thought that we knew what the roles of nuclear weapons were. There were credible arguments that nuclear weapons could be used in this or that strategic or tactical scenario. The fundamental problem right now is that nobody can think of a use for nuclear weapons other than in the case of the resurgence of the Soviet Union. The services are denuclearizing as fast as they can. The Navy is throwing nuclear weapons off their carriers, the Air Force is downloading them from every airplane they can, and the Army has gotten rid of all of theirs. Our military leaders don't see any credible scenario in which nuclear weapons can do anything—except, perhaps, for using an enhanced radiation weapon on a Sprint missile as a defense against theater ballistic missiles. Short of the resurgence of the Russians, no one in the military services sees a clear need for nuclear weapons.

It's worthwhile to consider what would happen if the United States renounced possession of all nuclear weapons

Harold Agnew: One doesn't know what's going to happen in China. They are selling long-range missiles to other nations, and we don't seem to be improving our relationships with China.

Paul White: Most people who pause and think about what it would mean to have a complete free fall of the nuclear-weapons establishment stop and say that we need to keep at least a few hundred to a couple thousand nuclear weapons. Nobody is quite sure just what the right number is, but everybody starts getting uncomfortable when we start talking about reducing the stockpile below a certain number.

Harold Agnew: You have to keep a large enough number so that the Kadaffis of the world don't get the idea that just by having a few nuclear weapons they can become your equal. You've got to have a factor of a hundred or so above any credible number that they could conceivably whomp up.

Carson Mark: The idea of fighting with nuclear weapons is going to remain in free fall. That's why the Navy can get rid of certain classes of weapons and why the Army doesn't know what to do with the ones they have. But still, we need to maintain a position of deterrence.

Stirling Colgate: It's worthwhile to consider what would happen if the United States renounced possession of all nuclear weapons. What would happen to Western culture? Could we persuade the French, for example, to give up their nuclear weapons?

Harold Agnew: No way! If only France had them, then in two weeks Germany would be saying, "Mein Gott, we're not going to let those frogs do that to us."

Stirling Colgate: Now, if it's absolutely impossible for us to get rid of all our nuclear weapons, we can start talking more sensibly about what we can do. The United States, for the foreseeable future, is going to be the custodian of nuclear power for the world, and that is the challenge the Laboratory and the nation must face.

Gene McCall: Part of our problem is that plans for deterrence are being developed by political scientists who are simply throwing out numbers. But there are quantitative ways of estimating how many nuclear weapons we need to maintain in the stockpile. You pick a possible enemy nation and ask how many weapons we need to destroy 80 or 90 percent of its national wealth if we are attacked. If you do that for the whole world in reasonable scenarios, you can come up with a number. Now, it's probably not the right number because you'll never think of all the situations that could occur, but at least it's a number that has a quantitative basis.

Merri Wood: Any estimates you make today are going to be different tomorrow and different five years from now. Furthermore, there will always be a need to make changes in the stockpile, either for strategic or tactical reasons or because we're worried that elements of the stockpile are turning into silly putty. Then, what Stirling said earlier becomes an issue, namely, that after a while we're not going to find anyone to certify the new nuclear weapons.

But the real problem, as I see it, is that we'll find far too many people who are willing to certify new or modified nuclear weapons based on very little data, or maybe no data. That's the issue. At the nitty-gritty, working level, how do you maintain the weapons designers' judgement and experience in the absence of testing? When I interviewed fifteen years ago for a position as a nuclear-weapons designer, people were saying there might be a comprehensive test ban any day, so all of us have thought about that possibility for a long time. The real need for nuclear tests is to validate designers' and engineers' judgement about phenomena that can't be tested without nuclear explosions. It's the designers' judgement that's really critical. You can't be too conservative, and you can't be too risky—it'll get you either way.

Merri Wood

The real need for nuclear tests is to validate designers' and engineers' judgement about phenomena that can't be tested without nuclear explosions. It's the designers' judgement that's really critical. You can't be too conservative, and you can't be too risky—it'll get you either way.

Gene McCall: I think you're exactly right. The scary part is that there will be no shortage of people who are willing to certify untested weapons, especially if they are certifying their own designs, or if they want to please someone in Washington. If the laboratories cannot conduct tests, the United States should consider the possibility of eliminating its capability to design and certify nuclear weapons.

Stirling Colgate: Try testifying to Congress about how you have certified a nuclear weapon's capability without really knowing what you're talking about. They will hammer you to pieces.

Harold Agnew: Now, if the military has no more future requirements, why do you need somebody to certify something that isn't going to be used?

Merri Wood: But there will always be future requirements as long as we are required to maintain a stockpile.

David Sharp: I always felt that the function of nuclear weapons was to have them, but not to use them. If we are actually forced to use nuclear weapons, deterrence has failed. The point is not to imagine all the possible scenarios in which nuclear weapons might be needed. Rather, we have to maintain the capability to design a weapon that works precisely because we cannot anticipate all those possibilities. And the only way to maintain that capability is to have a group of people who are actually doing something reasonable, credible, and defendable that will maintain their designer's edge and that doesn't depend solely on full-scale testing. I've heard it said how important full-scale testing is and I personally accept those arguments, but I don't think they are going to prevail during the next few years.

Dual-use technology is another route to maintaining the Laboratory's capabilities. We are rich in dual-use technologies that apply not only to conventional warfare but also to other areas. I am concerned about how to maintain a cadre of people who can design nuclear weapons when we need them and who, in the meantime, are doing something useful that will keep them intellectually alive.

David Sharp

Dual-use technology is another route to maintaining the Laboratory's capabilities. We are rich in dual-use technologies that apply not only to conventional warfare but also to other areas.

Stirling Colgate: I object to this notion that you have to find something interesting for the weapons people to do. All you have to do is to define the problem. The Lab has to address the stewardship of nuclear power. Once you've defined that, you have to have people here who are bright enough to address the matter. You don't have to tell them what to do.

Here's a technical example. With normal drilling rigs like those we use at the Nevada Test Site, it would be rather simple to place a command post or nuclear-tipped missile about 6 kilometers underground. I can imagine Saddam Hussein or someone else in a hostile nation saying, "Well, we've got a lot of oil rigs around here. Let's put more of our command bunkers and a whole batch of our new missiles from China down at six kilometers." Now what does it take to deter him from saying, "We're now sending off one nuke just to show we have the capability, and we have fifty more down there." Each time we tried to destroy one, we would need about a 40-megaton explosion. I claim we would be absolutely stymied and that he could just come up and say, "You can't do it, so just do what I tell you." Now, there's a problem, and we should be trying to solve it rather than talking about what we are going to find for all these weapons designers to do. They're bright enough to think of the problem; they're bright enough to think of the solution.

Gene McCall: I think the military answer to that problem is that we go after the doorways of those underground bunkers with conventional precision munitions.

Stirling Colgate: Until you've conquered a nation, do you actually believe you will know where all those bunkers really are?

Jim Smith: Another part of the stewardship role is stockpile maintenance and weapons dismantlement. If we want to reduce the stockpile by ten thousand warheads, who will take the parts and where will we put those critical masses anyway? Although the public will say, "Make them disappear," Congress will understand that we need to maintain confidence in the materials and that the explosives and plastics change with time in a weak-radiation environment. They also eventually become unavailable from the manufacturers That's why you've got to bring those materials up to date. We can in a short period of time go back and re-invent the nuclear expertise if we have to. But the most important thing is that we have a collection of people here who are doing the best science they can, who can respond in time of crisis, and who know how to handle plutonium, and so forth.

Gene McCall: Jim, your question about what do you do with weapons parts brings up an interesting dual-use technology. The Japanese, for example, are creating a program to burn plutonium. Why don't we design safe plutonium reactors that burn weapon parts? That's something we can do well.

Many of the old-time weapons people…form a "reserve" of talent. As time goes on, these people won't exist unless we can recruit smart people from the universities … . Our ability to attract new people into the weapons program must be maintained as we attempt to diversify.

Harold Agnew: The Committee on International Security and Arms Control of the National Academy of Sciences is looking into what should be done with returned plutonium, but no one on that committee has any expertise in reactors or anything else of that nature. The idea of using plutonium for reactor fuel has come up, but the DOE is a basket case on anything having to do with nuclear reactors. Clearly Los Alamos could get involved in that problem. You have a unique facilities for studying critical assemblies in the Nuclear Technology and Engineering Division, and you have other relevant expertise as well.

Ray Juzaitis: Political support for the Laboratory as an ongoing institution will ensure that the Lab itself continues and diversifies its activities. But an important issue is to ensure that our core weapon-design folks participate in that diversification. These are the folks who are going to define their own futures based on their special disciplines, the types of technical problems they are used to solving, and the technical culture they in which they are used to operating. That "culture" involves an integration of experimental data, theory, and large-scale computing. The synergies, and the diversification that is relevant to the preservation of nuclear weapon design competence, will not and cannot be defined by external forces but must be defined by the designers themselves. These people are used to working in close-knit teams. Since the beginning, teamwork has been a very important element in the nuclear-weapons design. It's important to keep this core team together and evolving by attracting new people.

We sometimes talk about the two cultures at the Lab—defined by nuclear weapons work and nonweapons work. If we diversify in a way that continually isolates the core group of nuclear-weapons people, we will lose the capability to find the synergies between the two communities. In the past the folks in the nuclear-weapons program diffused into the non-nuclear culture, and vice versa, but they were still familiar with the nuclear-weapons problems. Many of the old-time weapons people are still here, and they form a "reserve" of talent. As time goes on, these people won't exist unless we can recruit smart people from the universities to help work on the weapons problems. Our ability to attract new people into the weapons program must be maintained as we attempt to diversify. Otherwise, the Laboratory will surely lose its competence in nuclear weapons technology.

Harold Agnew: One advantage that the Laboratory has over universities in attracting personnel is that of maintaining continuity. A university's primary function is to teach students, and students stay a while and then go someplace else. So a university is a difficult place to maintain continuity of a team. I agree with Stirling's point that somebody has to be, for the world—or at least for the U.S.—a chief honcho of nuclear capability. The Lab should stake out that role and make it very clear that that's the role it wants to play.

Merri Wood: In the past people have been very eager to work in the weapons program. You didn't have to twist any arms.

You're always going to find people who want to know, to understand, and to contribute even if the program exists in a limited, controlled mode.

Stirling Colgate: A postdoc in the Lab's astrophysics group who has run out of his or her time and looks around at the possibility of either staying at Los Alamos or competing in the university market or somewhere else will find that a position in the weapons program looks mighty attractive.

Greg Canavan: Historically, the technical challenges and the physics of the weapons program have been very exciting. When I was a graduate student, I got interested in some of the physics to the point that one day I woke up and knew I was going to work on weapons physics.

I see a world in which people don't regard nuclear weapons as a part of the solution any more but, rather, as part of the problem. … It may become a real problem to maintain confidence in the nuclear-weapons program.

Stirling Colgate: The curiosity I see among astrophysics students about how a nuclear weapon works, what's going on, and what can be done is just tremendous. You're always going to find people who want to know, to understand, and to contribute even if the program exists in a limited, controlled mode.

Greg Canavan: Right now we have good people in the weapons program, and we're able to recruit more. But I see a world in which people don't regard nuclear weapons as a part of the solution any more but, rather, as part of the problem. Just this past year Livermore's symposium on high explosives was viewed by some as a school for would-be proliferators. I see a growing tendency at the Lab to sharpen the break between the people who work in weapons and those who work in non-weapons areas inside the Lab. That means diffusion of talents into and out of the weapons program will be more difficult in the future. In the next decade it may become a real problem to maintain confidence in the nuclear-weapons program.

Gene McCall: We've been talking mainly about weapons designers, but designers are the tip of the iceberg in the nuclear weapons program. What about the people doing the fabrication and diagnostics?

Merri Wood: Even the politicians recognize the role of diagnosticians in maintaining the stockpile. But when there's no nuclear testing, what role do the diagnosticians have? The physics of nuclear weapons operates in a unique regime of extreme physical conditions, and without tests, there's no place to practice.

Chick Keller: You observe solar eclipses.

Merri Wood: Those happen on a much longer time scale, minutes instead of microseconds, and you are flying around in an airplane or are on top of a mountain so you can adjust and calibrate your equipment right up to and after you take data. In a nuclear test you have to bury all of your equipment in a narrow, deep hole weeks before the test, it must work perfectly the first time, and it must continue to take good data for some period of time in the extremely hostile environment created by the nearby nuclear explosion. Post-shot calibration is, of course, not an option.

Chick Keller: But the physics and the diagnostics are not unrelated. I suggest that we keep the testing team together by doing a variety of related things. In the present situation we are unlikely to have the funding available to keep the team together for the primary mission alone. I used to be in J Division, the old nuclear-testing division, and I'd like to think of myself as part of that broader team. Some of us from that division are looking at the concept of dual-use technology and suggesting that a very large effort be developed that would join the global environmental studies the Lab is now doing with diagnostic capabilities, data integration, and so on. The effort would also include working with industry on sensors and other technologies that have dual uses. There have been many precedents for scientists working simultaneously in both the nuclear and non-nuclear areas, particularly scientists working in the testing program. We don't want to lose our testing capability because the world situation may change. But we can use those scientists in many other ways that relate back to testing.

The physics of nuclear weapons operates in a unique regime of physical conditions, and without tests, there's no place to practice.

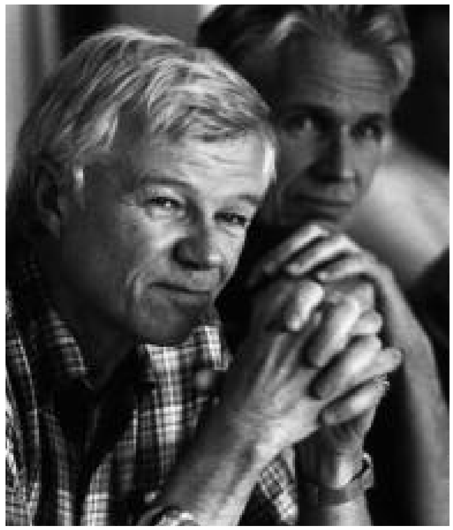
Stirling Colgate: There was a year when seven out of nine Los Alamos nuclear tests of weapons in Nevada were designed by scientists who came to X Division—the Applied Theoretical Physics Division—from astrophysics with the assurance that they could work quarter-time in astrophysics and publish that research. They have published and continue to publish frequent and important scientific results. It can't be all bad to be split a little bit.

Erica Jen: How true is it that weapons work is something of a one-way street in the sense that you are unable to publish your results in the open literature or present them at conferences and that you can't easily move back and forth between the classified and unclassified sectors? I think that the perceived lack of open scientific exchange, together with skepticism as to the scientific vitality of the weapons programs, present difficulties in attracting people from basic research.

Gene McCall: People do weapons work because that's what they want to do—not because they are trapped into it. And although they can't publish characteristics of a weapon, they can publish much of the basic development work. They simply choose not to publish.

Erica Jen: A while back there was an effort to start a peer-reviewed classified journal and to encourage people to publish classified results. Did that ever get off the ground?

Ray Juzaitis: There is a defense journal and there are classified conferences. So, those professional needs are met to some degree. However, the peer community is small, and results are often communicated directly to those who are interested. I would say that most people have not felt the need to publish in the open literature as long as there was assured technical vitality in the nuclear weapons disciplines.



Gene McCall

People do weapons work because that's what they want to do—not because they are trapped into it.

Merri Wood: A whole lot of work goes into solving specific nuclear design problems. The work is well documented, and we're very proud of the documentation. So our work is very satisfying. But if this country decides that it doesn't want any more nuclear designers and we have to go outside the Lab to find a job, then we'll have all this wonderful classified work that we can't show to anyone. I think I disagree with Gene. At least in my area of work, I can't publish the results. And the rules for classification are right on the money. I wouldn't want to publish that stuff. I don't want any proliferators getting their information from me. We do this kind of work at national laboratories for a reason. It doesn't fit in at universities or in industries. There's a reason we don't just run out and transfer this technology. We need this technology to build nuclear weapons, but it will not build a better toaster. It's interesting, challenging work. It's hard work, and we think it's important and useful work. But when it comes time to turning it into something else and still preventing proliferation, there are some real barriers.

Harold Agnew: It would be interesting if Edward Teller were at this discussion because he has been on a binge of saying, "Declassify *everything*! Secrecy is bad!"

When I was Lab Director, I know some people were really antagonistic toward the weapons program. I must say I have no sympathy with those people, and if I hadn't been so tolerant, I would have kicked them out. This Lab had a mission, it still has that mission, and it needs to be supported. That doesn't preclude the Lab from doing other things, but I don't think you should be ashamed of or apologize for what has been the primary mission of the Laboratory. And if you follow what Gene and Stirling were suggesting, that mission still requires a rather sizeable program, although the elements in that program may change. For example, I see the Lab making greater and greater contributions to the intelligence field and to efforts to prevent the proliferation of nuclear weapons.

Our government's ability to draw on people from the Lab to evaluate technology development and monitor the activities in other nations has been a real national asset and is appreciated very much by the various intelligence agencies.

The intelligence division at Los Alamos has recently done some spectacular things. Since the early 1960s I can remember observing overhead pictures of one particular facility in the Soviet Union, and just recently some people from the International Technology Division actually visited that facility! We thought that facility had a very different purpose, but now we know that it was part of the Soviet nuclear rocket program. They also learned of another enormous facility located in a huge mountain in Russia, which was designed to test full-scale nuclear reactors and even to blow them up. The Lab could start a joint program with Russia to use some of those facilities for nuclear research.

Greg Canavan: Although the role of nuclear weapons is declining, the weapons business in general is not winding down. The next decade may be relatively benign, but the ones after that look anything but gentle to me. There's enormous potential for military conflict at all levels around the world. And like it or not, the weapons enterprise, particularly in the area of conventional and smart munitions, is likely to expand. This Laboratory has very strong expertise in conventional explosives and a whole range of other related technologies. I believe the Laboratory should play an integrating role in the area of conventional and precision weapons technology. I see us applying our technology to new sensors and new means of manipulating information for the delivery of precision weapons.

The Department of Defense is also interested in simulations of virtual reality in which virtual prototypes of new hardware such as tanks, missiles, or new weapons concepts are created and their performance in battlefield situations is simulated on the computer. The idea is to use high-speed computation to test ideas and concepts even before the hardware is built.

The next decade may be relatively benign, but the ones after that look anything but gentle to me. There's enormous potential for military conflict at all levels around the world. And like it or not, the weapons enterprise, particularly in the area of conventional and smart munitions, is likely to expand.

The Lab has an enormous amount to offer in this area not just in terms of its big supercomputers but also in terms of people in the nuclear weapons program who for most of their careers have been working with codes of superhuman complexity. These people have developed a knack for the essentials of very complex simulations even when they don't understand every detail of the physics that is put into each code. We've already done such detailed simulations for the Strategic Defense Initiative Organization. The components of those simulations could be fit together in new ways that would allow us to test almost any new weapons or sensors or communication concept that the Defense Department would like to try.

Gene McCall: Virtual-reality simulations are now used to simulate what a pilot sees on the radar display of his aircraft. The system detects, or captures, a radar pulse coming out of an airplane, processes it, plays it back, and makes it look to the pilot as though he or she is flying over Albuquerque or Baghdad, for example. The system can be both a training aid and a maintenance aid.

Greg Canavan: It's not only a training and maintenance aid. It can also be used for testing and evaluating actual or virtual hardware in a simulated battle environment.

Gene McCall: Right. You can use this capability to simulate what it takes to destroy a tank, for example. And this type of simulation is relevant to an even larger and more pressing problem, namely, modeling the transportation infrastructure of this country. How does one do transportation modeling? How does one incorporate revolutionary new devices like listening systems into the transportation system? These kinds of problems are already being explored, and they are well-matched to the Lab's capabilities.

Chris Barnes: We already have two efforts in transportation along those lines. One is aimed at simulation-based virtual prototyping of the entire transportation system, and the second has to do with evaluating the design and architecture for intelligent-vehicle highway systems.

Stirling Colgate: In astrophysics we also like to model reality. One of the emerging reality games in astrophysics is the development of smooth-particle hydrodynamics codes or many-particle simulations (greater than 10 million) to model galaxy formation or supernova explosions. We need to try this new reality game on old nuclear-weapons designs and compare the results with old test data and with the results of our standard weapons codes. The resulting competition among various computational techniques will greatly improve our confidence in weapons simulations.

Greg Canavan: I was asked by a general on the Joint Chiefs of Staff to write a paper about what warfare would be like forty years from now. That's so far down the line that we can't predict what technologies will have emerged by then. But we do know that if we continue to improve computers by a factor of 2 every 2 years for the next 40 years—that's a factor of 2^{20}, or about a million—then the best computer will be roughly as smart as a chicken!

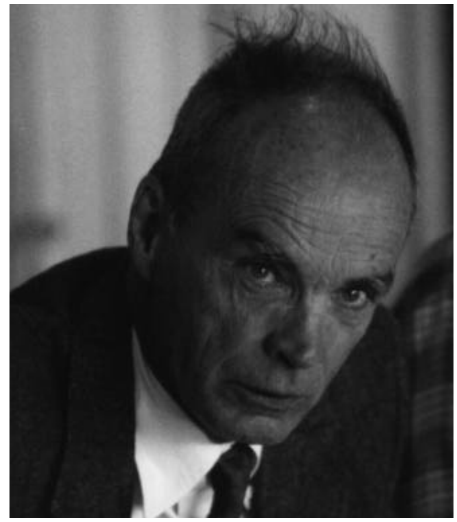
So even over the long term, the smartest military machine and also the smartest commercial planner is still going to be a person. The trick, then, will be to maximize the decision-making performance of individual persons: to get information to them, help them evaluate possibilities so they can make decisions, and get that information back to the battlefield or marketplace as quickly as possible. All the person in the loop is going to do is make decisions.

To maximize human performance in both the military and non-military sectors, we are going to see a convergence of biotechnology, computer hardware, and computer software on a time scale of thirty or forty years from now. Los Alamos has a significant research component in each of those three areas right now. And we need to keep investing in them because they are the wave of the future.

Stirling Colgate: The Lab is in a position to make many kinds of contributions to the military and the civilian sectors primarily because we have a strong basic-research effort. And the strength of that effort was built on the fact that it takes mathematicians and physicists to make a nuclear weapon. That core ability in math and physics interacts with other basic-research efforts at the Lab, whether it's the Human Genome Project or research on high-temperature superconductors or global climate modeling. The edge that the national labs have in math and physics is, I believe, our greatest strength. It means we have the presumption of being able to understand almost anything and the capability to offer that understanding to all areas of applied research. In our interactions with the agencies in Washington, we need to emphasize that we can do computing, mathematics, and physics. You can't find that depth and breadth of expertise in any other laboratory.

Gene McCall: The experimental component at this Lab is also very strong. We do very good experiments, and we do them in a hurry.

Stirling Colgate: I didn't mean to imply that physics is all theory, but the unique thing that has come out



Stirling Colgate

The Lab is in a position to make many kinds of contributions to the military and the civilian sectors primarily because we have a strong basic-research effort. And the strength of that effort was built on the fact that it takes mathematicians and physicists to make a nuclear weapon. The edge that the national labs have in math and physics is, I believe, our greatest strength.

Harold Agnew

The National Research Council had a panel on nuclear energy, … loaded with people from the establishment, and the report said everything was fine. No one was asking, "Is this really the best reactor technology?   What could we do in the future?" There's a tremendous role for that sort of adjudication, and…the labs have the expertise in the hard sciences do it honestly.

of the nuclear weapons program is physics, math, and, if you want, computing.

Gene McCall: And, Stirling, something else that is quite unusual about Los Alamos is the close and very positive association between theorists and experimentalists.

Stirling Colgate: We need to move the Department of Energy away from the idea that the Lab's role is a relatively narrow one, that of maintaining nuclear weapons. In addition we must emphasize the value of maintaining a significant level of scientific expertise that can be tapped for a very broad range of problems.

Harold Agnew: Having people in a multitude of disciplines who can jump on problems such as using image enhancement to investigate the Challenger disaster is a tremendous asset that the nation can use to great advantage. The Lab expertise can be applied to some really pressing problems, one of which might be the problem of waste disposal, not just nuclear waste but chemical and toxic wastes of all kinds.

Andy White: On a visit to DuPont a year or so ago, I heard the suggestion that the national labs should play the role of referee in waste-disposal problems and other controversial issues. The labs could be the final authority making technical decisions on things like deep-well injection of industrial waste.

Stirling Colgate: The strength of our technical culture should be applied to the entire business of risk assessment. The national labs are particularly situated to participate in this area.

Harold Agnew: That type of arbitration is certainly needed for nuclear energy. The National Research Council had a panel on nuclear energy, but the panel was loaded with people from the establishment, and the report said everything was fine. No one was asking, "Is this really the best reactor technology? What could we do in the future?" There's a tremendous role for that sort of adjudication, and as Stirling was saying, the labs have the expertise in the hard sciences do it honestly. But I want to add that I have a bachelor's degree in chemistry, so don't forget the chemists. Originally, all we could do with plutonium was take a glob, cast it, and machine it. We never made any improvements until the chemists and the physicists joined together to really understand how to handle these materials and develop the alloys.

Paul White: At this moment, the Laboratory has a unique opportunity to define its future, one that will be more closely tied to the economic security of the nation. Right now we are in the process of defining unique niches in the marketplace that—because of our history, the capabilities we have assembled here, and the way we do business—are particularly suited to us. Risk assessment, of nuclear reactor technology or of waste disposal options for example, might be one area. What are some others?

Andy White: Computational science is an area in which the Laboratory already excels, and it is an an important component for the Lab's success in the future. Our Advanced Computing Laboratory, or ACL, is designed to provide an integrated computational environment with sufficient computational power for solving the overwhelming scientific problems that are critical to the nation. These "Grand Challenge" problems relate to the environment, national security, economics, and so on, and they require computational resources significantly greater than the resourses generally available today. To attack these problems, we acquired a CM-5 massively parallel computer from Thinking Machines Corporation and integrated this state-of-the-art machine with appropriate networking, storage and retrieval, and visualization resources. Right now we are primarily involved in three "Grand Challenge" applications: global ocean modeling, multiphase flow in porous media, and molecular dynamics simulations of novel materials.

Working with others at Caltech, Jet Propulsion Laboratory, and the San Diego Supercomputer Center, we are part of one of the nation's five gigabit testbeds to investigate geographically distributed computing between large computational resources. Another exciting project that has recently been initiated is the formation of the Computational Testbed for Industry through which we are collaborating with many industrial firms on a variety of problems of mutual interest.

Greg Canavan: Computer simulation and information processing is an area where the Lab has an edge not only because of our tremendous computing power but also because we have the people who know how to harness and use that power. The applications are limitless. They range from the virtual-reality simulations I was talking about earlier to basic problems in fluid dynamics, elementary-particle physics and the evolution of the universe, and to information processing in the marketplace. New opportunities are opening up all the time, and Lab personnel have the know-how to take advantage of them.

Chris Barnes: I come from X-1, a group in the division that designs weapons. About three or four years ago a few of us decided to change directions and learn about neural nets and their application to information processing and massive data manipulation. It was a risky thing to do, but it's really paid off and now we're developing new kinds of neural nets and applying them to a whole range of problems. One relates to preventing proliferation of nuclear weapons through export control.

Preventing the export of equipment related to nuclear-weapons development requires sifting through a vast amount of data. An evaluation of the global data on purchasing, exporting, or licensing of certain types of equipment or combinations of equipment can provide clues about the extent of a nation's nuclear development. Neural nets are a perfect tool for analyzing that data automatically. Another example relates to the interest of the Internal Revenue Service in identifying multinational corporations that are shifting their books around to avoid paying their proper taxes. The types of information processing required for that problem is almost a perfect overlap with nonproliferation problems.

We originally got into this type of information processing because a large banking corporation wanted us to develop a system for doing portfolio management and for determining future profitability of credit-card accounts. Our work on neural nets clearly falls into this category of dual-use technologies. We are playing a useful role in the diversified weapons program as well as in the public and private sectors.

Andy White: Massive data manipulation is an important problem in research as well. We are starting to perform very large global climate simulations in connection with the Department of Energy's High Performance Computing and Communications and Computer Hardware, Applied Mathematics, and Model Physics programs. Soon one simulation will produce as much data as we've stored on the Common File System over the last sixteen years! The first problem is where do you store all the data, and thesecond problem is how do you analyze it and use it to understand what you've done. Developing tools to store, handle, and analyze massive amounts of data is a problem that cuts across disciplines, and the Laboratory is now making tremendous strides in this area.

Information highways are being built across the country, and Los Alamos is on the freeway, the electronic freeway.

Greg Canavan: Many of us worry that the cost of doing business at a laboratory located at the end of a road to nowhere will price us out of the market in some areas of research.

But that's not an impediment in the area of information processing.

David Forslund: Information highways are being built across the country, and Los Alamos is on the freeway, the electronic freeway. Geographic isolation is not a serious problem for the Laboratory in this information age.

Greg Canavan: Not only are we on the electronic freeway but also the value added to many products by our analysis can be so high that we could even live with the high costs.

Gene McCall: Everyone says the Lab should participate in making the U.S. economically competitive and that we should cooperate with industry. But we need a way of developing joint projects that will really make a difference. One approach would be to send 500 people from the Lab for one-year sabbaticals in industry. That way we would wind up with a cadre of people spread throughout the Lab who have some real understanding of current industry concerns.

Harold Agnew: It's also a good way to build up a large constituency around the country.

Chris Barnes: We recently started working on a Cooperative Research and Development Agreement with General Motors, and one of our guys has already talked about doing a sabbatical at GM. Since we already have a group working in this area, he'll have a place to come back to. I believe it's very important to build up a core group of experts at the Laboratory in a particular field before we send people on industrial sabbaticals.

Gary Doolen: One of the most exciting new directions at Los Alamos is nanotechnology, a field that includes research on three very novel areas: self-assembling computers with molecular components, tiny robots that incorporate biochemical sensors, and improved materials with very small grain sizes.

Nanotechnology research spans many different disciplines and will require a substantial investment over a number of years as well as coordination and focus on a scale beyond the range of normal university research. At present the federal government has taken no central-planning responsibility in this area. Some industries have initiated research in nanotechnology, but they are currently downsizing their efforts for economic reasons. In contrast, in 1992 the Japanese invested $200 million in this area because they recognize the potential for large long-term payoffs.

In 1991, several universities and large industries concluded that the Department of Energy sould coordinate a national effort in nanotechnology. Scientists at Los Alamos and Sandia are now working in this area, and each institution has an experimental and theoretical research and development initiative funded at the $1 million level. At Los Alamos, the support comes from the Laboratory Directed Research and Development Fund.

One major success, resulting from a collaboration among Los Alamos, Yale University, IBM, and the University of South Carolina, has been the design and production of conducting molecules that can perform the function of ordinary wires but are thousands of times smaller. On the ends of each molecule one can place chemical groups that bind tightly to selected metal contacts normally found in miniature circuits. The conducting molecules have been made in large quantities and tested for uniformity and integrity using atomic-force microscopes. When these molecules are poured onto a surface containing prearranged metal contacts, the molecules bind very strongly to the contacts; in other words, they self-assemble. Plans are being made to design transistor-like molecules that self-assemble in a similar fashion.

The hope is to combine these self-assembling molecular components with existing technology for electronic switching that is a thousand times faster than the technology used in today's most advanced computers. The combination is expected to yield self-assembling computing devices that are about the same size as today's chips but many times faster. Similar work in the areas of robotics and biochemical sensors also have large anticipated payoffs in both the civilian and military arenas. The Laboratory has unique research and management capabilities for this type of dual-use research, and many people believe Los Alamos should play a major role in nanotechnology.

Erica Jen: Nanotechnology is an example of a program that originated from the efforts of a few individuals who had an idea and the liberty to explore it. It's also a program with the potential to grow quite large, and everybody recognizes that the Lab has the responsibility as well as unique capabilities to develop and to support large programs. But there is also the question of support for small projects and individual researchers who are not tied to any particular program with measurable goals and product specifications and who are not responding to an already existing funding directive. Without such support, you can't expect freedom, diversity, or creativity. At present, individual researchers cannot help but see that decisions and directives at the Laboratory are often governed not by the scientific health of the research efforts but rather by short-term criteria heavily biased toward large-scale programmatic efforts and by fuzzily understood funding policies. The effect on morale and productivity has been disastrous. And I would argue that the nurturing of small projects with as yet no foreseeable connection to big projects is an approach that would enormously increase the institution's ability to make significant scientific and technological advances.

Gene McCall: The type of work you describe has in the past been funded on the budget noise of the large programs. That is why the inertial-fusion program, for example, did not come out of basic research. It came out of the weapons program and, in fact, out of the field-test division of the program. The magnetic-fusion program and the accelerator division came out of the weapons program.

Erica Jen: Still, big programs are much more likely to succeed at the Lab if the people who had the original idea are here and are communicating with other people around them. The work Chris Barnes is doing in applying neural nets to massive data-manipulation problems grew out of the fact that researchers at the Center for Nonlinear Studies were playing with these new ideas, in a rather theoretical and abstract framework, and their research inspired him to get involved. If that basic research program had been at, say, Berkeley rather than at Los Alamos, Chris Barnes would probably never have begun that work, and external researchers would have had little reason to turn to Los Alamos as the natural focal point for related programs. The same is true of Los Alamos research programs in applications of chaos theory and in lattice-gas simulations, which grew out of the research work of individuals here at the Lab, namely, Mitchell Feigenbaum and Brosl Hasslacher, respectively. So how can we continue to support the generation of new ideas?

Erica Jen

There is also the question of support for small projects and individual researchers who are not tied to any particular program with measurable goals and product specifications and who are not responding to an already existing funding directive. Without such support, you can't expect freedom, diversity, or creativity.

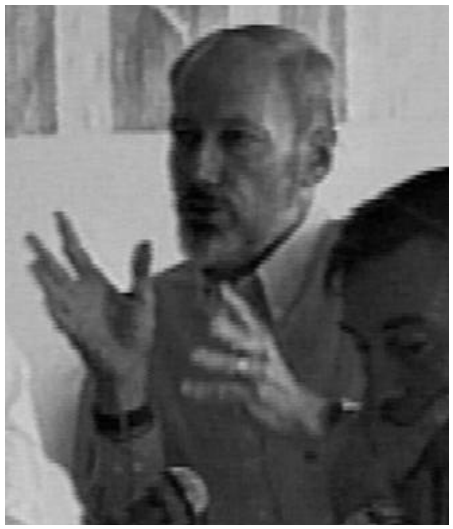
Andy White: It is a difficult problem, and it is a national issue, not just a Lab issue. There's a constant war between big centers and individual research projects. The National Science Foundation is studying the balance at the Foundation between basic and applied research and could make major changes in the current mix.

Gene McCall: That's why I think the Lab must protect the basic research we have by supporting that research within larger programs; otherwise we're going to lose our basic-research capabilities. If we identify basic research as a separate item and hang it out there, it's going to get cut off.

Greg Canavan: The steps the Department of Energy and the Lab have taken to assure Congress that they are in control of the Lab's programs are tending to undercut precisely those things that made Los Alamos a first-rate laboratory in the first place. In other words, accountability is tending to cut out cross-fertilization and make it harder for people to get together and to do innovative projects. In the ten or eleven years I've been here, I've seen a big loss in flexibility.

Harold Agnew: Is the size of the Laboratory part of the problem? It was pretty large when I was director but not as large as it is now. In the good old days we had a theoretical division, a physics division, and divisions for explosives chemistry, weapons tests, health, and maybe a few more. I don't know what the mechanism is now for gathering new ideas, starting new programs, or deciding you are going to send five hundred people out to industry to increase technology transfer. In the old days it was pretty much the division leaders who made the decisions. Carson Mark, the leader of the Theoretical Division, would decide what he was going to do, and Dick Taschek, the leader of the Physics Division, decided what he was going to do. When I became director, we would talk these things over once a month, on a Friday. I provided liquor, which I guess now would be considered scandalous, but it certainly allowed people to forget their inhibitions. We didn't have that many division leaders or associate directors, so everybody could come and speak out.

Chick Keller: I think many of our problems stem from the over-regulation of the Lab, which, in turn, stems from the fact that we are perceived as having no mission. If you have no mission, you've got to get something done, and the health and safety and financial people are told, "Get in line and help make it happen. Right now!" Those people have no vested interest in our productivity.


Chick Keller

I think many of our problems stem from the over-regulation of the Lab, which, in turn, stems from the fact that we are perceived as having no mission.

David Sharp: Right now a lot of the things that are costing us time and money are being imposed on us by people whose first interest is safety and environment *not* the scientific productivity of the Lab.

Jim Smith: Here's an example. Nineteen years ago I used to carry a gram or so of plutonium in my car from the CMR building down Diamond Drive to DP site. It was legal. I didn't break any rules. Today, we've got to pay for extra security, so plutonium workers cost $600,000 to $700,000 per person each year. And if you want to haul a few grams of plutonium down a road, you have to close the road and bring out the machine guns. All these changes came about because somebody was worried. Nothing ever happened the old way. I never diverted any plutonium. But people in Washington decided it had to be done "right." And right is unbelievably expensive these days.

Harold Agnew: The safety and environmental concerns are here to stay, although we may learn to prioritize them better. I am more interested in the ways in which the Lab is diversifying. The era of greater and greater funding is over. It's probably going to be less. If you want to do some of the new things you've been talking about, I think you have to say, "Look, we think the future will be much better if we do this, so we're going to cut out that." You don't necessarily have to get new money, but you do have to have people who are willing to give up what they are doing and start something new.

Gene McCall: I remember the story of someone once asking Harold how many people work at Los Alamos, and he said, "About half of them." Well, I would say that about half the people at the Lab are willing to change what they are doing as long as they perceive that the people up the line will support them in whatever they are asked to do.

Chris Barnes: The individual has to be willing to take some risk. In my section in the Inertial Fusion and


Jim Smith

I used to carry a gram or so of plutonium in my car from the CMR building down Diamond Drive to DP site. It was legal. I didn't break any rules. Today…if you want to haul a few grams of plutonium down a road, you have to close the road and bring out the machine guns. All these changes came about because somebody was worried. Nothing ever happened the old way. I never diverted any plutonium. But people in Washington decided it had to be done "right." And right is unbelievably expensive these days.

Plasma Theory Group, we had three or four internal proposals for new directions, which were fought for and finally agreed upon by the group. However, we had to take some risks to change fields from plasma physics to neural nets. After four years, it's finally beginning to pay off. Our section is funded, and soon we're likely to be over-funded. But there were a lot of people who thought we were totally nuts to go off and do neural-net work.

Harold Agnew: Flexibility needs to be built into the management as well as the staff, especially during times of change. The last years when I was here, I implemented a tenure limitation for group leaders, something that had never happened at Los Alamos. At first it was traumatic for this small community because people would say, "Gee, your daddy isn't group leader any more. He must have done something wrong." But I used the analogy of academic departments where the position of chairperson is not coveted by most people in the departments, and there's no social stigma for not being the chairperson any more. We also tried a tenure limitation on division leaders, but it was a little harder for them to accept, and we didn't implement it completely. Are there such limitations here now?

Basil Swanson: There's an experiment going on right now in the Isotope and Nuclear Chemistry Division in which group-leader appointments are limited-term appointments. Jill can describe how it works.

Jill Trewhella: I just recently moved to INC Division as leader of the Spectroscopy and Biochemistry Group. I was willing to take that position because it's a three-year limited-term appointment. Historically, INC Division has been strongly associated with the nuclear-weapons program, and until recently, 70 percent of its funds came from that program. Now only 30 percent of the funds are from weapons programs, and the division has successfully redirected its efforts into areas such as environmental restoration and basic science. That's been done by being flexible, looking to the future, and reorganizing the Division. The Division instituted limited-term group-leader appointments specifically to address the issue of flexibility and ensure that the Division responds to changes as needed. The group leaders are funded to spend 50 percent of their time doing research on programmatic efforts so that when they're finished with their three-year appointments, they can make an easy transition back to research without the trauma of having to start from scratch. It's worth noting that the Division has gone from a $35 million budget for the last fiscal year to a projected $45-million budget for the current fiscal year. So, there are new ways of doing things that can be very successful in the present environment.

Gene McCall: The key to maintaining the technical capability of the group leaders and division leaders is to keep them involved in technical work. Years ago even the division leaders did some technical projects. Nowadays, there's hardly even a deputy group leader who does a technical project. In X Division we've had three or four group leaders voluntarily leave their management positions over the last three years because their jobs were so far removed from the actual research.

Jill Trewhella: The reason group leaders in INC Division can keep doing research is that the bulk of the bookkeeping, administrative, and environmental, safety, and health activities are being handled by a separate business and operations unit. The group leader is thereby free to be a technical group leader and stay much closer to science.

I come from the biosciences, which is not a traditional area of strength at the Lab, so my colleagues and I have been forced to be very outward looking, to band together with whomever we can, and to develop joint initiatives such as the new initiative in structural biology.

Harold Agnew: Biosciences and materials science are areas that are really blooming in this country. In the San Diego area every other outfit is a bioscience something-or-other with a "tech" at the end of its name. Does that make your life easier? Is it like nuclear physics in 1943 when the whole world was our oyster?

Jill Trewhella: It helps and hurts at the same time. There is new money in biosciences, but as a result, everybody is jumping on the bandwagon—whether it's appropriate or not. So the quality control gets difficult at times. The Director wants to see biosciences dramatically increase at the Laboratory, but we have to step carefully and think very judiciously about what directions we promote because this is a relatively new area here.

Jill Trewhella

But the greatest strength of the national-lab environment is the opportunity for peer collaboration—the willingness of the people who are at your level to work with you and share ideas. In a university you're so scared, you want to get tenured, and you don't want anybody to write a paper with you unless that person is a student or a postdoc—that way you know your name will come first.

Harold Agnew: In the olden days we got our finger into the biosciences by irradiating rabbits with neutrons and stuff like that to study the effects of radiation. That was the only area of biological research we could justify because we were basically a weapons laboratory.

Jill Trewhella: The biosciences have grown a lot since then. We have initiatives in genome sequencing and informatics, structural biology, medical isotopes, medical applications of lasers, bioenergetics, and so on, and we're supported by funds from both the DOE and the NIH. As a biophysicist, I see the Laboratory as a glorious environment in which to work. Laser technology, high-field magnet technology, neutron scattering, and so on are all available to me here. At other institutions, I would never have had the opportunity to get anywhere near them.

But the greatest strength of the national-lab environment is the opportunity for peer collaboration—the willingness of the people who are at your level to work with you and share ideas. In a university you're so scared, you want to get tenured, and you don't want anybody to write a paper with you unless that person is a student or a postdoc—that way you know your name will come first.

Harold Agnew: You are so right. I'm an adjunct professor at the University of California, San Diego, and I find that nobody talks to anybody else at that institution—even in the same department sometimes.

Jill Trewhella: In the career that I've had at Los Alamos, it's not at all unusual for me to work and publish with two or three other career-equivalent people. It's wonderful.

Stirling Colgate: This is perhaps the most important aspect of the culture of the national labs. Peer collaboration is a way of life here, but it happens much less at the universities because of the initial pressure to attain tenure and the later pressure to obtain rewards, such as research support, based on individual creativity.

Larry Deaven: And at Los Alamos your peers may be other biologists or physicists, mathematicians, engineers, and chemists. I can point to two areas of biological analysis that emerged within the Life Sciences Division because of that kind of unique interaction. One was the development in the late 1940s of liquid scintillators for detcting ionizing radiation. Those scintillators enabled biologists to study metabolic processes at the molecular level and helped to redefine biology in terms of biochemical and biophysical principles. The second was the development of flow cytometers, which are instruments for sorting and analyzing cells. Those instruments are located in all major hospitals today. where they are used to diagnose cancer and AIDS. It was the interaction between life scientists and physicists here at the Laboratory that gave rise to these unique technologies.

In the early days of the Laboratory, much of the life-science research was involved with irradiating whole animals to refine our knowledge of radiation exposure and to establish radiation-protection standards. When I came to the Laboratory in 1971, we had a small group of cell biologists who were studying the effects of radiation on cells and subcellular particles. Our health-research unit was small compared with those in other national laboratories. However, during the next ten years, we very carefully positioned ourselves in the direction of modern biology, and for that reason, we have a large, multidisciplinary human-genome center at the Laboratory today.

Larry Deaven

It's clear that health-science research and especially preventive medicine are going to be major agendas for this country as the need for defense expenditures winds down.

It's clear that health-science research and especially preventive medicine are going to be major agendas for this country as the need for defense expenditures winds down. The kinds of discoveries being made in the human-genome project—new information on the genetic component of disease and new diagnostic tools—will have a big impact on medical science and the practice of medicine only if delivery systems can be devised to make them available in a cost-effective manner. I believe that the Laboratory has the expertise to help develop affordable delivery methods, and I think we should try to initiate new projects to address the problems.

Gene McCall: We should be a little cautious. We should remember a few phrases, such as synthetic fuels, renewable energy resources, solar energy for the masses, and others that ten years ago were the wave of the future but are hardly ever mentioned now by the mainstream funding agencies. We should build our own base in the biosciences, and we should avoid responding only to what looks fashionable right now.

David Forslund: The unique strength that this Lab brings to all these problems is the combination of ideas and talents from physics, chemistry, biology, materials science, and so forth. This broad combination of people and technology enables us to envision new areas of research and applications that are important to industrial problems. That's a strength we have over some of the industrial-research facilities, which tend to have a more narrow focus.

Larry Deaven: The Lab's diversity is surely paying off in the genome project. We have a robotics effort, we have a strong informatics effort, and we have some innovative work under way from the Chemical and Laser Sciences Division and the Physics Division in the direction of new methods for DNA sequencing. You find elements of those activities at any of the other genome centers, but ours is unique in terms of having all of those elements under one roof. And our center for human genome studies could be used as a model in developing other new initiatives in bioscience and biotechnology.

Alan Bishop

Our biggest challenge—perhaps even nationally—is just how to interface basic disciplines with application teams without using up the resource that generates advances in basic science and technology. Many scientists want to belong to a discipline, …and feel a part of its continuous growth and evolution.

Alan Bishop: Materials science is another area that is intensely interdisciplinary. In the manufacturing process, for example, if you don't have a controlled starting point and the right material that can get you to the desired end point, then there isn't a process.

So creating complex materials is an area in which Los Alamos can compete and is competing very well due to our ability to bring together different disciplines. For example, I'm a physicist, and I'm working very closely and effectively with Basil Swanson, a chemist, and his colleagues on new materials with intriguing electronic and light-emitting properties.

However, if you simply take people out of their disciplines and put them into a team to solve some specific problem, you create a very short-term fix, but you can easily jeopardize the technology base of that discipline, which is in fact the source of our ability to tackle complex problems—the base in physics, math, chemistry, and so on, that Colgate and others have emphasized.

Our biggest challenge—perhaps even nationally—is just how to interface basic disciplines with application teams without using up the resource that generates advances in basic science and technology. Many scientists want to belong to a discipline, and they want to be able to regularly come home to that discipline and feel a part of its continuous growth and evolution. They can be effective members of interdisciplinary teams, but they should be able to take the questions from those teams back to their own disciplines where their greatest strengths lie and where the questions can best be interpreted and solved. Trying to convert people from one discipline to another is very difficult, and the success rate is pretty low.

The area of materials science, in terms of both structural and electronic materials, is one where the Lab can have a major impact. And the impact will be felt on both the defense and the nondefense sides of the laboratory. Dual-use actually means something very directly in the area of materials science.

Paul White: Do we have some examples where work in that area has been reinforced by collaboration with industry?

Alan Bishop: I've been working with various companies for the last eighteen months or so, and it has been educational on both sides. At the beginning, high-level discussions are usually widget-driven. The company wants some particular device, and it wants our help in making it. But after a while, the stronger industrial firms—my own limited experience includes companies like Hewlett Packard and IBM—begin to recognize that what they really want is our depth of technology, our expertise in modeling and developing and testing new materials, and our ability to put appropriate science into large computer codes and then simulate effectively. That's the most positive kind of new direction. It's taken over a year to reach that point, and I hope it takes less time in the future, but we got there, and that is encouraging.

Paul White: Some people in the Clinton campaign proposed that the weapons labs devote 20 percent of their efforts to collaborations with industry. Where are we relative to that number?

Kay Adams: I'm the director of the Industrial Partnership Center at the Laboratory, and one of the biggest disappointments I've had is that the DOE expects us to have an economic impact on a company in one to two years, but a more appropriate and realistic time scale would be three to five years. There are companies that view Los Alamos like grandmother's attic where they can find all kinds of neat things to bid on. Well, we don't have much off-the-shelf technology, and I wouldn't want to see us emphasizing the development of off-the-shelf technology because then we'd be destroying our seed corn.

We need to keep replenishing our good ideas, to build synergism between new ideas and new applications, and to leverage the kinds of long-range projects we've been talking about today. We need to build up areas that are four to ten years away from having products. And no one in Washington at this point is ready to accept that as being a fundamental need.

Also, because of the fifty-fifty cost sharing of some programs, we can work only with big companies. But larger firms tend to be less innovative and less interested in taking risks to implement new technology than are some of the smaller companies. So the big firms come to the labs to look for a Band-Aid. Even though the help we offer might get us some brownie points, we lose intellectual property. We also get tied up in knots when we try to work with two or three different companies on the same problem. It would be better to have our technology open to the public so that more people could just run with it. That would probably do the economy more good.

We have not yet determined how the national labs can truly impact the area of economic competitiveness. And yet the DOE expects us to prove we can be outstandingly helpful to the economy in two to three years. We need more time to build on our strengths and show what we can really do.

Jim Smith: I'd like to expand on some of Kay's comments. The Laboratory is working with American Superconductor Corporation and Intermagnetics General, two small businesses, to manufacture wires from high-temperature superconducting materials under a Cooperative Research and Development Agreement, but we also do our own independent research. Right now it looks like we've gotten ahead of the companies, and that poses some interesting questions about intellectual-property rights, and the issues are very cloudy.

We've gotten two kinds of guidance from the DOE. Some people want us to work with industry, and they say that the collaboration is the final product. The people who have been around longer want to see us actually make the product because then they can judge whether we've performed well or poorly. Just what is our mission?

Judging our success in tech transfer is rather difficult because for now Washington can either ask a company if we've helped them, or they can ask us about the properties of the wire we've made. When the companies begin selling products, Washington may be in a better position to see whether we have helped industry.

Harold Agnew: Your funding agency can argue that the reason you got ahead is because you've used the goods from each of the three companies. That presents a real problem.

Kay Adams

We have not yet determined how the national labs can truly impact the area of economic competitiveness. And yet the DOE expects us to prove we can be outstandingly helpful to the economy in two to three years. We need more time to build on our strengths and show what we can really do.

Gene McCall: It is not to the Lab's advantage to deal with a hundred different companies even though it may be to the advantage of the individual Lab researcher. The Lab needs to work on the basis of some overall concept, and if working with a particular company fits into that, fine, do it.

Kay Adams: The basic premise of the Technology Transfer Initiative was that we would be working on dual-use technologies, which would build on our strengths in the defense arena and also have commercial impact. That's a very solid, fundamentally defendable premise. The implementation, however, has presented some significant problems. But we do have some positive models. The Defense Advanced Research Projects Agency and many other organizations in the government have put together groups that have worked with the national labs and industries for years, and they've been highly successful.

Stirling Colgate: DARPA funds research, but, in contrast, we *do* it. That's why, by comparison, working with the Lab presents a conflict of interest to industry.

Basil Swanson: Not just the Lab but the nation as a whole must identify long-term industrial needs because industry has gutted its labs and no longer has the capability to address long-term problems. Industry needs access to new technology, and I think this Lab can bring it to them.

Kay Adams: The reason for doing shorter-term projects is to build credibility. Right now some industries do not look on the national labs as being their friends or even as being a resource for them. They see the labs either as competitors or as being totally nonessential to anything that they are doing. So we have to show that we are willing to listen and that we can actually deliver more than was originally expected.

The work on high-temperature superconductors with Hewlett Packard and DuPont, the work on high-performance supercomputing with Cray Corporation, and the work on advanced numerical methods for oil recovery with Mobil Corporation have all been successful. We're building these alliances, and when people start to work with us on very short-term kinds of things, they start to see the potential that's here and become willing to take a little more risk and work on longer-term projects. Industry will come along. It's just a matter of getting them in the door and working with them. I'm more concerned with the Department of Energy and the national funding capability. We need a realistic way of defining success in the area of industrial collaboration.

When people start to work with us on very short-term kinds of things, they start to see the potential that's here and become willing to take a little more risk and work on longer-term projects.

Harold Agnew: It seems to me that the smaller companies really need you more than the larger firms. The labs are not competitors with the little guy because he hasn't gotten to the stage where he's really competing with anyone yet; he just has a dream.

Kay Adams: That's absolutely correct. Right now there are set-asides for small business, but they tend to be extraordinarily small. The entire program for the Technology Transfer Initiative is going to be $141 million and only $6 million has been put aside for small businesses. None of this money has yet been released.

Paul White: Perhaps what small business need now is the technological equivalent of the Agriculture Extension Service. That service was established in 1914 to work at the local or regional level to funnel the results of research done at land-grant colleges into the hands of farmers. In the technology arena, a similar program might channel the results of research at universities and government laboratories into the hands of small business.

Gene McCall: When I look at a possible future for the Laboratory, I see half of the Lab being in defense work with half of that being in nuclear weapons. The nuclear-weapons part interfaces in my mind with the economic competitiveness through the design of a safe reactor for the burning of plutonium from weapons. The non-nuclear part of weapons work interfaces with the modeling and improvement of the transportation system. We ought to work on the design and building of autonomous vehicles and the modeling of the battlefield and of defense systems.

On the more commercial side, we should work on solid-state lasers, which have applications in defense, biology, surgery, compact-disc players, and so on. The biotechnologies are a little bit separate but still seem to form a major thrust in the laboratory. To my mind, those areas, along with materials work, flesh out an appropriate future for the Lab. So far, we have not developed a collective vision for the Laboratory, like the one I just outlined, but once we establish firm priorities and say, "Here's what we are going to do," we can start interacting with those companies that match the Lab's vision of itself. If we work on piecemeal applications with industry, I don't think we will get consistently good performance from the Laboratory staff.

David Sharp: We do need a collective vision, but we don't need to emphasize large projects to the exclusion of small ones. In this context large means $100 or $200 million, and small means $10 million or less. The idea that the Laboratory's strength is geared toward large projects reminds me a little bit of a dinosaur's strength being geared toward eating large quantities of food. Focusing on finding the one large project is a silver-bullet approach to planning the Laboratory's future.

We have a great big problem, which is money, but that doesn't mean there's one great big solution. The solution is going to come in small pieces from many vigorous efforts in economic competitiveness, in biotechnology, and in the environment. Efforts that start small can grow and become extremely important to the Lab. GenBank, the national database for DNA-sequence data, grew to be a fairly substantial project, but it didn't start that way—it started with just a couple of people with vision. Small thrusts, which are only a few million dollars, can coalesce and reinforce each other to become a major thrust. But the attempt to start $100-million programs from scratch will be extremely difficult. Anything that big is going to be a line item in Congress, and every one of the fifty states is going to claim that it should have its piece.

Gene McCall: I don't agree. The burning of weapon plutonium is something we can do. We can come up with a standard, very safe reactor design, and that would be an appropriate large project for the Lab. Roughly 25 percent of the carbon dioxide in the world is the result of power production. As we go to more electric cars and electric vehicles to get away from the pollution and environmental problems caused by burning fossil fuels, the electricity has got to come from somewhere. Now if it comes from a nonpolluting source like nuclear reactors, we're all better off. But we need to show people that this is a nonpolluting, safe source of energy, and that it will not have a negative affect on the environment of the future. Such a project would be a $100-million project.

Jill Trewhella: The key to a healthy organization is a balanced approach. We need to pursue areas of strength and look for opportunities for large projects. Having a thousand little one-person efforts is ridiculous. At the same time, in the context of building on our strengths and looking for large programs, we need to encourage and provide an environment where small projects can flourish. We can do both. Moreover, we can't just respond to the national needs for the next decade and ignore what the national needs are going to be in twenty or forty year's time; we have to do both.

Robert Ecke: One big problem is the lack of a national science policy, so, as Gene was saying earlier, we're always in the position of reacting. Solar energy used to be big, and then a different administration came in and energy was not a big thing, so now we go off and do something else that's defined as a national need. It really is important for the nation to decide what it wants, and then we can respond in a reasonable way.

Paul White: But the nation is made up of people with many different interests, all of whom want different things, and all of whom read the newspaper and listen to the TV and the radio.

Jill Trewhella: And that's why we can't afford to be purely reactive. We have to, at some level, decide what we're good at and set our own directions. We have to be responsive but not reactive.

Robert Ecke: But if we decide we want nuclear reactors and the public doesn't want nuclear reactors…

Stirling Colgate: …then we have to keep leading and that's where advertising comes in. What percentage of industry's budget is spent on advertising? 10 percent? We need to explain the technical issues. We have to be a leader in the technical education of the general public and our political leaders. Leadership in technology starts with managing our nuclear arsenal, and it goes on to any other area that we might be working in. We should be setting the new directions for ourselves, we shouldn't be sitting around asking for them.

Harold Agnew: In the old days we talked to the Joint Committee on Atomic Energy. They had grown up in the nuclear-weapons business, so they understood it very well. The same was true of the Atomic Energy commissioners, who were really a cut above most people in the bureaucracy today. Some of them even had a strong technical background. It was easy to talk with them. They set the policies, but they relied upon the Lab directors for advice. If we had a good idea, it was easy to implement in policy. Nowadays everybody in the Congress wants to have a finger in everything, and if anybody tries to get something accomplished, they get investigated, or as we say, Dingelled, so people hide.

The American people are sending a message to the scientific community that they want help with major problems: health, the economy, the environment. I don't think this means, as some argue, that the scientists are out of the decision-making loop; rather, they have a mandate to consider the broader impact of their work.

David Sharp: We have to realize we are working in a different political environment. The American people are sending a message to the scientific community that they want help with major problems: health, the economy, the environment. This message has been picked up and has bedrock support across the political spectrum.

Thus, at the national level the relevance of proposed projects to specific national needs receives greater emphasis in the overall evaluation process. I don't think this means, as some argue, that the scientists are out of the decision-making loop; rather, they have a mandate to consider the broader impact of their work.

I review programs on applied math. Three, four, five years ago, a proposal by an applied mathematician would state the problem and the beautiful theorems he was going to prove and that would be sufficient for getting funding. Today those proposals contain a description of a whole project all the way from the mathematics problems down to the applied problems where that mathematics is going to be plugged in. The whole proposal is connected to a consortium with industry or some National Science Foundation interdisciplinary center. A lot of people are learning how to play this game of building integrated programs. I don't think they can do it quite as well as we do at Los Alamos, but they're learning because that is the way the funding of programs is being conducted.

Paul White: The mechanism for identifying national needs is different from what it was fifteen or twenty years ago, but that doesn't mean we can't influence what happens. It doesn't mean we can't lead. We have be active, and we have to lead in different ways. ■

The Participants

Kay V. Adams is director of the Industrial Partnership Center at the Laboratory. As manager of all of the Laboratory's industrial activities, she assists in the identification of key R&D technology areas that are of interest to industry and represents those activities to industries and organizations outside the Laboratory.

Harold M. Agnew was a member of the small group that worked with Enrico Fermi to inititiate the first nuclear-fission chain reaction at the University of Chicago in 1942. Shortly thereafter he joined Project Y at Los Alamos and worked on the development of the atomic bomb. He became leader of the Weapons Physics Division in 1964 and director of the Laboratory in 1970. Agnew is a recepient of the Fermi Award, the highest scientific award of the Department of Energy, in recognition of his "many contributions to nuclear physics and nuclear weaponry, and his forthright counsel to the government in the field of national security."

Chris Barnes joined the Laboratory in 1979 to work on inertial confinement fusion theory in the Applied Theoretical Physics Division. Barnes has worked in neural network and machine-learning applications for the past five years.

Alan R. Bishop has been involved in resesearch on condensed matter and nonlinear science since joining the Laboratory in 1979. He is currently leader of the Condensed Matter and Statistical Physics Group in the Theoretical Division.

Gregory H. Canavan came to Los Alamos in 1981 to build advanced lasers and to perform plasma-physics experiments in inertial fusion. For the last decade he has been active in developing and testing advanced concepts for strategic defense and future strategic forces.

Stirling A. Colgate was a staff physicist at Lawrence Livermore Laboratory for twelve years and president of New Mexico Institute of Mining and Technology before joining Los Alamos in 1976. In 1980 he became leader of the Theoretical Astrophysics Group. He is a Senior Fellow at the Laboratory and a member of the National Academy of Sciences.

Larry L. Deaven has been involved in biology and DNA research since he joined the Laboratory in 1971. He is currently the principle investigator of the National Laboratory Gene Library Project at Los Alamos and Deputy Director of the Los Alamos Center for Human Genome Studies.

Gary D. Doolen has been Acting Director of the Center of Nonlinear Studies since 1990. His research interests span nuclear physics, atomic physics, plasma physics, controlled thermonuclear reactors, lattice gas methods and nonlinear mathematics. Doolen is a past senior scientific editor for *Defense Research Review.*

Robert E. Ecke came to the Laboratory in 1983 and has worked in low-temperature physics and nonlinear science. He is currently in the Condensed Matter and Thermal Physics Group and for the past two years has been Acting Deputy Director of the Center for Nonlinear Studies.

David W. Forslund is a Laboratory Fellow, Deputy Director of the Advanced Computing Laboratory, and a Fellow of the American Physical Society. His areas of expertise span physics and computer science and include space plasma physics, magnetic fusion, laser fusion, massively parallel computing, and distributed computing.

Erica Jen is a mathematician in the Theoretical Division and Acting Deputy Director of the Center for Nonlinear Studies. Her research focuses on the mathematical analysis of discrete dynamical systems. Other current projects include electronic-preprint bulletin boards and educational programs in nonlinear science and complex systems.

Raymond J. Juzaitis has been with Los Alamos National Laboratory since 1979 and now serves as the Division Leader of the Applied Theoretical Physics Division. A member of the DOE Verification Panel and the American Nuclear Society, Juzaitis has received two Department of Energy Awards of Excellence for Nuclear Weapons Design.

Chick F. Keller has been involved in computer modeling of fluid dynamic processes at the Laboratory since 1967. He is currently Director of the Laboratory's Institute for Geophysics and Planetary Physics and a Principal Investigator on a joint project between the University of California and Lawrence Livermore and Los Alamos national laboratories on global climate modeling.

Carson Mark came to Los Alamos from Canada in 1945 as part of the British Mission collaborating on the Manhattan Project. He joined the Laboratory in 1946 and served as leader of the Theoretical Division from 1947 until his retirement in 1973. He currently serves as a Laboratory consultant.

Gene H. McCall has been involved in plasma, laser, and hydrodynamics research at Los Alamos since 1969. He was a founder of the Inertial Fusion Program at Los Alamos and leader of the Laser Division; he is now a Laboratory Fellow in the Applied Theoretical Physics Division. He has been an advisor to the Department of Energy and to the Department of Defense.

David H. Sharp joined the Laboratory in 1974 and now holds the position of Laboratory Fellow. His current research interests include the modeling of complex fluid flows and the formulation and analysis of gene regulation. Sharp is a Fellow of the American Association for the Advancement of Science and the American Physical Society.

James L. Smith joined the Laboratory in 1973 to study superconductivity and magnetism in the actinide elements. After the advent of high-temperature superconductivity in 1988, he joined the Superconductivity Technology Center as Chief Scientist. He is a Laboratory Fellow and received an E. O. Lawrence Award from the DOE in 1986.

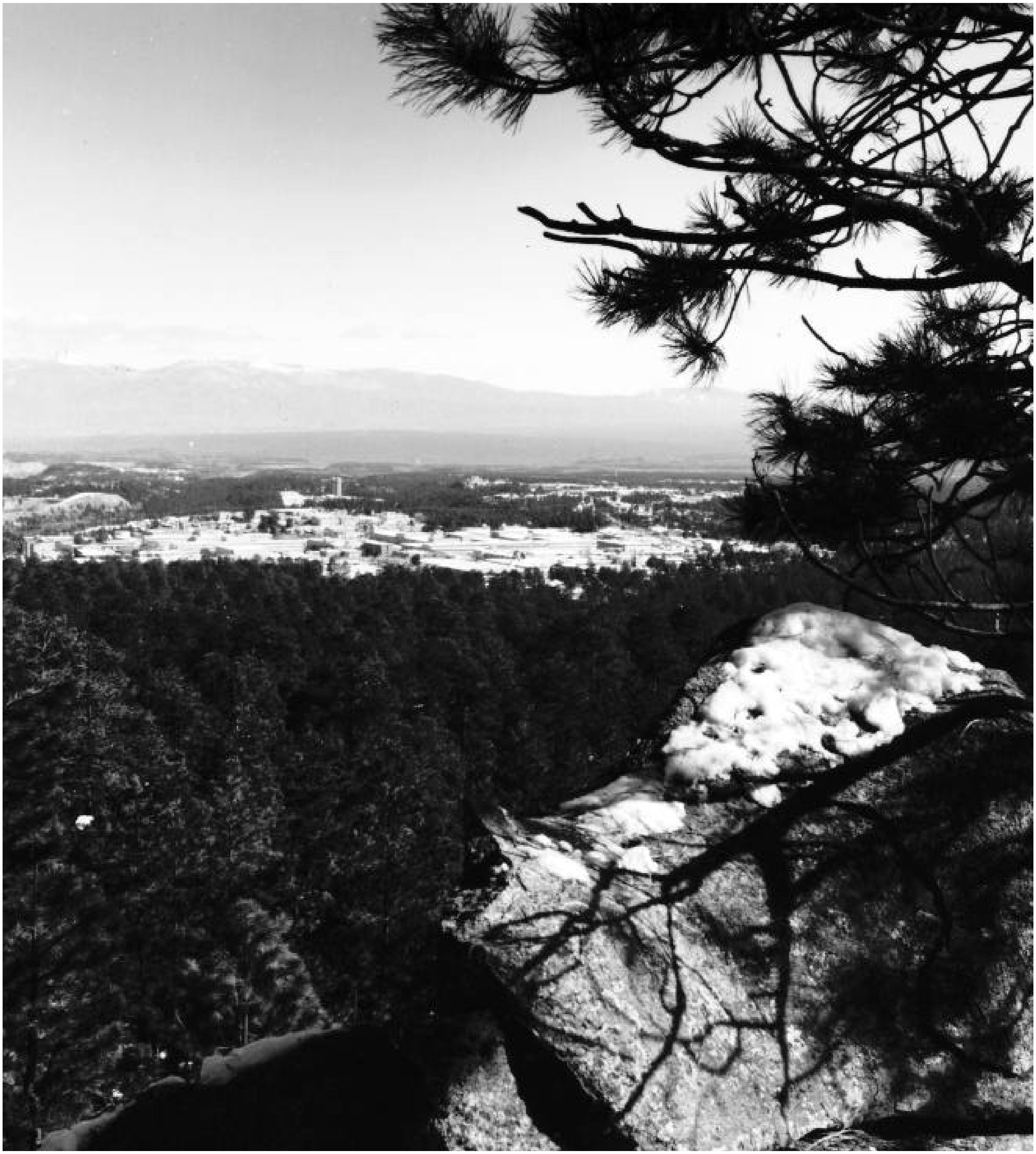
Basil I. Swanson has been involved in condensed-matter spectroscopy and advanced electronic and electro-optic materials since joining the Laboratory in 1980. He is currently a Laboratory Fellow and the Principal Investigator for Laboratory programs in low-dimensional mixed-valence solids and in spontaneous self-assembly approaches to advanced electro-optic and nanostructural materials.

Jill Trewhella joined the Laboratory in 1984 to further develop her interests in using physical techniques to study how biomolecular structure regulates or controls biological activity. She is currently Principal Investigator for NIH and DOE projects in structural biology and is leader of the Spectroscopy and Biochemistry Group in the Isotope and Nuclear Chemistry Division.

Andrew B. White, Jr., is Director of the Advanced Computing Laboratory and Program Director for High Performance Computing. He has been at the Laboratory since 1979. His leadership of the Laboratory's role in the Federal High Performance and Computing and Communications Program culminated in Los Alamos being designated as one of the two High Performance Computing Research Centers.

Paul C. White is Program Manager for Special Projects in the Nuclear Weapons Technology Office. He acts as a liaison between the Nuclear Weapons Program and the Nonproliferation and Arms Control Program, and serves as the Laboratory's technical representative to the DOE Steering Group, which coordinates U.S. assistance to Russia in the dismantlement of nuclear weapons. White was formerly leader of the Applied Theoretical Physics Division and Deputy Director of the Center for National Security Studies.

Merri M. Wood has worked in the Thermonuclear Applications Group since joining the Laboratory in 1979. Her work has included stockpile support, weapons physics, and advanced development. She has also been active in the Equal Employment Opportunity and Affirmative Action arenas.



What Is the Future of Los Alamos?

Hans A. Bethe

The Cold War is over. We no longer need to fear a sudden attack by Russian nuclear weapons. Many of those weapons are being dismantled, and we have offered to buy the fissile material contained in them.

Still, nuclear weapons exist in Russia and other countries. There is danger of proliferation. Therefore, we must for the foreseeable future maintain a stockpile of nuclear weapons and make sure that the weapons are in good condition. Los Alamos is very competent to carry out those tasks.

But no country is likely to challenge the United States' nuclear-weapons sophistication—now or in the future. Therefore, I see no need for further weapons development.

Of course, we should have tests of further improvements for the safety of nuclear weapons. But I think such tests can be accomplished by 1996, in compliance with the Hatfield amendment. Thus we have a great opportunity to participate in a comprehensive test ban, and steps in this direction should be taken by 1995 when the Nonproliferation Treaty comes up for review. The United States will be safer in a world with a comprehensive test ban and nonproliferation than in one with further nuclear weapons development.

Such a program leaves the majority of Los Alamos scientists free to pursue civilian research and development. There are many non-military challenges. During the Cold War this country as a whole and the Laboratory in particular have not thought much about civilian technology. That lack of attention is one important reason for the preeminence of Japan and Germany in modern industry.

The fellows of the Laboratory, with the help of many other staff members, have developed the Advanced Projects Initiative. They have selected about ten long-range projects on various technologies that are likely to become important in one or two decades and are especially suited to the special skills and experience of Los Alamos scientists.

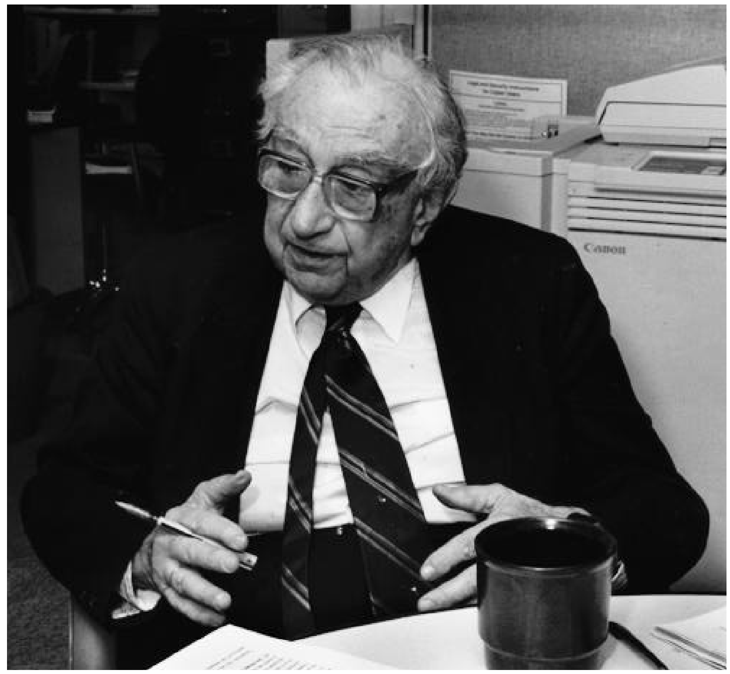
Rightly, the fellows have selected long-range projects. Industry must look to its bottom line and finds it difficult to engage in long-range projects, especially if success is not guaranteed. Universities cannot engage in very big projects, and here again is a chance for Los Alamos. Of course, these arguments should not prevent the Laboratory from working on special tasks for present needs, supported by appropriate contracts with specific industries, but the main concentration should be on long-range projects. If a program like that of the Laboratory fellows proceeds, it would be good to establish some kind of Industry Council, which might deliberate about long-range needs. Universities should also be appealed to for advice. As the work proceeds, new ideas and projects will emerge.

Los Alamos National Laboratory should go at these projects with the same enthusiasm as in World War II and as in the development of H-bombs in 1951–54. The ingenuity already exists.

There are signs that the government understands the need for civilian technology development at the National Laboratories. I hope this will translate into secure money support.

If all this comes to pass, Los Alamos National Laboratory will have a future as brilliant as its past fifty years. ■

The Laboratory of the Atomic Age



Edward Teller

All of us hope that, in the long run, the energy of the atomic n nucleus will be used only for peaceful purposes. After a half century of nuclear explosives and nuclear reactors, the practical effect of the former has been more important, both in our thinking and in our expenditures. The existence of nuclear weapons has had a decided influence on human affairs. Los Alamos, since its beginnings in March 1943, has been a unique place, distinguished both by its nature as a community and its continuing influence on world history. I will speak to several aspects of both characteristics.

In 1938, Enrico Fermi, my wife, Mici, and I were planning to drive from Stanford, California to the East Coast. I clearly remember Oppenheimer's suggestion that we stop on the way and spend some time in New Mexico. Even more particularly, I remember Fermi's remark that New Mexico would be an ideal place to develop weapons. That remark struck me as rather peculiar at the time because it wasn't until a year later that Fermi made the final decision to emigrate from Fascist Italy to the United States.

When we arrived in Los Alamos in late March of 1943, we found a striking contrast between the beautiful mountain surroundings and the dreary, green, barracks-like buildings in which we were going to live and work. Today, nothing is left of those early structures. After the Second World War, people were permitted to own the houses in which they lived, and the once uniform dwellings were replaced by a beautiful variety.

Los Alamos is different from all other communities not only in appearance but also in spirit. It was founded by scientists, and throughout its existence, has been led by scientists. The result has been an ongoing liberalism in the old sense of the word, that is, an unquestioned—perhaps even unnoticed—tolerance of widely different ideas. Another remarkable fact about Los Alamos is the uniformity of the standard of living of its residents. There are no very rich and no very poor. Perhaps this characterization of Los Alamos sounds too good to be true, but I think such praise is not unfounded. In large part, this liberal environment is attributable to the founder of Los Alamos, Robert Oppenheimer.

Los Alamos National Laboratory, together with its sister laboratory in Livermore, California, is currently near the center of a great controversy. On one side, it can be claimed that the efforts undertaken by the weapons laboratories won the Cold War and are ultimately responsible for the collapse of the Soviet Union. If even a small portion of this claim is justified, the implied effect on world history is most important—particularly because expenditures on nuclear weapons amounted to roughly only 3 percent of the United States military budget over the past half century and, even more so, because the Cold War was won without any significant loss of life. On the other side of this argument, it is claimed that we would be better off had we never created atomic bombs, and, now that the Cold War is over, we may forget about them forever.

I intend to address this controversy at some length as well as its relevance to the future of the weapons laboratories. But to discuss the future, we must first consider the past.

The most important accomplishment of Los Alamos was the construction of atomic bombs in the two years and four months from the founding of the Laboratory to the test in Alamagordo. Before this history-making period I had been working in the Manhattan Project laboratory in Chicago, which was code-named the Metallurgical Laboratory, and while there I saw a lot of my good friend Eugene Wigner. I had come to learn that he was almost always right, and he strongly advised me not to go to the new Los Alamos laboratory. The only difficulty, according to Wigner, was the production of the needed nuclear explosive material, that is, plutonium. Once we had enough of that, he asserted, it would be easy and obvious to put together an atomic bomb. For once Wigner was completely wrong.

Just a few weeks after we arrived in Los Alamos, Emilio Segré discovered spontaneous fission, a most important and unwanted source of neutrons. This discovery meant that as we tried to assemble the fissile material into a configuration that would result in a nuclear explosion, the neutrons from spontaneous fission would trigger a diverging chain of nuclear-fission reactions, and a premature explosion of far lower energy-yield, that is, a fizzle, would result. It was not many months afterwards that the complete concept of the solution appeared: an elaborate spherical assembly of the fissile material wrapped in "lenses" of chemical high explosive, the operation of which would result in substantial compression of the "incompressible" plutonium (or uranium).

The implementation of this concept required a great deal of refinement of both the techniques for handling chemical explosives and the calculations required for reliable estimates of the performance of the design. The result was the experimental production of unprecedented pressures, exceeding even those that we knew existed at the center of the earth. In a period of twenty-eight months, several new branches of experimental physics and numerous calculational techniques were opened up. All this was made possible by the skillful leadership of Oppenheimer.

That we "created" the atomic bomb is not an entirely correct statement. The atomic bomb had been long since predicted by Leo Szilard, and it would have been developed, in any case, in the next one to two decades. That the nuclear explosives were made available in time to write a favorable conclusion to the Second World War is the great accomplishment of Los Alamos.

In fact, the early availability of nuclear explosives and the subsequent possession of an overwhelming military force at a continuing cost of a small fraction of 1 percent of our gross national product (counting only the essential nuclear explosives on which this force was centered) have made it possible for the United States to retain great influence in world affairs during the following half century. It was, in my opinion, an unprecedented situation in history: that low-cost military power should become available and nonetheless not be used for conquest, or for the imposition of our wishes in general, but rather for the sole purpose of deterrence, stability, and peace on a global scale.

In the meantime, the development of nuclear weapons in the Soviet Union, Great Britain, France, China, India, Israel, and Pakistan is proof that the technology to make nuclear weapons is there to be used by anybody. The detailed facts about such developments have been kept secret and can therefore not be quoted. But in the special case of Iraq, a commission of the United Nations has investigated, in an open manner, the work of Saddam Hussein's regime. They found not only that Iraq was within a few years of having nuclear weapons but also that these developments had required the expenditure of billions of dollars and the work of more than twenty thousand remarkably well-trained Iraqis in addition to the importation of a great deal of equipment and supporting technology. The UN commission's findings have undermined two opposing myths: First, that a nuclear explosive could be secretly developed and completed in someone's garage, and second, that secrecy will stop the proliferation of nuclear weapons.

After the Second World War, Oppenheimer's slogan concerning Los Alamos was, "Let us give it back to the Indians." To his great credit Norris Bradbury, as the first postwar director of Los Alamos, prevented that from happening. As for the further development of nuclear explosives, Oppenheimer's attitude was summarized in the statement he made to me in the fall of 1945, "We did a wonderful job and it will be many years before anybody can improve on it."

After the war I left Los Alamos to go to the University of Chicago, but I came back for frequent visits. In the summer of 1946, I traveled to Albuquerque to participate in discussions with the military regarding further developments of nuclear weapons. The military stated their opinion in very clear terms: The weapon used in Nagasaki is exactly what is needed and no changes whatsoever are to be recommended. Fortunately, Bradbury and the other leaders at Los Alamos did not accept this opinion and instead worked on significantly reducing the weight of nuclear explosives without sacrificing any of their effectiveness. Without such a development our postwar weapons would have quickly become rather ineffective in comparison with the capabilities of other nations, particularly those of the Soviet Union, Britain, and France.

Thus, the question of whether the efforts of Los Alamos were still needed in the period after the Second World War has been clearly answered in a positive manner. I believe that the historical situation following the Second World War is, in some respects, comparable to the one we are facing at present. In the current post-Cold-War period, we cannot simply conclude that the weapons laboratories have become superfluous. Indeed, it is a fundamental characteristic of technology, particularly in modern times, that new possibilities continue to open up and, soon thereafter, are realized. While some people may believe we are fast approaching "the end of history," I still find myself in agreement with Plato: "Only the dead have seen the end of war."

Although Bradbury took a strong stand on advancing the development of fission weapons, he nevertheless considered the development of the hydrogen bomb to be either an impossibility or, at best, a challenge that would require many years of considerable effort. During virtually all of my second stay in Los Alamos (1949–1951), I worked diligently on planning the first hydrogen bomb and did not consider plans for a sister laboratory. But when the decision to continue a vigorous effort was reached only by a hairsbreadth following the successful test in the spring of 1951, I came to the conclusion that the creation of a second laboratory would serve the national interest by generating competition and maintaining mutual support. These comments notwithstanding, it should always be remembered that the first American hydrogen bomb was created and tested by Los Alamos and that those developments would not have been possible without the cooperation of many of the old-timers.

With the passage of many years, it has become quite clear that the hydrogen bomb played an important role in the national military posture of both the United States and the Soviet Union. I have had the privilege of meeting some of the Russian scientists who worked on nuclear weapons. Most important was Andrei Sakharov, who is credited with the development of the hydrogen bomb in the Soviet Union and who later became an exceptionally courageous advocate of civil liberties in that Communist regime. Sakharov confirmed that the development of the hydrogen bomb proceeded independently and almost simultaneously in the Soviet Union and in the United States. It has also become clear that in the age of fairly accurate, rocket-based delivery systems, hydrogen bombs with yields of many megatons are no longer the most effective weapons. The real significance of the development of the hydrogen bomb is not that it offers exceptionally high yields but rather that it affords many options in explosive power, size and shape, and effects.

I have used the conventional words "Cold War" in referring to the four decades from roughly 1950 to 1990. I believe, however, that the use of the word "war" in that context is an unjustified exaggeration and becomes more so when applied to the development of defenses against the most effective method of delivering nuclear weapons, namely by rockets.

In private conversations Russian scientists have recognized the United States' undoubted leadership in the area of strategic defense, the accuracy of which depends largely on the use of computers. The need for sophisticated computers may be the reason the Soviet government consistently opposed the development of strategic defense, even after President Reagan suggested a collaboration. A more open attitude became apparent, however, when the Soviet Union collapsed and President Yeltsin took over.

The most significant initiatives for developing defensive arms came from the two nuclear-weapons laboratories. Recently Los Alamos played an essential role in planning the adaptation of nuclear explosives for use in a defensive manner as well as adaptations of satellites in low Earth orbits for a variety of purposes, including not only the prompt reporting of the launching of missiles but also a warning of activities that indicate preparations for aggression. In addition, the adapted satellites can be used for observation of weather, prediction of natural catastrophes, and monitoring pollution on a global scale.

The Russians have shown particularly great interest in the last point. They have openly stated their lack of confidence in the evaluation of pollution and its effects on the former territories of the Soviet Union unless this evaluation is made or supported by authorities thoroughly different from those of the old regime, preferably having a major international component. Thus the competitive development of arms has led to methods of observation of the human environment that can be used in an internationally cooperative manner for important peaceful purposes.

Without the work of the two nuclear-weapons laboratories, the aims of the Soviet Union toward expansion and eventual domination would have proceeded in a more successful manner—these aims were basically a continuation of the policies of the tsars of Russia. Remarkably, the frustration of these aims, coupled with the corruption and ineptitude of the Soviet government, led to the abrupt end of the Soviet Union. It is also not a coincidence that this event followed just a few months after the United States demonstrated in the Gulf of Persia that high-accuracy systems can be used to defeat a big, well-trained army of a dictator at a minimal cost, over the span of a few days, with an incredibly small sacrifice in the lives of our soldiers and those of our allies. The weapons laboratories have thus contributed in more than one significant way to the dissolution of the great and terrible Soviet dictatorship—a victory achieved without war.

Even in the absence of terrible tensions between the Soviet Union and the rest of the world, nuclear weapons remain a reality. Complete elimination of these arms from the stockpiles of the United States and other important powers would merely encourage dictators of small countries to acquire these weapons, thus giving them opportunities for aggression beyond that possessed by their neighbors and also perhaps beyond that of all other nations.

We must also consider the proliferation of missiles of various ranges. Approximately twenty countries now possess this dangerous, rapid means of weapons delivery. These missiles can be used not only to carry nuclear weapons but also chemical or biological weapons and ordinary high explosives, as shown by Saddam Hussein.

The American Strategic Defense Initiative, identified in the minds of many people with nuclear-weapons-based confrontation, actually is planned to provide a defense against rockets carrying any of these means of swiftly executed aggression. Although it would not preclude all use of nuclear weapons, it would make swiftly performed aggression, based on rockets and perhaps based on aircraft as well, far more difficult.

What I have said already implies, to a considerable extent, my answer to the question: What should the weapons laboratories do in the future? For the sake of clarity and emphasis, I will give two direct answers to the question—one is a general answer, the other, a discussion of a specific possibility.

In the present period of "Cold Peace" (a "Hot Peace" would be an active cooperation among all nations for their mutual, general benefit), it is justified to cut back the expensive, routine peacetime activities of our armed forces as well as reduce the number of enlisted personnel, the quantities of military bases, and the amounts of stockpiled equipment and materials. It is, however, by no means justified to cut back research on future military capabilities. This should be clear to everyone from the indisputable fact that such research over the past half century made the real difference in winning the Cold War. It will play a similar role in maintaining peace during the next half century.

Furthermore, with little alteration, modern instruments of war can serve as tools for peaceful purposes. For example, expensive nuclear weapon-tipped missiles can be converted into delivery systems for sending observation satellites into orbit. Such satellites would give us warnings and detailed predictions of hurricanes, storms, and floods not only in the United States but throughout the world.

Assuming that such research continues as it should, it ought to be concentrated in the present weapons laboratories. Among alternate facilities neither those serving pure research, nor those serving industrial development, nor even the Defense Department's military laboratories will be appropriate for such research and development. The first two lack the necessary contacts with the military and also lack the tradition of R&D on military systems. The other military laboratories, on the other hand, exist primarily for the stepwise improvement of existing weapons based on the closely prescribed requirements of the Pentagon. Compared to Los Alamos and Livermore, they have played essentially no role in the radical improvements, such as nuclear explosives or the instruments of extreme accuracy that have been developed successfully for defense against missiles and to hit point-like military targets. In contrast, the nuclear-weapons laboratories have experience in all these fields. They are the appropriate places for the creation and development of the fundamental advances in military capability on which America's future safety will depend.

In the context of this general answer, I want to discuss the specific issue of whether the DOE weapons laboratories should devote attention to new types of nuclear weapons. We must first consider what kind of new nuclear weapons would be of interest. Through the present time, we have emphasized the development of nuclear weapons using the best of the possible nuclear weapons materials for a broad spectrum of military objectives necessary for the security of the nation and our allies. However, nuclear weapons design for the proliferant nation or terrorist might assume an entirely different character from that of the major nuclear powers, taking paths prescribed by the limited availability of weapons materials and more narrowly defined objectives. Our understanding of these different paths to nuclear weapons is by no means complete. Without a full knowledge about these alternative prospects, we might not be able to detect the development of a weapons stockpile because we might not recognize the technology or even understand the clues that our superior remote sensing systems might provide for us.

One of the more likely paths to nuclear weapons taken by a proliferant nation or terrorists might be through the use of spent reactor fuel. The present worldwide inventory of plutonium in spent fuel is about 1300 tons. If we accept the unclassified information in reports by several Los Alamos scientists, frightening nuclear weapons can be constructed from rather small quantities of plutonium from reactor spent fuel. If the amount required for a threatening nuclear weapon were 10 kilograms, the present worldwide inventory would allow the construction of 130,000 such weapons. Not only do we need to understand how weapons might be designed from plutonium or other materials derived from spent fuel, but we should also understand the various options for the disposition of this material that would prevent its use in weapons.

During the past few years the great changes in the world have produced a clear and urgent necessity to consider major changes in our defense strategy while evaluating possible technical accomplishments. Such a necessity replaces the relatively simple problem of what to do with the DOE weapons laboratories with the difficult task of planning in detail exactly what should be done in a new era. Fortunately, Los Alamos National Laboratory has the intellectual tradition and power to give the needed answers, if not in a perfect way, at least as well as can any other existing institution.

Finally, it should be remembered that, after the Second World War, Los Alamos initiated a number of projects that are not connected with nuclear explosives. Among these, I select one particular enterprise for comment both because of my own interest and because it may have an exceptionally great importance in the future.

In their first half century nuclear reactors have progressed greatly and now produce 17 percent of the world's electricity. In Japan and, even more so, France, the progress is far greater than average. Nuclear-reactor-derived energy is clearly needed if the energy demands of the world in the twenty-first century are to be fully satisfied in an environmentally acceptable fashion. Yet, in the United States, public opinion and thus political trends have brought this development to a full stop.

What is badly needed is not just a safe reactor but an obviously safe reactor, one whose safety is easily recognized, even by non-experts. There are many means to achieve such safety, and Los Alamos is working on one of them. The Los Alamos reactor falls just short of being critical, by producing in each neutronic generation only 90 or 95 percent of the neutrons needed to sustain the neutron population in a steady state. Therefore, the reactor obviously cannot work at all unless the missing 5 or 10 percent of neutrons are furnished by an accelerator. The electric power for operating the accelerator is in turn furnished by the reactor in such a way that a considerable amount of thermal power is left over for the generation of electricity. The means of ensuring that such a reactor will not become supercritical and "run away" are obvious and easily appreciated—the accelerator design could include an automatic shut-off mechanism.

Furthermore, in the Los Alamos reactor concept the low-pressure hot fuel would be cooled by pumping it to heat exchangers located outside the reactor core but within the same core containment. This technology parallels somewhat the molten-salt reactor developed at Oak Ridge. This approach simplifies the design of the reactor core and utilizes the heat more efficiently. The Oak Ridge reactor had the disadvantage that its breeding capacity was marginal, which in turn imposed constraints on its engineering. Such constraints are removed in the case of the proposed Los Alamos reactor by the presence of the neutron-generating accelerator.

A further important feature of the Los Alamos reactor concept is the continuous removal of the radioactive fission products. In this manner, the great amounts of radioactive materials are removed from the energy-rich, high-temperature portions of the reactor system. The consequences—and, more important, even the likelihood—of a nuclear accident are thereby minimized. Indeed, the main cause of legitimate concern about existing reactors is that means for eliminating an uncontrolled energy release by the reactor might fail, resulting in the scatter of a lot of radioactivity over considerable distances. The Los Alamos reactor concept either eliminates or greatly reduces the possibilites for such accidents.

Much remains to be done before such a reactor concept can become a reality. The need for the accelerator and for continuous fuel-reprocessing equipment could make the reactor economically unattractive. I hope that these objections can be overcome, in which case the reactor could break the deadlock that has prevented further construction of reactors in the United States during the past decade.

I cannot even attempt to discuss all the possibilities for future successes that may be achieved by Los Alamos National Laboratory. What assuredly is needed is thoughtful consideration of the support that may be available from Washington and an open-minded pursuit of the truly limitless possibilities presented to inquiring minds by scientific and technological revolutions.

The conditions clearly exist for surpassing in the next half century the history-making successes achieved in the Laboratory's infancy. Indeed, I have reasons both to criticize and also to agree with Oppenheimer's declaration (slightly paraphrased) of forty-seven years ago: "It will be a long time before anybody can do better." ■

Bombs Away?

W. Jack Howard

Last December 2 marked fifty years since the nuclear age began at the University of Chicago squash court where Enrico Fermi and his colleagues (including a youthful Harold Agnew) cautiously brought the first nuclear pile to criticality. That experiment and what followed at the national weapons laboratories played a major role in ending a war and preserving a shaky peace. Now freedom seems ascendant behind the former Iron Curtain, but even as the world celebrates the end of decades of cold war, those anxious years must be acknowledged as being vastly preferable to a third world war.

As we reflect on the ways in which technology has influenced the nation's yesterdays, our thoughts also turn to the future—to what is in store for our society and to the ways in which science might help shape our tomorrows. It is also a time to consider the destiny of the weapons laboratories and how they can help provide that science.

The laboratories' obligation to provide for the stewardship of nuclear weapons probably won't go away soon and entails more than a quality guarantee for the current stockpile. "Surety," one element of stewardship, requires a continuing sense of responsibility for safety, security, and use control as well as performance.

Since the future of nuclear weapons is a virtual unknown, the design technology must be kept modern, not just pickled-in-place with some magic preservative juice. The watch on proliferation itself will require an understanding of new weaponization options as the underlying science moves forward around the world.

Some of the weapons pledged for retirement by the former Soviet Union are no longer controlled by Moscow. Even the ones accounted for must be transported and dismantled without mishap. The United States has a vested interest in the entire dismantlement operation, and our congress has designated funds for this activity. Our concerns can best be addressed by making U.S. nuclear expertise available as well, and the Yeltsin government seems receptive to scientific help.

I've argued that nuclear weapons have solved problems in the past, but some people say that they will *be* the problem of the future. Either way, the weapons laboratories' abandonment of the technology now would be a repudiation of their legacy and would leave the nation at risk.

The national laboratories, then, must find ways not only to maintain a capability in nuclear weapons (for purposes as yet unspecified) but also to prepare for an important role in fulfilling other technical needs of the nation. It's not a matter of choice *between* swords and plowshares; it's keeping the proper blend, which will surely involve a jagged decline of the weapons portion.

The fifty years of Los Alamos, which this issue celebrates, have been a useful preparation for the challenge of purveying science and technology to a post-Cold-War world. I think the heritage of weapons work will prove to have advantageous carry-overs such as the following:

- We've been named "national laboratories." How reasonable it is, then, that the laboratories be engaged in the nation's work—civilian work in addition to, not instead of, the stewardship of U.S. nuclear weapons.

- The unusual breadth of disciplines at the weapons laboratories gives them a powerful capability to solve new and difficult problems. Their ownership by the government guards against self-serving solutions—they can seek an answer without concern about their sponsor's spin on the question.

- Weapons R&D taught all of us that (contrary to the timeworn aphorism) invention is the mother of necessity; something becomes desirable once it is shown to be attainable. We used to be handed "Military Characteristics" for new weapons that described things we already knew how to do. The "Civilian Characteristics" can follow the same route. (Isn't "tech transfer" usually done with just such an approach?)

- Others may judge the worth of a nascent technology by its potential dollar-return to the offeror. People in the national laboratories are more likely to feel fulfillment if science promises to provide for some national good or to overcome some threat to the well-being of society. Lab folks won't even be startled if they don't amass huge personal fortunes in the process of applying science.

- Although Los Alamos has not set the most vigorous example of marketing prowess, the national laboratories are not without such skills; and those skills are applicable in the civilian sector. We should understand by now, however, that not *every* effort, even those with overwhelming merit, will bring instant public recognition and approbation. Take WIPP—*PLEASE*.

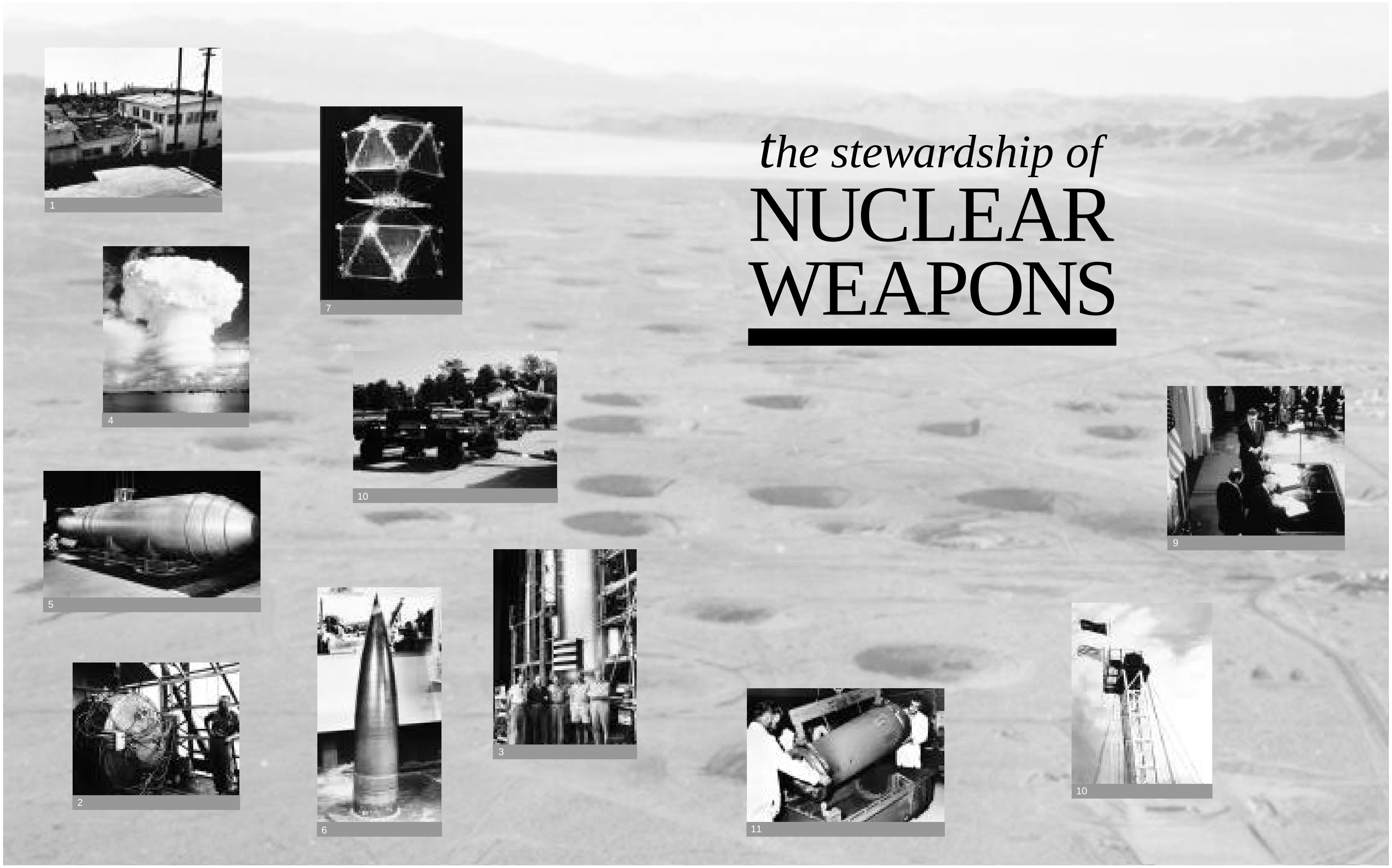
- The laboratories know how to build reliability into products whether they be destined for civilian or military use. They also understand "user-friendly," and if someone were ever to ask, they could even incorporate cost as a design criterion.

The world *might* come to "the Hill" seeking solutions, but it seems more likely to me that, with many inventions already in hand, Los Alamos will have to ferret out civilian or military customers who don't yet recognize their own "necessities."

As we reflect on the last half century, its achievements and even its disappointments, we can wish for an equally memorable next fifty years for Los Alamos National Laboratory. With some imagination, and if the needs and jargon of new customers can be accommodated, the future can be very bright indeed. ■

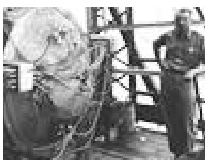
W. Jack Howard was Executive Vice President at Sandia National Laboratories when he retired in 1982 after thirty-five years of service. He served as Assistant to the Secretary of Defense for Atomic Energy from 1963 to 1966, and he received a Distinguished Public Service Medal from the Department of Defense in 1966. He was also a delegate to the Strategic Arms Limitation talks in Geneva during 1976.

the stewardship of NUCLEAR WEAPONS



The Stewardship of Nuclear Weapons

1 D building was the site of plutonium processing at Los Alamos during World War II. Plutonium was first manufactured in 1941 in cyclotrons. Its material properties were almost completely unknown at the start of the Manhattan Project. Los Alamos scientists were able to determine many of those properties in less than two years.



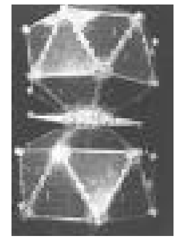
2 The “gadget,” the first nuclear test device, was assembled at the base of the tower at Trinity site on July 15, 1945. Norris Bradbury is standing at right. This was the prototype of the weapon dropped on Nagasaki on August 9, 1945.

3 The first thermonuclear device, Mike, is shown clad in cryogenic plumbing on Elugelab island in 1952 with Laboratory scientists and officials of American Car and Foundry, the company largely responsible for fabricating the device. The extensive cryogenic equipment required to liquefy the deuterium and tritium fuels in Mike gave rise to the modern cryogenics industry.

4 “Mike” was detonated on Halloween, 1952. Its yield was 10 megatons—an order of magnitude larger than the yield of the largest fission weapon of the time..

5 The first deliverable thermonuclear weapon, the Mark 17, was the largest weapon ever deployed by the U.S. military—it weighed 21 tons. It remained in the stockpile only two years before being replaced by smaller, lighter, and more easily handled

6 This 280-millimeter artillery shell enclosed a Mark 19 warhead, which was produced from 1955 to 1956. In reaction to the Korean War and the Soviet threat in Europe, the military saw a need for such tactical nuclear weapons.



7 The Vela satellites, deployed in support of the Limited Test Ban Treaty of 1963, carried sensors designed at Los Alamos to detect nuclear explosions in the atmosphere and in space. Between 1969 and 1972 Laboratory scientists found the first evidence of gamma-ray bursts in data from these satellites. The question of whether these extremely intense bursts originate inside or outside our galaxy is now hotly debated by astrophysicists in the Laboratory and around the world.

8 B-61 bombs are being loaded into a military aircraft. The B-61, which has been extensively modified since it was first introduced into the stockpile, is one of the most versatile bombs ever designed at Los Alamos.

9 President Ronald Reagan and Soviet Premier Mikhail Gorbachev signed the treaty eliminating intermediate-range nuclear weapons on December 8, 1987. This treaty was the first of a series that, together with the demise of the Soviet Union, ended the nuclear arms race.

10 The Joint Verification Experiment in August and September 1988 at the Nevada Test Site and the Soviet test site at Semipalatinsk demonstrated the effectiveness of CORRTEX (continuous reflectometry radius-time experiment), a system developed at Los Alamos to measure the hydrodynamic yield of an underground nuclear experiment. CORRTEX helped with the enforcement of the Threshold Test Ban Treaty of 1974.

11 A nuclear weapon is being unloaded for dismantlement. Los Alamos and the other weapons laboratories have responsibility for removing nuclear weapons from the stockpile and dismantling them.

Redefining the U.S. Nuclear Weapons Program and the DOE Nuclear Weapons Complex

John D. Immele and Philip D. Goldstone

Los Alamos: A Rich, Fifty-Year Heritage

As we look back on the fifty-year history of Los Alamos National Laboratory, we can be justifiably proud of the accomplishments that are the foundation of our rich heritage. While the nation faced World War II and then the Cold War, we developed nuclear and thermonuclear explosives. In the early years we were also instrumental in developing the manufacturing technologies employed at the production plants in the nuclear weapons complex.

Over the intervening decades we met the challenges of changing national security needs. As weapons-delivery systems changed and the need for lighter, smaller, and more specialized warheads became apparent, we developed weight- and size-optimized designs. As increased attention was directed toward warhead safety and security, we developed insensitive high explosives, fire-resistant weapon components, and other enhanced surety (safety/security) design features.

Once again the security needs of the nation are changing—this time in the most profound ways since the early days of the Cold War. Once again Los Alamos and the nuclear weapons program must respond by building on our rich heritage and unique scientific and engineering capabilities to meet the new challenges.

Changing Priorities for the U.S. Nuclear Weapons Program

The context of the U.S. nuclear weapons program has changed greatly, in a way that is profoundly affecting the goals of this program and the requirements placed on it. Foremost among these changes are the welcome collapse of the global military threat posed by the former Soviet Union, and the breakup of its old political structures. Simultaneous with these changes come major public concern about continuing problems within the U.S. economy and growing concern over the U.S. federal deficit. These economic issues have put increasing pressure on defense expenditures and have catalyzed a resurgent political emphasis on domestic policy.

While nuclear weapons will not disappear and deterrence remains an essential element of U.S. national security, the Soviet collapse has led to major reductions in the planned size of the U.S. nuclear force, the demise of "traditional" nuclear targeting strategies, and a shrinking and shifting rationale for deterrence. An immediate effect has been the massive pullback of forward-deployed tactical weapons, a relaxation of strategic alert, and the cancellation of all near-term weapon production. The reduced nuclear threat has also resulted in less public acceptance of the perceived risks of nuclear weapons ownership. Increased attention is therefore being given to nuclear weapons safety and to the environmental impacts of the nuclear weapons complex.

The United States must continue to have a credible nuclear capability for the foreseeable future, given the reality of continued nuclear weapons deployment by several other nations. On the other hand, the bilateral initiatives for a much-reduced stockpile (culminating in the recent START II agreement) have created an increasingly clear mandate for a very much smaller, as well as more environmentally sound, manufacturing and maintenance capability. These initiatives for stockpile reduction have resulted in a concomitant requirement for massive weapon dismantlement by the complex as well as significant reductions in the future need for new tritium production.

Though the end of the cold-war era is apparent, increased concerns about proliferation of nuclear and other weapons of mass destruction have been highlighted by revelations of Iraq's nuclear weapons development, and activities in North Korea are raising additional questions. These concerns have been made poignant by a resurgence of ethnic conflicts; meanwhile Russia, with its nuclear weapons and expertise, continues to skirt economic collapse. The specter of a wider number of states or groups possessing nuclear weapons is a frightening one, and the questions of how to prevent, detect, and mitigate such proliferation are vital.

Within the nuclear weapons program and nuclear weapons complex there have been distinct but synergistic roles played by the national laboratories and by the production complex. The weapons laboratories (Los Alamos, Lawrence Livermore, and Sandia National Laboratories), which carry out the research, development, and testing for U.S. nuclear weapons, are responsible for weapons design, engineering, certification, safety, and security. Their expertise supports nonproliferation activities, including analysis and assessment of emerging foreign nuclear technology, and supports our emergency-response capability. The technical expertise and judgement available in the nuclear weapons laboratories proved vitally significant in helping to uncover and evaluate the Iraqi program. The laboratories' unique nuclear expertise and technology will be used to guide the restructuring of the production complex.

The production complex (for example, Savannah River, Rocky Flats, Pantex) has been responsible for weapons component and material production, material processing, and weapons maintenance and dismantlement. Both the laboratories and the production complex have had roles in materials management and surveillance (monitoring) of the stockpile.

The changes in the budgetary, environmental, and national security arenas over the last five to ten years have produced significant changes in the weapons program and the complex. An increasing awareness of the impact of human activities on the environment has led to both regulatory and cultural change. Faced with significant environmental or safety concerns, some major, one-of-a-kind production facilities have had their operations curtailed for extended periods; some, such as Rocky Flats, will never resume their previous production role. Reduced production requirements and cost concerns have halted the development of a new production reactor to supply tritium. Furthermore, total defense funding is being significantly reduced, and there are increasing demands on the DOE's nuclear-defense-related funds—for reconfiguration of the nuclear weapons complex, environmental restoration, waste management, and increased attention to the environment, safety, and health in operations.

Motivated in part by the reduced threat, as a response to concerns for the future of the nonproliferation treaty, and in light of the Russian and French moratoria on nuclear testing, in late 1992 Congress passed the Hatfield amendment to limit and eventually end U.S. nuclear testing. The laboratories' level of Weapons Research, Development, and Test (RD&T) activity had already been reduced by about 38 percent between 1987 and 1992–93. The rate of nuclear testing was reduced by more than half in the same period—even before accounting for the Bush administration's restrictions of fall 1992 and the subsequent passage of the Hatfield amendment. Cost savings are still being actively sought in RD&T activities as well as by reconfiguring the production complex.

The nation is now approaching a critical juncture in which appropriate decisions must be made to transition the nuclear weapons complex and program to a new equilibrium that is designed to effectively support the new priorities and expectations, while assuring the quality the nation needs in its nuclear capability. The laboratories will, of necessity, play a key role in this transition and a vital one in the new state of the complex that emerges.

New Goals for the U.S.Nuclear Weapons Program

We must define a new national nuclear weapons capability, including a new state of the DOE nuclear complex, that is consistent with the following long-term goals and requirements.

With the drawdown and aging of the nuclear force, it will increasingly be the competence and capability of the RD&T laboratories, and the competence of the U.S. nuclear complex to produce, modify, and maintain weapons, that will represent deterrence. In other words, "deterrence by capability" will increase in strategic importance relative to deterrence by targeted nuclear forces. Thus the U.S. must maintain a technological nuclear edge—defined by system effectiveness, not large numbers—and we must provide for the stewardship of the ongoing stockpile. The principal elements of the DOE's nuclear stewardship will be stewardship of technology and stewardship of nuclear materials.

The ongoing nuclear force will be based on a much-reduced stockpile of no more than 2500 to 5000 weapons. The number of distinct weapons systems in the stockpile will be similarly reduced. The emphasis of the program will therefore be to assure adequate safety, security, reliability, and flexibility of these remaining forces. There will be few new starts of weapons development programs—none in the short term, though safety modifications to existing weapon systems will be pursued. In the long term, any new weapons development programs will be primarily driven by aging of stockpiled weapons to the point where reliability or safety is no longer acceptable or (potentially) by the desire to incorporate significant safety improvements.

There will be a permanent shift in emphasis within the complex from production of nuclear materials to management and control of nuclear materials and waste. There will also be limited needs for perishable materials such as tritium. Only a very limited fabrication capability will be needed, small compared to the previous capability.

A New Equilibrium for the Redefined Weapons Program and Complex

We suggest that following a transition period a new and different "equilibrium state" will emerge, both for the nation's nuclear-security posture and for its nuclear weapons program. Below we outline our vision of the principal elements of this new equilibrium.

Stewardship, safety and security, and prototyping. The research and development program must center around stewardship of the remaining stockpile, providing both expert judgement regarding the safety, security, reliability, and vulnerability of U.S. weapons as well as the ability to address problems that may arise within the stockpile, particularly as it ages. Although new production will be rare, the need for replacements in the stockpile will be inevitable. Under the anticipated test-ban regime, we must still be able to guarantee the safety and performance of these replacements. We must also maintain the ability to manufacture them in a new minimum production complex.

The value of improvements in safety and security was highlighted by two recent government studies initiated by Congress. However, we anticipate that those improvements will be introduced in a "graded" approach paced by need and affordability rather than by the availability of one or another new component. Certification of improved systems is in many cases tied to nuclear testing, particularly when these improvements involve changes in explosives or materials components. If the Congressional mandate to limit testing to the fiscal years 1994 through 1996 stands, we hope in that period to develop and test prototypes of potential back-ups to current systems, which are as robust as possible to uncertainty under a test ban and which have a full complement of modern safety features.

We propose that in the absence of new weapons production, the development of prototypes is an effective way to maintain active technological competence in weapon design, engineering, and production technology. To be effective, prototyping must include both component and integrated testing so that the actual performance of the prototype can be reflected back to its designers and engineers. Above-ground testing without a nuclear explosion will provide many of the needed benchmark experiments, though the full benefit of prototyping for sustaining design judgement and engineering competence would not be obtained without underground tests. Secretary of Defense Les Aspin as well as the Armed Services Committees have suggested that a similar process be used to retain capability and technological expertise in conventional weapons systems.

Active technological expertise is the foundation of stewardship. In addition to providing expert judgement regarding the safety, security, reliability, and vulnerability of U.S. nuclear weapons, that expertise will be called upon to provide assessments of the nuclear forces of other nations. As the national security context continues to shift, the weapons complex will be also be called upon to evaluate some limited technological options—not for tomorrow but for ten or twenty years from now when delivery systems begin to face technological obsolescence.

Nonproliferation/Counterproliferation. During the 1990s security concerns will shift from bilateral arms-control treaties to multilateral control of proliferation. Arms-control, verification, and intelligence efforts will increasingly overlap. The U.S. and former Soviet Union will be securing and dismantling much of the extant stockpile of nuclear weapons. Large amounts of nuclear material will be removed from weapon systems and must be safeguarded in a way that provides international confidence. The world will face an increasing threat of nuclear weapon use from new sources, and the diffusion of missiles and chemical-biological weapons will compound international stability concerns. The Nonproliferation Treaty and International Atomic Energy Agency (IAEA) will be central elements in the diplomatic and political efforts to avert and mitigate proliferation.

The redefined nuclear weapons program will increasingly be called upon to contribute to nonproliferation and arms-control efforts. We will continue to provide qualified teams for nuclear-emergency or accident-response contingencies. We must apply all necessary technical skills to help in the prevention, detection, and mitigation of weapons proliferation. Considerable nuclear-weapons expertise will be needed to safeguard the storage and handling of nuclear materials, to assess foreign technology, to verify treaties, and to provide advanced computing methods and advice related to export controls. The weapons laboratories' capability to field complex physical measurements in difficult environments will be used to detect key indicators of the intent to proliferate. One such technology is LIDAR (Light Detection and Ranging), which has been applied by Los Alamos in a wide range of environmental and atmospheric sampling programs. In addition, the threat of increased terrorism calls for the development of new technologies to aid the intelligence community.

Predictive capabilities, above-ground experimentation, and the issue of nuclear testing. Underpinning stewardship and the ability to support continuing national security needs is the maintenance of nuclear-design competence. To continue to provide this competence now that plans are under way to end tests involving nuclear explosions, the nuclear-weapons-design activities at the laboratories are emphasizing greater predictive calculational capabilities and "above-ground experiments" (AGEX), that is, physics and materials experiments that do not involve a nuclear explosion. The laboratories will use these above-ground experiments in relevant physical regimes to exercise nuclear weapons design expertise and weapons technology. An appropriate suite of complementary experimental capabilities will be needed, since all of the physics aspects of a nuclear explosion cannot be produced simultaneously without an underground test.

We have recommended that a minimum program of underground testing be retained to provide confidence in weapon reliability and quality, to maintain expertise, and to investigate technical options or design modifications for prototype weapons. However, because of the end of the Cold War and the hope of further discouraging proliferation, the nation plans to end nuclear testing as a matter of policy, and we do not anticipate a change in that policy unless there is substantial provocation. When testing is in fact ended, the laboratories will rely on the above-ground capabilities as their principal experimental resource to address technical issues, maintain expertise, and validate theoretical models and calculations used to predict weapon behavior. This will ameliorate but not prevent a decline in weapon expertise and judgement. The best available high-performance computational capabilities, including massively parallel computer architectures, will be utilized to enable more accurate and complete simulations and design codes. These codes will be extensively tested against available nuclear-test data and the results of above-ground experiments to provide the best possible predictive capability for weapon reliability and safety.

Consolidation of the RD&T complex. With the recognition that maintaining active competence in people and confidence in equipment is essential, the infrastructure of the complex will be consolidated and reconfigured to reduce cost and reflect new goals. It is important that this consolidation continue to integrate design, materials, and experimental capabilities at a common location to preserve program quality, and it should try to preserve the current architecture of two weapons-design laboratories for interlab peer review. However, to retain this architecture while reducing RD&T funding, the weapons infrastructures of these laboratories would have to be significantly supported by activities related to stockpile support, environmental restoration, and waste-management roles.

Nuclear weapons are complicated systems, and a *very* high value is placed on their performing when desired, not accidentally. In many areas of weapons technology, repetitive statistical testing is not possible, and national security precludes open technical exchange. Further, the nuclear testing of weapons designs is expected to be prohibited by the late 1990s. To maintain reliability and preserve quality, some form of intellectual competition and peer review is essential. In the U.S. this has historically been provided by two independent nuclear design laboratories and one warhead systems engineering laboratory. Whatever the future form of the weapons program, adequate intellectual competition and peer review must be maintained through some appropriate mechanism.

Minimum Complex 21. "Complex 21" is the designation for the downsized and cleaner nuclear weapons complex that will meet the needs of the twenty-first century. The safe storage of plutonium and enriched uranium (either as dismantled weapons components or in other forms) is a dominant requirement of the new complex; another is processing dismantled material. While sealed weapon components can be stored for many decades, some of these units will have to be reprocessed as they age. Also, the world community will likely press for permanent storage (as vitrified waste) or energy conversion (via reactor or accelerator burning) of excess Russian and U.S. fissile materials. The capability to fabricate a modest number of new warheads or remanufacture those in the enduring stockpile will be optimally located at the chosen nuclear-materials storage and processing site. (One way of assessing the needed capacity for fabrication is to compare the number of weapons in the long-term stockpile with a typical weapon lifetime. From this basis we can estimate a need for about 100 to 200 units per year—down by an order of magnitude from peak Cold-War production rates!) We expect that Complex 21 will be a radical departure from the present complex: some of today's plants will not have direct counterparts in Complex 21, though their technologies will live on.

In the future the traditional distinction between responsibilities of the production complex and the design laboratories will become somewhat more diffuse. While processing and fabrication for the stockpile will be based within the manufacturing and storage facilities, each such technology, for example plutonium processing and fabrication, will be based upon an R&D capability at the laboratories. Modern manufacturing and process technology will be developed at the laboratories to minimize waste and worker exposure and to resolve environmental and safety concerns. Fabrication of some nonnuclear components and the few weapon prototypes necessary to support the nation's weapons RD&T program will be accomplished by the national laboratories. This evolution of the national laboratories' responsibilities will allow the nation to effectively maintain both its research and development capabilities and an essential back-up processing and fabrication capability.

Although we expect no new production of plutonium or highly-enriched uranium, Complex 21 will eventually include a limited new tritium-production capability to replace current reactors. In addition, the nation will continue to move toward a strategy that satisfactorily manages long-lived radioactive wastes from defense and other sources.

Environmental Management. Similarly, Minimum Complex 21 must address those DOE sites at which weapons production took place in the previous four decades. Their environmental restoration and closure poses immense technological and financial challenges. Vitrification and storage of high-level waste at Yucca Mountain; the testing and safe operation of WIPP; the management and cleanup of the Hanford storage tanks; liquid effluent cleanup at numerous sites; acceptable disposal of mixed waste; effective long-term environmental monitoring; and residue elimination, cleanup and decommissioning of Rocky Flats are some of the large—and costly—environmental hurdles the DOE has yet to clear.

In the new configuration of the complex, increased investment in environmental science and technology will enable the DOE to more affordably address its own environmental responsibilities and comply with regulations. Such investment will be used to reduce environmental risk on a national scale, mitigate industry's cost for environmental compliance (now over $100 billion per year), help nurture a competitive, high-technology environmental industry, and develop improved foundations for regulatory policies.

Based on a model already implemented at Los Alamos, the laboratories are using a risk-based, cost-benefit analysis (for example, Multi-Attribute Utility Theory) to develop priorities regarding compliance agreements. This approach should be adopted nationwide so that DOE, DOD, and the EPA Superfund resources are applied to the most urgent problems. Local communities need to be increasingly involved in the weighting factors for such analysis. Field sensors developed at Los Alamos for monitoring emissions from facilities are providing an improved and less costly basis for environmental assessment. Such sensors, when extended to a national scale, could provide a better assessment of national environmental issues, and if they are extended to satellite-based capabilities long used by the weapons laboratories, such remote sensing capabilities could be the basis of a worldwide environmental network.

Laboratory capabilities in accelerator design and nuclear-materials processing are being applied to two outstanding issues: safe disposal of long-lived nuclear waste, and the safe, economical production of the tritium needed for the ongoing stockpile. Tritium has a 12-year half-life; therefore, unlike other nuclear materials it cannot be stored or reused indefinitely. Accelerator production of tritium is attractive because it does not produce transuranic waste or build up a large mass of other waste products. It will offer an economical and technically attractive source of tritium as the stockpile is reduced. Accelerator technology also offers an effective, economical means to transmute long-lived, high-level waste and actinides (such as plutonium waste) into shorter-lived waste that decays in roughly 100 years rather than 10,000 years. Such transmutation would dramatically ease the requirements on geological nuclear-waste repositories. The laboratories will continue to support current waste-management efforts such as the Yucca Mountain project; however, accelerator transmutation of waste may produce a major shift in waste-management strategy and could also be used for the conversion of weapons-grade plutonium.

Integration of RD&T with broader national missions—Technology Transfer and Conventional Defense. The RD&T laboratories will be integrated with broader national missions to provide leverage and to make more efficient or dual use of the advanced-technology capabilities primarily developed and maintained for the weapons-program mission. Cooperation and partnerships with U.S. industry will become a routine and significant element of the laboratories' activities, enabling transfer and application of unclassified advanced technology with the aim of assisting U.S. competitiveness and addressing demanding civilian problems such as environment and infrastructure. The 1992 DOE Defense Critical Technologies Plan identifies many such opportunities. For example, the laboratories' computational-science capabilities are already benefitting industry in the areas of combustion modeling, oil-well logging, and simulating performance of complex mechanical systems. They are also being applied to health research through such projects as the HIV database and the modeling of complex biological molecules. We can make a substantial difference to U.S. economic competitiveness by designing an "information interstate highway system" to link government and industrial assets in every state. Mechanisms to expand the ability of the laboratories to develop high-leverage interactions with industry will be encouraged.

Similarly the laboratories' technologies will continue to be tapped by the Department of Defense for conventional (non-nuclear) defense. The synergism between nuclear and non-nuclear work has already been demonstrated in explosives development, armor/anti-armor, advanced munitions, and computer simulations of performance, safety, and lethality. This linkage will be continued as a strategic element in maintaining an effective defense R&D base. In addition, many of the technologies being developed for advanced manufacturing and waste remediation will have significant benefit to the DoD.

John D. Immele (left) is currently Associate Director for Nuclear Weapons Technology at the Laboratory. He received a B.S. in chemistry from the University of Illinois in 1969 and a Ph.D in nuclear chemistry from the University of California, Berkeley, in 1972. Postdoctoral positions in theoretical physics at the University of Munich and the University of Maryland were followed by various staff positions at Lawrence Livermore National Laboratory culminating in his appointment as Deputy Associate Director of its Nuclear Design program. He joined Los Alamos National Laboratory in 1988. Immele's views on nuclear-weapons issues have appeared in various public media.

Philip D. Goldstone (right) is currently the Chief Scientist for the Inertial Confinement Fusion (ICF) and High Energy Density Physics programs at the Laboratory. He received B.S. and M.S. degrees from the Polytechnic Institute of Brooklyn in 1971 and 1972 and his Ph.D. in physics from the State University of New York at Stony Brook in 1975. After joining the Laboratory for a postdoctoral appointment in experimental nuclear physics, he stayed to work on shock hydrodynamics, high-energy-density physics using lasers, and the physics of laser-driven ICF. From 1981 until 1989 he led the Laser-Matter Interaction and Fusion Physics group and from 1986–89 was also the program manager for ICF experiments. From 1989 to 1992 he served on the staff of the Associate Director for Nuclear Weapons Technology, providing support and advice on research issues.

Los Alamos: The Ongoing Commitment

Our laboratory has grown and matured with the nation's security needs. Our fiftieth anniversary coincides with sweeping changes in national requirements for military and economic security. We have outlined an effective and achievable vision for the reconfiguration of the weapons complex that emphasizes the critical importance of maintaining the intellectual basis for stewardship as the size of the nuclear force decreases. The key issues that will determine the future of the program are summarized in the accompanying box. Our vision has been put forth as part of our ongoing commitment to the nation. Whatever decisions are made about the future of the weapons complex, Los Alamos will be at its hub to ensure the integrity of the U.S. nuclear capability. ■

Key issues framing today's decisions

The United States must maintain the necessary capability in nuclear weapons technology to provide stewardship of the remaining stockpile, provide technological option and assessment capability, and provide the capability to maintain, modify, and produce weapons when necessary. Technological competence and capability is not sustained simply by maintaining existing, deployed systems. Historically, it has been sustained by a "natural process" brought about by the development/test/production cycle. This is similar to the dilemma now faced by the DoD in maintaining its technical base. An active strategy should be developed to provide for long-term competence and for confidence in the safety, reliability, and relevance of the nation's nuclear force.

The laboratories will be an integral—in fact central—element of the redefined weapons complex. They will also play a vital role in providing the research and technology needed to enable the transition to it. Cost-effective stewardship of the nation's nuclear capability cannot be obtained simply by reconfiguring and downsizing today's capabilities; R&D investments must be made to make it possible and to bound its cost. The laboratories will have to invest in and develop capabilities ranging from more adequate above-ground facilities, to demonstration of lower-waste, lower-cost plutonium processing and fabrication technologies, to more cost-effective technologies for supporting environmental-management goals for the complex.

Although Congress has adopted plans to phase out nuclear testing, it should be recognized that such testing has been a vital element in maintaining long-term competence and confidence in the safety and performance of the nuclear force. Safeguards and investments in appropriate facilities for above-ground simulations and experiments are crucial to ameliorating the effects on technical competence of stringent test limits or a test ban. In particular, investments in improved hydrodynamic testing capabilities and in computational enhancements are now urgent.

Innovation and competence in weapons science and technology areas are critically bolstered by the continued presence at the laboratories of a range of program activities that is broader than their central nuclear-weapons mission. This diversity also enhances the potential for technology transfer and commercialization.

A more than one-third reduction through 1993 (and perhaps an additional 20 percent reduction in 1994) in the level of effort in research, development, and testing has made laboratory consolidation an increasingly important issue. Future RD&T funding is projected to fall below the critical level at which substantial changes in the present architecture of the RD&T complex must be examined. If so, we must consider a careful shift of laboratory missions in a way that preserves the necessary expertise and facilities, provides adequate mechanisms for intellectual competition and peer review, and maintains the quality of such a reduced program.

The national interest places very high value on slowing or preventing proliferation of weapons of mass destruction. Competent weapons-program expertise is required to develop detection and materials-accountability technologies as well as to assess intelligence data.

The DOE must make the transition to a Minimum Complex 21 that supports the new goals and requirements outlined in the main text. The expertise of the weapons Laboratories should be an integral component of this transition. Despite the termination of plutonium fabrication at Rocky Flats, the U.S. will still need to reestablish a long-term facility for nuclear-materials storage and processing as well as a small but flexible fabrication capability to support the stockpile. The laboratories should provide and demonstrate technology for the future U. S. plutonium capability as well as technologies for cleanup of Rocky Flats.

There will likely be ongoing public concern regarding nuclear waste and related issues. The potential of accelerator alternatives for tritium production, waste transmutation, and conversion of weapons-grade plutonium should be aggressively evaluated as possible ways to ameliorate and benefit from these issues economically. □

Key issues framing today's decisions

The United States must maintain the necessary capability in nuclear weapons technology to provide stewardship of the remaining stockpile, provide technological option and assessment capability, and provide the capability to maintain, modify, and produce weapons when necessary. Technological competence and capability is not sustained simply by maintaining existing, deployed systems. Historically, it has been sustained by a "natural process" brought about by the development/test/production cycle. This is similar to the dilemma now faced by the DoD in maintaining its technical base. An active strategy should be developed to provide for long-term competence and for confidence in the safety, reliability, and relevance of the nation's nuclear force.

The laboratories will be an integral—in fact central—element of the redefined weapons complex. They will also play a vital role in providing the research and technology needed to enable the transition to it. Cost-effective stewardship of the nation's nuclear capability cannot be obtained simply by reconfiguring and downsizing today's capabilities; R&D investments must be made to make it possible and to bound its cost. The laboratories will have to invest in and develop capabilities ranging from more adequate above-ground facilities, to demonstration of lower-waste, lower-cost plutonium processing and fabrication technologies, to more cost-effective technologies for supporting environmental-management goals for the complex.

Although Congress has adopted plans to phase out nuclear testing, it should be recognized that such testing has been a vital element in maintaining long-term competence and confidence in the safety and performance of the nuclear force. Safeguards and investments in appropriate facilities for above-ground simulations and experiments are crucial to ameliorating the effects on technical competence of stringent test limits or a test ban. In particular, investments in improved hydrodynamic testing capabilities and in computational enhancements are now urgent.

Innovation and competence in weapons science and technology areas are critically bolstered by the continued presence at the laboratories of a range of program activities that is broader than their central nuclear-weapons mission. This diversity also enhances the potential for technology transfer and commercialization.

A more than one-third reduction through 1993 (and perhaps an additional 20 percent reduction in 1994) in the level of effort in research, development, and testing has made laboratory consolidation an increasingly important issue. Future RD&T funding is projected to fall below the critical level at which substantial changes in the present architecture of the RD&T complex must be examined. If so, we must consider a careful shift of laboratory missions in a way that preserves the necessary expertise and facilities, provides adequate mechanisms for intellectual competition and peer review, and maintains the quality of such a reduced program.

The national interest places very high value on slowing or preventing proliferation of weapons of mass destruction. Competent weapons-program expertise is required to develop detection and materials-accountability technologies as well as to assess intelligence data.

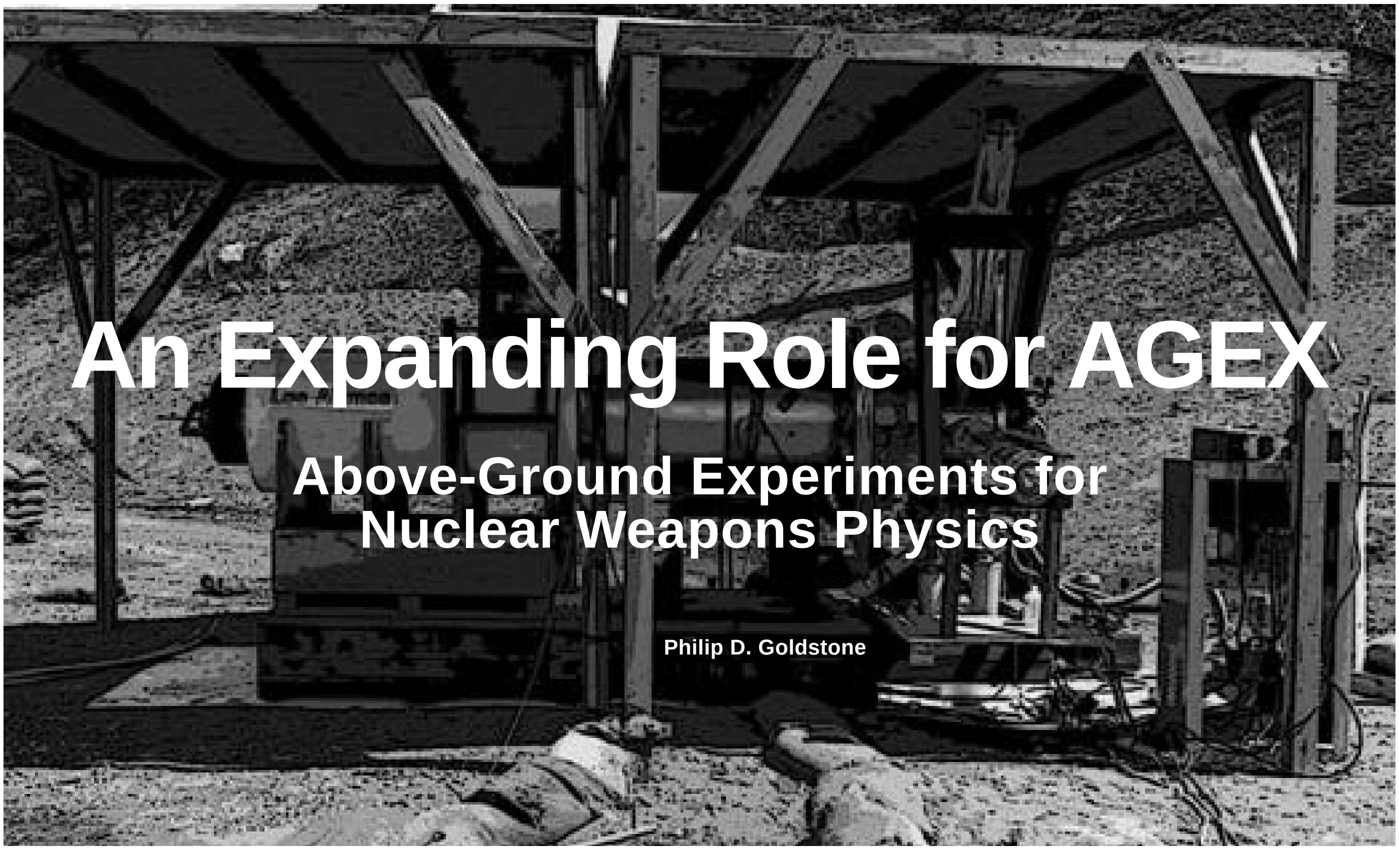
The DOE must make the transition to a Minimum Complex 21 that supports the new goals and requirements outlined in the main text. The expertise of the weapons Laboratories should be an integral component of this transition. Despite the termination of plutonium fabrication at Rocky Flats, the U.S. will still need to reestablish a long-term facility for nuclear-materials storage and processing as well as a small but flexible fabrication capability to support the stockpile. The laboratories should provide and demonstrate technology for the future U. S. plutonium capability as well as technologies for cleanup of Rocky Flats.

There will likely be ongoing public concern regarding nuclear waste and related issues. The potential of accelerator alternatives for tritium production, waste transmutation, and conversion of weapons-grade plutonium should be aggressively evaluated as possible ways to ameliorate and benefit from these issues economically. □

An Expanding Role for AGEX

Above-Ground Experiments for Nuclear Weapons Physics

Philip D. Goldstone



Procyon, the high-explosive pulsed-power generator at Los Alamos

For the last fifty years U.S. expertise in nuclear-weapons design and engineering was a natural by-product of an active development and testing program. With the end of the cold war, the development of agreements to retire most of the U.S. and former Soviet weapon stockpiles, and plans to phase out U.S. nuclear testing by 1996, the way the nuclear weapons program has operated is changing dramatically. However, nuclear weapons will remain a presence in the world for the foreseeable future, and the U.S. will retain a smaller but very real nuclear capability. We at the weapons laboratories will therefore have a vital responsibility—that of long-term stewardship of the stockpile, including continued assurance of reliability and safety. Since the weapons remaining in the stockpile will age, eventually we will have to provide competent assessments of the need to modify or replace aging weapons or components. This role will continue indefinitely into the future and beyond the career lifetime of many of today's weapon scientists.

Imagine, twenty years from now, a military commander or a member of a stockpile-surveillance team noticing that one of the weapons being stored has changed in appearance. They will want to know, "is this still safe, and would it work if needed?" They will call the Laboratory and ask the experts regarding this weapon. Will they be able to rely on the answer they get? A competent technical capability—including the honed expertise and judgement of weapon designers and engineers—is the principal element required for the assurance and stewardship of whatever nuclear stockpile the nation retains. At Los Alamos, where over 80 percent of the weapons that are expected to remain in the stockpile were designed, this responsibility is particularly important.

Without nuclear testing we will simply not be able to maintain today's level of competent judgement. The challenge we now face is to continue to exercise our weapons R&D expertise so that we retain—as much as possible—the ability to assess our own systems (and the nuclear capabilities of others), if we are without any fully integrated nuclear tests and without ongoing development programs. It will be necessary, but not sufficient, for the scientists and engineers involved in this enterprise to do research; the research itself must stress their judgement in weapons science and engineering. Furthermore, to prevent losing capability within a generation, these research programs must be technically interesting to attract and train new scientists and engineers.

In the absence of underground testing, a high premium will be placed on predictive computational capability for design and engineering of weapons. If our computer-simulation capabilities were perfect, then, in principle, actual testing would not be required for us to be confident of the performance of an aging, modified (for improved safety, for example), or even redesigned device. Computational-simulation capabilities are simply insufficient for this challenge, despite many years of development and the ongoing revolution in high-performance computing.

Therefore, part of the strategy for meeting the dilemma posed by a cessation of nuclear testing is to develop higher-performance computing capabilities and to exploit those capabilities by running more accurate and more predictive codes.

Although the weapons-design codes may be viewed as an "archive" where design physics is contained for future designers to use, continued comparison of code predictions against experiment—*reality*—is required to provide validation of the codes and the designer's judgement. The physics, materials behavior, and engineering associated with nuclear weapons is extremely complex. In a matter of moments, weapon components are brought from their normal physical state to the most extreme conditions found in the solar system. It is the act of continually testing both the codes and the designers against experiment that develops and maintains expertise and that prevents theory and reality from diverging. As long as a simulation code contains implicit approximations and assumptions, its value is intimately tied to the judgement of those who use it and interpret its output. A divergence of theory and reality can (and often does, in many human endeavors) result in a false sense of confidence in computation or design judgement that is potentially disastrous. This is the principal underlying reason for regularly performing nuclear tests, even if we could anticipate never having to modify a stockpiled weapon again. Therefore, in the absence of underground testing, appropriate "above-ground experiments" (AGEX) must be vigorously pursued.

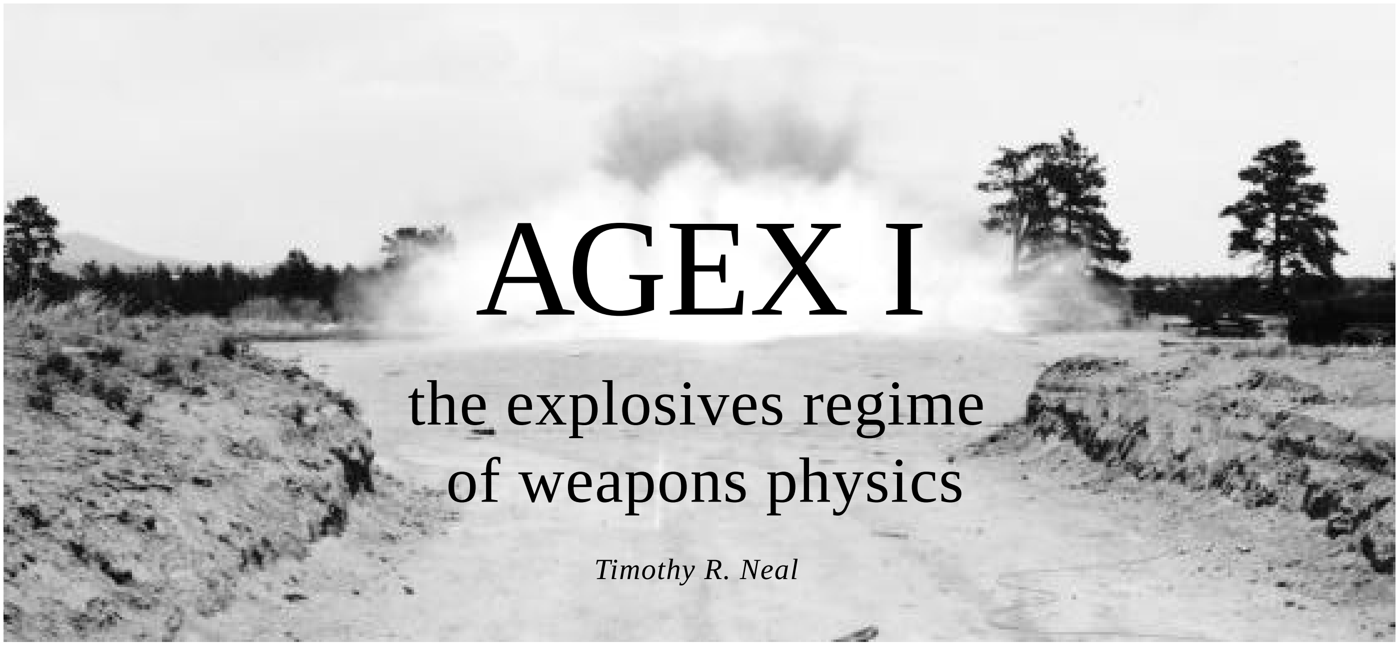
The role of experiments is, first, to provide the physics data and the underlying physical models to improve the predictive nature of the codes. Second, and possibly more important, experiments provide ongoing validation and exercise of weapons capability and judgement. A proper mix of experiments will even actively involve the fabrication, engineering, quality assurance, and fielding skills vital to weapons capability. Third, only a technically stimulating research program that combines experiment with theory is capable of attracting and retaining quality scientists and engineers in the weapons effort. Finally, in its own right, experimental capability—the ability to measure and interpret data, whether that data derives from an underground weapons test, an above-ground experiment, or from other sources such as nonproliferation activities—is an essential component of U.S. nuclear weapons stewardship.

The following two articles discuss the two physical regimes that must be addressed in experiments related to weapons design. The "explosives regime" (or AGEX I) includes the physics and chemistry of explosives as well as the behavior of matter subjected to the pressures, shocks, and temperatures that may be achieved with typical high-explosive configurations. For example, the behavior of heavy-metal assemblies compressed by high explosives is an important element of weapon design and engineering.

Above-ground experiments directly applicable to many (though not all) important aspects of the high-explosives regime of weapon design are possible without a nuclear explosion. The principal example is that of hydrodynamic testing, including flash radiography of high-explosives-driven assemblies. The DARHT (Dual-Axis Radiographic Hydro Test) facility is the single most important new AGEX capability for the future confidence in the U.S. stockpile. Explosives characterization as well as hydrodynamic testing is discussed in "AGEX I—The Explosives Regime of Weapons Physics."

The second physical regime that must be addressed by experiments is the "high-energy-density" (or AGEX II) regime, which is typically achieved only after the initial production of nuclear energy. It involves temperatures from tens to thousands of electron-volts or eV (1 eV is equivalent to 11,600 kelvins) and pressures greater than 10 megabars (1 megabar equals 1 million atmospheres). Technical issues in the high-energy-density regime include radiation flow and the interaction of radiation with matter; radiation hydrodynamics and hydrodynamics at extreme pressures; nuclear cross sections and neutron interactions; and the behavior of dense plasmas.

Whereas the explosive regime can be more or less directly accessed at full scale, achieving the high-energy-density regime over anything approaching the full spatial and time scales of a weapon would imply an extraordinary amount of energy, comparable to that of a nuclear explosion. Although this is clearly neither achievable nor desirable in an above-ground experiment, various pieces of the physics can be accessed and studied by using an appropriate set of specialized facilities. The efforts to develop and use such facilities for the study of the high-energy-density regime are discussed in "AGEX II—The High-Energy-Density Regime of Weapons Physics." ■



AGEX I

the explosives regime of weapons physics

Timothy R. Neal

The history of explosives research and above-ground experimentation for nuclear weapons began with the Manhattan Project. During the hectic, almost frantic, war days at Los Alamos, it became clear that, if possible, the fissionable material in the weapon should be plutonium. It was equally apparent that the critical mass of plutonium needed to produce a nuclear explosion would have to be assembled in the weapon through a spherical implosion driven by powerful explosives (Figure 1). Thus from the beginning the development of nuclear weapons was intimately connected with and dependent on developing fabrication, quality-control, and inspection technology for high explosives (explosives with energies greater than that of TNT). Initial experiments in the spring and summer of 1943 revealed, among other things, that for the weapon to work the design of the explosive charges and the timing of their detonation would have to achieve a precision hitherto not contemplated. The achievement of those goals left Los Alamos, at the end of World War II, uniquely in possession of the most advanced explosive-fabrication technology on earth and a mission to make nuclear weapons safer and more efficient—a mission that has continued into the present.

For a long period of time, the work on weapons implosions has utilized conventional plastic-bonded high explosives, which could be precisely machined. Improvements were continually made to increase the accident resistance of these materials. The emphasis on safety in nuclear weapon research led to the development of insensitive high explosive (IHE) at Los Alamos. During the 1970s the Laboratory pioneered the use of IHE in nuclear weapons designs, which dramatically decreased the possibility that the explosives would detonate during accidental insults. Most modern weapons are designed to incorporate insensitive explosives. An IHE—such as triaminotrinitrobenzene (TATB)—can be dropped from great heights and will shatter but not explode. If exposed to fire in an accident, TATB will burn, but it is extremely unlikely to undergo a transition from burning to deflagration or detonation. Even when exposed to high temperature, extreme pressures, or shocks, these materials resist explosion. Thus, they can be handled quite safely with simple precautions.

In addition to safety, the stability and reliability of nuclear weapons in the nation's stockpile have been ongoing concerns. Scientists and engineers have continued to study the compatibility of materials contained in weapons during long-term storage and to develop new materials for weapons components. The development of new materials has even led to applications in the commercial sector. For example, a high explosive developed in the weapons program, nitrotriazolone (NTO), is under consideration for use as a gas producer in automobile air bags.

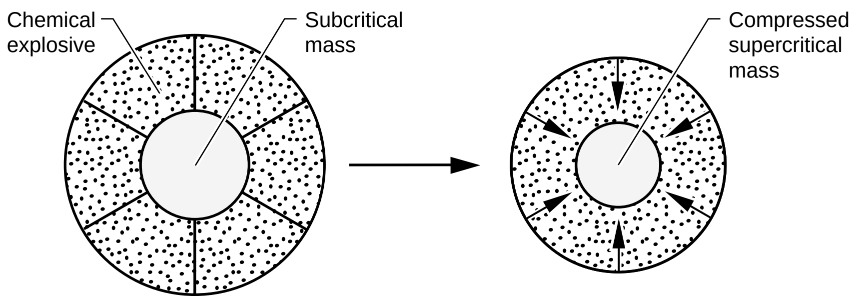


Figure 1. Explosive-driven Implosion
Explosion of a fission weapon is initiated by the implosive force generated by the detonation of a layer of high explosive surrounding the fissile fuel. The detonating high explosive compresses a subcritical mass of fissile material to form a supercritical mass that then rapidly releases nuclear energy through an uncontrolled fission chain reaction.

Research on Safety and Performance of High Explosives

The end of the Cold War has led to increased emphasis on safety. An overriding worry is that an accident might cause the explosive in a nuclear weapon to release its energy, thus causing the assembly of a critical mass and the production of some sort of nuclear yield. Even if a nuclear yield is totally averted through inherent design features, the explosive-energy release might still disperse radioactive plutonium across the countryside. Nuclear weapons have long been designed to avoid or drastically reduce such threats. For example, all weapons in the stockpile are inherently “one-point” safe; that is, the initiation of the explosive at some random point will not produce a nuclear yield. Weapons have also been tested against the raging inferno of a jet-fuel burn to assure their safe response should, for example, a bomber loaded with nuclear weapons catch on fire. However, during the Cold War, as we stood eyeball to eyeball with the Soviets, certain low risks were considered to be more tolerable. Now that the Soviet threat is retracted and our current intent is to dismantle or store needed nuclear arms rather than brandish them, the public deserves even greater assurances about safety. Accident analyses have therefore been extended to address extremely low-probability accidents. Complex, multiple-accident scenarios now being considered include the possibility that after a bomber loaded with nuclear weapons catches on fire, another large plane crashes into it. Can the new “wooden” insensitive high explosives withstand both the high temperature and the severe impact that would be involved in such an accident?

In order to predict the response of explosives in various accident scenarios, research has been under way to further understand the detonation process in high explosives. Unlike gasoline, which must be mixed with the oxygen in the air in order to burn completely and rapidly, high explosives contain enough oxygen to undergo extremely rapid and complete exothermic (heat-producing) chemical reactions. The high explosive is said to undergo detonation if the chemical reaction propagates by compressing the material ahead of it and reaches 90-percent completion within a few millionths of a second. Such rapid reactions produce strong shock waves.

The detonation of a high explosive is typically initiated by a small shock wave that strongly compresses the explosive at a point, causing it to heat up and burn. The exothermic chemical reaction happens so rapidly that the pressure of the reaction products compresses the fuel around it causing that fuel, in turn, to heat up and react, and so the detonation proceeds to spread out from the point of initiation just like a spherical wave. This compression-driven reaction travels at supersonic velocities and is called a detonation wave. The leading edge of the detonation wave is a shock front; that is, there is a discontinuity in pressure, temperature, and density across the front. The pressures built up in the gaseous reaction products behind the shock front are typically on the order of a few hundred thousand atmospheres, and the temperatures are typically between 2000 and 4000 kelvins.

Most accidental insults to a nuclear weapon would not produce shock waves that could initiate the detonation of high explosives. However, exposure to fire along with the impact of a crash might initiate a deflagration, a burn front that propagates by heat conduction rather than compression and therefore proceeds about a million times more slowly than a detonation. A deflagration in explosives and propellants might, however, build up into a full-scale detonation.

The deflagration-to-detonation transition is a significant safety consideration in all industrial, military, and nuclear weapon applications of high explosives and propellants. A comprehensive study of this problem involving a consortium of university and government laboratory participants is under way, and the results of the study are being incorporated into engineering codes for predictive design and safety assessment of nuclear weapons. When the deflagration-to-detonation process is properly understood, we can effect safety measures to guard against even a low-risk accident.

The most important thrust of current explosives research is to develop better models of deflagration and detonation through a combination of experimental and theoretical work. Many advances were achieved in modeling the detonation of conventional high explosives. The fact that chemical reactions in these materials can be considered to occur instantaneously simplifies the modeling of the detonation wave. In contrast, the reaction times of insensitive high explosives are slower and seem to depend on their location inside the explosive charge. Thus the modeling of detonations in IHE has been a far more difficult problem. Through a very strong experimental program scientists have been able to confirm theoretical predictions concerning the behavior of insensitive high explosives, in particular, that reaction rates are strongly accelerated by increases in temperature and pressure. The results of these experiments on reaction rates have been used to develop more precise models of the initiation and detonation of insensitive high explosives and to better understand the effects of reaction rates on the sensitivity of the explosive to heat and impact.

Good models of deflagration and detonation are essential because the set of possible accidents is too broad to test each directly. Growing computing capabilities make it possible to use basic models to simulate the behavior of explosives even in complex geometries. Thus the wave of the future emphasizes carefully selected benchmark experiments to characterize explosive behavior followed by the linking and extension of those results through numerical simulations on supercomputers.

Los Alamos scientists are extending their historic mission in high explosives research to discover at the molecular level what an explosive is and how it works. This fundamental research enlists sophisticated spectroscopic experimental techniques to learn what holds the explosive molecules together, how they come apart during initiation and detonation, and how the released energy builds up the pressure and temperature of the gaseous reaction products so they can do useful work (for example, drive the implosion of a metal

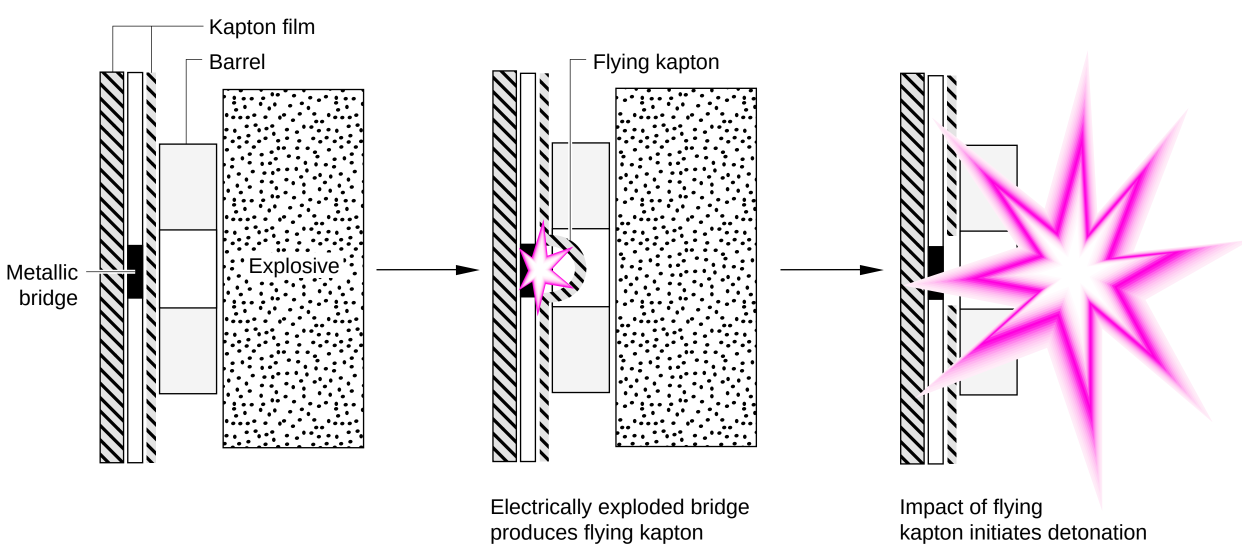


Figure 2. The Slapper Detonator

A detonator initiates detonation of a high explosive by creating a short, high-pressure pulse in the explosive. Illustrated here is the operation of a standard detonator called the slapper. An intense pulse of electrical energy causes the metallic bridge to burst. The burst drives the kapton film down a short barrel that cuts the film like a cookie cutter. When the piece of flying kapton hits the explosive, it generates a sufficiently strong pressure pulse to cause the explosive to detonate.

sphere). Such studies improve the ability to predict how a new explosive will behave and may also lead to an improved first-principles approach for prescribing explosives with specific desired characteristics. Collaboration of theorists and organic chemists at Los Alamos has recently led to the discovery of a new class of insensitive high explosives that, unlike previous explosives, are very rich in nitrogen and contain much less carbon and oxygen. The first of these to be synthesized, LAX-112, is less sensitive than TNT and produces a more powerful detonation than TATB. Work is continuing to find an explosive within the new class that is even more insensitive but retains the high performance of LAX-112.

The design and engineering development of systems to initiate the detonation of high explosives components is also a part of explosives work. These initiation systems use electrical capacitor discharge units to explode a bridgewire and thereby create a high-pressure pulse in a small region of the explosive. Recent advances in initiation systems include improved safety features. For example, the requirements for stored electrical energy are much lower, and traditional exploding bridgewires have been replaced by flying slappers (Figure 2). Emphasis on improved safety has also led to the development of safer explosives for initiation systems.

The next generation of initiation systems will be based on even safer detonators. Laser-driven slappers or direct optical deposition will create the shock waves that initiate detonation. In more traditional systems metal wires were coupled to the detonator; hence, those wires could feed electrical pulses to the detonator from lightning or other accident sources. In contrast, the new optical power sources cannot be triggered by external sources in any accident scenario.

The disposal of high explosives and propellants that are removed from weapons systems and the environmental redemption of waste from explosive and propellant manufacturing are technologies of prime current interest. Research and development on safe, environmentally acceptable methods for explosive and propellant disposal are under way at Los Alamos and involve interdisciplinary collaborations among many parts of the Laboratory. Methods ranging from base hydrolysis to biological degradation and supercritical water oxidation are under investigation. In the latter, the explosive is broken down into innocuous gases that can then be released.

Above-Ground Hydrodynamic Experiments with High Explosives

The demonstrations in 1943 that an explosive-driven implosion of a metal sphere or cylinder was possible opened up to study the behavior of matter under the extreme pressures, shocks, and temperatures generated by high explosives. This specialized science is termed hydrodynamic testing because solids and metals seem to flow like liquids when driven by the detonation of high explosives.

Firing sites are the laboratories for hydrodynamic testing. Because each experiment self-destructs during a test, the entire experiment must be rebuilt before it can be repeated. Scientists therefore cast about continually for ways to obtain more experimental data from a single experiment. In spite of the commonplace descriptor, "hydrodynamics," nothing is taken for granted. Every aspect of the broad subject of detonations, the interaction of gaseous explosive products with inert materials, and the possible effects of material strength on the resulting flows is examined extensively. In a common type of experiment, a metal plate is placed in contact with the high explosive, and the high explosive is detonated with the goal of determining how effective it is at pushing on the metal plate. The pressure exerted by the detonating explosive is typically about a million times greater than atmospheric pressure—much higher than the yielding strength of any ordinary metal—and causes the metal to move rapidly, covering distances of a few millimeters in a millionth of a second. Early diagnostics consisted of electronic gauges and high-speed optical motion picture cameras that took a few pictures at the rate of a million pictures per second.

In addition to the experiments with metal plates, experiments were also carried out on weapons assemblies containing surrogates for the fissile material. Such experiments allowed measurements to be made on the early stages of implosion. The results were then used to calibrate computer simulations of weapons implosions that included the behavior of the fissile material.

In the 1960s, a major new diagnostic was added to the repertoire—flash radiography. The technique involves the use of a high-energy electron beam to produce extremely short-duration bursts of x rays. During a hydrodynamic test a single x-ray burst passes through the rapidly moving test object and is recorded on film. The resulting x-ray image

Figure 3. PHERMEX
PHERMEX and its associated image-analysis tools have been continually upgraded and maintained as the premier high-energy radiographic facility in the world. A radio-frequency linear accelerator directs a pulse of 30-MeV electrons to a tungsten target where the energy of the electrons is converted into bremsstrahlung radiation. This burst of x rays is used to make radiographic images of hydrodynamic tests involving high explosives. The photograph shows the thick cylindrical reinforced concrete bunker that houses and protects PHERMEX from the blasts generated during a hydrotest. Woven-steel blast mats covering one end of the bunker are adjacent to the explosive firing point. Electrical signals generated by the hydrotest begin their journey to the recording equipment underground in the structure shown in the lower portion of the photograph.

of the test object effectively "freezes" the motion of explosive-driven weapon components. Such radiographs are analyzed, in great detail, to determine whether the behavior of the weapon components agrees with theoretical predictions.

The machine called PHERMEX (pulsed high-energy radiographic machine emitting x rays) was built mainly for such weapons-system hydrodynamic testing—or hydrotesting, as we call it (Figure 3). PHERMEX was the country's first such facility and was, in certain respects, ahead of its time. It contains a large radio-frequency linear accelerator that produces a beam of relativistic electrons with energies of 30 MeV. The beam is directed at a tungsten target where the energy of the electrons is converted into bremsstrahlung radiation, most of it in the x-ray range. Through continual redesign and upgrade programs PHERMEX and its associated image-analysis capabilities have remained the premier high-energy radiographic capability in the world. Because flash radiography does not perturb the experiment in any way, it yields an accurate measure of whether the explosive performance matches theoretical and engineering predictions.

At first flash radiography stood alone as an isolated diagnostic. But because of the high cost of such experiments, electronic and optical diagnostic capabilities were soon added to the PHERMEX firing site. Thus began our current approach to hydrotesting: diagnose each experiment as thoroughly as possible to get the most return for the investment and to maximize the understanding of total system behavior.

This philosophy continues into the future with the construction of the new DARHT (dual-axis radiographic hydrodynamic test) Facility. Dual axis means that the facility has two x-ray machines that produce x-ray bursts from two directions (Figure 4). At present the images are captured on x-ray film or specialized storage media residing in a recoverable cassette that wards off blast and shrapnel damage. Only a tiny fraction of each x-ray burst actually penetrates the hydrotest object to record an image on the detector, so extensive image-analysis techniques are needed to quantify the resulting pictures. If the two bursts are generated at different times, the resulting images allow determination of velocities of the material in the interior of the test object. As an alternative, the two pictures can be taken at the same time but from different positions to give a "stereoscopic" view that yields a type of three-dimensional image. Finally, there is the option of orienting one x-ray machine to one area of a hydrotest to obtain the best possible resolution and orienting the other machine to a completely different area for similar reasons.

The biggest advance in measurement techniques in the last decade has been the development of quantitative radiography. Radiographs are no longer just pictures of items going hither and yon with distance scales superimposed for measurements. Radiography is now able to determine the density of compressed materials, the location of material interfaces to submillimeter precision, and the computer-assisted tomographic (CAT) reconstruction of interior sections of a distorted object. The latter process is analogous to the CAT scans used in the medical field (Figure 5).

Progress has also been made in other types of diagnostics. Electronic measurements have now attained temporal resolution of a billionth of a second, and hundreds of them may be made during a single hydrotest. Ultrafast color motion-picture cameras are now joined by electronic cameras that are over ten times faster. Lasers are being used as interferometers to precisely measure the velocity of surfaces (see "Line-imaging Laser Interferometer for Measuring Velocitues"). The laser light can be transmitted and returned to detectors through fiber optics, a method that allows measurements to be made in hard-to-reach places. Laser interferometers have traditionally been used to measure the velocities of only a single point on a surface. With the help of image analysis techniques, measurements can now be made along an entire line of a test object. Measurement along a line within an axially symmetric object translates into a continuous high-precision velocity map of an entire surface. In another technique microwaves are propagated through microminiature cables that bend around obstructions. The crushing of these cables as the detonation proceeds yields microwave interferometric measurements of the positions and velocities of shock and compression waves. A selection of these techniques is regularly applied to each hydrotest to measure the position, velocity, and condition of material surfaces as well as the propagation and pattern of wave-like disturbances.

The thrust for the future in hydrotesting is increased precision and all-encompassing diagnostics. Because nuclear testing will no longer be available as the final arbiter, our computational models and codes must be tightly tied to those phenomena that can be measured. In addition, those measurements must be made more universal to elucidate not just the behavior of surfaces but that of interiors as well. Even surface measurements must attain a new degree of sophistication that will yield information about temperature and material breakup.

The explosives AGEX activities must rise to a new role in the nuclear defense activities of the Labo-

Figure 4. Plan of the Proposed DARHT Facility

A major new initiative in the explosives AGEX program is the design and construction of a dual-axis radiographic hydrodynamics test facility. This high-intensity flash x-ray test site will contain two high-energy linear induction accelerators at right angles to each other. The x-ray bursts will be ten times more effective than those available at PHERMEX, enabling flash radiography of dense objects. The two distinct x-ray bursts will be used to generate radiographs of a single hydrotest at different times and with orthogonal views. The extension to the rounded end of one of the machine buildings will contain extensive capabilities for optical diagnostics. The electronic diagnostic equipment for DARHT, like that for PHERMEX, will be located underground near the firing site.

Line-imaging Laser Interferometers for Measuring Velocities

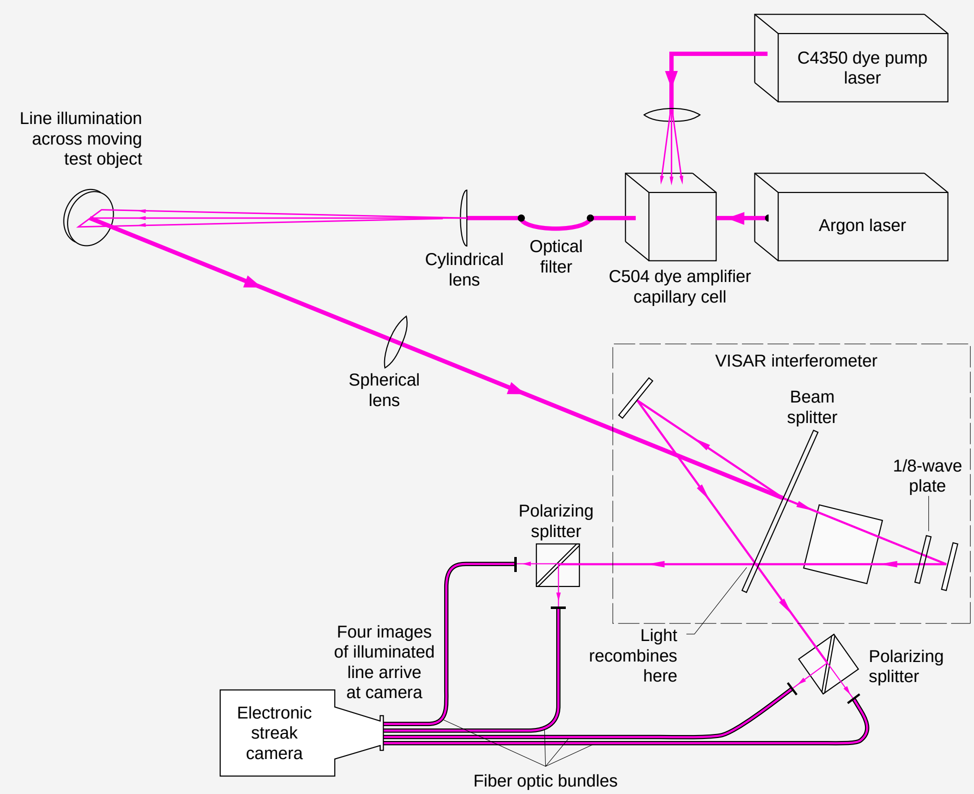
Willard F. Hemsing

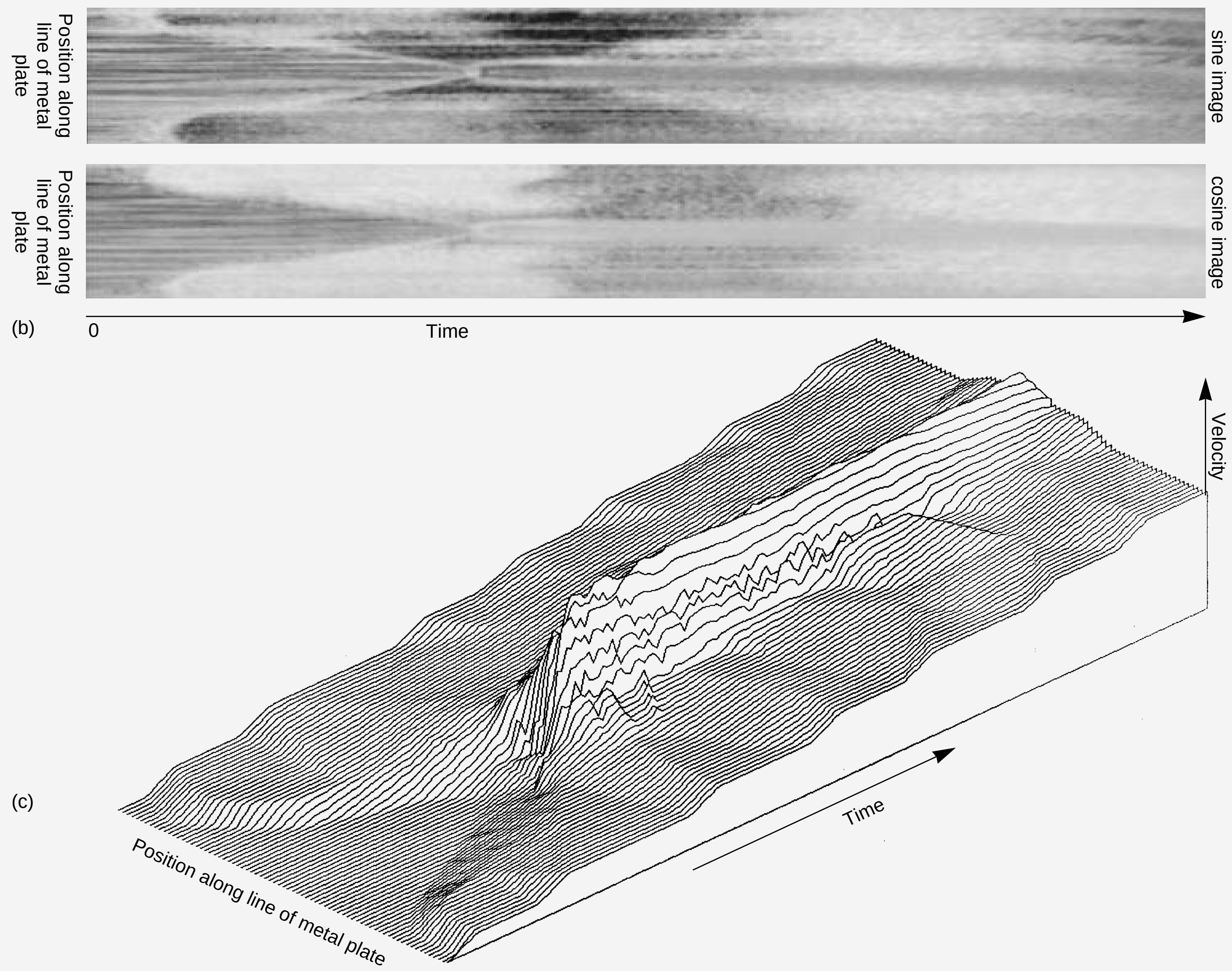
Hydrodynamic tests create hostile conditions in which high pressures can easily compress solids and accelerate materials to velocities of several kilometers per second. Among the advanced diagnostics for hydrodynamic tests at the Laboratory is our line-imaging VISAR (Velocity Interferometer System for Any Reflector). The VISAR measures the velocities of points along an illuminated line on a fast-moving test object. The instrument exploits the fact that when laser light is reflected from a moving surface, the wavelength of the light is Doppler-shifted in proportion to the velocity of the point that reflects it. The VISAR employs optical interference to generate bright and dark bands of light called interference fringes. The fringes oscillate between bright and dark as the test object accelerates. The VISAR measures velocity by accurately determining the number of whole and partial oscillations that occur as the test object accelerates. Its useful product is a continuous velocity history for all the points that are visible in the image.

(a) Our line-imaging VISAR uses a cylindrical lens to focus laser light onto a line on the test object. Conventional optics image the illuminated line through a special wide-angle Michelson interferometer, where a retardation plate delays the vertical polarization component of one beam by a quarter of a wavelength. As a result, when the beams are recombined to produce interference, the fringes of the vertical polarization component are shifted and their oscillations lag behind those of the horizontal component. Specifically, the intensities of corresponding points in the horizontal and vertical components depend on the sine and cosine, respectively, of the velocity at each point on the target. Polarizing beam splitters separate the horizontal polarization component from the vertical component where light exits from each side of the interferometer. This separation produces two pairs of images of the interference intensities along the illuminated line. The two images for each polarization are simply negatives of each other.

Fiber-optic bundles transmit the four images to the photocathode of an electronic streak camera. The

(a)





camera rapidly sweeps the images across a charge-coupled device that digitizes them into a microcomputer. Later, we subtract one image of each polarization from its negative to double the signal and cancel optical noise. Analysis of the images yields the velocity histories of many points in the line as a continuous function of time.

The VISAR's sensitivity to acceleration, instead of to velocity alone, best accommodates measurements of velocities from 100 meters per second to over 20 kilometers per second. Its recording time can vary from milliseconds to nanoseconds; the length of the line it observes can range from 0.3 to 30 millimeters across the target surface. Because it records pictures with their great capacity to store information, our line-imaging VISAR can capture many times more data than conventional VISARs. We have found its ability to simultaneously record large quantities of information relating different points on a test object extremely advantageous. This is most useful in measurements in which velocity gradients are important, and in tests that destroy expensive hardware, especially when test-to-test variations are important. Although our line-imaging VISAR is versatile, its use is precluded when smoke blocks its optical path or when the test-object surface loses light reflectivity.

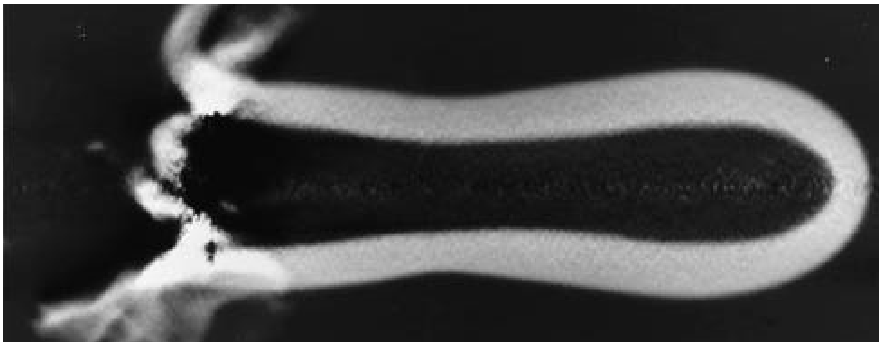
(b) The sine and cosine interference images from an experiment in which two converging detonation waves, produced by an explosive initiated at two separate points, drove a metal plate. Triangles extending across the left third of the images are the edges of interference fringes as they responded to the acceleration of the plate. A change from dark to bright, corresponding to an increase in velocity of 200 meters per second, is visible in the cosine image.

(c) An isometric plot of velocity, deduced from the photograph in (b), as a function of position along the illuminated line and time. The "cliffs" at the lower left indicate the acceleration of the metal as it was driven by the two converging pressure waves. The ridge extending from the center to the upper right is a region of high velocity caused by the pressure enhancement where the waves collided. □

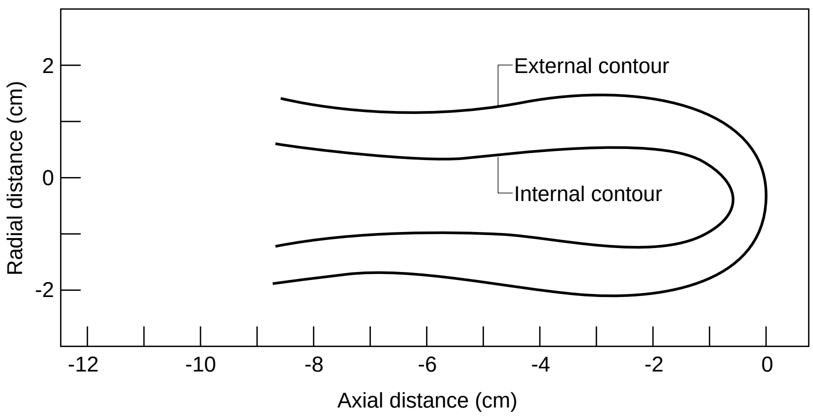
ratory and of the nation. Our capabilitites in explosives characterization, hydrodynamic modeling, and technology development are a special resource to the national materials science community, to U.S. industry, and to the conventional defense community. They are a unique and critical resource to the nuclear weapons community. As availability of under-ground nuclear testing fades, above-ground hydrotesting will become the keystone for nuclear weapon design, qualification, and safety assessment. ■

Figure 5. Quantitative Radiography

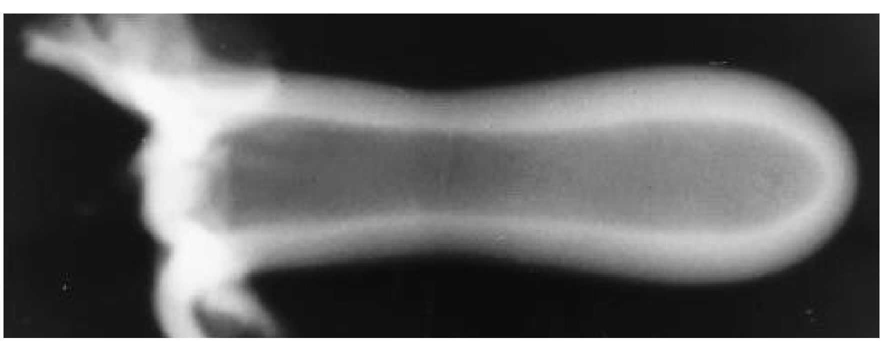
When metals are subjected to the shock pressures and temperatures created by the detonation of high explosives, they seem to flow like liquids. This figure shows images of an explosively formed penetrator made of copper during its high-velocity (2.4 kilometers per second) flight. The penetrator was originally a cone-shaped piece of copper backed by high explosive. The force of the high-explosive detonation shaped the copper into the form shown here.



(a) This radiograph is the average of four different radiographic films of the penetrator in flight. Of interest here is the detailed shape of the inner cavity. The lighter areas represent greater material thickness.



(b) The line drawings of the internal and external contours of the penetrator were estimated by a least-squares fitting of an analytical model to the x-ray film densities. In the forward portion of the penetrator, where axial symmetry is high, the edges of the contours are thought to be accurate to within 0.2 millimeter.

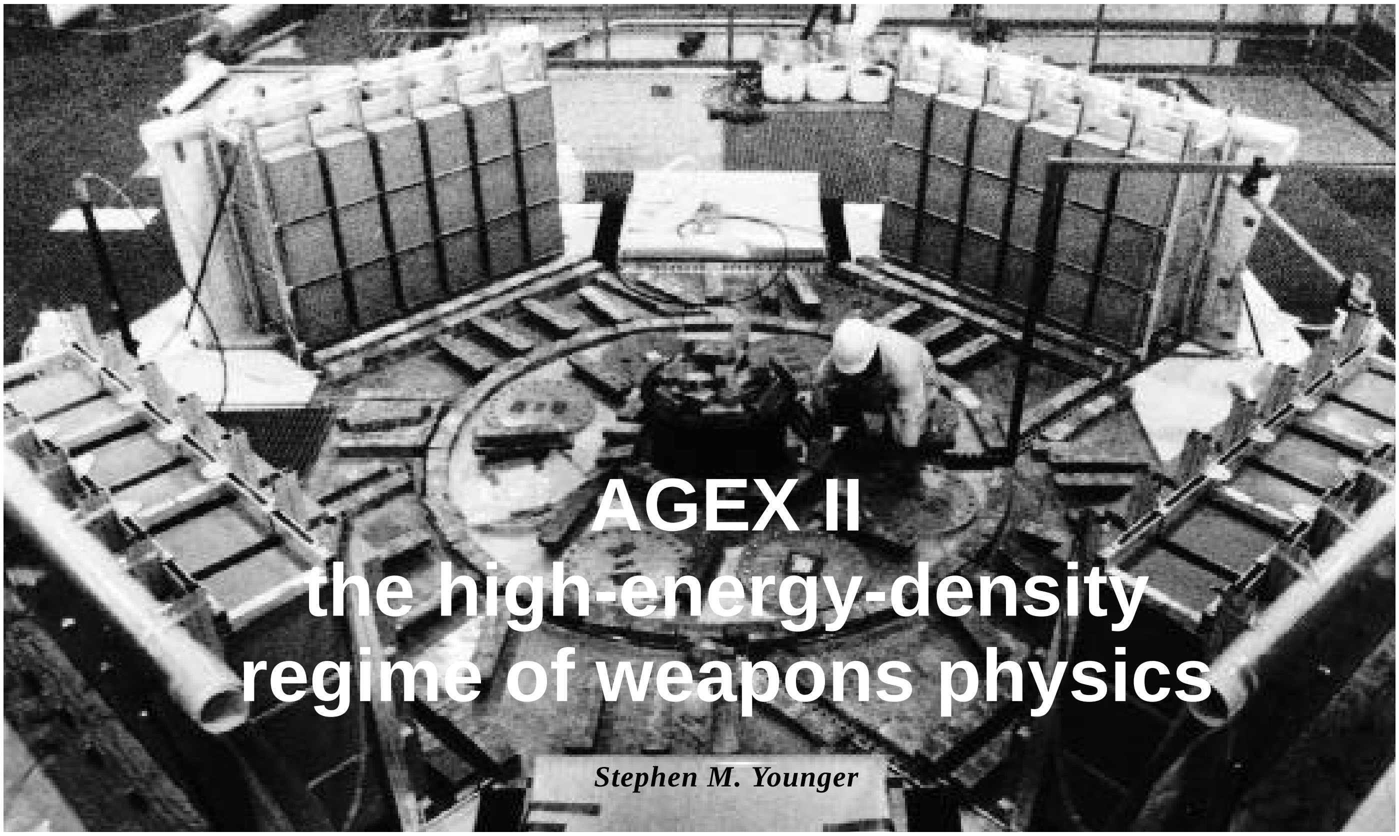


(c) This cross-sectional view of the penetrator is a computer-assisted tomographic (CAT) reconstruction of the interior of the penetrator made from a high-quality radiograph like the one shown in (a). The gray scale represents material density. The combination of good edge location and density reconstruction results from a high-quality original radiograph and excellent image-analysis capabilities. The knowledge of both edge location and density variation is critical to the interpretation of hydrodynamic experiments.

Timothy R. Neal has been Division Leader of Expolosives Technology and Appliations since 1991. He joined the Laboratory in 1967 as a staff member with the Flash Radiograph Group. In 1979 he served as Program Manager for the Confined Testing Program, and in 1980, he was Associate Division Leader for Dynamic Testing. From 1981 until February 1990 he served as Group Leader for Hydrodynamics, where he oversaw the consolidation of groups involving flash radiography, image analysis, and hydrodynamics. He served as Adjunct Associate Professor of Physics at New Mexico State University, instituted the continuing U.S./United Kingdom exchange in weapons hydrodynamics and the U.S./France exchange in image analysis, and was instrumental in developing the Dual Axis Radiographic Hydrodynamics Test (DARHT) construction project.

AGEX II
the high-energy-density regime of weapons physics

Stephen M. Younger



The Pegasus II capacitor bank for pulsed-power experiments

Nuclear explosives achieve higher temperatures and pressures than any other object in our solar system. Pressures in excess of ten million atmospheres, temperatures over 1000 electron volts (1 eV corresponds to 11,600 kelvins) and very high densities typify a nuclear explosion. Under these conditions even the heaviest atoms are almost completely ionized, and neutron radiation is so intense that higher-order nuclear processes (such as multple capture) become common. Our knowledge of such extreme energy-density conditions has been gained through a combination of theoretical calculations and experiments performed on actual nuclear explosions. With the reduction in the number of underground nuclear tests, however, our access to these unique conditions has been sharply reduced, and by the end of 1996 it will disappear entirely. There is an urgent need to develop laboratory techniques that will allow us to simulate the conditions found in a nuclear explosive both to provide more accurate information on the physics of matter at high energy density and to provide a vehicle for continued development of the special skills required to maintain an understanding of nuclear weapons.

The Physics of Nuclear Explosives

Nuclear weapons are very complex devices. During the high-explosive phase of the weapon, materials are subjected to pressures of several hundred kilobars and reach temperatures of several eV. These conditions are reproducible in the laboratory and a great deal of data are available to describe material response and hydrodynamic processes at these pressures. (See “AGEX I—The Explosives Regime of Weapons Physics.”) When the fissionable material in the weapon reaches a critical mass, however, a chain reaction occurs, which causes the rapid generation of energy. This chain reaction occurs on a time scale short compared to the ratio of the size of the device to the sound speed, so the material does not have a chance to expand during the energy-generation phase. Since the energy cannot go into kinetic energy, it goes into thermal energy, raising the temperature of the material to extraordinary values and thus raising the pressure to many millions of atmospheres.

Laboratory studies of the properties of high-energy-density matter face two major challenges: First, one must reproduce the very high densities and temperatures typical of a nuclear explosion. Second, one must be able to probe the conditions in the sample, usually via an x-ray burst (to probe atomic properties) or a pressure pulse (to probe material equations of state). The energy required to heat a sample is roughly given by $(3/2)nkT$, where n is the density of particles (nuclei plus ionized electrons), k is Boltzmann's constant, and T is the temperature. A simple calculation shows that normal-density uranium at 1 keV has an energy density of about 500 megajoules per cubic centimeter. Even for a sample 1 millimeter across the net energy required is 500 kilojoules, a substantial amount for laboratory experiments. Also, in contrast to fissioning metals, which generate heat internally, laboratory samples must be heated by an outside energy source. The heating takes several nanoseconds, long enough for the sample to begin to disassemble. The resulting density and temperature gradients complicate the interpretation of the experiment. Reducing the density allows one somewhat greater flexibility, since hydrodynamic tampers can be used to keep the material from expanding during the experiment. Unfortunately, low-density samples lack some of the unique aspects of dense plasma. The relevant figure of merit for dense matter is the coupling parameter, Γ, the ratio of the average electrostatic energy between neighboring ions to the average thermal kinetic energy. For low Γ thermal processes dominate and the plasma behaves as an ensemble of individual particles. For high Γ the electrostatic force dominates, the plasma becomes "stiff," and it can even condense into a solid phase. The goal of high-energy-density physics is to produce a sample dense enough to resemble a strongly coupled plasma yet hot enough for the level of ionization to be representative of the material in a nuclear explosive. This requires both raw energy, to heat a sample of significant size, and power, to rapidly heat the sample before it expands to low density.

Diagnosing a high-energy-density plasma is also challenging. No material probe can withstand the conditions of a hot dense plasma, so remote measurements are essential. X rays, either those emitted by the plasma itself or those absorbed when an intense probe signal is passed through the plasma, can reveal much about the atomic properties of the material. Strong shock waves can be launched into the sample to determine its equation of state via the measurement of shock and particle velocities. Such experiments are made more complex by the presence of the hydrodynamic tamper or other artifacts of the plasma-containment mechanism.

Examples of High-Energy-Density Physics

	The Sun	Jupiter	High explosives	Lasers	Pulsed power	Nuclear explosions
Temperature (eV)	10^3	1	~1	>100	100	$>10^3$
Pressure (atm)	10^9	10^6	10^5	10^8	10^7	$>10^7$
Density (g/cm^3)	10	1	1	100	10	>10

No single above-ground experimental facility can simultaneously reproduce all of the relevant conditions found in a nuclear explosion. At Los Alamos we have assembled a broad array of high-energy-density facilities, including pulsed-power machines, lasers, and the LAMPF accelerator, that allow us to access a broad range of high-energy-density conditions for the study of physics relevant to nuclear explosives.

Athena: Pulsed Power for High-Energy-Density Physics

In Greek mythology Athena was the goddess of wisdom who carried the thunderbolts of Zeus. At Los Alamos Athena is the program that uses pulsed-power technology to explore high-energy-density physics in support of the nuclear weapons program. The advantage of pulsed power for high-energy-density physics is that many megajoules of energy can be stored in very compact devices and then rapidly delivered to an experiment. The Athena program uses two methods to generate intense electrical pulses: a large capacitor bank called Pegasus II and a high-explosive pulsed-power generator called Procyon.

Capacitor-bank pulsed power. The Pegasus II capacitor bank consists of 144 capacitors wired in parallel and arranged around a central target chamber. Over the course of several minutes, a high-voltage power supply charges the capacitors. Pegasus

II can reach a maximum voltage of about 100 kilovolts and store up to 4.3 megajoules of electrical energy. At full voltage, the power supply is disconnected, and the stored electrical energy is rapidly discharged into a target located in a central chamber. Depending on the switching, the discharge time can be from 0.3 to 6 microseconds. At peak current, around 10 megamps, the power flow in Pegasus II exceeds that produced by the electrical generating capacity of the United States.

The target in Pegasus II experiments is typically a hollow metal cylinder, several centimeters in diameter and a few centimeters tall, oriented with its axis connnecting the two current-carrying electrodes of the capacitor bank. A current I flowing in the cylinder produces a magnetic field

$$B = \mu I/2\pi r$$

(where $\mu = 4\pi \times 10^7$). The interaction of this magnetic field with the drive current results in an inward pressure,

$$P = B^2/2\mu,$$

that causes the rapid implosion of the cylinder. When the Pegasus II capacitor bank at full charge is discharged through a gold-foil cylinder with outer radius of 2 centimeters and wall thickness of 0.1 millimeter, the cylinder implodes in about 6 microseconds. The peak pressure is 1.1 megabar and the peak implosion velocity is about 1.3 centimeters per microsecond.

Pegasus II can be used for three main classes of experiments. Simply discharging the bank through a cylindrical target provides a test bed for studying implosion hydrodynamics and material properties. To study the growth of hydrodynamic instabilites, asymmetric or otherwise deliberately perturbed cylinders are imploded and the results compared to numerical simulations that use various theoretical models for instability growth. The time sequence in Figure 1 shows a side view of a very thin-walled cylinder as it implodes during a Pegasus experiment. The onset of small-scale instabilities, their evolution to larger-scale perturbations, and finally the complete breakup of the imploding shell are clearly visible. We have done a number of experimental studies to assess how instability growth depends on wall thickness, material, and other implosion parameters.

In the second class of experiments, a sample is placed inside the imploding cylinder to study its material properties under extreme pressures and temperatures. An imploding cylinder can compress a sample to a pressure of several megabars. At this extreme pressure the atomic structure of the sample becomes distorted. The forced overlap of atomic orbitals causes changes in the transport rates of heat, charge, and photons. Serious research on those transport phenomena in hot dense matter is just beginning to be pursued. Since the implosions are nearly adiabatic and create high-pressure conditions that last on the order of a microsecond, Pegasus II is an ideal test bed in which to study the structure and dynamics of matter under extreme conditions.

In the third class of experiments, very fast implosions are used to generate intense x-ray bursts. These bursts are needed for a wide range of experiments on radiation transport and the interaction of x rays with matter. When the cylindrical foil implodes to its axis, the kinetic energy of collapse is converted to thermal energy, and the matter in the foil becomes a hot plasma. The matter then rapidly cools by emitting an intense burst of x rays. Kinetic energy is proportional to the square of the velocity, so implosions at higher velocities produce more x rays. Since the magnetic pressure during the implosion is independent of the mass of the foil, the implosion velocity can be increased by reducing the thickness of the foil. Cylindrical

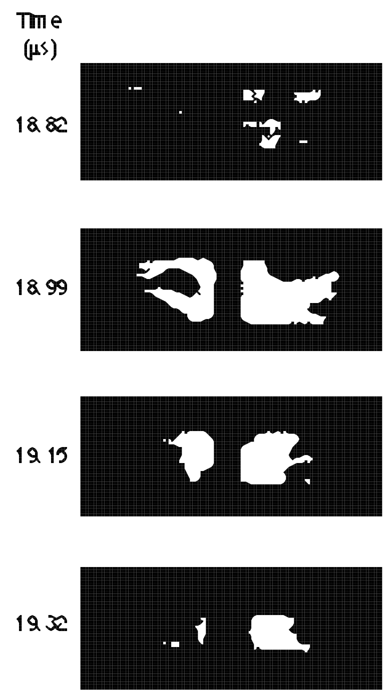


Figure 1. Implosion of a Thin Foil Cylinder at Pegasus II
A side view of the implosion of a 2500-angstrom-thick, 5-centimeter-radius, aluminum cylinder on the Pegasus capacitor bank is shown at various times, with an arbitrary zero time. Note the onset of short-wavelength perturbations at early times, followed by the growth of longer-wavelength perturbations, and eventually the complete breakup of the aluminum foil.

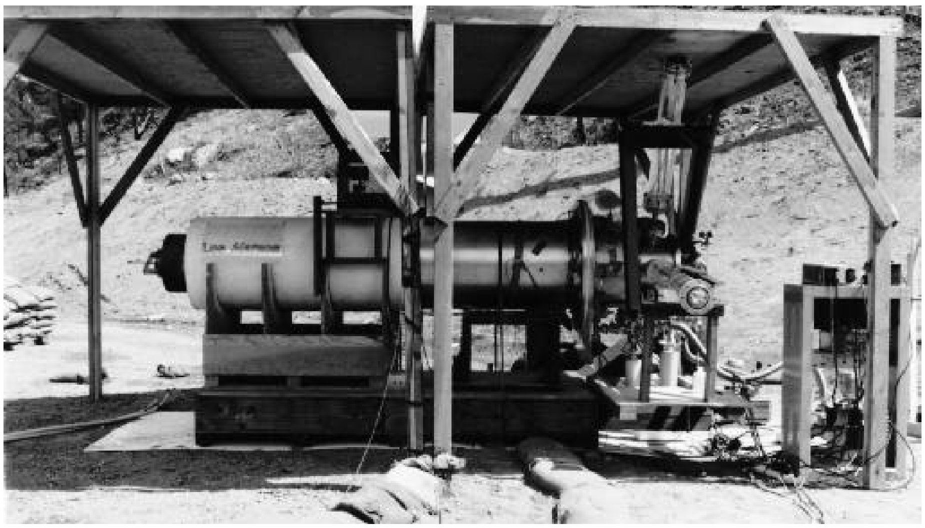


Figure 2. The Procyon High-Explosives Pulsed-Power System
The white cylinder at left is a Mark IX high-explosive generator. To the right is an explosive opening switch, and at the extreme right are the plasma flow switch and the implosion target. The clear tubes extending from the target contain diagnostic equipment.

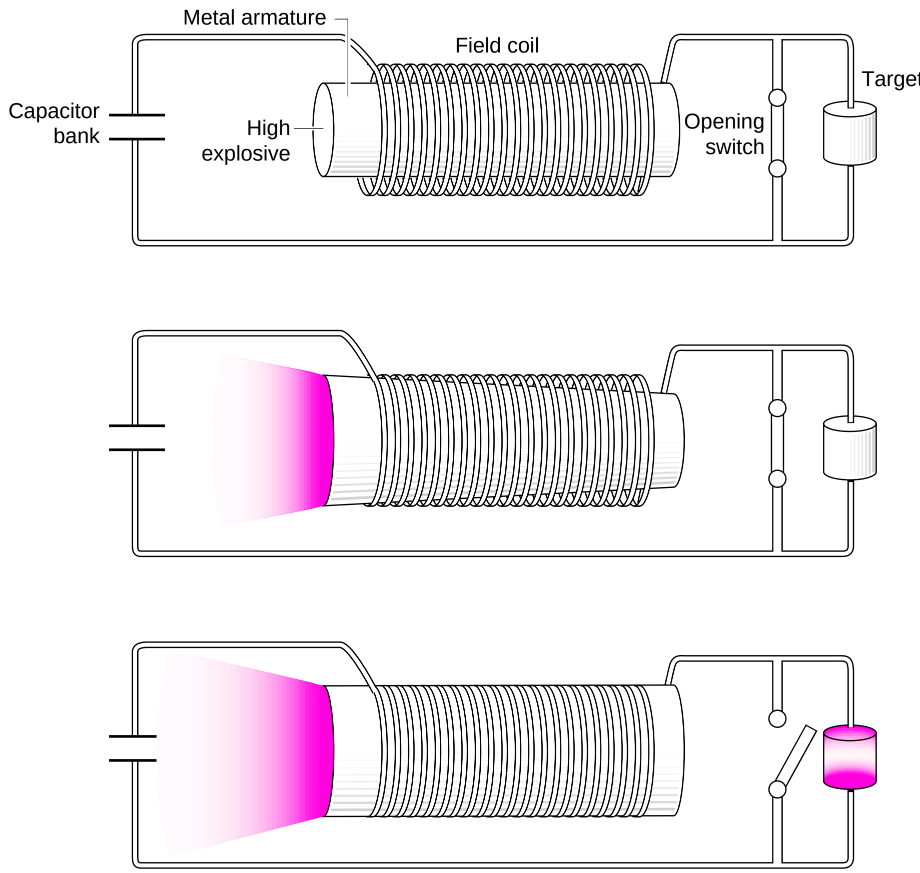
foils as thin as 2500 angstroms (less than the wavelength of visible light) have been imploded on Pegasus II to velocities in excess of 10 centimeters per microsecond. However, the onset of hydrodynamic instabilities limits the usefulness of very thin foils. The very rapid acceleration endured by these foils enhances the growth of Rayleigh-Taylor instabilities. The foils can become so unstable that they break up before reaching the axis, in which case various pieces of the foil arrive at different times. This prolongation of the arrival of the imploding shell results in a lower effective energy density in the plasma and hence a softer x-ray spectrum. Also, if the implosion velocity is so high that the implosion is over before the capacitor bank has the opportunity to deposit all of its energy in the target, then the resulting x-ray burst will be less intense. Two techniques are available to overcome these limitations. First, by using special switching techniques to reduce the rise time of the bank from several microseconds to several hundred nanoseconds, we can make energy deposition in the cylindrical foil more efficient. Second, by increasing the radius of the foils, we can extend the implosion time. We hope that an optimal combination of the two techniques will maximize the coupling of the energy in the capacitor bank to the kinetic energy of the foil while minimizing the growth of deleterious hydrodynamic instabilities.

High-explosive pulsed power. Pegasus II produces pressures of megabars and energy densities of hundreds of kilojoules per cubic centimeter in a volume of about a cubic centimeter. The production of significantly higher pressures in a similar-sized volume would require storing tens to over a hundred megajoules of electrical energy in a very large capacitor bank. This can be an expensive proposition. Fortunately there is a relatively low-cost alternative, namely, the amplification of electric power pulses with high explosives. At Los Alamos we have developed a series of high-explosive pulsed-power generators that have produced currents as high as 150 million amperes in compact, relatively low-cost units. The present device, called Procyon (Figure 2), can deliver more than 1 megajoule of energy into an implosion.

Figure 3 illustrates the operation of a high-explosive pulsed-power generator. A small capacitor bank sends a current pulse through a coil wound loosely about a copper cylinder that is filled with high explosives. This current creates a magnetic field in the gap between the coil and the copper cylinder. As the magnetic field reaches its peak value, the high explosive in the copper cylinder is detonated. The cylinder expands, and as it closes the gap between itself and the coil, it squeezes the magnetic field into a smaller and smaller volume and thereby increases the magnetic-field energy. At maximum field compression the switch shown in the figure is opened, allowing the field energy to be extracted in the form of a greatly amplified current pulse that flows through the target.

Even though high-explosive pulsed-power generators can produce tens or even hundreds of megajoules of electrical energy in a single pulse at affordable prices, the pulse is several microseconds long, so the power eventually delivered to a target is at most a few tens of terawatts (1 terawatt = 10^{12} watts). At present it is difficult to compress this energy into a much shorter, higher-power pulse, which would be useful for the production of intense x-ray bursts or ultrahigh pressures,

Figure 2. Operation of a High-Explosive Pulsed-Power System



As the capacitor bank discharges through the field coil, the current in the coil generates a magnetic field between the coil and the metal armature. The opening switch is in the closed position, preventing current from flowing to the target.

The high explosive is detonated at one end and expands the armature. The magnetic field is squeezed between the expanding armature and the field coil and greatly increases in magnitude. In this way high-explosive energy is converted into magnetic-field energy. The field energy, in turn, amplifies the current.

At peak compression of the magnetic field the switch is opened and a greatly amplified current pulse flows through the target.

although a variety of options are being evaluated. To reach significantly higher powers, one must employ other technologies more suited to short-pulse generation. Chief among those are high-power lasers.

Lasers: Higher Energy Densities for Shorter Times

Lasers can produce very short, high-power pulses and direct them into small volumes to create very high energy densities. But the current high-power lasers are very expensive energy sources and maintain high energy densities in a target for only a nanosecond or so and over volumes only of the order of a cubic millimeter. In spite of these limitations, high-power lasers have proven to be very versatile in the study of high-energy-density physics.

Trident is a neodymium-doped glass laser at Los Alamos that delivers two simultaneous pulses, each 100 picoseconds long and carrying 100 joules of energy. The laser consists of a very low-energy oscillator that forms the laser pulse, a series of rod amplifiers that increase the energy in the pulse to about 1 joule, and a set of disk amplifiers that provide the final amplification. The pulses are directed into a target chamber outfitted with a wide array of diagnostics, including x-ray and optical spectrometers, framing cameras, and streak cameras.

A Trident 100-picosecond laser pulse focused into a volume a few hundred microns in diameter yields an energy density of over 1 megajoule per cubic centimeter. Although the energy density is higher than that produced in experiments using Pegasus II, our 4-megajoule capacitor bank, the temporal and spatial scales of the experiments are much smaller and very sophisticated diagnostics are required to acquire data. The inertial-fusion program has made impressive progress in diagnostics development, so that it is now possible to obtain x-ray images of experiments with spatial resolution of less than 5 microns and temporal resolution of less than 100 picoseconds.

Trident was designed to be an easy-to-use tool for high-energy-density physics. It can deliver laser pulses with a wide variety of lengths and shapes for different experiments. Trident also has a small third laser beam, which is used to create a short x-ray pulse next to the target. X radiographs of evolving experiments can be obtained from the x-ray pulses and are particularly useful for the study of high-pressure hydrodynamics. Trident pulses, when applied to appropriate targets, can produce shock-wave pressures of several megabars and x-ray pulses of moderate temperatures.

Still higher temperatures over somewhat larger volumes can be obtained on the Nova laser at Lawrence Livermore National Laboratory. Nova, the largest glass laser in the world, produces pulses of up to 40 kilojoules in one nanosecond. We have fielded a number of experiments on Nova related to radiation hydrodynamics and x-ray-driven implosions.

How far can one go in increasing energy density by shortening the pulse length of the laser and reducing the size of the focused optical spot? Another laser at Los Alamos, Bright Source II, is providing the answer. Bright Source II pushes the limits of energy density by directing a relatively small amount of energy (only a quarter of a joule at present, although a 10-joule machine is on the horizon) into an incredibly short pulse that can be focused down to only a few microns. Bright Source II pulses last less than 300 femtoseconds, so that even though it is moving with the speed of light, a pulse is only about 900 microns long. The focused pulses have intensities of more than 5×10^{18} watts per square centimeter, well above pulse intensities produced by Trident or even Nova. The impact of a laser pulse on the surface of a target sample creates pressures of more than 1 gigabar, but only for about one picosecond, after which the sample expands under thermal pressure. (It is interesting to note that the radiation pressure—the pressure due to the momentum of the light itself—is 1 gigabar, which is comparable to the induced thermal pressure in the target.) During such a short pulse the atoms in the target do not have a chance to equilibrate and may not approximate fully the equilibrium conditions found in a nuclear explosion. Nevertheless, Bright Source II can heat thin solid foils to keV temperatures, creating a high-density and very hot plasma. Hence this laser can be used to probe the structure and dynamics of matter at conditions that approach those found in a nuclear explosion. The hot plasma cools both by expansion and by the emission of x rays. Figure 4 shows a typical x-ray spectrum from an aluminum sample illuminated by a Bright Source II pulse. Up to 1 percent of the incident laser energy is converted to line radiation around 2 keV. The line radiation is useful for studying the interaction of x rays with matter.

The extremely short pulses available from the Bright Source II laser also provide an effective means to study very rapid processes, such as transient chemical reactions. In typical chemical detonations several transient molecular species such as OH radicals persist only for a short time but are important in determining the overall energy balance in the detonation products. An experiment is currently underway at Bright Source II to measure the OH radical in a forced detonation—the first such measurement of its kind for an explosive process.

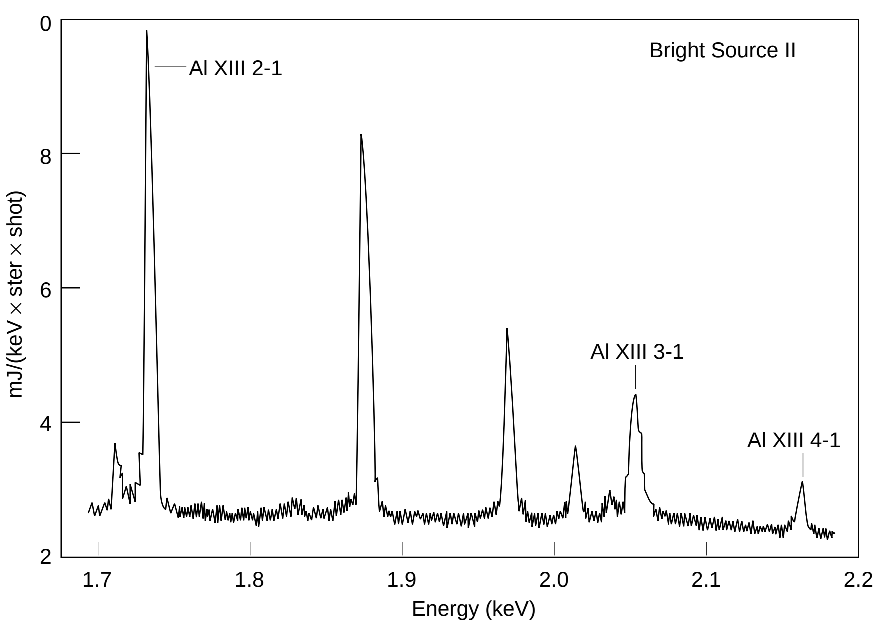


Figure 4. X-Ray Spectrum Induced by a Pulse from Bright Source II
The x-ray spectrum results from the impact of a 0.25-joule, 300-picosecond pulse from the Bright Source II laser on an aluminum foil. The intense lines serve both as a diagnostic of the conditions in the dense radiating plasma and as valuable probes for use in other experiments.

Nuclear Physics at LAMPF

Moving up again in the energy scale, we encounter nuclear energy densities—where the relevant energy parameters are not kiloelectronvolts as in plasmas but megaelectronvolts.

The formation of a critical mass during the detonation of a nuclear explosive and the attendant chain reaction result in an intense neutron burst. Neutrons interact with nuclei through a complex set of scattering and capture processes, some leading to the production of additional neutrons and/or the initiation of fission and others to the production of stable isotopes. To model the dynamics of fission in weapons, we must have accurate descriptions of all of the dominant neutronics processes. The knowledge of weaker processes (such as those involving transient nuclear states) can provide valuable diagnostics on the progress of the nuclear burn and contribute to radiochemical analyses of nuclear explosions. The Laboratory has conducted an extensive series of experiments on nuclear physics important for weapons at LAMPF and other nuclear facilities.

LAMPF is the most powerful accelerator in the world. Although some machines accelerate charged particles to higher energies, none is capable of delivering as many particles per unit time to the target as LAMPF. This capability is important when one wants to study weak processes, including the study of higher-order nuclear cross sections. In addition to the accelerator itself, the LAMPF facility includes several target areas. The areas of particular concern for the weapons program are the Weapons Neutron Research (WNR) facility and the Manuel Lujan, Jr. Neutron Scattering Center (LANSCE).

LAMPF has been extensively used by the weapons program. Fundamental aspects of fission have been studied by examining the relative timing of fission and neutron emission in fissioning nuclei. Angular distributions of neutrons, gamma rays, and fission decay products have been measured to determine the energy and momentum balance in fission. Detectors used in nuclear tests have been calibrated on LAMPF to provide absolute measurements of the neutron flux from the nuclear device. We are currently evaluating new techniques for using proton and neutron sources to image dynamic phenomena in opaque samples.

The Future

The next several years promise to be among the most interesting and productive ever for high-energy-density physics. We have assembled an array of facilities to investigate a wide range of physics issues of importance to the nuclear weapons program. The structure and transport properties of hot dense matter will be systematically studied in a regime where single-atom theories break down and many-body effects are important. The interaction of strong shock waves and x-ray pulses with matter will continue to be studied with the aim of providing quantitative data for use in our computer models of nuclear explosions. Experimental data on hydrodynamics and hydrodynamic instabilities will allow us to validate increasingly sophisticated algorithms in new computer codes, particularly those that will need to be developed to exploit the promise of massively parallel computers.

Each of our capabilities can be extended to higher energies for even more interesting applications. The next advance in Laboratory capacitor banks is Atlas, a 25-megajoule machine that will permit us to study high energy densities over tens of cubic centimeters. The Procyon high-explosive pulsed-power generator will be followed by a more advanced system that will deliver in excess of 200 million amperes. Bright Source III is being designed to produce focused intensities over 10^{20} watts per square centimeter to permit the study of multiphoton x-ray interactions. This intensity is high enough to rip apart the vacuum in the electrostatic field near a nucleus to create electron-positron pairs, literally creating matter from energy. ■

Stephen M. Younger is the Program Director for Inertial Confinement Fusion and High Energy Density Physics at the Laboratory. He received a Ph.D. in theoretical physics from the University of Maryland in 1978. From 1974 to 1982 he performed theoretical studies of atomic processes at the National Bureau of Standards in Washington, D.C., concentrating on electron scattering from atoms and ions. From 1982 to 1989 he was a staff member and group leader at Lawrence Livermore National Laboratory, where he specialized in advanced thermonuclear weapons design, x-ray lasers, electromagnetic weapons, and other defense programs. In 1989 he came to Los Alamos and in 1991 he was named to his present position. In addition to working on defense issues, Younger has continued to contribute to fundamental atomic and plasma physics. He is a Fellow of the American Physical Society and has served on numerous government panels and committees.

Line-imaging Laser Interferometers for Measuring Velocities

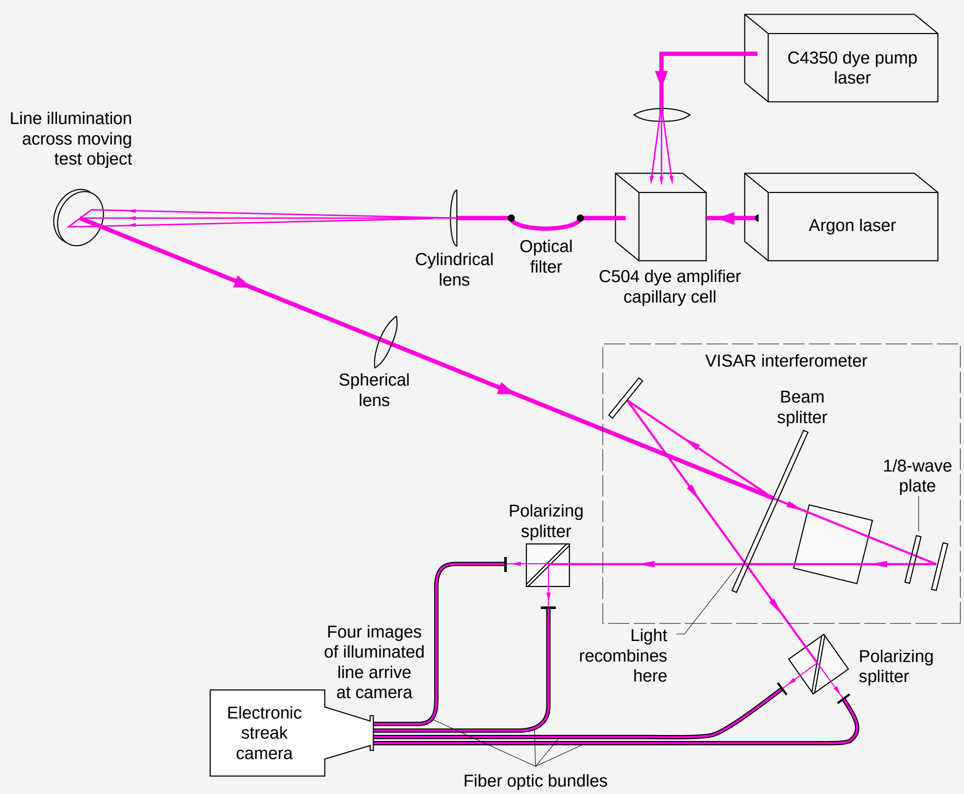
Willard F. Hemsing

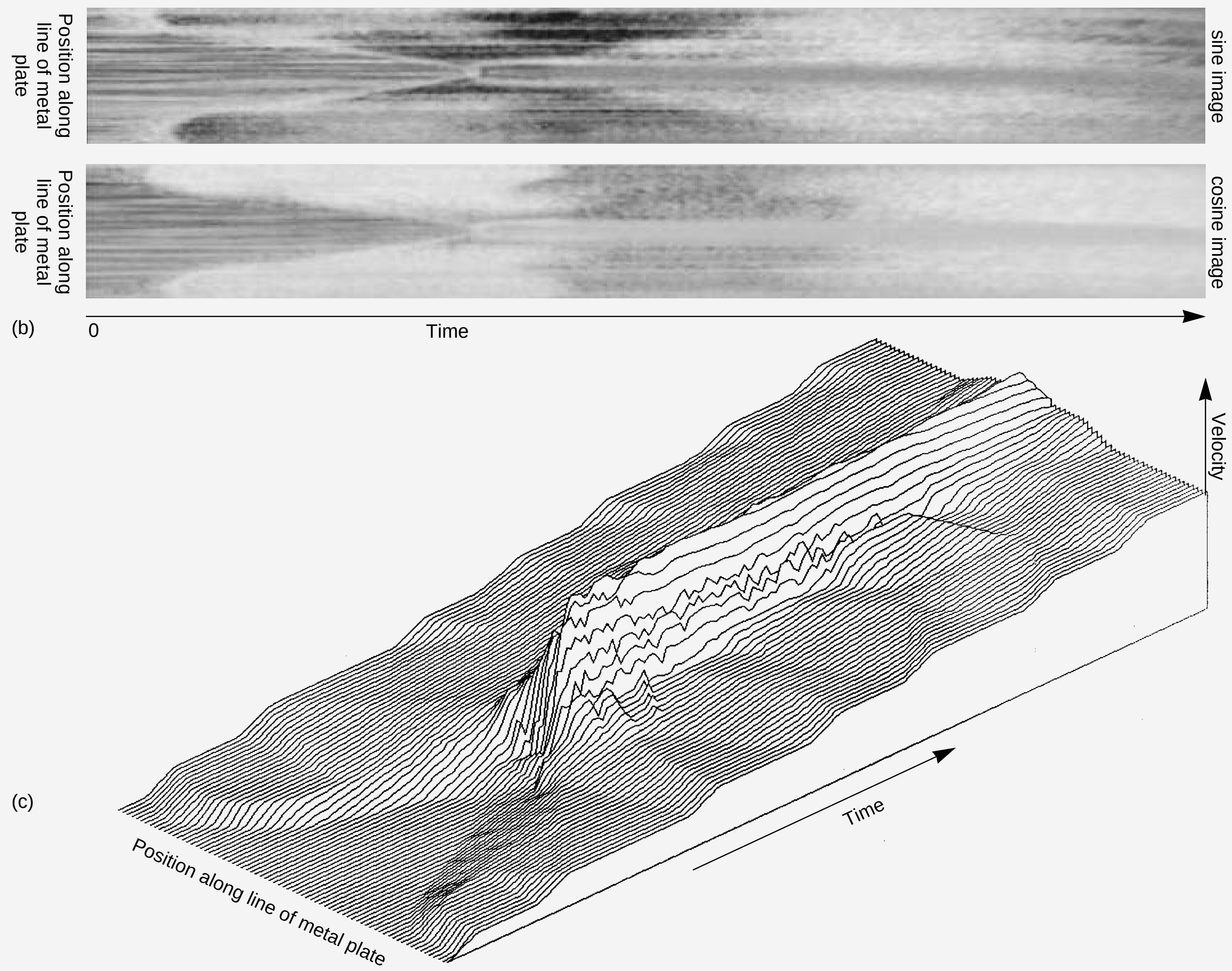
Hydrodynamic tests create hostile conditions in which high pressures can easily compress solids and accelerate materials to velocities of several kilometers per second. Among the advanced diagnostics for hydrodynamic tests at the Laboratory is our line-imaging VISAR (Velocity Interferometer System for Any Reflector). The VISAR measures the velocities of points along an illuminated line on a fast-moving test object. The instrument exploits the fact that when laser light is reflected from a moving surface, the wavelength of the light is Doppler-shifted in proportion to the velocity of the point that reflects it. The VISAR employs optical interference to generate bright and dark bands of light called interference fringes. The fringes oscillate between bright and dark as the test object accelerates. The VISAR measures velocity by accurately determining the number of whole and partial oscillations that occur as the test object accelerates. Its useful product is a continuous velocity history for all the points that are visible in the image.

(a) Our line-imaging VISAR uses a cylindrical lens to focus laser light onto a line on the test object. Conventional optics image the illuminated line through a special wide-angle Michelson interferometer, where a retardation plate delays the vertical polarization component of one beam by a quarter of a wavelength. As a result, when the beams are recombined to produce interference, the fringes of the vertical polarization component are shifted and their oscillations lag behind those of the horizontal component. Specifically, the intensities of corresponding points in the horizontal and vertical components depend on the sine and cosine, respectively, of the velocity at each point on the target. Polarizing beam splitters separate the horizontal polarization component from the vertical component where light exits from each side of the interferometer. This separation produces two pairs of images of the interference intensities along the illuminated line. The two images for each polarization are simply negatives of each other.

Fiber-optic bundles transmit the four images to the photocathode of an electronic streak camera. The

(a)



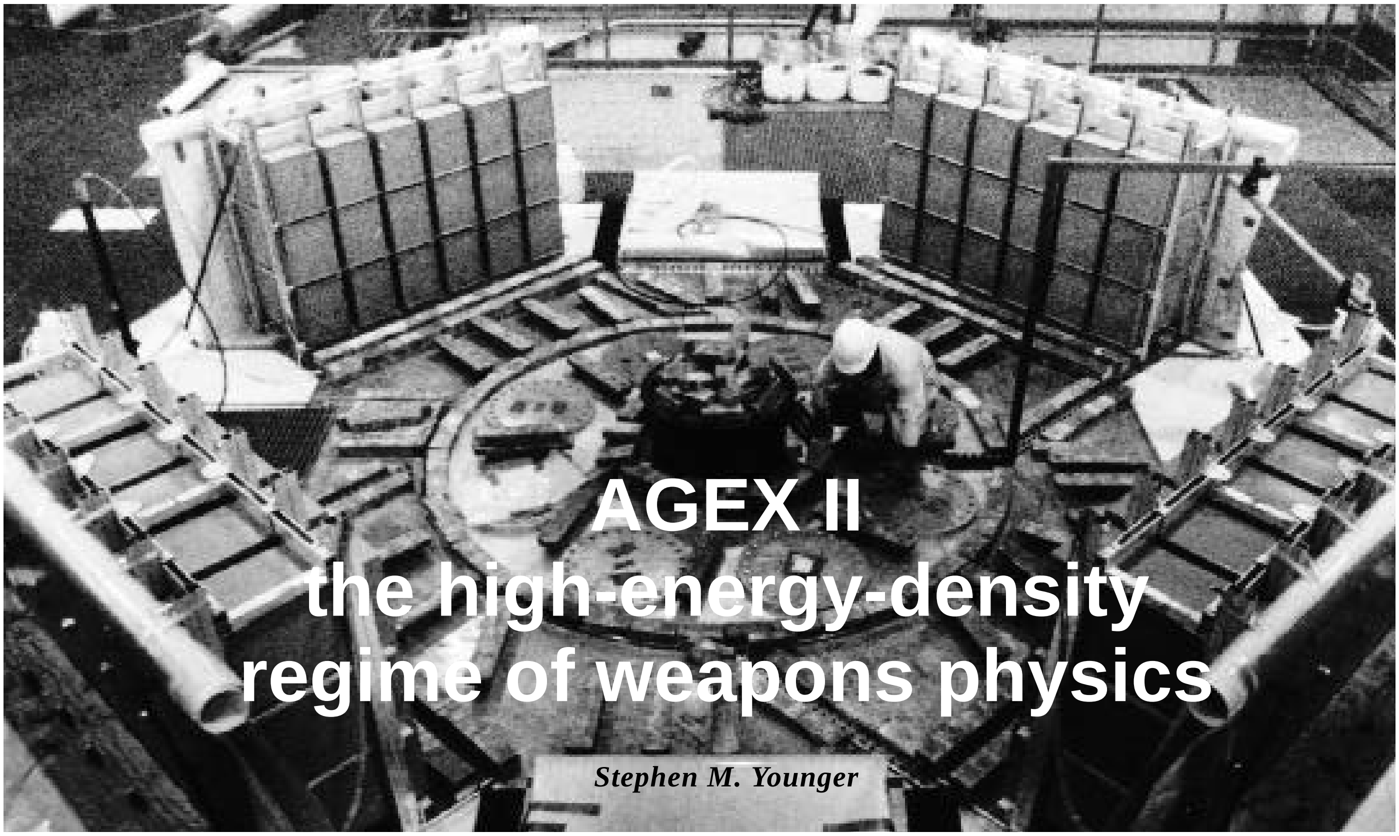


camera rapidly sweeps the images across a charge-coupled device that digitizes them into a microcomputer. Later, we subtract one image of each polarization from its negative to double the signal and cancel optical noise. Analysis of the images yields the velocity histories of many points in the line as a continuous function of time.

The VISAR's sensitivity to acceleration, instead of to velocity alone, best accommodates measurements of velocities from 100 meters per second to over 20 kilometers per second. Its recording time can vary from milliseconds to nanoseconds; the length of the line it observes can range from 0.3 to 30 millimeters across the target surface. Because it records pictures with their great capacity to store information, our line-imaging VISAR can capture many times more data than conventional VISARs. We have found its ability to simultaneously record large quantities of information relating different points on a test object extremely advantageous. This is most useful in measurements in which velocity gradients are important, and in tests that destroy expensive hardware, especially when test-to-test variations are important. Although our line-imaging VISAR is versatile, its use is precluded when smoke blocks its optical path or when the test-object surface loses light reflectivity.

(b) The sine and cosine interference images from an experiment in which two converging detonation waves, produced by an explosive initiated at two separate points, drove a metal plate. Triangles extending across the left third of the images are the edges of interference fringes as they responded to the acceleration of the plate. A change from dark to bright, corresponding to an increase in velocity of 200 meters per second, is visible in the cosine image.

(c) An isometric plot of velocity, deduced from the photograph in (b), as a function of position along the illuminated line and time. The "cliffs" at the lower left indicate the acceleration of the metal as it was driven by the two converging pressure waves. The ridge extending from the center to the upper right is a region of high velocity caused by the pressure enhancement where the waves collided. □



AGEX II the high-energy-density regime of weapons physics

Stephen M. Younger

The Pegasus II capacitor bank for pulsed-power experiments

Nuclear explosives achieve higher temperatures and pressures than any other object in our solar system. Pressures in excess of ten million atmospheres, temperatures over 1000 electron volts (1 eV corresponds to 11,600 kelvins) and very high densities typify a nuclear explosion. Under these conditions even the heaviest atoms are almost completely ionized, and neutron radiation is so intense that higher-order nuclear processes (such as multple capture) become common. Our knowledge of such extreme energy-density conditions has been gained through a combination of theoretical calculations and experiments performed on actual nuclear explosions. With the reduction in the number of underground nuclear tests, however, our access to these unique conditions has been sharply reduced, and by the end of 1996 it will disappear entirely. There is an urgent need to develop laboratory techniques that will allow us to simulate the conditions found in a nuclear explosive both to provide more accurate information on the physics of matter at high energy density and to provide a vehicle for continued development of the special skills required to maintain an understanding of nuclear weapons.

The Physics of Nuclear Explosives

Nuclear weapons are very complex devices. During the high-explosive phase of the weapon, materials are subjected to pressures of several hundred kilobars and reach temperatures of several eV. These conditions are reproducible in the laboratory and a great deal of data are available to describe material response and hydrodynamic processes at these pressures. (See "AGEX I—The Explosives Regime of Weapons Physics.") When the fissionable material in the weapon reaches a critical mass, however, a chain reaction occurs, which causes the rapid generation of energy. This chain reaction occurs on a time scale short compared to the ratio of the size of the device to the sound speed, so the material does not have a chance to expand during the energy-generation phase. Since the energy cannot go into kinetic energy, it goes into thermal energy, raising the temperature of the material to extraordinary values and thus raising the pressure to many millions of atmospheres.

Laboratory studies of the properties of high-energy-density matter face two major challenges: First, one must reproduce the very high densities and temperatures typical of a nuclear explosion. Second, one must be able to probe the conditions in the sample, usually via an x-ray burst (to probe atomic properties) or a pressure pulse (to probe material equations of state). The energy required to heat a sample is roughly given by $(3/2)nkT$, where n is the density of particles (nuclei plus ionized electrons), k is Boltzmann's constant, and T is the temperature. A simple calculation shows that normal-density uranium at 1 keV has an energy density of about 500 megajoules per cubic centimeter. Even for a sample 1 millimeter across the net energy required is 500 kilojoules, a substantial amount for laboratory experiments. Also, in contrast to fissioning metals, which generate heat internally, laboratory samples must be heated by an outside energy source. The heating takes several nanoseconds, long enough for the sample to begin to disassemble. The resulting density and temperature gradients complicate the interpretation of the experiment. Reducing the density allows one somewhat greater flexibility, since hydrodynamic tampers can be used to keep the material from expanding during the experiment. Unfortunately, low-density samples lack some of the unique aspects of dense plasma. The relevant figure of merit for dense matter is the coupling parameter, Γ, the ratio of the average electrostatic energy between neighboring ions to the average thermal kinetic energy. For low Γ thermal processes dominate and the plasma behaves as an ensemble of individual particles. For high Γ the electrostatic force dominates, the plasma becomes "stiff," and it can even condense into a solid phase. The goal of high-energy-density physics is to produce a sample dense enough to resemble a strongly coupled plasma yet hot enough for the level of ionization to be representative of the material in a nuclear explosive. This requires both raw energy, to heat a sample of significant size, and power, to rapidly heat the sample before it expands to low density.

Examples of High-Energy-Density Physics

	The Sun	Jupiter	High explosives	Lasers	Pulsed power	Nuclear explosions
Temperature (eV)	10^3	1	~1	>100	100	$>10^3$
Pressure (atm)	10^9	10^6	10^5	10^8	10^7	$>10^7$
Density (g/cm^3)	10	1	1	100	10	>10

Diagnosing a high-energy-density plasma is also challenging. No material probe can withstand the conditions of a hot dense plasma, so remote measurements are essential. X rays, either those emitted by the plasma itself or those absorbed when an intense probe signal is passed through the plasma, can reveal much about the atomic properties of the material. Strong shock waves can be launched into the sample to determine its equation of state via the measurement of shock and particle velocities. Such experiments are made more complex by the presence of the hydrodynamic tamper or other artifacts of the plasma-containment mechanism.

No single above-ground experimental facility can simultaneously reproduce all of the relevant conditions found in a nuclear explosion. At Los Alamos we have assembled a broad array of high-energy-density facilities, including pulsed-power machines, lasers, and the LAMPF accelerator, that allow us to access a broad range of high-energy-density conditions for the study of physics relevant to nuclear explosives.

Athena: Pulsed Power for High-Energy-Density Physics

In Greek mythology Athena was the goddess of wisdom who carried the thunderbolts of Zeus. At Los Alamos Athena is the program that uses pulsed-power technology to explore high-energy-density physics in support of the nuclear weapons program. The advantage of pulsed power for high-energy-density physics is that many megajoules of energy can be stored in very compact devices and then rapidly delivered to an experiment. The Athena program uses two methods to generate intense electrical pulses: a large capacitor bank called Pegasus II and a high-explosive pulsed-power generator called Procyon.

Capacitor-bank pulsed power. The Pegasus II capacitor bank consists of 144 capacitors wired in parallel and arranged around a central target chamber. Over the course of several minutes, a high-voltage power supply charges the capacitors. Pegasus

II can reach a maximum voltage of about 100 kilovolts and store up to 4.3 megajoules of electrical energy. At full voltage, the power supply is disconnected, and the stored electrical energy is rapidly discharged into a target located in a central chamber. Depending on the switching, the discharge time can be from 0.3 to 6 microseconds. At peak current, around 10 megamps, the power flow in Pegasus II exceeds that produced by the electrical generating capacity of the United States.

The target in Pegasus II experiments is typically a hollow metal cylinder, several centimeters in diameter and a few centimeters tall, oriented with its axis connnecting the two current-carrying electrodes of the capacitor bank. A current *I* flowing in the cylinder produces a magnetic field

$$B = \mu I/2\pi r$$

(where $\mu = 4\pi \times 10^7$). The interaction of this magnetic field with the drive current results in an inward pressure,

$$P = B^2/2\mu,$$

that causes the rapid implosion of the cylinder. When the Pegasus II capacitor bank at full charge is discharged through a gold-foil cylinder with outer radius of 2 centimeters and wall thickness of 0.1 millimeter, the cylinder implodes in about 6 microseconds. The peak pressure is 1.1 megabar and the peak implosion velocity is about 1.3 centimeters per microsecond.

Pegasus II can be used for three main classes of experiments. Simply discharging the bank through a cylindrical target provides a test bed for studying implosion hydrodynamics and material properties. To study the growth of hydrodynamic instabilites, asymmetric or otherwise deliberately perturbed cylinders are imploded and the results compared to numerical simulations that use various theoretical models for instability growth. The time sequence in Figure 1 shows a side view of a very thin-walled cylinder as it implodes during a Pegasus experiment. The onset of small-scale instabilities, their evolution to larger-scale perturbations, and finally the complete breakup of the imploding shell are clearly visible. We have done a number of experimental studies to assess how instability growth depends on wall thickness, material, and other implosion parameters.

In the second class of experiments, a sample is placed inside the imploding cylinder to study its material properties under extreme pressures and temperatures. An imploding cylinder can compress a sample to a pressure of several megabars. At this extreme pressure the atomic structure of the sample becomes distorted. The forced overlap of atomic orbitals causes changes in the transport rates of heat, charge, and photons. Serious research on those transport phenomena in hot dense matter is just beginning to be pursued. Since the implosions are nearly adiabatic and create high-pressure conditions that last on the order of a microsecond, Pegasus II is an ideal test bed in which to study the structure and dynamics of matter under extreme conditions.

In the third class of experiments, very fast implosions are used to generate intense x-ray bursts. These bursts are needed for a wide range of experiments on radiation transport and the interaction of x rays with matter. When the cylindrical foil implodes to

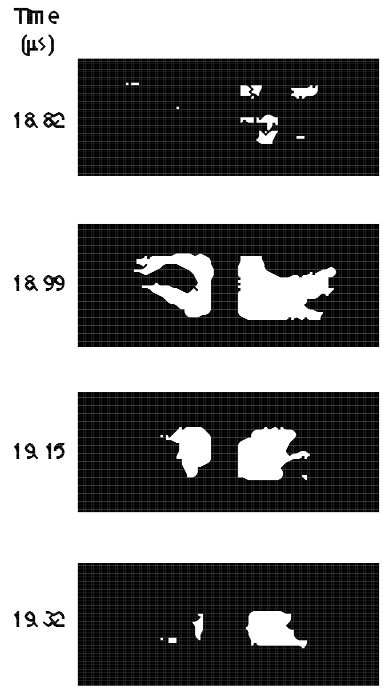


Figure 1. Implosion of a Thin Foil Cylinder at Pegasus II
A side view of the implosion of a 2500-angstrom-thick, 5-centimeter-radius, aluminum cylinder on the Pegasus capacitor bank is shown at various times, with an arbitrary zero time. Note the onset of short-wavelength perturbations at early times, followed by the growth of longer-wavelength perturbations, and eventually the complete breakup of the aluminum foil.

its axis, the kinetic energy of collapse is converted to thermal energy, and the matter in the foil becomes a hot plasma. The matter then rapidly cools by emitting an intense burst of x rays. Kinetic energy is proportional to the square of the velocity, so implosions at higher velocities produce more x rays. Since the magnetic pressure during the implosion is independent of the mass of the foil, the implosion velocity can be increased by reducing the thickness of the foil. Cylindrical

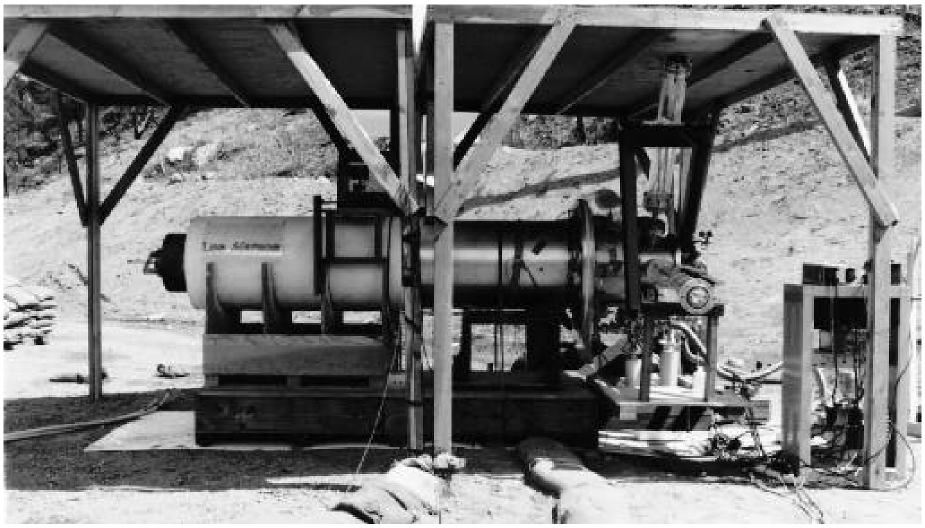


Figure 2. The Procyon High-Explosives Pulsed-Power System
The white cylinder at left is a Mark IX high-explosive generator. To the right is an explosive opening switch, and at the extreme right are the plasma flow switch and the implosion target. The clear tubes extending from the target contain diagnostic equipment.

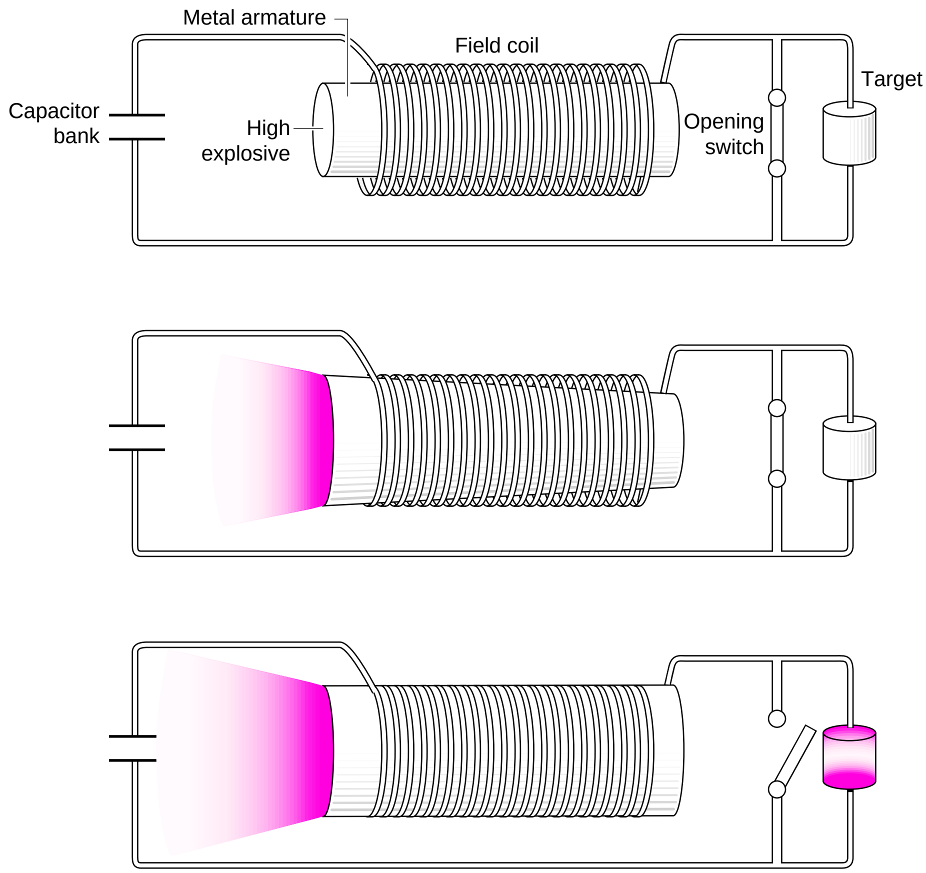
foils as thin as 2500 angstroms (less than the wavelength of visible light) have been imploded on Pegasus II to velocities in excess of 10 centimeters per microsecond. However, the onset of hydrodynamic instabilities limits the usefulness of very thin foils. The very rapid acceleration endured by these foils enhances the growth of Rayleigh-Taylor instabilities. The foils can become so unstable that they break up before reaching the axis, in which case various pieces of the foil arrive at different times. This prolongation of the arrival of the imploding shell results in a lower effective energy density in the plasma and hence a softer x-ray spectrum. Also, if the implosion velocity is so high that the implosion is over before the capacitor bank has the opportunity to deposit all of its energy in the target, then the resulting x-ray burst will be less intense. Two techniques are available to overcome these limitations. First, by using special switching techniques to reduce the rise time of the bank from several microseconds to several hundred nanoseconds, we can make energy deposition in the cylindrical foil more efficient. Second, by increasing the radius of the foils, we can extend the implosion time. We hope that an optimal combination of the two techniques will maximize the coupling of the energy in the capacitor bank to the kinetic energy of the foil while minimizing the growth of deleterious hydrodynamic instabilities.

High-explosive pulsed power. Pegasus II produces pressures of megabars and energy densities of hundreds of kilojoules per cubic centimeter in a volume of about a cubic centimeter. The production of significantly higher pressures in a similar-sized volume would require storing tens to over a hundred megajoules of electrical energy in a very large capacitor bank. This can be an expensive proposition. Fortunately there is a relatively low-cost alternative, namely, the amplification of electric power pulses with high explosives. At Los Alamos we have developed a series of high-explosive pulsed-power generators that have produced currents as high as 150 million amperes in compact, relatively low-cost units. The present device, called Procyon (Figure 2), can deliver more than 1 megajoule of energy into an implosion.

Figure 3 illustrates the operation of a high-explosive pulsed-power generator. A small capacitor bank sends a current pulse through a coil wound loosely about a copper cylinder that is filled with high explosives. This current creates a magnetic field in the gap between the coil and the copper cylinder. As the magnetic field reaches its peak value, the high explosive in the copper cylinder is detonated. The cylinder expands, and as it closes the gap between itself and the coil, it squeezes the magnetic field into a smaller and smaller volume and thereby increases the magnetic-field energy. At maximum field compression the switch shown in the figure is opened, allowing the field energy to be extracted in the form of a greatly amplified current pulse that flows through the target.

Even though high-explosive pulsed-power generators can produce tens or even hundreds of megajoules of electrical energy in a single pulse at affordable prices, the pulse is several microseconds long, so the power eventually delivered to a target is at most a few tens of terawatts (1 terawatt = 10^{12} watts). At present it is difficult to compress this energy into a much shorter, higher-power pulse, which would be useful for the production of intense x-ray bursts or ultrahigh pressures,

Figure 2. Operation of a High-Explosive Pulsed-Power System



As the capacitor bank discharges through the field coil, the current in the coil generates a magnetic field between the coil and the metal armature. The opening switch is in the closed position, preventing current from flowing to the target.

The high explosive is detonated at one end and expands the armature. The magnetic field is squeezed between the expanding armature and the field coil and greatly increases in magnitude. In this way high-explosive energy is converted into magnetic-field energy. The field energy, in turn, amplifies the current.

At peak compression of the magnetic field the switch is opened and a greatly amplified current pulse flows through the target.

although a variety of options are being evaluated. To reach significantly higher powers, one must employ other technologies more suited to short-pulse generation. Chief among those are high-power lasers.

Lasers: Higher Energy Densities for Shorter Times

Lasers can produce very short, high-power pulses and direct them into small volumes to create very high energy densities. But the current high-power lasers are very expensive energy sources and maintain high energy densities in a target for only a nanosecond or so and over volumes only of the order of a cubic millimeter. In spite of these limitations, high-power lasers have proven to be very versatile in the study of high-energy-density physics.

Trident is a neodymium-doped glass laser at Los Alamos that delivers two simultaneous pulses, each 100 picoseconds long and carrying 100 joules of energy. The laser consists of a very low-energy oscillator that forms the laser pulse, a series of rod amplifiers that increase the energy in the pulse to about 1 joule, and a set of disk amplifiers that provide the final amplification. The pulses are directed into a target chamber outfitted with a wide array of diagnostics, including x-ray and optical spectrometers, framing cameras, and streak cameras.

A Trident 100-picosecond laser pulse focused into a volume a few hundred microns in diameter yields an energy density of over 1 megajoule per cubic centimeter. Although the energy density is higher than that produced in experiments using Pegasus II, our 4-megajoule capacitor bank, the temporal and spatial scales of the experiments are much smaller and very sophisticated diagnostics are required to acquire data. The inertial-fusion program has made impressive progress in diagnostics development, so that it is now possible to obtain x-ray images of experiments with spatial resolution of less than 5 microns and temporal resolution of less than 100 picoseconds.

Trident was designed to be an easy-to-use tool for high-energy-density physics. It can deliver laser pulses with a wide variety of lengths and shapes for different experiments. Trident also has a small third laser beam, which is used to create a short x-ray pulse next to the target. X radiographs of evolving experiments can be obtained from the x-ray pulses and are particularly useful for the study of high-pressure hydrodynamics. Trident pulses, when applied to appropriate targets, can produce shock-wave pressures of several megabars and x-ray pulses of moderate temperatures.

Still higher temperatures over somewhat larger volumes can be obtained on the Nova laser at Lawrence Livermore National Laboratory. Nova, the largest glass laser in the world, produces pulses of up to 40 kilojoules in one nanosecond. We have fielded a number of experiments on Nova related to radiation hydrodynamics and x-ray-driven implosions.

How far can one go in increasing energy density by shortening the pulse length of the laser and reducing the size of the focused optical spot? Another laser at Los Alamos, Bright Source II, is providing the answer. Bright Source II pushes the limits of energy density by directing a relatively small amount of energy (only a quarter of a joule at present, although a 10-joule machine is on the horizon) into an incredibly short pulse that can be focused down to only a few microns. Bright Source II pulses last less than 300 femtoseconds, so that even though it is moving with the speed of light, a pulse is only about 900 microns long. The focused pulses have intensities of more than 5×10^{18} watts per square centimeter, well above pulse intensities produced by Trident or even Nova. The impact of a laser pulse on the surface of a target sample creates pressures of more than 1 gigabar, but only for about one picosecond, after which the sample expands under thermal pressure. (It is interesting to note that the radiation pressure—the pressure due to the momentum of the light itself—is 1 gigabar, which is comparable to the induced thermal pressure in the target.) During such a short pulse the atoms in the target do not have a chance to equilibrate and may not approximate fully the equilibrium conditions found in a nuclear explosion. Nevertheless, Bright Source II can heat thin solid foils to keV temperatures, creating a high-density and very hot plasma. Hence this laser can be used to probe the structure and dynamics of matter at conditions that approach those found in a nuclear explosion. The hot plasma cools both by expansion and by the emission of x rays. Figure 4 shows a typical x-ray spectrum from an aluminum sample illuminated by a Bright Source II pulse. Up to 1 percent of the incident laser energy is converted to line radiation around 2 keV. The line radiation is useful for studying the interaction of x rays with matter.

The extremely short pulses available from the Bright Source II laser also provide an effective means to study very rapid processes, such as transient chemical reactions. In typical chemical detonations several transient molecular species such as OH radicals persist only for a short time but are important in determining the overall energy balance in the detonation products. An experiment is currently underway at Bright Source II to measure the OH radical in a forced detonation—the first such measurement of its kind for an explosive process.

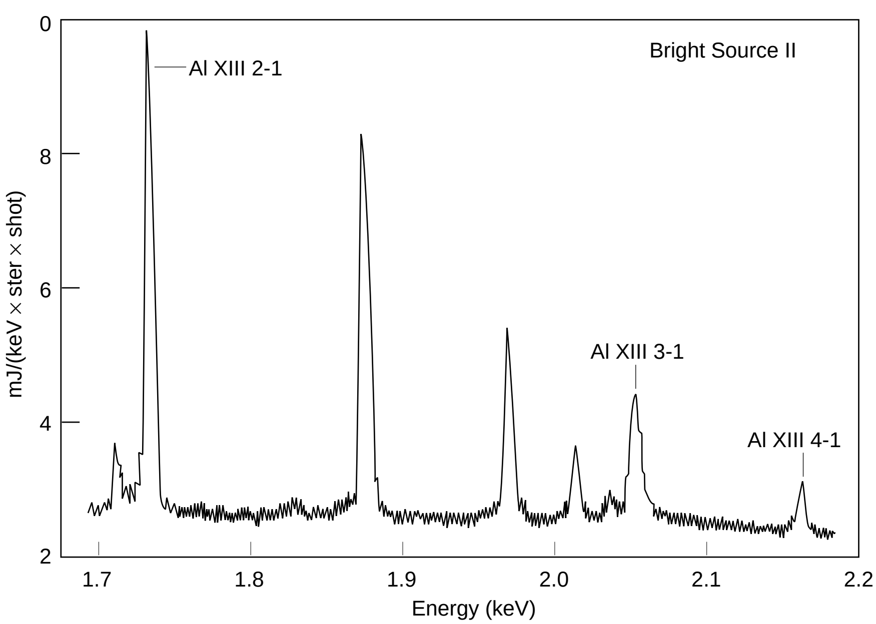


Figure 4. X-Ray Spectrum Induced by a Pulse from Bright Source II
The x-ray spectrum results from the impact of a 0.25-joule, 300-picosecond pulse from the Bright Source II laser on an aluminum foil. The intense lines serve both as a diagnostic of the conditions in the dense radiating plasma and as valuable probes for use in other experiments.

Nuclear Physics at LAMPF

Moving up again in the energy scale, we encounter nuclear energy densities—where the relevant energy parameters are not kiloelectronvolts as in plasmas but megaelectronvolts.

The formation of a critical mass during the detonation of a nuclear explosive and the attendant chain reaction result in an intense neutron burst. Neutrons interact with nuclei through a complex set of scattering and capture processes, some leading to the production of additional neutrons and/or the initiation of fission and others to the production of stable isotopes. To model the dynamics of fission in weapons, we must have accurate descriptions of all of the dominant neutronics processes. The knowledge of weaker processes (such as those involving transient nuclear states) can provide valuable diagnostics on the progress of the nuclear burn and contribute to radiochemical analyses of nuclear explosions. The Laboratory has conducted an extensive series of experiments on nuclear physics important for weapons at LAMPF and other nuclear facilities.

LAMPF is the most powerful accelerator in the world. Although some machines accelerate charged particles to higher energies, none is capable of delivering as many particles per unit time to the target as LAMPF. This capability is important when one wants to study weak processes, including the study of higher-order nuclear cross sections. In addition to the accelerator itself, the LAMPF facility includes several target areas. The areas of particular concern for the weapons program are the Weapons Neutron Research (WNR) facility and the Manuel Lujan, Jr. Neutron Scattering Center (LANSCE).

LAMPF has been extensively used by the weapons program. Fundamental aspects of fission have been studied by examining the relative timing of fission and neutron emission in fissioning nuclei. Angular distributions of neutrons, gamma rays, and fission decay products have been measured to determine the energy and momentum balance in fission. Detectors used in nuclear tests have been calibrated on LAMPF to provide absolute measurements of the neutron flux from the nuclear device. We are currently evaluating new techniques for using proton and neutron sources to image dynamic phenomena in opaque samples.

The Future

The next several years promise to be among the most interesting and productive ever for high-energy-density physics. We have assembled an array of facilities to investigate a wide range of physics issues of importance to the nuclear weapons program. The structure and transport properties of hot dense matter will be systematically studied in a regime where single-atom theories break down and many-body effects are important. The interaction of strong shock waves and x-ray pulses with matter will continue to be studied with the aim of providing quantitative data for use in our computer models of nuclear explosions. Experimental data on hydrodynamics and hydrodynamic instabilities will allow us to validate increasingly sophisticated algorithms in new computer codes, particularly those that will need to be developed to exploit the promise of massively parallel computers.

Each of our capabilities can be extended to higher energies for even more interesting applications. The next advance in Laboratory capacitor banks is Atlas, a 25-megajoule machine that will permit us to study high energy densities over tens of cubic centimeters. The Procyon high-explosive pulsed-power generator will be followed by a more advanced system that will deliver in excess of 200 million amperes. Bright Source III is being designed to produce focused intensities over 10^{20} watts per square centimeter to permit the study of multiphoton x-ray interactions. This intensity is high enough to rip apart the vacuum in the electrostatic field near a nucleus to create electron-positron pairs, literally creating matter from energy. ■

Stephen M. Younger is the Program Director for Inertial Confinement Fusion and High Energy Density Physics at the Laboratory. He received a Ph.D. in theoretical physics from the University of Maryland in 1978. From 1974 to 1982 he performed theoretical studies of atomic processes at the National Bureau of Standards in Washington, D.C., concentrating on electron scattering from atoms and ions. From 1982 to 1989 he was a staff member and group leader at Lawrence Livermore National Laboratory, where he specialized in advanced thermonuclear weapons design, x-ray lasers, electromagnetic weapons, and other defense programs. In 1989 he came to Los Alamos and in 1991 he was named to his present position. In addition to working on defense issues, Younger has continued to contribute to fundamental atomic and plasma physics. He is a Fellow of the American Physical Society and has served on numerous government panels and committees.

Déjà vu all over again

Houston T. Hawkins

The Manhattan Project grew out of a chilling intelligence assessment by scientists—many of whom would later work at Los Alamos—that the Third Reich was actively pursuing the development of an atomic explosive. Indications were that research was being carried out by a team headed by Werner Heisenberg in the Reich Research Council, which reported to Field Marshall Hermann Goering. Development of an atomic warhead for the German V-2 rocket had the real potential of changing the course—and probably the outcome—of the war.

Tracking the Third Reich program instantly became the single most important intelligence task for the United States and Great Britain. Unfortunately, when the question of the status of the Third Reich atomic-explosive program was posed to the Office of Special Services (OSS)—the forerunner of the Central Intelligence Agency—the probable response from General Donovan, the head of the OSS, was, "What is an atomic explosive?" Information necessary to make such an assessment was so compartmentalized that even Vice President Truman did not know of the existence of our program until he became president following the death of President Roosevelt. Albert Gore, Sr., who as a congressman was told by Speaker of the House Sam Rayburn to hide millions of dollars in the budget for a "special project," did not know or dare ask about the project for which the money was appropriated.[1]

Therefore, the function of assessing the status of the Third Reich program had to be transferred to Los Alamos because only scientists actually working on our atomic-explosive project had the requisite clearances and knowledge to make the crucial judgements demanded by the enormity of the threat.

A team of scientists was assembled at Los Alamos and charged with providing the required assessments and tracking of the Third Reich program. Relying on technical literature published by the Germans even in the throes of World War II, on information collected by the Alsos Mission,[2] and on contacts that a few Los Alamos scientists, such as Niels Bohr, had had with Heisenberg, the team determined that the Germans had grossly overestimated the amount of highly-enriched uranium required for an atomic explosive and had overestimated the thermal-neutron cross section of graphite. The team suggested that the Germans would possibly pursue heavy water in lieu of graphite as the neutron moderator in their plutonium-production reactor.

[1]Private communication, 1986. Senator Albert Gore, Sr., said that money for the Manhattan Project was dispersed throughout the federal budget. He had no idea what the money was for until he read about the use of the atomic bomb in Japan. He opined that the project was possible only because of the trust and discipline that existed in the House of Representatives at that time.

[2]The Alsos Mission was established by General Leslie Groves within his Intelligence Department to collect information on German nuclear-physics programs. The mission operated within European areas liberated by Allied Forces. For more information about the Alsos Mission, see *Alsos* by Samuel A. Goudsmit (Henry Schuman, Inc., 1947).

As a result, the heavy-water plants in Nazi-occupied Norway were targeted for destruction by the Royal Air Force and British commandos. There was some uncertainty as to whether the Germans had knowledge of the use of plutonium in atomic explosives or even knew of its existence. Therefore the destruction of the heavy-water plants was only added insurance that the Third Reich never would be able to develop a plutonium-production capability.

Inspection of German facilities after the war by intelligence officers indicated that the Los Alamos team had provided very accurate assessments of the Third Reich effort. According to Samuel A. Goudsmit, historian of the Alsos Mission, because of the lack of progress and direction, support for Heisenberg's efforts by the Third Reich had waned substantially about the same time that our program was going into high gear.

In 1990, as American troops were being assembled for deployment to the Persian Gulf, we were again confronted with the chilling prospect that a despotic regime was on the verge of acquiring one or more nuclear weapons. Like Adolf Hitler, Saddam Hussein—the regime's ruler—had already demonstrated a capability to deliver weapons of mass destruction[3] and had demonstrated a resolve to use such weapons even against his own people. In the words of Yogi Berra, "It was déjà vu all over again."

Coming to grips with that prospect was a frantic process made all the more urgent by the impending deployment of U.S. military forces and compounded by the dearth of information available on the Iraqi nuclear-weapons program. However, before U.S. forces landed in Saudi Arabia, the U.S. intelligence community, relying heavily on assessments from Los Alamos and Livermore national laboratories, had reached a general consensus that the Iraqis were still within several months to a year of having a nuclear weapon. Barring the diversion of highly-enriched uranium from their research reactors at Tuwaitha, the Iraqis probably did not possess enough plutonium or enriched uranium to actually build a nuclear weapon. Moreover, inspections by the International Atomic Energy Agency found the reactor fuel still in place.

However, our assessment of the status of the Iraqi program—although technically accurate—proved more an example of good fortune than an example of good intelligence. Unknown to us at the time our assessment was made was the sheer magnitude of the nuclear-weapons program being carried out by the Iraqis—in violation of the Nonproliferation Treaty—under the cover organization Petro Chemical 3 (PC-3).

Borrowing the technology behind the "calutrons" developed early in the Manhattan Project, PC-3 had built an enrichment facility and was in the process of separating weapon quantities of highly-enriched uranium. In essence, these separators, which the Iraqis called Baghdadtrons, were large mass spectrometers capable of deflecting uranium ions of differing masses into graphite collectors. The basic technology, called electromagnetic isotope separation, had been abandoned by the U.S. as a means of separating large amounts of uranium because of its relative inefficiency and high operating costs. However, the technique had provided top product enrichment for the uranium used in the "Little Boy" device.

Considering the size of the Iraqi nuclear-weapons program, had the Iraqi invasion of Kuwait not occurred and had that invasion not precipitated a military response by the Allies, Iraq probably would today possess material for one or more nuclear weapons, forever altering the strategic situation in the politically volatile Middle East.

Since Desert Storm, Los Alamos scientists have served on several inspection teams under United Nations sponsorship and have played a significant role in developing our understanding of the scope and nature of the Iraqi program to build nuclear weapons. Currently, as part of an ongoing effort, the Laboratory is heavily involved in developing new nonproliferation-monitoring methods and negotiating more effective agreements on dual-use technologies. This process goes on today.

The major lesson learned in Desert Storm is that because of the growing availability of plutonium and enriched uranium throughout the world and because of the proliferation of nuclear-weapons know-how, we cannot afford to enter future conflicts blind to the realities of any nuclear-weapons programs in the area of conflict. We cannot have any future déjà vu interrupted with an inopportune flash of blue light. It is this realization that has provided the real spur to nonproliferation initiatives and programs within the DOE and other federal agencies. ■

[3]Interestingly, the Iraqui AL ABBAS and AL HUSSAIN missiles were derivatives of the Soviet SCUD missile that was, in turn, a derivative of the Third Reich's WASSERFALLEN surface-to-air missile.

Houston T. Hawkins is leader of the Laboratory's International Technology Division. He joined the Laboratory in 1988 after twenty-five years of service in the U.S. Air Force.

Proliferation Challenges in Perspective

Joseph F. Pilat

A New Look at an Old Problem

Are we entering a world in which there will be "bombs for all"? If soundings of the media and the academic and policy communities are to be believed, we are indeed entering a brave new world of nuclear proliferation. In the view of the *Economist*:

> Fears of a nuclear Armageddon have dominated the past half-century. Unhappily, despite the end of the Cold War and the cascade of weapons cuts announced by America and Russia, the fears are still there. Twenty years ago, when efforts began to ban new bomb-builders, pessimists predicted that by now there would be 20 or 30 thrusting new nuclear powers (besides the famous five: America, Russia, Britain, France and China.) They were wrong. Twenty months from now, however, their gloomy prediction could start to come true. The spread of the bomb poses the biggest threat to the fragile new world order. Coping with it will take cash, cooperation and strong nerves. It is barely do-able, and time is short.*

The issue of nonproliferation is by no means new. For decades the threat of proliferation has been recognized as a fundamental challenge to regional and international security, and for over forty years the United States has been committed to the nonproliferation of nuclear weapons as a fundamental national-security and foreign-policy objective. However, the post–Gulf War revelations of the advances achieved in the Iraqi nuclear-weapon program as well as questions about the inheritance of the former Soviet Union's nuclear expertise, personnel, and materials have recently highlighted the threat of spreading nuclear capabilities.

Before developments in Iraq and the former Soviet Union galvanized concern about nuclear proliferation, a growing recent interest in nonproliferation was being fed by fears of chemical and biological weapons as well as missile proliferation. It had generally been assumed that a mature, functioning nuclear-nonproliferation regime had succeeded reasonably well in stemming proliferation—with the exception of a few rogue states that refused to accede to the Treaty on the Nonproliferation of Nuclear Weapons (NPT) and to adopt comprehensive International Atomic Energy Agency (IAEA) safeguards. The NPT and the IAEA safeguards are the centerpieces of the international nonproliferation regime. The NPT was concluded in 1968 and came into force in 1970. The large number of signatories to the NPT (over 150) make it the most widely adhered to arms-control treaty in history. The objectives of the NPT are to prevent the spread of nuclear weapons to states that do not already possess them; to ensure the fullest cooperation in the peaceful uses of nuclear energy in a manner consistent with the objective of nonproliferation; and to encourage arms-control efforts in both the nuclear and non-nuclear arenas. International safeguards, as set forth in agreements negotiated with the IAEA by NPT parties, are applied to all source or special fissionable materials with the aim of preventing the diversion of nuclear energy from peaceful uses to nuclear weapons or other nuclear explosive devices.

In recent years we have seen an interest in the development of chemical and biological weapons, particularly among nations the United States views as potential adversaries. Perhaps because our worst fears were never realized, chemical-weapons capabilities do not appear as threatening as they did prior to the Gulf War. Biological weapons, although a frightening prospect, are difficult to weaponize and employ effectively. With the new variables they bring to the equation, however, chemical and biological weapon programs complicate the nuclear-proliferation issue, as does the proliferation of delivery systems and other advanced conventional capabilities. The linkages among the types of proliferation are an obstacle in dealing with proliferation in the Middle East and other conflict-prone regions, and the mix of these capabilities is making proliferation where it

*"Bombs for All?" *The Economist,* March 14, 1992, 15.

is occurring more militarily significant as weapons of mass destruction are mated to delivery and support systems.

In contrast to perceived wisdom, however, the problem of the proliferation of weapons of mass destruction is not spiralling out of control, and the issue is unlikely to be as central to U.S. policy as during the period immediately after the Second World War. The stakes were higher then—the United States was the only nation with a demonstrated nuclear capability and the Soviet Union and other great powers were still only potential proliferators. Once Britain, France, China, and, in particular, the Soviet Union joined the U.S. as nuclear powers, the immediate focus shifted from potential proliferation to potential confrontation, and the arms race and the Cold War gained the spotlight.

For the foreseeable future at least, the threat of a nuclear confrontation will not approach that posed by the Soviet Union during the Cold War. With the decline of the Soviet threat, however, the global situation has become far less stable, and nonproliferation is likely to have a higher priority than in the last twenty-five years. Although we may be surprised in the future about one or another country's interests or achievements in developing nuclear-weapons technology, currently only a few countries that have an undisputed capability and a limited number of countries that have programs or display an interest. Of these states, those whose possession of nuclear weapons would be most destabilizing and most threatening to U.S. interests are not on the verge of developing nuclear weapons—either because their indigenous capabilities are virtually nonexistent, for example Libya, or because because of international pressures or actions, as in the case of Iraq.

Developments in the former Soviet Union, however, could fundamentally change the present calculus. In addition, China's willingness to export advanced military capabilities has been a serious irritant in the past, but the disintegration of China, a possibility when the old leadership changes, could result in problems similar to those now possible in the region of the former Soviet Union. In this context, persistent public reports of Iran's acquisition of former Soviet weapons and China's willingness to assist proliferators are disturbing.

The Fall of the Soviet Union Gives Rise to New Threats

The recent collapse of the Soviet Union represents an emerging challenge having the potential to undermine the nonproliferation regime more severely than any that has commanded attention in the last four decades. High-level attention is of course being given to issues directly surrounding nuclear inheritance in the old Soviet Union. Nuclear weapons, nuclear scientists, engineers and technicians, and other nuclear capabilities may flow out of Russia, Ukraine, Belarus, and Kazakhstan and be "imported" to other countries. Further unrest in the former Soviet Union, including the possible disintegration of Russia, may result in nuclear theft or sabotage, and terrorism or the deliberate use of a nuclear weapon by one of the post-Soviet states or elements within them cannot be ruled out.

There are limits to U.S. influence, but such actions as the recent nuclear-arms-control initiatives by former President Bush, diplomatic moves to dissuade the non-Russian successor-states from retaining nuclear forces, and the establishment of centers in Russia and the Ukraine to staunch the emigration of former Soviet nuclear specialists share these objectives.

Although nuclear inheritance of former Soviet weapons appears to be the most immediate and striking problem posed by the end of the Cold War and the collapse of the Soviet Union, it is perhaps not the most important. The international nuclear-nonproliferation regime as it exists was based on, or reflected in, a number of fundamental assumptions: stable but confrontational East-West relations; U.S.-Soviet cooperation in an area of mutual interest; and dominant U.S. and Soviet influence in international fora and in international relations including Soviet ability to control exports and proliferation behavior within its sphere of influence. All of these assumptions are now in question.

One positive effect of the end of the Cold War has been a greater prospect for dealing with problem countries whose past efforts to obtain nuclear weapons exploited Cold War diplomacy. Nonetheless, the potential negative consequences of proliferation arising from the demise of the Soviet Union could far outweigh this benefit. The successor states to the Soviet Union appear unable to play the old Soviet role in promoting nuclear nonproliferation.

The old Soviet export-control structure has broken down, and it is by no means clear how a system based on a totalitarian regime will be replaced in a situation approaching anarchy. Clearly, the old order cannot be restored, and a new export-control regime will have to mirror those in the West and have all their weaknesses but perhaps not all their strengths. Shadowy private organizations in the former Soviet Union with power, money, and influence have the view that everything is for sale and that the most salable items are those that are militarily sensitive. The migration of Soviet nuclear specialists may not be resolvable, despite current efforts.

But these are only aspects of a much broader problem. Nuclear capabilities exist outside of Russia, in the Caucasian and Central Asian republics, and could find their way to the Middle East and South Asia, or result in new nuclear states in the former Soviet Union, for example, in Armenia. Russia itself could disintegrate, with perhaps tragic consequences. And China could pose the same set of problems in coming years, when the leadership changes. However apocalyptic these problems may appear, it is important to realize that they are now largely prospective—they may never be fully realized. In Russia, at least, there is a strong awareness of the scope of the problem and a sound recognition of mutual interests and the need for cooperation.

These dangers are serious, but, again, they remain largely prospective. If these dangers are not realized, the near-term proliferation threat will largely be limited to those developing countries that have been of concern over the last ten to twenty years, particularly countries in the Middle East, South Asia, and Northeast Asia. Currently, the list seems to be declining rather than growing due to positive developments in areas such as Latin America and Africa. A number of states on these continents, including Argentina, Brazil, and South Africa, are now disavowing nuclear-weapons interests or programs. In the longer term, developments in the international-security environment and in the international arms and high-technology markets could lead to either an expansion or contraction of the level of threat we face today. Trends are moving in both directions at present.

The Nonproliferation Regime under Attack

The new challenges to the international nuclear nonproliferation regime made evident by post-Soviet nuclear developments, along with those in Iraq, have led to a growing concern among many observers that the regime will be ineffective in the face of emerging, dynamic complications. Critics have argued that Iraq has demonstrated that the NPT and IAEA safeguards have been proven ineffective, that export controls have failed, and that Iraq's example will lead other states to proliferate. Indeed, one critic argues:

> The allied raids on Iraq's nuclear facilities and infrastructure not only set back the Iraqi program but destroyed once and for all the fiction created in the public mind over many years by artful propaganda and obscuration that the safeguards regime and current international export controls provide an effective barrier to proliferation. It is to be regretted that this message about the regime's weakness, delivered by the Israelis in 1981, was buried at that time by an avalanche of criticism of Israel by the nations now in the Desert Storm coalition. The NPT regime itself provides no early red flag indicating that one of its members has begun marching down the road toward weapon production, even if no treaty violations have yet occurred.*

It is argued that the present nonproliferation regime is too narrowly focused and thus not responsive to states that are proceeding with dedicated nuclear weapons programs; that undeclared programs are not covered by current verification mechanisms; and that when such programs are based upon a sophisticated, indigenous defense industrial base and on both legal and illegal imports of dual-use items, they are not susceptible to even vastly strengthened export controls. It is also argued that the regime, which was designed during the Cold War and reflects the mutual interests and influence of the United States and the former Soviet Union, is ill-equipped to deal with problems of the post–Cold War period including those arising from the breakup of the Soviet Union.

While these assertions cannot be dismissed, they appear somewhat exaggerated, and the reports of the death of the nonproliferation regime are premature. It is clear that the regime is challenged, and in the next several years we shall see whether it meets its challenges. In this period we will see whether the extraordinary nonproliferation measures being taken in Iraq will continue—whether the international community continues sanctions and long-term monitoring of Iraq's military-industrial infrastructure, or whether it will allow Iraq to resume its weapon programs, which will surely occur once UN activity ceases. This will determine what lessons will be drawn by potential proliferators

*Leonard Weiss, Tighten up on nuclear cheaters, *The Bulletin of the Atomic Scientists* 47 (May 1991) 11-12.

from Iraq's behavior and its consequences in the next year or two—that is, whether proliferators will believe the consequences of their activities will be acceptable or unacceptable if they are caught.

In addition, North Korea's withdrawal from the NPT could have disasterous consequences for the treaty. If the North Koreans decide to remain within the treaty, the administration of IAEA safeguards will be difficult. In any case, the IAEA will be administering safeguards under difficult circumstances in South Africa. If the safeguards are not seen as credible and effective, the IAEA and the NPT will be severely damaged, reinforcing the negative impact of the Iraqi case. There will be uncertainties about stocks of weapon-usable materials and, perhaps, weaopns in South Africa. It will be critical to monitor whether safeguards are effectively administered in South Africa by the IAEA and to pay close attention to the international response to any problems that might arise.

Finally, the behavior of Israel, India, and Pakistan, as well as other countries demonstrating proliferation potential will be critical in this period. Will these states continue to pursue their nuclear programs quietly, or for one reason or another, will their programs become overt?

These are the most pressing of the challenges now facing the international community. The NPT is up for extension in 1995, and all these issues could have an influence on the outcome. Other problems with adverse implications for the regime and the extension of the NPT may also arise, including nuclear-weapon tests by proliferators; overt weapons declarations by non-nuclear NPT states; further safeguards violations; nuclear theft, sabotage, or terrorism; and differences over implementation of the Nonproliferation Treaty, particularly its provisions regarding arms control (Article VI).

U.S. Nonproliferation Policy

In the face of new threats and a changing global environment, the first response of the United States will be a renewed commitment to strengthening existing nonproliferation mechanisms. The U.S. will continue to rely on both multilateral and unilateral approaches, using political incentives, technological constraints, bilateral export controls, and multilateral treaties. The indefinite extension of the NPT and strengthening the IAEA-administered system of safeguards against diversion of nuclear material from civil to military uses will receive strong U.S. support, as will the implementation of the Treaty of Tlatelolco, intended to create a nuclear-weapon-free zone in Latin America.

U.S. nonproliferation policy will also be pursued through diplomatic efforts. These include: consultation and cooperation among the major nuclear suppliers aimed at the implementation of export controls, upgrading the existing lists of controlled items that trigger IAEA safeguards and encouraging new nuclear suppliers to accept responsible nuclear-export policies. In addition, the U.S. will maintain—or more often attempt to begin—a dialogue on nuclear issues with non-NPT states through which we can express our concerns and encourage broader application of IAEA safeguards. The US has already begun and will continue to work to build a nonproliferation consensus, particularly during and after international crises that throw a spotlight on proliferation issues.

The United States has been and will remain committed to reducing motivations for acquiring nuclear explosives. To this end, the U.S. will continue to seek to improve regional and global stability, to strengthen alliance systems, and to promote the legitimate security interests of states through economic and security assistance in some cases and by other means. Of course, each of these objectives has other defense and diplomatic rationales, which at times work at cross-purposes with nonproliferation.

Export controls will remain an essential element of U.S. nonproliferation policy. U.S. nuclear export controls are designed to prevent the spread of nuclear weapons as well as to facilitate cooperation with other nations in peaceful uses of nuclear energy, such as electricity generation, agricultural research, and medical applications. It is a longstanding policy objective of the U.S.—the legacy of Atoms for Peace in the 1950s—to pursue peaceful nuclear cooperation while avoiding the dangers to international peace and security arising from nuclear-weapon proliferation.

A similar approach is being pursued in the areas of chemical weapons, biological weapons, and missiles.

A New Approach

It is essential to continue ongoing efforts to strengthen existing U.S. nonproliferation measures across the board, ranging from diplomacy and intelligence to export-control arrangements and treaties. Howev-

New Technologies in Support of Nonproliferation

For many years Los Alamos National Laboratory has had an active arms-control program with two main goals: to provide technology for verifying compliance with arms-control treaties, generally bilateral, and to support international activities in nuclear-materials control. But the world of the 1990s demands a considerable broadening of this charter. Technologies are needed for the deterrence and detection of worldwide tendencies toward the proliferation of weapons of mass destruction.

Initially viewed as a component of the Laboratory's Arms Control and Intelligence programs, nonproliferation support has recently emerged as a major new Laboratory initiative under the auspices of the Department of Energy's Office of Arms Control and Nonproliferation. A vastly increased role is anticipated for Los Alamos in activities relating both to monitoring and preventing proliferation, primarily of nuclear weapons but also of other weapons of mass destruction and their means of delivery. This role will utilize the Laboratory's expertise in space-based monitoring of weapons programs and in materials control and accounting as well as its premier capability in the nuclear-weapons program.

The historical bases for such Laboratory activities are the existing programs to analyze weapons programs overseas and to provide assessments of their motivations and their technical capabilities. Los Alamos has also contributed to national and international efforts to safeguard special nuclear materials such as plutonium and enriched uranium. All International Atomic Energy Agency inspectors have been trained at Los Alamos, and the equipment used by IAEA inspectors to monitor activities of nuclear facilities has largely been developed here and at Sandia National Laboratory. To assist in limiting the spread of technologies used to produce nuclear weapons, Los Alamos and other laboratories have provided technical expertise for national and international export controls, including technical advice on revising or updating international lists of controlled items for the Zangger Committee, the Nuclear Suppliers' Group, and other multilateral bodies.

The first major arms-control activity at Los Alamos was the design and preparation of the Vela satellite in 1960, which was used to detect atmospheric nuclear-weapons tests. To verify compliance with the Limited Test Ban Treaty (1963), Los Alamos collaborated with Sandia to provide spaceborne instrumentation for detecting nuclear tests either in the atmosphere or in outer space. These activities illustrate the Laboratory's characteristic role of utilizing the most advanced technologies on short notice in the space environment and adapting them to a variety of launch vehicles. Today such capabilities transfer directly into nonproliferation-related functions such as the detection of x-ray, gamma-ray, radio-frequency, neutron, and charged-particle radiations from nuclear detonations. In support of programs such as the detection of directed-energy weapons tests, they also supply data on natural and artificial space radiation. In 1993 the ALEXIS satellite is scheculed for launch into orbit to provide improved determination of the low-energy x-ray environment in space.

The Department of Energy's Accident Response Group (ARG) and Nuclear Emergency Search Team (NEST) are kept in readiness to respond in the event of a nuclear threat or nuclear weapon accident. Los Alamos provides both technical experts and equipment to these groups. This operational-readiness capability proved important in preparing DOE inspectors (from Los Alamos and other laboratories) who supported the IAEA and UN Special Commission inspections in Iraq. Maintenance of these emergency-response assets and capabilities will become increasingly important.

To meet emerging threats of proliferation, the laboratories are responding with vigorous R&D programs across the board. Ongoing safeguards R&D programs are devoted to developing varius radiation detection and measurement devices, engineering new hardware and software to do the necessary measurements, and designing complete safeguards systems integrated with physical protection and process operations. Current and future safeguards technologies and techniques will enable us to begin to properly address the new problems of nuclear weapons and special nuclear materials raised by the collapse of the Soviet Union and the ongoing radical reductions in nuclear arms. Topics of concern include commodity export, technical-data transfer, consulting, and guidance on foreign visits.

Recent R&D activity in export control at the Laboratory has focused on developing a Proliferation Information Network, an on-line interactive database system to centralize proliferation data and provide analysis tools. Although the network is currently providing export-license information for government agencies and the national laboratories, there is a real possibility to expand the scope of the data in the system. Integration of these data into actionable intelligence poses challenging problems in data transfer, display, and administration. R&D programs in this area are focusing not only on technical solutions such as improved data links, pattern recognition, and anomaly detection but also on the administrative challenges of the compartmentalization of information for security reasons.

Monitoring activities include satellite systems for wide-area detection of suspicious activities. For some years the Laboratory has been involved in projects to analyze observables resulting from hydrodynamic shock propagation and from surface ground motion near an underground nuclear test. These programs have recently become parts of an Integrated Geophysics Program to investigate the entire range of phenomena by means of which an underground explosion couples its energy into detectable signals such as seismic or acoustic waves. Historically, the major goal has been to verify nuclear-test-ban treaties. However, applications of most interest in the future may be the detection and identification of covert nuclear-weapons tests by new nuclear-weapon states who may or may not be signatories to the Nuclear Nonproliferation Treaty. Longer-range goals could include the use of combined seismic and acoustic data to estimate yields of any detected tests.

Additional monitoring activities involve high-resolution instruments that remotely detect effluents in chemical plumes, local sensors to monitor activities in production facilities, and portable systems that detect the presence of weapons or special materials. Because of its extensive experience in developing and fielding nuclear sensors, the Laboratory can make significant contributions to this important area of arms control. Nondestructive testing expertise and facilities at the Laboratory continue to play and important role in the development of on-site inspection tools that may include, among others, radiation detection, radiography, or acoustic resonance. One of the promising technologies in this arena is LIDAR (light detection and ranging). The DOE weapons laboratories can assist the military with analyses of military vulnerabilities and response options and help define and develop future technologies based on current NEST capabilities.

Emerging proliferation threats require active programs to develop the knowledge, technologies, and capabilities to prevent the spread and use of nuclear weapons and other weapons of mass destruction. Nonproliferation and arms control are expected to be among the most rapidly growing Los Alamos programs through the 1990s and into the next century. The Laboratory has demonstrated for more than thirty years an ability to use its expertise in nuclear weapons and related technologies to address and solve challenging problems in these important fields. As requirements become more clear in the world that emerges from the incredible changes of the last several years, we anticipate that the Laboratory's special capabilities for fast response to critical technical problems will continue to play a major role in ensuring U.S. national security. □

er, in the current political climate those traditional responses are no longer wholly adequate. New approaches, from regional arms control to military options, are being considered. The United States is undertaking unilateral actions including further limitations on nuclear testing and a cutoff in U.S. fissile-materials production. A nuclear no-first-use policy may also be considered. Such actions on the part of the U.S. are designed to set an example, and these "arms control" approaches to nonproliferation have primarily been put forward in the context of strengthening the NPT. Whatever their security rationale, however, such trade-offs are unlikely to affect the behavior of proliferants or to have a decisive impact on the future of the NPT. Nonetheless, there is a widespread belief that they will.

Other responses, some of which are already being pursued and all of which have advocates within the U.S., could become important for nonproliferation in the 1990s and into the twenty-first century:

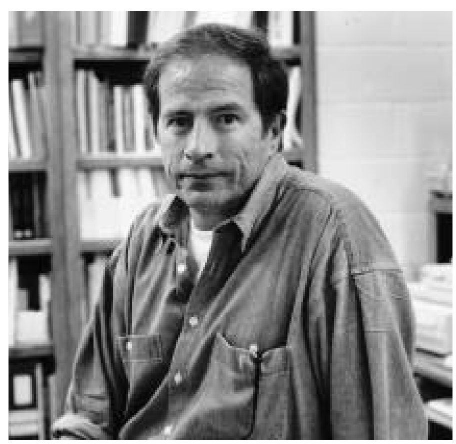
- building on the UN role in Iraq to ensure future UN action in the event of proliferation activities and to improve safeguards by utilizing UN inspection precedents;
- developing some type of embargo or sanction regimes to address noncompliant behavior of proliferators, perhaps on the basis of UN activity in Iraq;
- promoting new regional treaties, confidence-building measures, and monitoring and compliance arrangements to complement the global system in regions with particularly vexing proliferation problems;
- developing ballistic-missile defenses and a wide range of potential capabilities including accident response;
- and moving to a policy of managing proliferation with political, diplomatic, economic, and other instruments in cases where prevention fails.

The U.S. will also explore the difficult, but perhaps necessary, avenue of unilateral or internationally sponsored military action against proliferators.

Such measures, if realized, could augment but not replace the old approaches. More radical measures are unlikely to be seriously posed and considered unless there were some extraordinary event, such as a failure to extend the NPT in 1995. Even now the revival of the Baruch Plan, which called for strict international control over all nuclear activities, the conclusion of some overarching nonproliferation treaty, and other grandiose concepts are being put forward. Although in the minority, some see a regime of sanctions sponsored by the UN Security Council as an attractive alternative to the NPT/IAEA system. The appearance of such a variety of alternatives indicates that the problems and loopholes in the old regime are increasingly being recognized. Whatever else might be proposed or adopted, the U.S. will continue to seek to strengthen traditional elements of U.S. nonproliferation policy.

As new approaches to new proliferation threats are considered, it is important to recognize there is now no consensus on a major restructuring of the regime. However, the strong domestic and international support for strengthening the regime provides opportunities for the Laboratory to help to address emerging proliferation challenges. Los Alamos, along with the other DOE weapons laboratories, has long been involved in U.S. and international nonproliferation efforts. The laboratories' nuclear-weapon expertise provides unique capabilities for assessing foreign programs and intentions; for developing technology to detect, monitor, and respond to proliferation; and for operational support in national and international emergencies. (See "New Technologies in Support of Nonproliferation.")

There are no "silver bullets" to use in response to proliferation. However, technologies have in the past and can in the future enhance nonproliferation efforts. There is a strong requirement for action, a requirement demanding patience and vigilance over the long term. ■



Joseph F. Pilat, a member of the staff of the Laboratory's Center for National Security Studies, is spending 1993 at Cornell University as a Visiting Associate Professor in the Department of Government and a Visiting Scholar in the Peace Studies Program. He has participated in reviews of the NPT and was involved in the Open Skies negotiations. Pilat was a Senior Research Associate at the Congressional Research Service and has taught at Georgetown University, where he was a Philip E. Mosely Fellow at the Center for Strategic and International Studies. He has written widely on defense and national security issues.

New Technologies in Support of Nonproliferation

For many years Los Alamos National Laboratory has had an active arms-control program with two main goals: to provide technology for verifying compliance with arms-control treaties, generally bilateral, and to support international activities in nuclear-materials control. But the world of the 1990s demands a considerable broadening of this charter. Technologies are needed for the deterrence and detection of worldwide tendencies toward the proliferation of weapons of mass destruction.

Initially viewed as a component of the Laboratory's Arms Control and Intelligence programs, nonproliferation support has recently emerged as a major new Laboratory initiative under the auspices of the Department of Energy's Office of Arms Control and Nonproliferation. A vastly increased role is anticipated for Los Alamos in activities relating both to monitoring and preventing proliferation, primarily of nuclear weapons but also of other weapons of mass destruction and their means of delivery. This role will utilize the Laboratory's expertise in space-based monitoring of weapons programs and in materials control and accounting as well as its premier capability in the nuclear-weapons program.

The historical bases for such Laboratory activities are the existing programs to analyze weapons programs overseas and to provide assessments of their motivations and their technical capabilities. Los Alamos has also contributed to national and international efforts to safeguard special nuclear materials such as plutonium and enriched uranium. All International Atomic Energy Agency inspectors have been trained at Los Alamos, and the equipment used by IAEA inspectors to monitor activities of nuclear facilities has largely been developed here and at Sandia National Laboratory. To assist in limiting the spread of technologies used to produce nuclear weapons, Los Alamos and other laboratories have provided technical expertise for national and international export controls, including technical advice on revising or updating international lists of controlled items for the Zangger Committee, the Nuclear Suppliers' Group, and other multilateral bodies.

The first major arms-control activity at Los Alamos was the design and preparation of the Vela satellite in 1960, which was used to detect atmospheric nuclear-weapons tests. To verify compliance with the Limited Test Ban Treaty (1963), Los Alamos collaborated with Sandia to provide spaceborne instrumentation for detecting nuclear tests either in the atmosphere or in outer space. These activities illustrate the Laboratory's characteristic role of utilizing the most advanced technologies on short notice in the space environment and adapting them to a variety of launch vehicles. Today such capabilities transfer directly into nonproliferation-related functions such as the detection of x-ray, gamma-ray, radio-frequency, neutron, and charged-particle radiations from nuclear detonations. In support of programs such as the detection of directed-energy weapons tests, they also supply data on natural and artificial space radiation. In 1993 the ALEXIS satellite is scheculed for launch into orbit to provide improved determination of the low-energy x-ray environment in space.

The Department of Energy's Accident Response Group (ARG) and Nuclear Emergency Search Team (NEST) are kept in readiness to respond in the event of a nuclear threat or nuclear weapon accident. Los Alamos provides both technical experts and equipment to these groups. This operational-readiness capability proved important in preparing DOE inspectors (from Los Alamos and other laboratories) who supported the IAEA and UN Special Commission inspections in Iraq. Maintenance of these emergency-response assets and capabilities will become increasingly important.

To meet emerging threats of proliferation, the laboratories are responding with vigorous R&D programs across the board. Ongoing safeguards R&D programs are devoted to developing varius radiation detection and measurement devices, engineering new hardware and software to do the necessary measurements, and designing complete safeguards systems integrated with physical protection and process operations. Current and future safeguards technologies and techniques will enable us to begin to properly address the new problems of nuclear weapons and special nuclear materials raised by the collapse of the Soviet Union and the ongoing radical reductions in nuclear arms. Topics of concern include commodity export, technical-data transfer, consulting, and guidance on foreign visits.

Recent R&D activity in export control at the Laboratory has focused on developing a Proliferation Information Network, an on-line interactive database system to centralize proliferation data and provide analysis tools. Although the network is currently providing export-license information for government agencies and the national laboratories, there is a real possibility to expand the scope of the data in the system. Integration of these data into actionable intelligence poses challenging problems in data transfer, display, and administration. R&D programs in this area are focusing not only on technical solutions such as improved data links, pattern recognition, and anomaly detection but also on the administrative challenges of the compartmentalization of information for security reasons.

Monitoring activities include satellite systems for wide-area detection of suspicious activities. For some years the Laboratory has been involved in projects to analyze observables resulting from hydrodynamic shock propagation and from surface ground motion near an underground nuclear test. These programs have recently become parts of an Integrated Geophysics Program to investigate the entire range of phenomena by means of which an underground explosion couples its energy into detectable signals such as seismic or acoustic waves. Historically, the major goal has been to verify nuclear-test-ban treaties. However, applications of most interest in the future may be the detection and identification of covert nuclear-weapons tests by new nuclear-weapon states who may or may not be signatories to the Nuclear Nonproliferation Treaty. Longer-range goals could include the use of combined seismic and acoustic data to estimate yields of any detected tests.

Additional monitoring activities involve high-resolution instruments that remotely detect effluents in chemical plumes, local sensors to monitor activities in production facilities, and portable systems that detect the presence of weapons or special materials. Because of its extensive experience in developing and fielding nuclear sensors, the Laboratory can make significant contributions to this important area of arms control. Nondestructive testing expertise and facilities at the Laboratory continue to play and important role in the development of on-site inspection tools that may include, among others, radiation detection, radiography, or acoustic resonance. One of the promising technologies in this arena is LIDAR (light detection and ranging). The DOE weapons laboratories can assist the military with analyses of military vulnerabilities and response options and help define and develop future technologies based on current NEST capabilities.

Emerging proliferation threats require active programs to develop the knowledge, technologies, and capabilities to prevent the spread and use of nuclear weapons and other weapons of mass destruction. Nonproliferation and arms control are expected to be among the most rapidly growing Los Alamos programs through the 1990s and into the next century. The Laboratory has demonstrated for more than thirty years an ability to use its expertise in nuclear weapons and related technologies to address and solve challenging problems in these important fields. As requirements become more clear in the world that emerges from the incredible changes of the last several years, we anticipate that the Laboratory's special capabilities for fast response to critical technical problems will continue to play a major role in ensuring U.S. national security. □

Russian-American Collaboration

the transition to peacetime work

Stephen M. Younger

John Shaner of Los Alamos; 88-year-old Yulii B. Khariton, Chief Scientist at Arzamas-16; Alexander Pavlovskii; and Los Alamos Director Sig Hecker in Rus-

The week of November 11, 1992 was a historic occasion at Los Alamos. After a year of intense preparation, which involved cutting through bureaucratic impediments and political hesitation, members of the nuclear-weapons program at Los Alamos sat down with our counterparts from the former Soviet Union to talk science and to explore the possibilities of conducting joint experiments on topics of scientific interest. By the end of the week, we laid down plans and signed an agreement, which was subsequently approved and funded by our respective governments, to allow scientists who have worked for decades on opposite sides of the superpower struggle to join together in collaborative ventures of mutual benefit.

Joint ventures of this nature had been proposed informally as early as 1989, but the possibility took on increased significance after the dissolution of the USSR in December of 1991. Government officials in the United States became concerned about the stability of the former Soviet Union's nuclear-weapons establishment and the future of their weapons scientists. Clearly, the scientists needed support in order to carry out the expected reduction of the nuclear-weapons stockpile. Stable support would also lessen the possibility that their emigration to foreign nations would open up opportunities for nuclear proliferation.

In response to these concerns and at the request of the Department of Energy, we at Los Alamos and scientists at Lawrence Livermore National Laboratory began to investigate how we might help our Russian counterparts. Through a series of visits beginning in November, 1991, it became clear that the establishment of scientific collaborations on concrete, well-defined, non-military projects of mutual interest would not only engage the minds of the ex-Soviet weapons scientists but also shore up their legitimacy and stature in the eyes of their government and lead to the stable funding of the nuclear-weapons laboratories in the former Soviet Union.

The significance and urgency of beginning such collaborations should not be underestimated. United States visitors to Russia have been told that unless something significant happens soon, some nuclear-weapons scientists may be forced to leave Russia in order to support their families. Pay for even senior researchers is sporadic and decimated by inflation. Everyone farms for survival, and the first signs of malnutrition are becoming evident among the populations of the hitherto secret, and still closed, cities that are the sites of the nuclear-weapons laboratories. Although the United States cannot hope to maintain entire cities, we can have a major influence on the

“IN HOPE OF A WORLD WITHOUT ‘*US*’ and ‘*THEM*’”

Sig Hecker with Boris Litvinov, the premier Russian nuclear weapons designer at Chelyabinsk-70, in February 1992.

Front page of December 12, 1991, newspaper from Arzamas-16 discusses visit by Nerses Krikorian (l) and Dan Stillman of Los Alamos (c) with Alexander Pavlovskii (r). This was the first visit by Western scientists to a nuclear-weapons laboratory in the former-Soviet Union.

Sig Hecker and Livermore Director John Nuckolls bid farewell to Russian scientists at Arzamas-16 airport, February 1992.

Irv Lindemuth (l) and Bob Reinovsky (r) with Pavlovskii at the Zababakhin Scientific Talks in Kistym, Russia (near Chelyabinsk-70) prior to visiting Arzamas-16 (January 1992).

relatively small group of experts (perhaps a few thousand) who have an intimate knowledge of nuclear weapons and who thus constitute the greatest threat to nuclear nonproliferation.

As in the United States, the former Soviet Union had two major nuclear weapons-design laboratories: the All-Russian Institute of Experimental Physics in Arzamas-16, which is approximately 400 kilometers east of Moscow, and the All-Russian Institute of Technical Physics in Chelyabinsk-70, which is at the base of the Ural Mountains. Arzamas-16 was the birthplace of the first Russian fission and thermonuclear weapons and is sometimes called “Los Arzamas” by the Russians to highlight its similarity to Los Alamos. During nearly five decades of Soviet nuclear-weapons development, many of the brightest graduates from universities all over the USSR were assigned to those two institutes and given the best equipment and living conditions available. In addition to nuclear-weapons development, the Russian weapons institutes had extensive basic-research programs, which in some areas were ahead of anything known in the West.

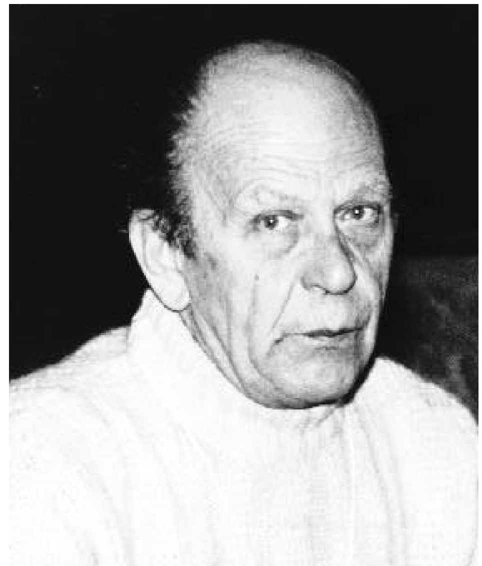
One particular technology developed at Arzamas-16 is related to the production of very-high magnetic fields and electric currents. In the Soviet Union, this work was started by Andrei Sakharov in 1951 as part of a program to achieve thermonuclear fusion. As a privileged member of the Institute, Sakharov was able to direct some of the most creative young people entering the laboratory to develop high-magnetic-field generators; among that group was Alexander Pavlovskii. During the 1950s and 1960s this program made extraordinary strides. In fact, some Russian achievements as far back as 1967 exceed the best capabilities of Western laboratories even today. Although high-magnetic-field generators were originally developed for possible military applications, Sakharov soon recognized that strong electric currents and magnetic fields offer fascinating opportunities for basic scientific research into the atomic structure of solids, high-pressure chemistry, and even high-energy particle accelerators. This interest remained with Sakharov throughout his life. Just before his death he told a colleague that seeing the development of this technology was "his fondest wish."

During more than forty years Alexander Pavlovskii and Vladimir Chernyshev led research teams in the fields of pulsed power and high magnetic fields. The technology they have developed is now routinely used to produce magnetic fields of 17 megagauss. They have also demonstrated routine use of 200-megajoule pulsed-power generators and have performed up to 200 high-energy pulsed-power experiments per year. These superb experimental capabilities have led to a myriad of applications in atomic and plasma physics, microwave generation, lasers, high-pressure chemistry, hydrodynamic studies, and more. Max Fowler pioneered the development of pulsed-power experiments at Los Alamos beginning in 1952, and he made many significant achievements. As a result, Los Alamos has the only U.S. counterpart to many components of this massive Russian program.

Chernyshev (l), Younger (c), and Pavlovskii (r) sign agreement for collaboration.

Gradually, the Soviets working in this area convinced their government to allow carefully screened publications of their work and even to allow their participation in some international conferences, but before 1987 they were not permitted to tell us where they lived or worked. After the fall of the Soviet Union, their participation in conferences increased dramatically and the Russians began revealing the remarkable breadth and depth of their accomplishments. Because of our contact with them at conferences and through publications, Los Alamos scientists were the first Westerners to be invited to visit Arzamas-16 for scientific discussions. Through a series of four technical meetings in 1992, some in Russia and some in the United States, a collaboration has been established between Los Alamos and "Los Arzamas" to jointly explore the basic scientific applications of high magnetic fields and electric currents.

The specifics of the initial collaboration are set down in an agreement signed at Los Alamos by Pavlovskii, Chernyshev, and me. The agreement calls for conducting experiments in Los Alamos and Arzamas-16 using teams from both laboratories. In the first experiment, Los Alamos scientists will participate in the design, assembly, and testing of a superhigh-current electromagnetic generator perfected by Chernyshev. This experiment will be conducted in Russia in July, 1993. Next, four ultrahigh magnetic generators developed by Pavlovskii will be purchased, transported to the United States, and used to study the properties of high-temperature superconductors in experiments to be conducted in Los Alamos. Both sides expect that the scientific data derived from this work will foster more ideas for further collaboration. We believe that agreements such as the one signed at Los Alamos represent historic opportunities to "beat swords into plowshares" by encouraging scientists to turn their attention from weapons development toward scientific applications that promote economic growth. It is appropriate that this first step took place at Los Alamos, the birthplace of the atomic age, and involved Arzamas-16, the institute that developed the first Russian nuclear weapons. ■



Interview with Alexander Ivanovich Pavlovskii

At the end of the intense week-long meeting at Los Alamos in November 1992 among scientists from Arzamas-16, Los Alamos, and Sandia, we met with the head of the Russian delegation, Alexander I. Pavlovskii, to talk about his experiences as a nuclear-weapons scientist in the former Soviet Union. Pavlovskii had been a protegé of Nobel Peace Prize winner Andrei Dmitrievich Sakharov. At the time of our conversation, he was Deputy Chief Scientist and Head of the Fundamental and Applied Physics Department of the All-Russia Scientific Research Institute of Experimental Physics at Arzamas-16, Russia.

Two translators were present: Elena Panevkina, who was by Pavlovskii's side at all meetings with non-Russian speaking scientists, and Eugene Kutyreff from the Laboratory's International Technology Division. We thank both of them for their patience and endurance.

Just as we were preparing to send this interview to Pavlovskii for his review, we learned of his sudden death on February 12, 1993. We were honored to have met him and moved by the candor and depth of feeling he expressed during our interview. Many scientists at Los Alamos knew Pavlovskii well, and we hope they will find this interview a fitting memorial to an exceptional man.

Los Alamos Science: Tell us how you got into science and how you came to work in a nuclear-weapons laboratory.

Pavlovskii: It's hard to say how I became a scientist because when I was young, you could not decide on a career by yourself—it was something that was decided for you.

In the late 1940s I attended Kharkov University, which is in Kharkov in the Ukraine. It houses the well-known Physicotechnical Institute. Lev D. Landau and Evgenii M. Lifshitz, known so well in the West for their physics textbooks, worked together at that institute during the 1930s before Landau moved to Moscow. During the third year of my university studies, I began to work at the Physicotechnical Institute. At the time, the Director was Kirill Dmitrievich Sinelnikov, one of those physicists who had gone through Cambridge University in England. I was very fortunate to have been associated with him and with Lifshitz, Alexander Il'ich Akhiezer, Giorgii Nikolaevich Flerov, and many other very interesting people.

Los Alamos Science: How did you become involved in nuclear-weapons work?

Pavlovskii: In my time, the Soviet Union had a specialized program for bringing people into a variety of types of work. After you finished at the institute or the university, you would be given your choice of several places where you might want to work. When I picked Arzamas, I didn't really know what was going on at that institute. I had a romantic notion of something new, something different. Sinelnikov didn't really want me to leave the Physicotechnical Institute. He wanted me to stay and work on the equivalent of a master's degree, but after I met in Moscow with Flerov, who was working at Arzamas-16 at that time, I was determined to go there.

My dissertation work had been in classical physics. I had been doing experimental work on the process of fluorescence. So when I arrived in Arzamas I realized that I needed to change my specialty a little bit, and I began studying nuclear physics. My professors at the Kharkov Institute had taught us well, so I was quite prepared to learn new things.

Almost immediately I became involved in studying physical processes for the development of the first Soviet hydrogen bomb, the father of which is, of course, Andrei Dmitrievich Sakharov. I was involved primarily in studying elementary nuclear cross sections and effective cross sections in special assemblies of materials. These experiments were absolutely necessary in order to understand the physics of nuclear fission and neutron transport.

I was in the department headed by Flerov, and I also worked with Sakharov and many other great men who are too numerous to mention. From the point of view of a scientist, it was a very unusual time, and a very interesting and productive time. All the questions facing us were brand new. In hindsight, other people may look at the development of thermonuclear weapons somewhat differently, but back then all of us were quite sure that in order to preserve peace and maintain a stable environment in the world, this type of weapon was absolutely necessary. As Andrei Sakharov himself said, it was "to prevent temptations."

Our scientists thought about defense in just the same way as the Americans. We thought the development of the hydrogen bomb was a very important undertaking. We worked very, very hard, sometimes around the clock. And through this labor we were able to resolve some very interesting scientific questions. Remember, our country had just undergone tremendous destruction during World War II, and the resources available to us were not nearly as good as those available to the American scientists. Therefore, for us to solve the same problems you were solving required a maximum effort on our part with a minimum of expenditures on materials. It was a time of great tensions but also one of great accomplishment.

Paradoxically, during this period of strenuous demands, the level of intellectual life was also very high. We not only read a great deal of artistic literature, we also looked at all the new inventions and new scientific discoveries. When we had free time, we were involved in sports. Many of us were interested in theater, and at every opportunity during our trips to Moscow we went to plays. In those first few years at Arzamas, I saw more shows than I have seen over the course of the rest of my life.

Los Alamos Science: How far is Arzamas from Moscow?

Pavlovskii: It's about 200 miles or an hour's plane ride. The Institute was necessarily isolated because of the work that was going on there, but the work sent us to Moscow, Leningrad, which is now St. Petersburg, and other cities, so we traveled quite often.

Los Alamos Science: Is it true that scientists were able to fly to Moscow with their families just to attend the theater or enjoy some leisure time?

Vladimir Chernyshev (l), Max Fowler (c), and Alexander Pavlovskii (r) at Los Alamos.

Pavlovskii: Yakov Zeldovich and Yuli Khariton were absolutely enamored of the theater, and they would sometimes take time off to visit Moscow and attend the theaters.

Los Alamos Science: Arzamas, it seems, had a collection of great physicists just as Los Alamos did during the Manhattan Project.

Pavlovskii: Yes, we had many renowned physicists, but also some young, inexperienced specialists; I was among the latter group. When I arrived at Arzamas in 1951, Andrei Sakharov was 30 years old, the Director of the Laboratory was 31 or 32 years old, and most of the rest of us were 23 to 25 years old. Together we made the first hydrogen bomb and the first thermonuclear weapons.

Los Alamos Science: Would you remind us of the dates of the major Soviet weapons developments?

Pavlovskii: Our first test of an atomic bomb was in 1949, the first hydrogen bomb was in August 1953, and in November 1955 we tested our first thermonuclear bomb.

Los Alamos Science: There was a time when the Americans believed the Russians were six months ahead in developing a thermonuclear device, and the government poured money into Los Alamos to accelerate our research.

Pavlovskii: I don't know whether we were months ahead of you, but the benchmark steps that I listed were not experimental tests, they were full-scale nuclear-weapons tests. If you've read Sakharov's memoirs, you would remember that he referred to the first hydrogen bombs as Idea 1 and Idea 2. Sakharov mentioned these early ideas because he was a very proper, very good person, and he wanted to demonstrate the creative contributions of Vitaly Ginzburg. The third idea, which led to modern thermonuclear weapons, really belongs to Sakharov. But as Sakharov himself has written, he and Zeldovich both developed Idea 3 together. It was essentially the same as the Ulam/Teller idea. What's interesting is that it doesn't really matter what country you were from; the physics is the same, the logical steps of science were identical, and, obviously, you had to come to the same conclusions.

Los Alamos Science: Perhaps you would like to say more about the conditions under which you were working compared with those of the American scientists.

Pavlovskii: There was a very big difference during the 1950s because we had very little laboratory equipment. We had to develop amplifiers, discriminators, and other electronics needed for our experiments from spare parts left over from the equipment of some Canadian radio stations that had been lend-leased during World War II. Despite the poor conditions, our specialists were still able to come up with electronic equipment that performed no worse than that of our American counterparts. The lack of equipment meant we had to apply more intellectual skill. It is a much more difficult task to develop the needed apparatus than to order it from some company.

The amplifiers we worked on, for example, were done using Elmore's book, a standard electronics text from those years. We needed to measure very fast processes in gas dynamics, and special amplifiers had to be built to record the signals. You eventually developed the PHERMEX radiographic facility in the United States. But in those days we just didn't have the capability of making that kind of equipment. Nonetheless, we achieved some very-good-quality results just by modifying other equipment and using materials that were available. There are many similar examples that I could mention.

Los Alamos Science: During that same period, you started to work on the types of pulsed-power experiments that are being discussed at this week's meeting with the American scientists. Is that correct?

Pavlovskii: While we were developing the first thermonuclear devices, we began working on many types of problems because the intellectual level of scientists at Arzamas was very high. As happened in America, all the great scientists in our country came to work on the project, but there was a slight difference—in the United States, they collected all of the best scientists from the rest of the world; we were limited to those in our country.

Los Alamos Science: Not all the great scientists who worked on the Manhattan Project worked on the hydrogen bomb. Many of them left Los Alamos after 1945.

Pavlovskii with John Immele, director of the Nuclear Weapons Program at Los Alamos.

Pavlovskii: Many of our best physicists left as well, but Sakharov, for example, stayed and worked for quite a long time, until 1968. During the years that he worked at Arzamas, we were extremely productive. Sakharov was a man who thought about many ideas at once. Even before he came to Arzamas in 1950, he had already begun to study cosmology and a variety of problems in quantum electrodynamics, such as the Lamb shift. He had also proposed an idea for muon-catalyzed fusion in 1948. The idea is to create molecular ions, each consisting of two deuterons held together by a muon rather than by an electron. Since the muon is 209 times heavier than the electron, it draws the two deuterons 209 times closer to each other. In fact, the two positively charged deuterons are so close together that they can fuse by quantum tunneling, form a helium atom, and release a great deal of energy in the process. He put forth the idea of using muons as passive catalysts for fusion in 1948.

In about 1950 or 1951 he and Igor Tamm proposed another idea for controlled thermonuclear fusion, namely, the magnetic thermonuclear reactor, the design of which is the progenitor of the present tokamak designs for magnetic-fusion reactors. At about the same time, he started talking about the design of magnetocumulative generators, which could concentrate magnetic field lines into small volumes using explosive compression. His original interest was to use the generator to achieve thermonuclear fusion. A little-known fact is that he later proposed using lasers to implode spherical targets filled with thermonuclear fuel and achieve thermonuclear fusion that way. That idea came to him immediately after scientists in the United States announced the invention of the ruby laser in 1960. I remember his great excitement when he returned from Moscow after conducting several seminars on the possibilities of laser fusion.

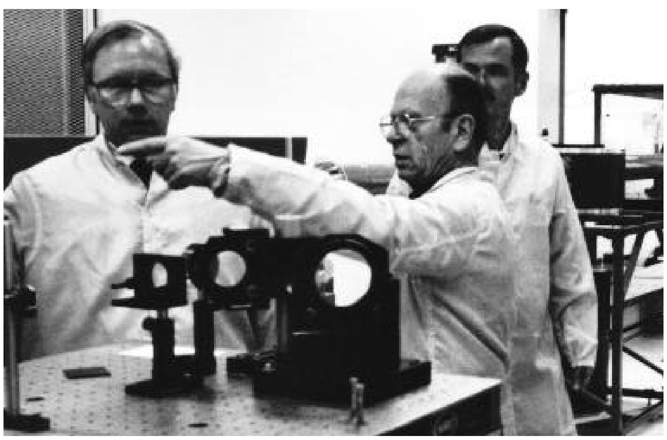
Bob Gibson with Pavlovskii at Trident laser in Los Alamos.

Los Alamos Science: How did the idea of using muons to catalyze fusion occur to him so early?

Pavlovskii: It is difficult to describe the logic of Sakharov's thinking. Many of us had great difficulty understanding even the simplest things he said. We had many seminars at Arzamas, during which Zeldovich used to serve as an interpreter between Sakharov and the rest of us. Sakharov made quantum leaps in his logic, and his words needed to be translated into much simpler language in order for everybody else to understand what he was talking about.

Los Alamos Science: Is Sakharov a hero to you?

Pavlovskii: That is a difficult question to answer. When I listed the various areas of work that Sakharov was involved in over a very short time, it was to show his unique genius and his ability to think in parallel. For example, when I would talk with him, he would stick his chin on his palm and appear to be listening intently. Then suddenly I would become acutely aware that he was thinking about something entirely different even though he never lost the train of thought of our conversation. We still have on hand his notes and papers on the magnetocumulative generators, and on the back side of these pages are some of his calculated estimates of the masses of the quarks. So, in addition to his very high moral character, he was a very unusual person. You almost had the feeling that he had mystical powers, that you were dealing with someone who lived, in part, on another plane of existence.



Steve Younger and Pavlovskii at the Trident laser.

Los Alamos Science: How did you feel when in the late 1950s Sakharov objected to the continuation of nuclear testing and then in 1968 called for disarmament and a rapprochement between the United States and the Soviet Union?

Pavlovskii: Well, my previous comments have only skimmed the surface of Sakharov's personality. He also had a peculiarity of looking at everything with a degree of innocence. He had this habit of thinking about society and government in very idealistic—you might even say naive—terms. That's why the first book he published after he left Arzamas and started public work was devoted precisely to what he called convergence, or the rapprochement between the socialist and capitalist systems.

When Sakharov was at Arzamas, I always looked at him as more of a scientist than a political man. Given the tense time and all the work we were doing, most of us didn't have any time to think about politics. It is difficult to judge whether it would have been better if he had simply remained a scientist rather than getting involved in political activities because in both fields he accomplished a great deal. The things he talked about publicly throughout his lifetime were very unusual, and the same can be said for his ideas in science. Not long before his death we met and discussed magnetocumulative experiments in great detail. But he seemed unable to separate his public life from his scientific life. Politics just seemed to have enveloped him fully.

Los Alamos Science: In the early 1970s you must have had a chance to read his report to the Central Committee of the Communist Party on the convergence of two global systems. That essay was published abroad in 1968 as "Reflections on Progress, Peaceful Coexistence and Intellectual Freedom." In it Sakharov warned of the dangers of thermonuclear war, he condemned dogmatism, terror, and Stalin's crimes, and he urged democratization and the convergence of political systems as the way to avoid global destruction.

Pavlovskii: Many people at Arzamas-16 had copies of his book in the '70s, and they read and discussed those issues. Sakharov never made a secret of the fact that he had such ideas. He always spoke freely about them, and because they were at a popular level, they could be discussed by everyone.

Los Alamos Science: In the Soviet Union were scientists pressured to work on nuclear weapons or could they make that choice independently?

Pavlovskii: For the most part, although some would tell you differently now, most people went to work on nuclear weapons voluntarily. There were some who were motivated by the slightly higher salaries we received, but the difference in payment was not very great. The majority of the scientists at Arzamas wanted to prove their abilities to their peers. In those years the climate was set for us by the great men with whom we worked. Although they were working primarily on weapons, their initiatives in other areas of science were usually well supported. For those who were truly interested in science, Arzamas provided a unique opportunity to do research in many interesting fields.

Los Alamos Science: So, if you worked on weapons, then you had other opportunities?

Pavlovskii: Yes, in those years we began to study many fields and trends of science that are still interesting today. The work on magnetocumulative generators is a good example. Sakharov wanted to use magnetic-flux compression, or "magnetocumulation," as we called it, in developing impulse accelerators that could create beams of elementary particles with high energies and high intensities. It could also be used to create super-strong magnetic fields and strong currents for a period of time long enough to study material properties under extreme pressure and extreme magnetic fields.

I became involved in that work from the beginning, and it has continued through all these years. Very similar work has been going on at Los Alamos under the direction of Max Fowler since 1953, and now we are here at Los Alamos, together with our American colleagues, discussing the possibilities for collaborative work in this area.

Los Alamos Science: Has basic research remained a strong effort at Arzamas?

Pavlovskii: Most of the people in our institute continue to do basic research as well as weapons research, and they take their basic research work very seriously. But recently it has become very difficult for us again, in some respects as difficult as it was when we first started. We lack a good industrial base, the ability to obtain equipment, and so on, just as in the early days.

Los Alamos Science: What is the economic situation of the scientists at Arzamas? Previously you were treated rather well by the State, and now?

Pavlovskii: Now that we are becoming a market economy, we don't understand what money means anymore. All the gradations have disappeared. There are now, in essence, only two categories: first, the workers, a group that is shrinking rapidly, and second, those engaged in the resale of goods, which doesn't involve any intellectual or physical prowess. In the second category are people who earn tens to hundreds of times more than the real workers. So the weapons scientists have no real privileges anymore. In fact our situation is even worse—it's turned 180 degrees.

Russian-American meeting in University House in Los Alamos.

Los Alamos Science: Will the weapons scientists be able to maintain their integrity in the present environment?

Pavlovskii: Within the overall spectrum of scientists in our country, you're always going to find a small group who could be bought or sold or coerced or subverted to become involved in other activities. The relatively young physicists are in the most vulnerable position, and there are those who want to leave the country. It is a very difficult time.

Los Alamos Science: So who is protecting and safeguarding the nuclear-weapons establishment?

Pavlovskii: The first line of defense is the people who designed the weapons in the first place. Those people are truly interested in science, and they have a rather high moral character. They understand the situation in the world today, and they know that they must play an extremely important and responsible role in controlling nuclear weapons. But it's hard to accomplish for various reasons.

During this meeting, in the evenings, I made some toasts, and we talked about the fact that laboratories such as ours, weapons laboratories, are beginning to have closer ties to one another. This development is not happening because of the will of individual scientists, but rather because of the collective feeling that we need to help stabilize the situation. There is also the need to find a colleague with whom you are of one opinion and can talk to in the same language, the language of science. That's a very important aspect of any type of collaborative efforts on the part of scientists. There are no limitations or difficulties in collaborations among scientists. It is natural. The politicians and those higher up in the power structure, however, have reservations.

Los Alamos Science: So, in your country what is the attitude about both the weapons and the weapons scientists?

Pavlovskii: It is a rapidly changing situation. At first those people who are now in power and who call themselves the democrats, despite the fact that they are still the old Communist Party functionaries, were saying, "Away with everything! Do away with these scientists who are studying weapons, do away with nuclear weapons, do away with nuclear testing!" That is an indication of the types of things that are going on in our country. But then, let's ask the question: If we do away with weapons scientists, what are we going to do with the weapons? There are many examples of that kind of question.

So basically the people in Arzamas are put in the position of trying to find contacts on their own. We feel very little involvement from the government, although they have supported us in this initiative to develop collaborations with Los Alamos. The contacts, however, were made at the lab-to-lab level. For example,


Alexander Pavlovskii

here at Los Alamos, I was very familiar with Max Fowler, Dennis Erickson, and others and with their work on pulsed power and high magnetic fields. We've met at international conferences, we've read each other's publications, and as a result we came up with the idea of starting collaborative efforts. Then, in February 1992 we had an exchange of delegations at the director level of our respective laboratories. First, the directors of our two nuclear weapons laboratories visited Los Alamos and Livermore, and then the Los Alamos/Livermore delegation came to Arzamas and Chelyabinsk. A memorandum was signed, stating that we had agreed to engage in collaborative efforts on a wide range of scientific topics. In my opinion, those topics are very important. They include questions on nonproliferation, storage of nuclear weapons, destruction of weapons, ecological problems, and nuclear-energy safety concerns. I'm not going to list them all, but there are a great many scientific areas that we want to talk about.

Los Alamos Science: Are you free to talk about anything, even if it's about nuclear weapons?

Pavlovskii: With whom?

Los Alamos Science: The American weapons scientists.

Pavlovskii: No. In fact, neither we nor our American counterparts will discuss, for example, bomb designs with each other because we have an obligation to safeguard that knowledge. Just as in the United States, the Russian government has issued an entire series of rules and regulations about safeguarding nuclear weapons.

Los Alamos Science: People here are worried that other countries, such as Libya or Iran or Iraq, may make overtures to the Russian weapons scientists. Do you know of any such instances?

Pavlovskii: No, I don't know of any actual instances.

Los Alamos Science: What was the reaction in the former Soviet nuclear weapons community to President Reagan's Strategic Defense Initiative proposal?

Pavlovskii: We perceived there were some scientific opportunities in SDI, but as for its chances of success, you should perhaps ask the American authors of SDI. When I was a young man, I used to play with electron beams and shoot things with them, but that was when I was a boy.

Los Alamos Science: The Russians weren't worried about SDI?

Pavlovskii: No. I had some conversations about SDI with American representatives when I was at the Beams '92 conference in Washington at the end of May. They had asked me to give a talk, and afterwards I asked them several questions about SDI, and I understood from their answers that SDI is a dead end.

Los Alamos Science: Were the Russians working on similar technology?

Pavlovskii: Sure, there was some work going on in this area, but it was looked at from a different perspective—more as research-type work. It has many purposes, so it's not really a total waste of time. It opened up some interesting directions of research.

Los Alamos Science: In the research areas that you shared this week, did you feel that the Russian work was superior to the American work?

Pavlovskii: That's really not the proper question. We've always had very talented mathematicians and theoreticians in Russia. For experimentalists, however, life is a little harder. In addition to paper and pencil, they need all kinds of equipment, and in our country there is a lack of necessary equipment. Nevertheless, in many areas they were able to overcome these difficulties and come up with some very good results. So let's consider the work that was discussed at the meetings we had here this week. While American scientists have an obvious interest in helping Russian scientists, the fact remains that in the magnetic-field-generator work, for example, we have been able to produce the highest magnetic fields in the world. And this is not the only area in which we've been able to achieve a lot with little.

Los Alamos Science: Are your best students going into physics?

Pavlovskii: Unfortunately not. The process is different now. Given the circumstances in our country, the morale of our people more or less dictates where the youth goes. The young people who are subverted by cars, by toys as it were, see it would be far too hard to get those things through a career in science. They want instant gratification. They go into what we call "business." The new word for it is business, but I think the old word for it was primitive speculation, black markets—let me finish, hold on—I don't think it's just a Russian problem, it's now becoming a world-wide problem. In many respects, the level of intellectuality has dropped significantly all over the world.

Los Alamos Science: Once again, science in the United States is becoming dependent on foreigners. In many areas of science, the ranks are being filled by students and postdoctoral fellows from other countries.

Pavlovskii, Mikhail Dolotenko, and Alexander Petrukhin overlooking Pueblo Canyon.

Pavlovskii: One of the first things I noticed when visiting the United States was that about half of the best mathematicians in American institutes are Russians. I'm not saying that the directors of your institutes are all going to be foreigners, but when I ask American scientists about the quality of the Russians who are working, teaching, and studying at American universities and institutes, they always express the highest opinion of their Russian colleagues. ■

SCIENCE AND INNOVATION *at Los Alamos*

.

Fred Reines (left) helps lower Wright Langham into a detector similar to the one used by Reines to detect neutrinos for the first time. The active medium of the detector was a liquid scintillator developed by F. Newton Hayes for assays of large biological samples. The availability of liquid scintillators led to the whole-body counter, a device for monitoring the amount of certain radionuclides in the bodies of workers exposed to radioactive materials. Wright Langham was one of the world's experts on the metabolism of plutonium.

Lattice-gas hydrodynamics, a discrete model for fluid flow, was invented by Brosl Hasslacher at Los Alamos with U. Frisch and Y. Pomeau. This novel formulation provides a fast, efficient, reliable method for simulating the Navier-Stokes equations and two-phase flow. A modification by Ken Eggert and coworkers is now being applied to model flow through porous media, a problem of great interest to oil companies.

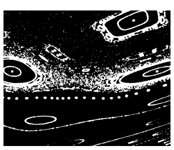
Norman Doggett and Judy Tesmer examine a gel at the Laboratory's Center for Human Genome Studies. The Human Genome Project, a joint DOE-NIH effort, was largely conceived at a DOE meeting in Santa Fe in 1986. Researchers at the Los Alamos Center developed a widely used technique for fingerprinting DNA, discovered the human telomere (the sequence at the ends of every human chromosome), are developing physical maps of several human chromosomes, and are preparing chromosome-specific libraries of clones, which are extremely useful in physical-mapping projects.

The Perhapsatron, the doughnut-shaped discharge tube shown here, was built by Jim Tuck in the early 1950s at Los Alamos. It marked the beginning of the United States effort to achieve controlled thermonuclear fusion by confining a hot, dense plasma with magnetic fields. Now the United States is part of an international effort to reach breakeven conditions in magnetically confined fusion.

An international team, including Los Alamos scientists, prepares to take measurements in a borehole in Greenland to look for deviations from Newton's law of gravity. The experiments involved both particle physicists and geophysicists, and led to further measurments in the ocean that found no deviation.

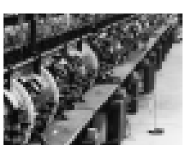
Norris Bradbury (left) and Stan Ulam (right) are shown at the Rover test site. Project Rover, a joint effort of the Laboratory and the Air Force begun in 1955, successfully developed technology for nuclear rocket propulsion suitable for sending large manned or unmanned payloads to planets beyond Mars or outside the solar system. The high-specific-impulse exhaust is generated by passing hydrogen through a very high-temperature nuclear reactor. Russian scientists picked up this technology when it was declassfied and have now offered to share their further developments with us.

The Bright Source, an intense, short-pulse excimer laser, was developed by Los Alamos scientists to study weapons physics, ultrafast chemical reactions, and atoms under extremely high electric fields. Lasers for fusion, isotope separation, environmental monitoring, and nuclear nonproliferation have also been developed at Los Alamos.



Complex shapes, generated by iterating a simple area-preserving transformation starting from fifty points in the plane, illustrate the groundbreaking work of Mitchell Feigenbaum on the emergence of complex, chaotic behavior in simple deterministic physical and mathematical systems. Feigenbaum's research in the late 1970s was firmly within the tradition that began with the "numerical experiments" of Fermi, Pasta, and Ulam in 1952 and continues today at the Laboratory's Center for Nonlinear Studies.

Los Alamos, under a funds-in agreement with Phillips Petroleum Company, is recording microseismic activity to map rock fractures in the giant Ekofisk oil field 250 miles off the coast of Norway. The formations are subsiding after significant petroleum production. The measurements will assist Phillips in developing production strategies. The techniques for mapping rock fractures were originally developed as part of the Laboratory's geothermal-energy research projects.

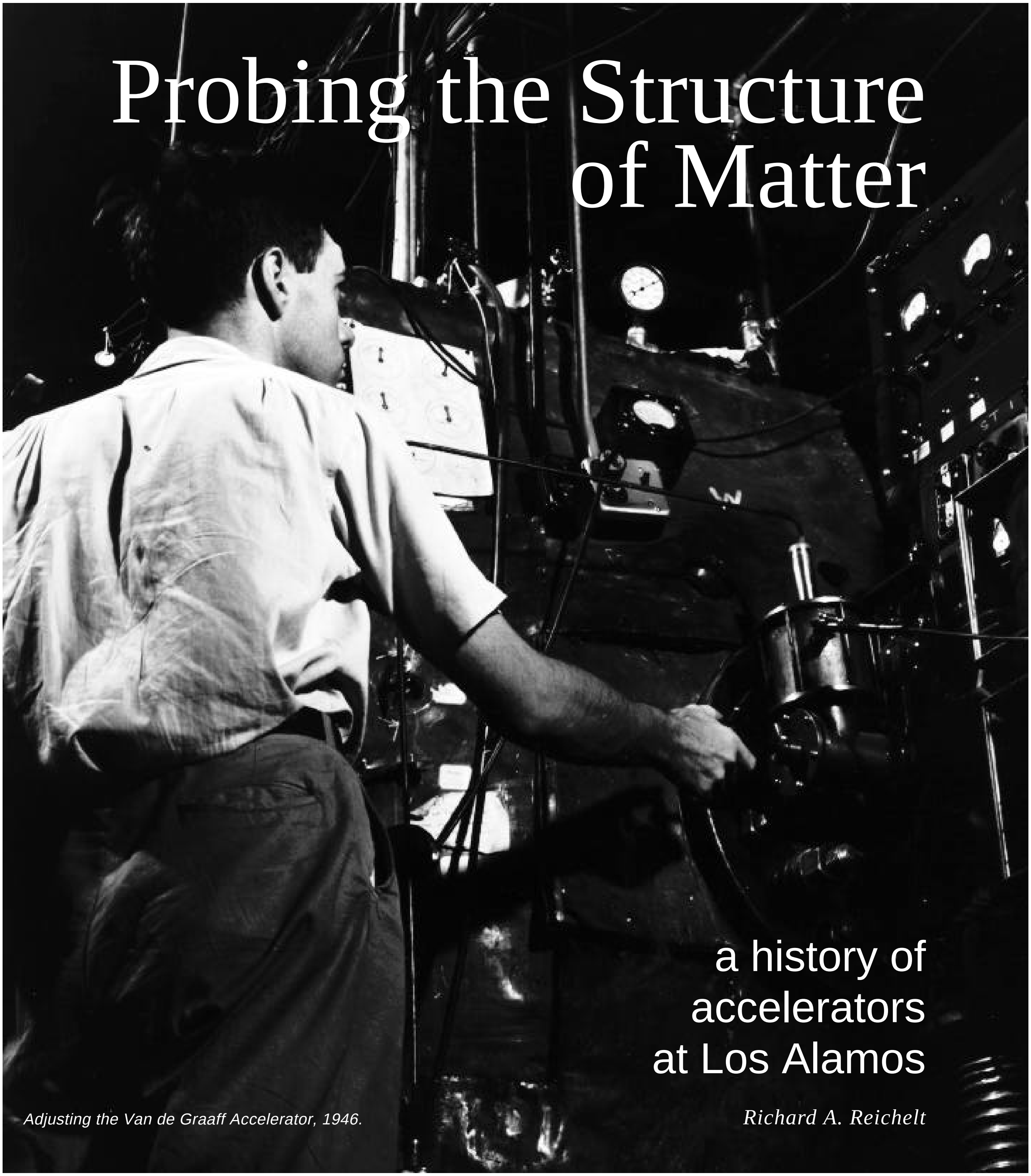


Side-coupled cavities at the Los Alamos Meson Physics Facility accelerate highly intense pulses of protons to 800 MeV . Laboratory scientists invented the side-coupled cavity in 1967 specifically for use in LAMPF, but it also revolutionized the design of x-ray-therapy machines and other medical linear accelerators.

In 1947 Enrico Fermi used this trolley, an analog Monte Carlo computer, to simulate the paths of neutrons (solid lines) in a weapon or a reactor. Nick Metropolis, Stan Ulam and John von Neumann developed the first digital Monte Carlo programs for the Manhattan Project. Later Monte Carlo developments at the Laboratory include the Metropolis method and the MCNP code to simulate neutron and particle transport, which is widely used in radiation safety, nuclear safeguards, reactor design, fusion engineering, medical radiation therapy and diagnostics, and even oil-well logging. The lattice simulations of quantum field theory discussed by Rajan Gupta in this issue also employ the Monte Carlo method.

The MANIAC I, the second large-scale electronic computer, was designed and constructed at the Laboratory under the direction of Nick Metropolis and became operational in 1952. It used vacuum tubes for logic and cathode-ray storage tubes for memory (each connected to a neon light to show whether it stored a 1 or a 0). The MANIAC I was used for hydrodynamic and other weapons-design calculations, a variety of scientific problems, and the development of a high-level programming language and operating system.

Probing the Structure of Matter



a history of accelerators at Los Alamos

Richard A. Reichelt

Adjusting the Van de Graaff Accelerator, 1946.

The history of Los Alamos is intimately linked to machines. Machines of all types—from reactors to computers to lasers—have been indispensable in advancing scientific knowledge at the Laboratory. Among all of those machines, accelerators occupy a special position. During World War II accelerators provided vital information to scientists designing the first nuclear weapons, and after the war some of those machines were used as tools of basic scientific research. Then, in the early 1970s, a half-mile-long accelerator called LAMPF (formally named the Clinton P. Anderson Meson Physics Facility) was completed after a decade of planning and construction. Over the years LAMPF has been the workhorse for many Laboratory programs, including programs in nuclear physics, weapons research, neutron scattering, radioisotope production, and pion cancer therapy. Accelerators for more specialized purposes have also been developed or improved by Los Alamos scientists.

The constant effort to upgrade and redesign accelerators has often led to unexpected results. New technological spin-offs, new vistas of scientific research, have opened up as one generation of machines replaced another. This photoessay briefly traces the evolution of accelerators at Los Alamos and examines their different applications. In the last section the promise of a future generation of machines is examined.

First a few words of explanation. Most accelerators work according to the same basic principle. An electric field is applied to a stream of charged particles (typically electrons or protons) and accelerates the particles to greater and greater energies. Those accelerated particles can then strike a target and interact with the target atoms. If the particles have a high enough energy, they interact with the nuclei of the atoms, often yielding "secondary" particles whose identities depend on the energy of the primary particles and on the target material. Sometimes the accelerated particles are only a means of producing the secondary particles (neutrons and pions are examples of secondary particles). Although the basic principle of particle acceleration is simple, its execution is not. Designing an accelerator requires creativity and ingenuity, and operating an accelerator can be an art in itself.

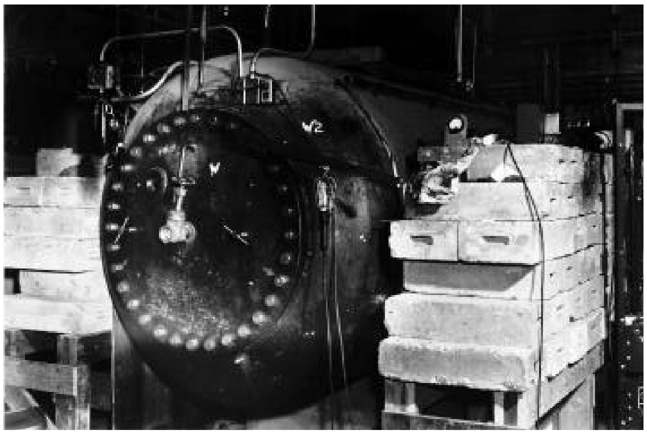
Technical Building Z. *This shack was built for the Cockcroft-Walton accelerator in 1943.*

Most modern accelerators are classified as either linear accelerators (linacs) or circular accelerators. The linac design, in which the accelerated particles move in a straight line, was first developed by Rolph Wideröe in 1928 and later refined by Luis Alvarez. The circular machine, in which the particles move in a circular path, was conceived by Ernest Lawrence in 1929. He called his invention the cyclotron.

World War II Accelerators

J. Robert Oppenheimer, director of Project Y during World War II, was responsible for bringing accelerators to Los Alamos. He believed that only by consolidating machines and scientists in one location could a nuclear bomb be speedily built. And speed was important—the Germans were thought to be well advanced in developing a nuclear bomb. Oppenheimer envisaged concentrating the work of designing and building the the first nuclear bombs in one location so that ideas and findings could easily be exchanged. Accelerators, as suppliers of nuclear data, were urgently needed to provide an experimental foundation for the work.

Thus, in the spring of 1943, four bulky machines were transported to a remote New Mexico location from



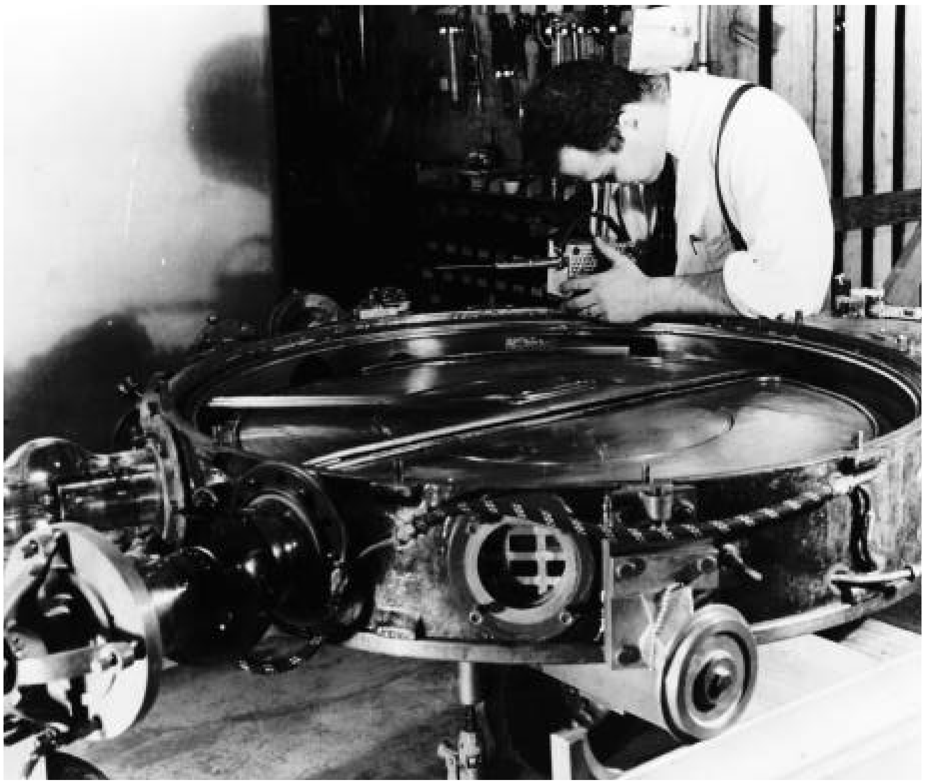
The Wisconsin Short Tank. *One of five accelerators commandeered for use at Los Alamos during World War II, the Short Tank was an improved version of a Van de Graaff accelerator and was designed mainly by Joseph McKibben at the University of Wisconsin. The lead shielding and concrete blocks surrounding the pressure vessel served as radiation protection.*

universities across the country. To throw curious observers off the track, the accelerators followed a circuitous route. They were first diverted to a medical officer in St. Louis and then shipped in boxcars to Santa Fe. Finally the accelerators were moved on flatbed trucks to Los Alamos, just as some of the world's most eminent scientists were beginning to gather there. Massive steel pressure tanks for some of the machines arrived during technical conferences in mid April.

The accelerators, it was hoped, would help scientists tackle two major challenges. The first was to determine the critical masses of the proposed nuclear fuels—plutonium (^{239}Pu) and uranium highly enriched in the isotope ^{235}U. (The critical mass is the smallest mass of nuclear fuel of a certain shape necessary to sustain a chain reaction.) Both fuels were so scarce that direct measurements of the critical mass were impossible; production techniques for the fuels were only just being developed. The second challenge was to find a way of preventing a "fizzle," or predetonation, in the plutonium bomb—a problem arising from the spontaneous fission of fuel impurities. Of the two problems predetonation was the more intractable and the source of pessimism about the feasibility of a plutonium bomb.

The accelerators made it possible for scientists to determine the critical masses for each proposed bomb design. The machines supplied neutrons for studying the neutron interactions involved in an explosive fission chain reaction. At that time neutron-induced fission—the process by which a neutron causes the nucleus of a heavy atom to split and release energy and more neutrons—had not been studied at all the relevant neutron energies. Several months earlier, in December, 1942, Enrico Fermi had induced the first fission chain reaction at the University of Chicago using a "pile" consisting of enriched uranium interleaved with graphite moderators. The moderators were designed to slow down neutrons and make the fission reaction more efficient. But fission induced by fast neutrons was expected to be less efficient than fission induced by slow neutrons. In an explosive chain reaction there could be no moderators; the neutrons would emerge from fission at high energies—between 0.1 and 3 MeV (1 MeV = 1 million electron volts)—and travel unimpeded until they collided with other heavy nuclei. How would the fast neutrons interact with heavy nuclei? What percentage of those neutrons would

The Illinois Cockcroft-Walton Accelerator. *This accelerator was used by John Manley and his group to investigate the efficacy of different metals as a "tamper"—a liner surrounding the nuclear explosive that acts as a neutron reflector and makes the explosive chain reaction more efficient. Gold, uranium, platinum, tungsten, and other metals were investigated as possible tampers.*



A Portion of the Harvard Cyclotron. *A group led by R. R. Wilson investigated nuclear reactions induced by the low-energy neutrons provided by this 42-inch cyclotron.*

cause fission? How many neutrons would be released in a fission reaction caused by a fast neutron? What percentage would be reflected back into the fissile material from a metal liner—or "tamper"—surrounding the core? As sources of neutrons with various energies, accelerators were just the tools scientists needed in order to study the nuclear reactions relevant to weapons physics.

Within a few months of their arrival, all four accelerators were producing neutrons in hastily erected wooden buildings. A Cockcroft-Walton accelerator requisitioned from the University of Illinois produced 2.5-MeV neutrons. Two Van de Graaff accelerators from the University of Wisconsin produced neutrons with energies between a few hundredths of an MeV and several MeV. And a cyclotron from Harvard University produced neutrons with even lower energies. Together the accelerators produced neutrons with energies spanning the pertinent energy spectrum.

The summer of 1944 signaled an abrupt shift in the program at Los Alamos. In mid April Emilio Segré and his band of graduate students had discovered some bad news—that ^{240}Pu (an isotopic impurity present in ^{239}Pu) had a high cross section for spontaneous fission. The original plan for initiating an explosive fission chain reaction in either a plutonium or a uranium bomb had been to use a gun mechanism to fire one subcritical chunk of nuclear fuel into another subcritical chunk. But if the fuel was plutonium, the neutrons produced by the spontaneous fission of ^{240}Pu were likely to initiate a chain reaction in the two chunks before the gun mechanism could bring them together into a supercritical mass. The result would be a fizzle. Therefore a gun-type plutonium weapon was out of the question. So in August the Laboratory threw its resources into achieving criticality in a plutonium weapon by implosion, that is, by detonation of a layer of conventional explosives surrounding a subcritical sphere of plutonium. The idea was that the inward force of the conventional explosion would compress the plutonium sphere and thereby create a supercritical mass, provided the implosion was sufficiently symmetric. The implosion option had been pursued previously—early experiments had involved detonating TNT wrapped around iron pipes—but implosive forces of the required symmetry had not been achieved.

To help in the design of an implosion weapon, a betatron—a circular electron accelerator— was procured in December, 1944. Pulses of x rays produced by the betatron were used to obtain a sequence of images of a sphere of mock fuel as it was being imploded by a particular configuration of high explosive. This diagnostic technique, along with others, helped solve the problem of uneven collapse in the implosion weapon.

On July 16, 1945, a plutonium weapon was tested near Alamogordo, New Mexico. On August 6 a uranium weapon, Little Boy, was dropped on Hiroshima, Japan, and on August 9 a plutonium weapon, Fat Man, was dropped on Nagasaki. Soon after those bombing attacks Japan surrendered unconditionally. Oppenheimer's idea of consolidating scientists and machines at Los Alamos had simultaneously unleashed a horribly destructive force and helped to bring about an end to World War II.

Postwar Developments

Norris Bradbury, who became director in October, 1945, had the job of charting the direction the Laboratory would take in the postwar decades. His primary concern was to transform an institution that had been built for a short-term purpose—designing and building nuclear bombs—into an institution with long-range purposes and goals. Should Los Alamos become a nuclear-bomb factory? Or should it cease weapons work altogether? Or should it build a foundation of basic scientific research with weapons applications? Recognizing the coming competition with the Soviet Union and wanting to avoid "technological surprise," Bradbury (along with others) decided in favor of the last option. He also believed that such a program of basic scientific research would help to retain the cadre of talented individuals whose mentors had been among the brightest scientists of the twentieth century. As part of the experimental foundation for that new program, three wartime accelerators were purchased by the government—the Short Tank, the Cockcroft-Walton, and the cyclotron. The Long Tank was returned to the University of Wisconsin.

A high-energy Van de Graaff accelerator, a vertical model designed by Joe McKibben, was built to replace the Long Tank; it provided monoenergetic neutrons with energies up to approximately 8 MeV. Those high-energy neutrons and the 14-MeV neutrons provided by the Cockcroft-Walton were used to study neutron interactions relevant to nuclear fusion. The old Harvard cyclotron was upgraded into a variable-energy cyclotron that could accelerate different kinds of charged particles. Additionally, the cyclotron group developed a special camera that recorded the angular distribution of the accelerated particles after they had been scattered by the nuclei within a particular target element. Such data provide information about the energy levels of the target nucleus. Scientists' wives, many with university degrees, were enlisted to scan the photographs, becoming a team of first-rank nuclear spectroscopists.

In August, 1949, the Soviets exploded their first nuclear bomb. President Truman subsequently announced that the United States was embarking on a program to build a variety of nuclear weapons, including a fusion, or thermonuclear, bomb. Such a bomb utilizes a fission bomb to trigger the fusion of deuterium and tritium nuclei; its explosive yield is many times higher than that of a fission bomb alone. Actually, research on a fusion weapon had been pursued at Los Alamos continuously since the war years. Those early efforts were vital to the success of tests, called the Greenhouse series, that led to the first thermonuclear reaction—the George shot—in 1951.

Diagnostic tools for weapons were also developed during the early Bradbury years. Two electron linacs were built to provide radiographs of the implosion process. That work eventually led to the construction in 1963 of PHERMEX (pulsed high-energy radiographic machine emitting x rays), a huge electron accelerator, housed in a concrete bunker, that generates x rays by accelerating an electron beam onto a tungsten target. The x-ray bursts are sent through model weapons at a remote blasting site and provide three-dimensional pictures of imploding spheres. PHERMEX was also used to study fluid dynamics and the behavior of matter under extreme, shock-driven conditions. The origins of PHERMEX were in the pioneering World War II work done with the betatron. Still in operation, PHERMEX has recently been used to study the strength of ceramic tank armor.

Ten years after the war Los Alamos was still the foremost nuclear-physics laboratory in the world. Bradbury's program of basic scientific research had fed into applied fields in many ways—nuclear spectroscopy, optical modeling, and the thermonuclear weapon. But the world was catching up, and the wartime accelerators were not state-of-the-art for studying nuclei and nuclear forces. In 1946 the first proton linac—built at the University of California by Luis Alvarez, a physicist who had been involved in the nuclear-bomb project—had became operative. A linac accelerates charged particles with a series of electrical "pushes," each of which increases the energy of the particles by an amount that is small compared to the total energy gain desired. Alvarez used radar oscillators developed during the war to produce the accelerating electric field—a radio-frequency oscillating electric field—in a single long resonant cavity. Along the length of the cavity were forty-five "drift tubes" that prevented the protons from being decelerated during the negative phase of the electric field. The Alvarez design would have tremendous implications for accelerator development at Los Alamos.

An Aerial View of LAMPF in 1983.

LAMPF

LAMPF is a half-mile-long linear accelerator built atop a narrow mesa not far from Los Alamos. In 1983 Louis Rosen, the chief architect of LAMPF, described the original motivation for building the massive accelerator as follows:

T*he most fundamental reason we advanced [for requesting funds of Congress] stemmed from a belief, still held today, that eventually this country and the entire industrialized world will be forced, whether they like it or not, to a nuclear-energy economy. … We are simply running out of conventional organic sources* *of energy. … If a nuclear-energy economy is inevitable, how can we make it as efficient and safe as possible? With advanced nuclear technology. The reason we are doing so badly now with power reactors is that we didn't start with enough technology. And because technology is the child of science, we need a strong science base. Without that base, technology cannot advance and will soon dry up. We felt that the need for science as the basis of technology was the most compelling reason for Los Alamos to engage, at a very high level, in basic nuclear science.*

Rosen first proposed the idea of building the accelerator in 1962. In a memo to J. M. B. Kellogg, then leader of the Physics Division, he sketched the scientific importance of a "meson factory"—a new, high-intensity proton accelerator that would supply an abundance of pi mesons to study nuclear interactions and the structure of nuclei. Pi mesons, or pions, are short-lived particles that can be created by firing protons accelerated to nearly the speed of light at light-element targets. Since most accelerators at that time were being designed to achieve higher energies per particle rather than higher beam intensities (or numbers of particles per unit time), Rosen thought that a meson factory could open an entirely new realm in nuclear physics.

The memo eventually reached Bradbury and began to generate enthusiasm among scientists at Los Alamos. A small group of experimental and theoretical physicists began looking at two possible accelerator designs—the cyclotron and the linac. Both designs had disadvantages and drawbacks, but the scientists eventually decided in favor of the linear accelerator. The cyclotron couldn't achieve the necessary proton energy without an unacceptable loss in beam intensity; a then unheard-of intensity of 1 milliamperes was desired.

Victory Day: June 9, 1972. *Louis Rosen and others watch control instruments as the LAMPF linac produces its first beam of 800-MeV protons.*

The accelerator was envisaged in three stages. In the first stage a Cockcroft-Walton would accelerate the protons to a low energy (0.75 MeV). In the second stage an Alvarez-type drift-tube linac would accelerate the protons further to a medium energy (100 MeV). The third and final stage would accelerate the protons to 800 MeV. But no one yet knew how to accelerate protons to energies higher than 200 MeV.

In the face of considerable skepticism from outside experts, a small team of scientists led by Darragh Nagle and Ed Knapp began the search for a suitable acceleration scheme. They investigated many different cavity designs, including pillbox structures, cloverleaf structures, and slow-wave helices. Then, in early 1965 they perfected a design, known as the side-coupled cavity, that could accelerate high-intensity beams of protons to 800 MeV.

In 1967 a working prototype, called the Electron Prototype Accelerator, demonstrated the viability of the new cavity design. The demonstration had immediate practical consequences. Several manufacturers of x-ray machines for the medical community began incorporating the side-coupled cavity design into their new models. The result was smaller, more efficient machines. Today the side-coupled cavity is recognized as having revolutionized x-ray-therapy machines and other medical linear accelerators. Incidentally, the side-coupled cavity was never patented by its inventors.

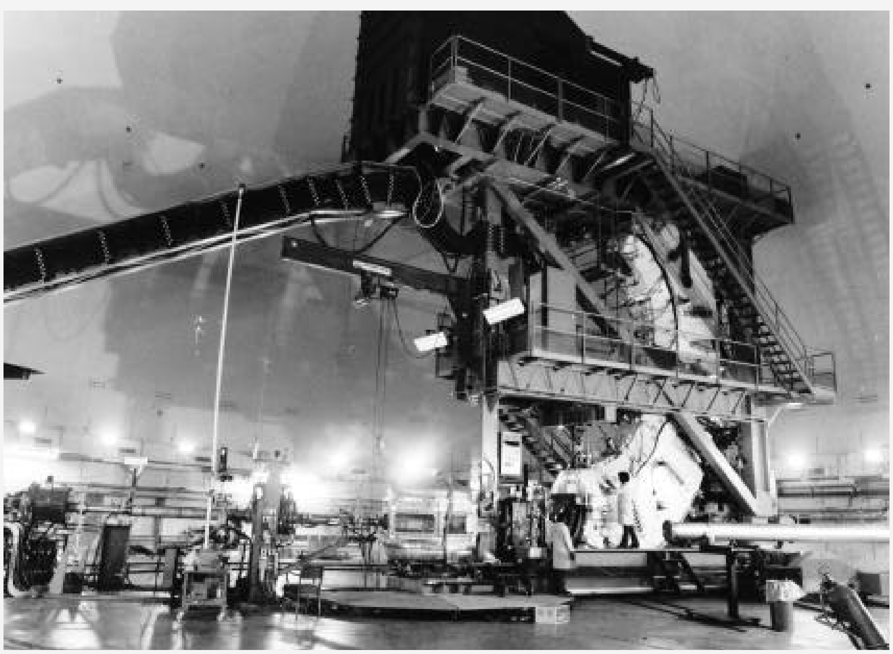
The groundbreaking ceremony for LAMPF was held on February 15, 1968. Four years later, after numerous Congressional funding battles led by Rosen, the facility was completed. LAMPF is operated as a national user facility; that is, beam time is shared among many individuals from both the United States and abroad. An international committee of experts evaluates requests for beam time and advises LAMPF's director on their scientific merit. The accelerator has spawned a remarkable number of research programs. In the area of pure science, LAMPF acts as a unique bridge between the particle-physics and nuclear-physics communities, enabling each to understand the other's methodologies and modes of thought (see "Medium-Energy Physics at LAMPF").

Medium-Energy Physics at LAMPF

Mikkel B. Johnson

The LAMPF accelerator provides primary beams of protons and negatively charged hydrogen ions as well as secondary beams of neutrons, pions, muons, and neutrinos. The uniqueness of LAMPF as an experimental facility derives from the high intensity of those beams and the consequent capability to exploit rare reactions to answer specific questions about particles and nuclei. Additionally, the high resolution of the spectrometers and other detectors available at LAMPF make precision measurements feasible. As a result, over the last twenty years LAMPF has helped to open up an entirely new field of basic research—medium-energy nuclear physics.

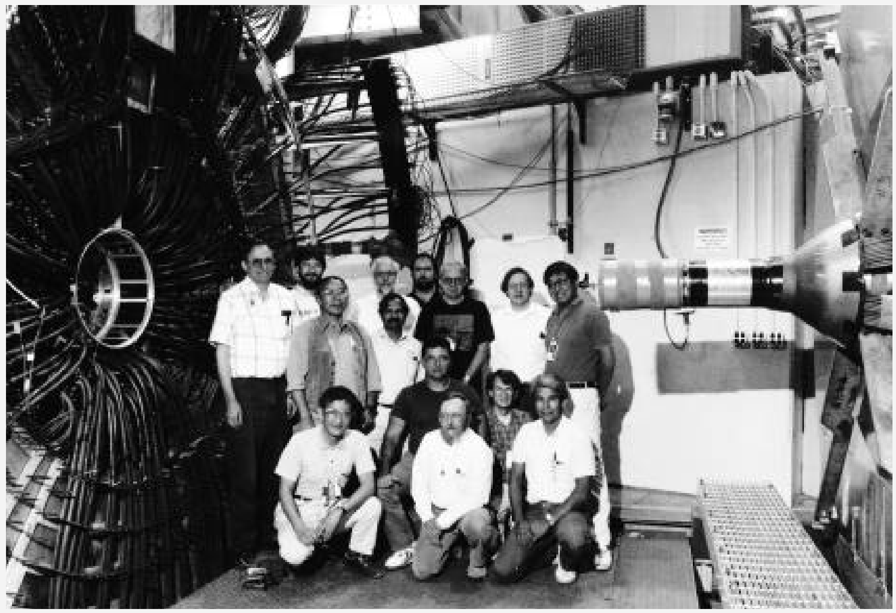
By using the tools available at LAMPF, nuclei can be explored in new ways. Experiments with muons, pions, and nucleons have quantified the size, shape, and composition of nuclear states and measured the response of nuclei to the addition of charge, energy, and momentum, as well as quanta of vibrational and rotational excitation. Some of the most interesting results have been obtained when the incoming particles cause the nuclei to respond in certain extreme ways.



The High-Resolution Proton Spectrometer at LAMPF. *This instrument, known as the HRS, is used to measure cross sections for elastic and inelastic scattering of protons from nuclei. Careful measurements with the HRS of the spin dependence of cross sections led to the widespread acceptance of relativistic descriptions of nuclear dynamics. HRS results also stimulated theoretical and further experimental investigations of how the nuclear medium affects the nucleon-nucleon interaction.*

For example, in a pion double-charge-exchange experiment, nuclei can be forced to accept two units of charge at one time. Measurements of the dependence of the scattering cross section on the final state of the nucleus have led to a strikingly detailed picture of correlations in the motion of neutrons and protons. In addition, double-charge-exchange experiments have uncovered new modes of motion of nuclei in which two patterns of vibration coexist and have also led to the discovery and study of new nuclear species. In experiments with nucleon beams, nuclei have been forced to accept a small amount of energy and at the same time a large momentum in various spin configurations hitherto incapable of being distinguished. For reasons that are not completely understood but are being actively sought, theories that work well in more normal situations have been discovered to break down under those unusual conditions. And state-of-the-art studies in atomic physics have addressed previously inaccessible regions of the spectrum with a unique technique that combines a beam of laser light and a beam of negatively charged hydrogen ions.

Additionally, various scattering experiments and measurements of the decay products of muons and pions have focused on the nature of the underlying strong, weak, and electromagnetic interactions. Precision measurements with muons have yielded new insights into quantum electrodynamics. A systematic program extending over many years has completely mapped out the character of the nucleon-nucleon interaction over the entire energy range of LAMPF and has thus provided a bank of basic data for theorists and laid the foundation for interpreting nucleon-nucleus scattering experiments. An experiment using the world's most intense source of very-low-energy neutrons has led to development of a completely new technique for detecting a signature of breakdown of a fundamental symmetry in nuclear forces. The technique, which uses properties of complex nuclei to magnify the signal for the breakdown of parity (the mirror symmetry between left and right) by a factor of about 1 million, has uncovered unexpected results and opened a rich area of exploration. Other investigations of fundamental interactions include the scattering of neutrinos from various targets. Neutrinos interact so weakly that even experiments using the high-intensity neutrino beam available at LAMPF require several years to complete. One such experiment has provided the only available measurement of electron neutrino–electron scattering. The experimental results showed that the interplay predicted by the standard model between the charged and neutral parts of the electroweak interaction was indeed a reality. Therefore, since only the neutral part of the weak force is involved in the interaction of electrons with the muon neutrino or the tau neutrino, the interaction of the electron neutrino with electrons is fundamentally different from the interaction of the other neutrinos with electrons. That difference provides a possible explanation for the observed shortfall in electron neutrinos coming from the sun. Yet other experiments at LAMPF hunt for breakdowns of the standard model. Although none has been detected, the searches at LAMPF for the decay of a muon into an electron and gamma rays, which would herald such a breakdown, have consistently led the world in sensitivity.



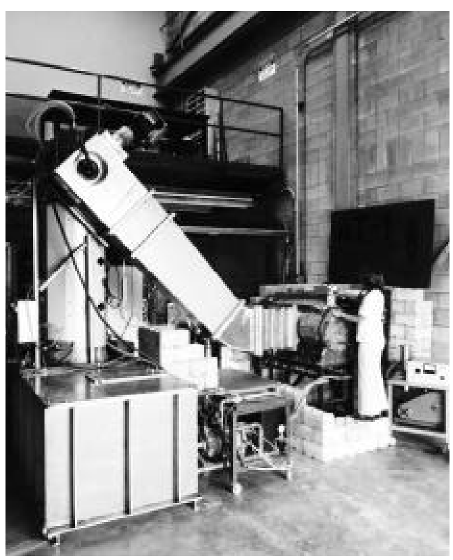
The MEGA Detector at LAMPF. *This detector will be used in a search, scheduled to begin this summer, for the decay of a muon into an electron and a gamma ray. The occurrence of that reaction would signify a breakdown of the standard model. The sensitivity of the MEGA detector to the decay is two orders of magnitude greater than that of detectors used in previous searches.*

Experiments such as those mentioned above constitute some of the highlights of the contribution of LAMPF to nuclear science. They have provided answers to many specific questions and at the same time have paved the way for a slow but very important transformation in the way nuclear physicists think about their subject. Before the era of the meson factory, nuclei could be largely understood as a collection of nucleons undergoing nonrelativistic motion and interacting through potentials. That picture is no longer adequate to describe what the medium-energy beams "see" of nuclei. To understand the new data, the catalogue of constituents of nuclei has been enlarged to encompass mesons and excited states of nucleons themselves. Additionally, the picture of the dynamics of their motion has changed. Relativity can no longer be ignored, and interactions must be described in terms of the coupling of mesons to nucleons. Even today the picture is continuing to evolve as particle and nuclear physicists realize deeper connections between their once quite distinct fields. □

Mikkel B. Johnson, a Laboratory Fellow, has pursued research in theoretical nuclear physics at the Medium Energy Physics Division since 1972.

From the beginning it was recognized that LAMPF could be a source of particles other than pions, in particular, of neutrons. Neutrons are produced at LAMPF through “spallation”—a nuclear reaction in which neutrons are knocked loose as a heavy nucleus breaks apart, or spalls, after being impacted by energetic protons. Consequently, in the early 1970s the Weapons Neutron Research Facility was built as an adjunct to LAMPF. A unique feature of the facility is a 30-meter-diameter ring of magnets called the Proton Storage Ring, a device that combines a long train of short proton pulses into a single equally short but much more intense proton pulse. Any 800-MeV proton that enters the magnetic field created by the magnets is forced to travel around and around the same circular path. Furthermore, the time required for an 800-MeV proton to travel once around the circular path is the same as the time between the short pulses that make up the long train of pulses from LAMPF. Therefore, at the instant that one proton pulse has traveled once around the circular path, a second pulse enters the magnetic field and melds with the first. Similarly, at the instant that the two melded pulses have traveled once around the circular path together, a third pulse enters the magnetic field and melds with the other two. The process is allowed to continue until all the short pulses in a pulse train have been combined into a single intense pulse. The intense pulse is then “kicked” out of the magnetic field and aimed at a tungsten target. Reactions between the protons and nuclei within the target create an intense burst of neutrons with a wide range of energies—valuable tools for studying weapons physics.

But other uses also were planned for the neutron pulses. Neutrons, unlike x rays, are scattered hardly at all by the electrons of atoms, but they are scattered by atomic nuclei. And those scatterings provide information about the structures of solid materials and of large molecules, including biological molecules, in solution. For that reason the neutron is an invaluable tool in condensed-matter physics, materials science, and biophysics. In 1986 the Department of Energy, acting in concert with various national committees, designated the WNR’s neutron source as a national user facility for neutron scattering. The new facility is known informally as LANSCE and formally as the Manuel Lujan, Jr. Neutron Scattering Center. Interestingly, it is now the research at LANSCE that might be the key to LAMPF’s future, as will be discussed below.



The PIGMI accelerator.

New Accelerator Technologies

Harold Agnew became Laboratory director in 1970. As a result of LAMPF’s successes, Agnew decided that research on accelerator technology should receive special emphasis. A new division, called the Accelerator Technology Division, was formed in 1978. AT Division was tasked with developing new accelerator designs, and the older Medium Energy Physics Division continued operating the LAMPF accelerator, developing user programs, and pursuing nuclear- and particle-physics research.

Under the leadership of Ed Knapp, AT Division soon built prototypes of two innovative linear accelerators with medical and industrial applications. One, called PIGMI (pion generator for medical irradiations), had the potential of leading to a small accelerator—about 500 feet long—for use in cancer-treatment programs. The other was built for the Fusion Materials Irradiation Test Facility at the Hanford Site in Richland, Washington. That prototype was designed to produce neutrons with which to test different wall materials of planned fusion reactors. Both the PIGMI and FMIT accelerators employed a new acceleration device called a radio-frequency quadrupole cavity (RFQ).

The history of the RFQ is instructive because it demonstrates the way attempts to upgrade a technology can lead to new and unforeseen scientific developments. The RFQ was originally conceived by the Soviet physicists Kapchinskii and Teplyakov in 1969. It remained unknown in the West until 1977, when a Russian-educated Czech refugee named Joe Manca began working at the Laboratory. Both he and Don Swenson, who had learned about the RFQ at a Russian international conference, kept telling colleagues about the new, very efficient device for accelerating charged particles to

An RFQ Cavity. *The accelerator is shown here in front of the accelerator it may someday replace—the LAMPF Cockcroft-Walton.*

low energies. At first few people took the pair seriously, but development work at the Laboratory by Dick Stokes and others proved them to be correct. The first RFQ outside of the Soviet Union was first operated as part of the PIGMI prototype in February, 1979.

The RFQ marked an abrupt departure from previous low-energy acceleration devices. It could accelerate almost 100 percent of the beam from an ion source. And unlike the drift-tube linac, which employs a magnetic field to focus a beam, the RFQ employs rapidly alternating electrical fields to both focus and "bunch" the beam. The well-defined bunches are completely matched for follow-on acceleration devices.

Soon the RFQ played a role in a growing technological field—space weaponry. During the buildup of the Soviet arsenal in the 1970s, space weapons were contemplated as a means of non-nuclear defense. Could a beam of laser light or a beam of neutral hydrogen atoms be fired thousands of miles through space to destroy an enemy ballistic missile? The Laboratory, working on an Army project called Whitehorse, considered the possibility of neutral-particle beams as defensive weapons. Then, in the mid 1970s, using back-of-the-envelope calculations, some physicists at Los Alamos began examining the theoretical feasibility of laser-beam and neutral-particle-beam weaponry. When in 1975 John Madey and coworkers at Stanford University built a device coined the free-electron laser, its implications for non-nuclear defense were immediately recognized. Madey's device used an accelerator to create a laser beam by moving electrons past an array of magnets called a wiggler. The resulting laser beam could be tuned to any desired frequency. Interested in Madey's discovery, a small group of Los Alamos scientists designed a similar device and built a prototype of a small free-electron laser in the basement of the Laboratory's Physics Building.

In a televised speech in 1983, President Reagan called on the nation's scientific community to begin a program that would enable the U.S. to "intercept and destroy strategic ballistic missiles before they reached our own soil or that of our allies." The new program, called the Strategic Defense Initiative, was to employ two technologies that were already under development at the Laboratory: the free-electron laser and the neutral-particle beam.

At Los Alamos the first challenge in building a neutral-particle-beam weapon was to determine how launch and space environment would affect the performance of such a weapon. Consequently, on July 13, 1989, an Aries rocket with a special experimental payload called BEAR (beam experiments aboard a rocket) was launched at White Sands Missile Range. When the rocket reached its apogee, the accelerator fired a neutral-particle beam at different orientations to the earth's magnetic field. The BEAR experiment, which was a collaboration between Los Alamos scientists and industry, demonstrated that a neutral-particle beam is unaffected by the space environment and that a compact accelerator can survive launch. A key component of the payload was an RFQ.

The Ground Test Accelerator is the next step in developing neutral-particle-beam technology. In an in-

Aries Rocket at White Sands Missile Range. *This type of rocket carried into space an accelerator for producing a neutral-particle beam.*

Ground-penetrating Radar Units. *These units are effective at testing the integrity of shallowly buried fuel-storage tanks. Depth penetration is up to tens of meters in lossy soils and up to a kilometer under ice.*

dustrial partnership with Grumman Corporation, the Laboratory is building a large facility to test whether such a device can do its intended job in space. And Los Alamos and Livermore national laboratories are contributing to development of free-electron-laser weapons by providing technical expertise for the Ground-based Free-Electron Laser Project being undertaken at White Sands Missile Range.

Into the Future

Prompted by the end of the Cold War, Siegfried Hecker, director since 1986, has begun to reshape the direction of the Laboratory. Although the stewardship of nuclear weapons and other defense interests remain the highest priorities, many of the new programs are aimed at enhancing the United States's economic and industrial competitiveness. Efforts to apply accelerator technology to production of silicon chips, transmutation of nuclear waste, developing sources for neutron scattering, and cleanup of hazardous waste are being pursued. Also being investigated are advances in accelerator technology through the use of superconducting materials and microwaves.

AT Division, under Stanley Schriber, has begun a gradual shift from nuclear science toward materials science. Accelerators that were developed for SDI turn out to have many applications in materials research. For example, efforts to upgrade the present generation of free-electron lasers have great potential for processing of materials, surfaces, and chemicals. One example is silicon-chip production. The so-called Advanced Free-Electron Laser is a compact, portable free-electron laser that can be tuned to very small wavelengths—in the extreme ultraviolet—to etch silicon chips. The eventual goal of the program, led by

Rich Sheffield, is to produce a gigascale chip, one including about a billion components. A free-electron laser might be used also to liquefy methane gas at remote refineries in Alaska or Australia to enable safe transport to stateside refineries. In October, 1992, the world's first portable free-electron laser—not much larger than the bed of a pickup truck—produced its first beam at Los Alamos. Other SDI spin-offs include focusing and alignment devices for accelerators, new telescope technologies, and technologies that support the Superconducting Super Collider planned to be built near Dallas, Texas.

SDI accelerator developments have also made possible a variety of future applications that require average powers much higher than the power of LAMPF, presently the highest-power accelerator in the world. A conceptual design for the Accelerator Production of Tritium (APT) system is being developed by a team involving Los Alamos, Sandia, and Brookhaven National Laboratories and six companies. Tritium is an essential component of U.S. nuclear weapons; since it decays it must be regularly replaced. The APT system could provide a non-reactor source of tritium for the future. Related applications of high-power accelerators could include destruction of excess plutonium, nuclear-waste transmutation, and electric-power production; these applications are being evaluated by the National Academies of Science and Engineering. (See "Acceleratorbased Conversion of Surplus Plutonim.")

Other efforts by AT Division exploit new developments in materials science. One group, led by Joe DiMarco, is investigating the fabrication of accelerator cavities from niobium, a material that, when cooled to temperatures below 9 kelvins, becomes superconducting. The cost of niobium itself and of developing the fabrication technology might be offset by savings in radio-frequency-power costs. Another group, led by Bob Hoeberling, is developing a ground-penetrating radar that is capable of testing the integrity of buried gasoline and waste tanks—a potentially invaluable tool in hazardous-waste cleanup.

The trend toward materials science may also affect LAMPF. As discussed in "Neutrons in Our Future: A Proposed High-Flux Spallation Neutron Source," the Laboratory hopes to upgrade LAMPF so that it could produce more intense neutron pulses. The Laboratory's neutron scattering center would then be competitive with the most advanced neutron-scattering facilities in the world and would have significant implications for the United States's ability to develop new technologies in materials science. ■

The Clean Room at Accelerator Technology Division. *Superconducting accelerator cavities of niobium are fabricated under conditions of extreme cleanliness.*

Acknowledgments

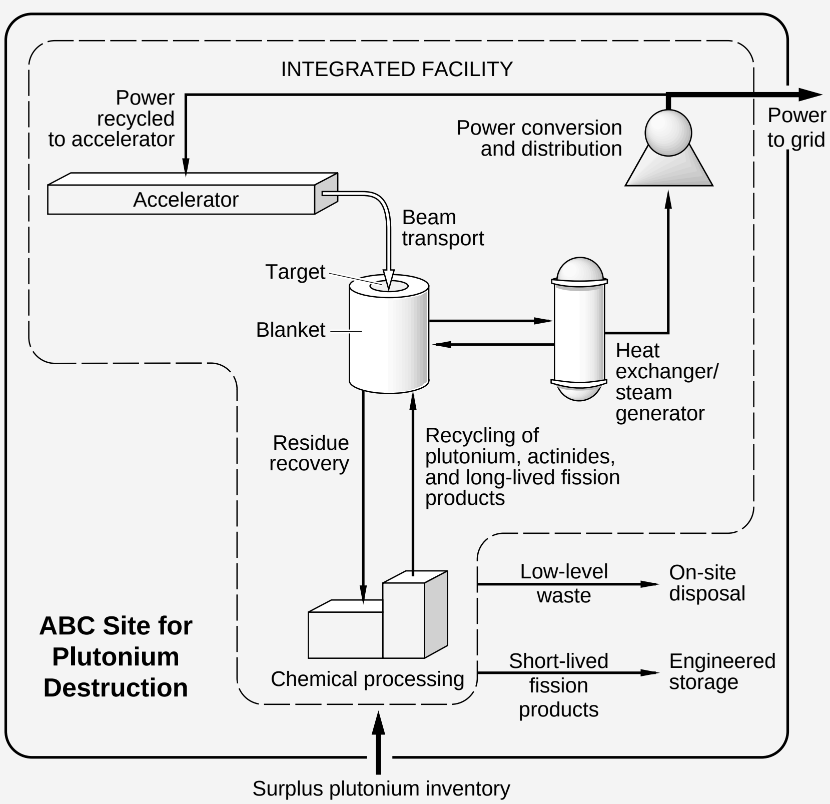
I wish to thank Joe McKibben, Louis and Mary Rosen, Ed Knapp, Stanley Schriber, Dave Schneider, Tom Wangler, Joe Di Marco, Mike Lynch, Bob Hoeberling, Rich Sheffield, Mikkel Johnson, Ed Arthur, T. J. Trapp, Rulon Linford, Lew Agnew, Bob Seidel, Jim Stovall, Nancy Shera, and Walt Reichelt for their insights about accelerators and their help in writing this article.

Accelerator Based Conversion of Surplus Plutonium

Reductions in the number of nuclear weapons in the United States and the former Soviet Union have resulted in tons of weapons-grade plutonium that need to be disposed of safely. Storage facilities are needed for the near term, but the ultimate disposition is also important. Some solutions involve converting the plutonium to a form similar to other high-level wastes destined for geologic repositories, such as spent reactor fuel and glassified wastes. Those forms would be substantially more proliferation-resistant than the present concentrated form. Even so, plutonium originating from weapons programs and the larger, growing quantities of plutonium from commercial spent fuel would continue to present a proliferation nuisance.

The Accelerator Based Conversion (ABC) technology under investigation at the Laboratory and illustrated in the figure could be used to destroy plutonium from both weapons and commercial reactors. The technology is being designed to transmute the “dominant” long-lived radioactive products generated during plutonium consumption (those that are most difficult to dispose of safely) and to generate electric power from the heat released by the various conversion processes. Initially, ABC systems could destroy the plutonium returned from the weapons program. They could also reduce the long-term toxicity of existing defense wastes destined for a geologic repository. In the longer term ABC plants could consume plutonium, other actinides, and dominant long-lived radioactive waste present in spent fuel from nuclear reactors. Accelerator-based conversion systems would transmute these long-lived radioactive materials into stable or short-lived fission products. The controlled consumption of plutonium afforded by ABC technology could thus provide an international method to reduce opportunities for proliferation.

As shown in the figure, the ABC system uses a proton beam from the accelerator to produce an intense neutron source at the target. The blanket surrounding the target contains plutonium and other actinides that are to be destroyed. The neutrons from the target are moderated, or slowed down, in the blanket, where they induce fission of the unwanted materials, which, in turn, releases more neutrons. Some of these neutrons are captured in the nuclei of long-lived fission products and thereby transmute those nuclei to short-lived or stable products. The intense flux of thermal neutrons allows the ABC system to have lower inventories of actinides and fission products for a given burn rate of those materials than other proposed systems for burning plutonium and long-lived wastes. Further advantages include smaller end-of-life inventories and potential safety enhancements. The fast burn-up of material in the ABC method requires frequent chemical processing to remove the stable and short-lived products for disposal. The unfissioned actinides, including plutonium, and dominant long-lived fission products are returned to the blanket for further exposure to the high neutron flux. The addition of accelerator-produced neutrons to the blanket not only ensures that adequate numbers of neutrons are available to transmute all of the unwanted materials but also provides for subcritical operation in the blanket and therefore prompt control of fission reactivity. This type of control may prove to be particularly advantageous in designs involving very high neutron fluxes and continuous flow of material through the blanket. The heat generated by fission in the blanket is converted to electric power. Some of this electric power can be used to run the accelerator, and the rest can be made available to the electric-power grid. □

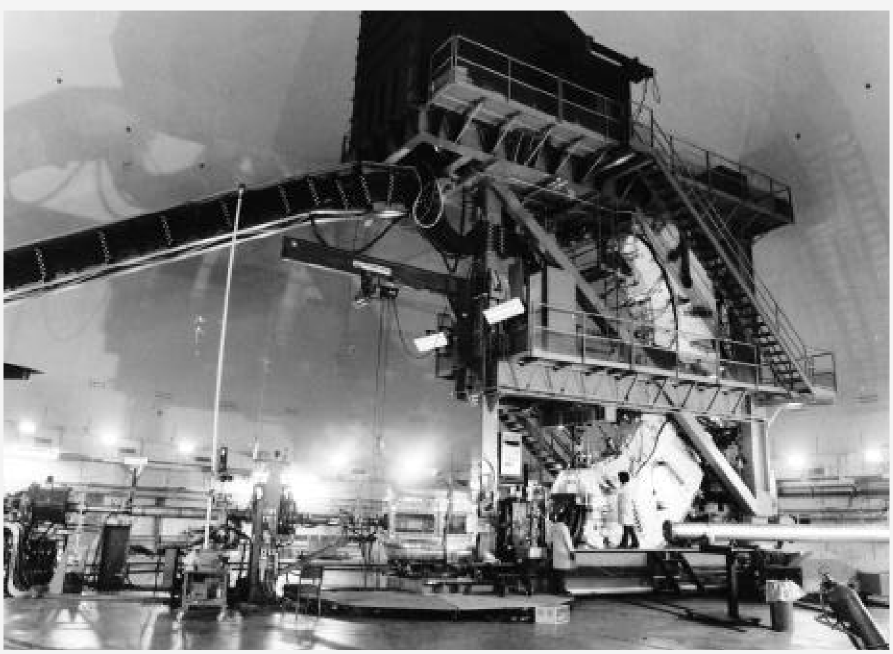


Medium-Energy Physics at LAMPF

Mikkel B. Johnson

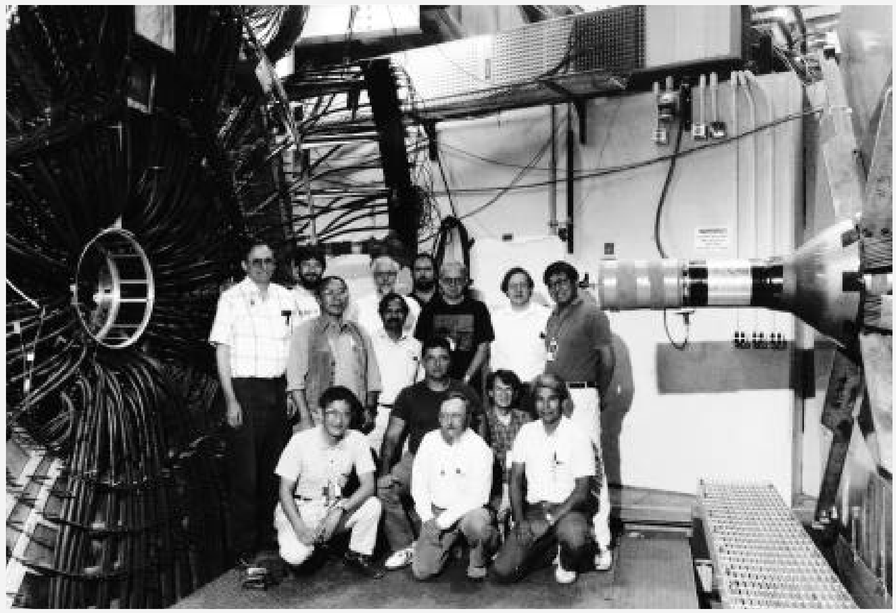
The LAMPF accelerator provides primary beams of protons and negatively charged hydrogen ions as well as secondary beams of neutrons, pions, muons, and neutrinos. The uniqueness of LAMPF as an experimental facility derives from the high intensity of those beams and the consequent capability to exploit rare reactions to answer specific questions about particles and nuclei. Additionally, the high resolution of the spectrometers and other detectors available at LAMPF make precision measurements feasible. As a result, over the last twenty years LAMPF has helped to open up an entirely new field of basic research—medium-energy nuclear physics.

By using the tools available at LAMPF, nuclei can be explored in new ways. Experiments with muons, pions, and nucleons have quantified the size, shape, and composition of nuclear states and measured the response of nuclei to the addition of charge, energy, and momentum, as well as quanta of vibrational and rotational excitation. Some of the most interesting results have been obtained when the incoming particles cause the nuclei to respond in certain extreme ways. For example, in a pion double-charge-exchange experiment, nuclei can be forced to accept two units of charge at one time. Measurements of the dependence of the scattering cross section on the final state of the nucleus have led to a strikingly detailed picture of correlations in the motion of neutrons and protons. In addition, double-charge-exchange experiments have uncovered new modes of motion of nuclei in which two patterns of vibration coexist and have also led to the discovery and study of new nuclear species. In experiments with nucleon beams, nuclei have been forced to accept a small amount of energy and at the same time a large momentum in various spin configurations hitherto incapable of being distinguished. For reasons that are not completely understood but are being actively sought, theories that work well in more normal situations have been discovered to break down under those unusual conditions. And state-of-the-art studies in atomic physics have addressed previously inaccessible regions of the spectrum with a unique technique that combines a beam of laser light and a beam of negatively charged hydrogen ions.



The High-Resolution Proton Spectrometer at LAMPF. *This instrument, known as the HRS, is used to measure cross sections for elastic and inelastic scattering of protons from nuclei. Careful measurements with the HRS of the spin dependence of cross sections led to the widespread acceptance of relativistic descriptions of nuclear dynamics. HRS results also stimulated theoretical and further experimental investigations of how the nuclear medium affects the nucleon-nucleon interaction.*

Additionally, various scattering experiments and measurements of the decay products of muons and pions have focused on the nature of the underlying strong, weak, and electromagnetic interactions. Precision measurements with muons have yielded new insights into quantum electrodynamics. A systematic program extending over many years has completely mapped out the character of the nucleon-nucleon interaction over the entire energy range of LAMPF and has thus provided a bank of basic data for theorists and laid the foundation for interpreting nucleon-nucleus scattering experiments. An experiment using the world's most intense source of very-low-energy neutrons has led to development of a completely new technique for detecting a signature of breakdown of a fundamental symmetry in nuclear forces. The technique, which uses properties of complex nuclei to magnify the signal for the breakdown of parity (the mirror symmetry between left and right) by a factor of about 1 million, has uncovered unexpected results and opened a rich area of exploration. Other investigations of fundamental interactions include the scattering of neutrinos from various targets. Neutrinos interact so weakly that even experiments using the high-intensity neutrino beam available at LAMPF require several years to complete. One such experiment has provided the only available measurement of electron neutrino–electron scattering. The experimental results showed that the interplay predicted by the standard model between the charged and neutral parts of the electroweak interaction was indeed a reality. Therefore, since only the neutral part of the weak force is involved in the interaction of electrons with the muon neutrino or the tau neutrino, the interaction of the electron neutrino with electrons is fundamentally different from the interaction of the other neutrinos with electrons. That difference provides a possible explanation for the observed shortfall in electron neutrinos coming from the sun. Yet other experiments at LAMPF hunt for breakdowns of the standard model. Although none has been detected, the searches at LAMPF for the decay of a muon into an electron and gamma rays, which would herald such a breakdown, have consistently led the world in sensitivity.



The MEGA Detector at LAMPF. *This detector will be used in a search, scheduled to begin this summer, for the decay of a muon into an electron and a gamma ray. The occurrence of that reaction would signify a breakdown of the standard model. The sensitivity of the MEGA detector to the decay is two orders of magnitude greater than that of detectors used in previous searches.*

Experiments such as those mentioned above constitute some of the highlights of the contribution of LAMPF to nuclear science. They have provided answers to many specific questions and at the same time have paved the way for a slow but very important transformation in the way nuclear physicists think about their subject. Before the era of the meson factory, nuclei could be largely understood as a collection of nucleons undergoing nonrelativistic motion and interacting through potentials. That picture is no longer adequate to describe what the medium-energy beams "see" of nuclei. To understand the new data, the catalogue of constituents of nuclei has been enlarged to encompass mesons and excited states of nucleons themselves. Additionally, the picture of the dynamics of their motion has changed. Relativity can no longer be ignored, and interactions must be described in terms of the coupling of mesons to nucleons. Even today the picture is continuing to evolve as particle and nuclear physicists realize deeper connections between their once quite distinct fields. □

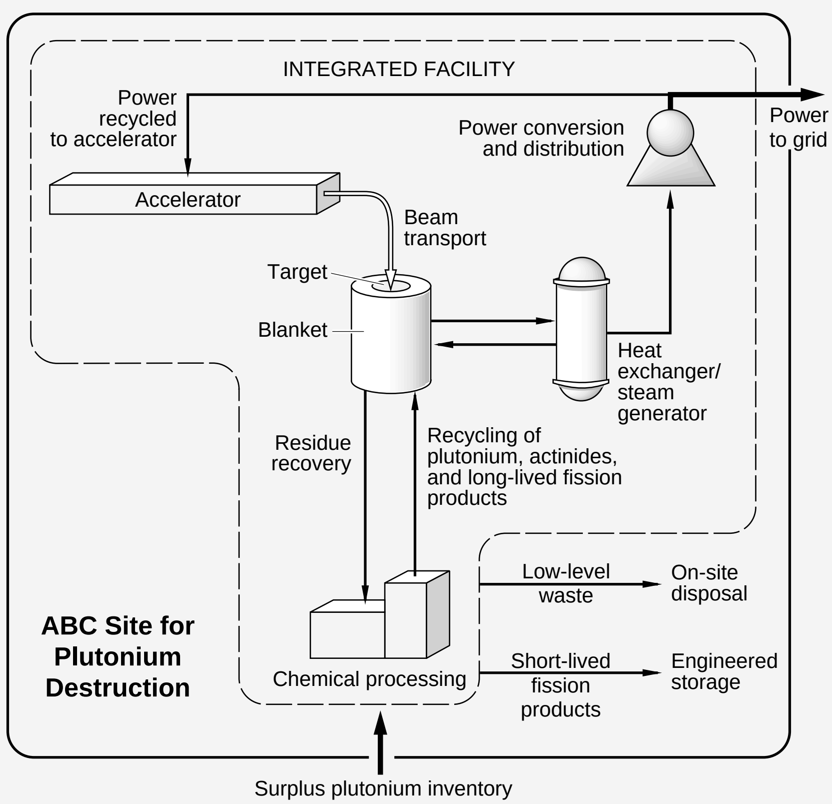
Mikkel B. Johnson, a Laboratory Fellow, has pursued research in theoretical nuclear physics at the Medium Energy Physics Division since 1972.

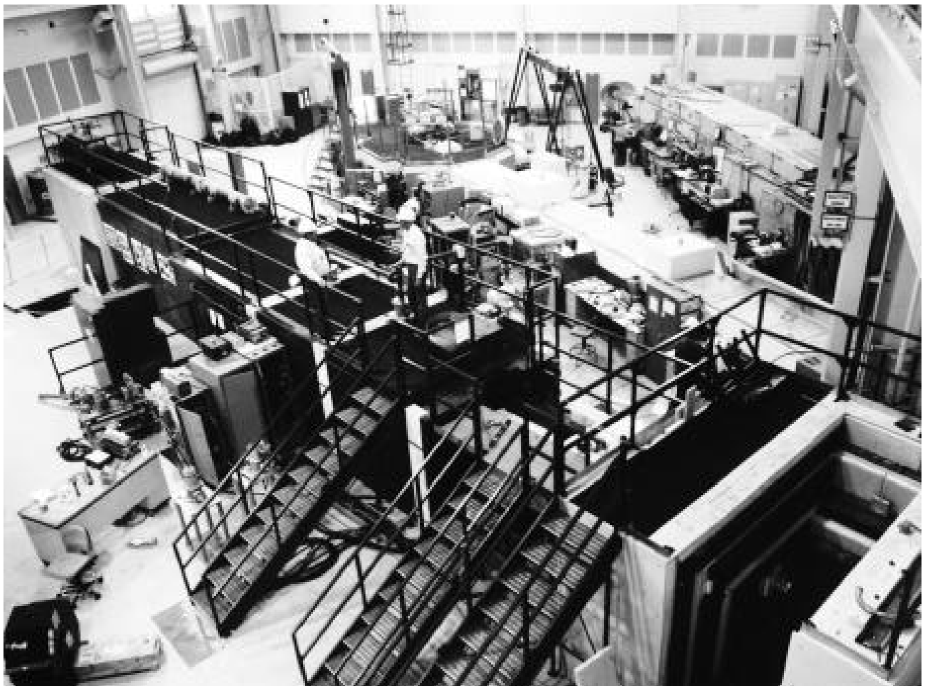
Accelerator Based Conversion of Surplus Plutonium

Reductions in the number of nuclear weapons in the United States and the former Soviet Union have resulted in tons of weapons-grade plutonium that need to be disposed of safely. Storage facilities are needed for the near term, but the ultimate disposition is also important. Some solutions involve converting the plutonium to a form similar to other high-level wastes destined for geologic repositories, such as spent reactor fuel and glassified wastes. Those forms would be substantially more proliferation-resistant than the present concentrated form. Even so, plutonium originating from weapons programs and the larger, growing quantities of plutonium from commercial spent fuel would continue to present a proliferation nuisance.

The Accelerator Based Conversion (ABC) technology under investigation at the Laboratory and illustrated in the figure could be used to destroy plutonium from both weapons and commercial reactors. The technology is being designed to transmute the "dominant" long-lived radioactive products generated during plutonium consumption (those that are most difficult to dispose of safely) and to generate electric power from the heat released by the various conversion processes. Initially, ABC systems could destroy the plutonium returned from the weapons program. They could also reduce the long-term toxicity of existing defense wastes destined for a geologic repository. In the longer term ABC plants could consume plutonium, other actinides, and dominant long-lived radioactive waste present in spent fuel from nuclear reactors. Accelerator-based conversion systems would transmute these long-lived radioactive materials into stable or short-lived fission products. The controlled consumption of plutonium afforded by ABC technology could thus provide an international method to reduce opportunities for proliferation.

As shown in the figure, the ABC system uses a proton beam from the accelerator to produce an intense neutron source at the target. The blanket surrounding the target contains plutonium and other actinides that are to be destroyed. The neutrons from the target are moderated, or slowed down, in the blanket, where they induce fission of the unwanted materials, which, in turn, releases more neutrons. Some of these neutrons are captured in the nuclei of long-lived fission products and thereby transmute those nuclei to short-lived or stable products. The intense flux of thermal neutrons allows the ABC system to have lower inventories of actinides and fission products for a given burn rate of those materials than other proposed systems for burning plutonium and long-lived wastes. Further advantages include smaller end-of-life inventories and potential safety enhancements. The fast burn-up of material in the ABC method requires frequent chemical processing to remove the stable and short-lived products for disposal. The unfissioned actinides, including plutonium, and dominant long-lived fission products are returned to the blanket for further exposure to the high neutron flux. The addition of accelerator-produced neutrons to the blanket not only ensures that adequate numbers of neutrons are available to transmute all of the unwanted materials but also provides for subcritical operation in the blanket and therefore prompt control of fission reactivity. This type of control may prove to be particularly advantageous in designs involving very high neutron fluxes and continuous flow of material through the blanket. The heat generated by fission in the blanket is converted to electric power. Some of this electric power can be used to run the accelerator, and the rest can be made available to the electric-power grid. □





Experimental hall at LANSCE

NEUTRONS in our future

a proposed high-flux spallation neutron source

Roger Pynn

There is a paradigm in scientific research that repeats itself continually—the discovery that earned yesterday's Nobel Prize becomes the tool for today's research. Take x rays, lasers, and transistors, for example. Each was worth a Nobel in its day, and each is now found not only in almost every research laboratory but also in hospitals, supermarkets, and homes. The same paradigm applies to neutrons. Discovered by James Chadwick in 1932, these neutral particles were the stuff of esoteric research until fast fission and politics combined to make them central players in the Los Alamos story. Nuclear reactions in which neutrons participate are at the heart of all of the nuclear weapons designed here and elsewhere. Other neutron reactions—those in which neutrons are scattered rather than absorbed by nuclei—are the basis for the use of neutrons as probes of the structures of materials. That area of research, referred to simply as neutron scattering, is an important part of today's scientific agenda, which stresses industrial competitiveness and quality of life. To design new and improved materials for industrial applications, scientists build on their understanding of existing materials, a large part of which comes from information about their structures. Neutron scattering provides that information, often in situations where other techniques fail.

Successful neutron-scattering experiments require large number of neutrons to be directed at a sample because only a small fraction of the neutrons are scattered. The first neutron sources that were sufficiently intense for such experiments were nuclear reactors, and neutron scattering began as a parasitic activity at research reactors that were built in the 1950s to obtain data for nuclear-power programs. To this day the most productive neutron-scattering program is to be found at a reactor—the Institut Laue Langevin (ILL) in Grenoble, France. However, the situation is changing. A newer technique for producing neutrons at proton accelerators rather than nuclear reactors is fast becoming competitive. The technique, proton-induced spallation of heavy-metal nuclei, is currently the basis of the neutron source at the Laboratory's Manuel Lujan, Jr. Neutron Scattering Center and will remain the basis of a more intense neutron source that the Laboratory hopes to build. An upgrade of the LAMPF proton accelerator will make the more intense neutron source possible—which brings us back once more to our paradigm. LAMPF was built more than twenty years ago to study nuclear reactions that involve energetic protons or pions. Now, one of those reactions, proton-induced spallation, may be the basis for a new neutron-scattering facility.

The success of neutron scattering and its continuing importance are a result of several properties of the neutron. Because of its neutrality and the weakness of its interactions with matter, the neutron—unlike x rays or light— can penetrate deeply into solids and liquids and provide information about bulk, as opposed to surface, structure. In addition, because neutrons are scattered by both the nuclei and the unpaired electrons in matter, they provide information about both atomic and magnetic structure. The thermal neutrons generated by nuclear reactors or spallation sources have energies that are comparable to those of vibrating or diffusing atoms in solids. Therefore neutrons can probe not only the equilibrium positions of atoms in solids but also temporal structural changes. Because the neutron-scattering power of atomic nuclei varies erratically and often considerably with atomic number, neutrons can often distinguish between neighboring elements and can easily distinguish the lightest element, hydrogen, even in the presence of much heavier elements. The latter property makes neutrons a particularly powerful probe of biological molecules and man-made polymers, both of which contain substantial amounts of hydrogen.

For more than forty years neutron scattering has played an indispensable role in studies of condensed matter, providing essential information about materials as different as antiferromagnets, ribosomes, and shape-memory alloys. Often the information has been unobtainable by other means. Even a partial list of contributions from the past decade is impressive. During that period neutron scattering revealed the structure of the first high-temperature superconductors; the structure and excitations of buckminsterfullerenes, or bucky balls; the conformation of molecules in a polymer melt; the interfacial structure of artificially produced polymeric and magnetic layers; the structure and dynamics of new catalysts; the spin dynamics of highly correlated electron systems; and the condensate fraction in superfluid helium. It is safe to say that a large part of the conceptual and theoretical underpinning of the modern theory of solids would be unverified and incomplete without neutron scattering. And without that knowledge our current technology could not exist.

LANSCE has made its share of contributions during the five years it has been operating. The discovery by Gregory J. Kubas of the Laboratory's Inorganic and Structural Chemistry Group that certain metal complexes can coordinate molecular hydrogen is widely regarded as one of the most significant developments of the 1980s in inorganic chemistry. Studies at LANSCE of the vibrational and rotational dynamics of those dihydrogen ligands have provided insight into the nature of this unique chemical bond—the first known example of stable intermolecular coordination of a sigma bond to a metal. The system mimics a catalytic reaction "frozen" in an intermediate state of a type that is usually too ephemeral to study and understand. The dihydrogen ligand is important in catalysis because it can easily exchange hydrogen with other ligands in a complex. It is conceivable, for example, that hydrogen could be added to other ligands such as ethylene (catalytic hydrogenation) at a much lower cost in energy than the 104 kilocalories per mole required to break the hydrogen-hydrogen bond of uncoordinated molecular hydrogen.

Polymers and other macromolecules absorbed at solid or fluid surfaces have many applications to a wide variety of technologies. They are a means for achieving colloidal stabilization in water-treatment schemes, ceramic processing, inks, and fuels; they are used for mechanical protection of solids against friction and wear in motors and computer disks; and surface-active molecules at liquid-liquid interfaces are used to clean up oil spills and to enhance emulsification and blending. The variation of polymer density close to an absorbing surface had been studied theoretically but was difficult to study experimentally until neutron reflection provided the answer. Work at LANSCE verified theoretical predictions for the profile of the "polymer brush" formed by the stretching of polymer molecules away from a solid surface into a surrounding fluid and provided a characterization of the "polymer mushrooms" that occur as the grafting density of the absorbed polymers (the number of attached polymers per unit area) is decreased.

As the transportation industry struggles to improve fuel efficiency, it is turning increasingly to new composite materials—such as aluminum reinforced with silicon-carbide particles—that provide the dual advantages of strength and lightness. To understand the mechanisms of failure of such materials and to assess their lifetimes in real components, it is important to understand the residual stresses induced in the materials during fabrication. Depending on their distribution, such stresses can be devastating—aircraft fuselages have disintegrated in flight and railroad tracks have cracked and caused train crashes—or beneficial—wine barrels have been held together by metal hoops for centuries. Unfortunately no conventional technique for measuring residual stress, such as strain-gauge sectioning or hole drilling, is truly nondestructive. Over the last five years, neutron diffraction has proved to be a unique, nondestructive alternative and has been systematically exploited at LANSCE. Our work on composite materials has allowed sophisticated computer models for stresses—residual stresses as well as stress induced by applied load—to be verified and, in some cases, improved.

In spite of the successes and acknowledged importance of neutron scattering, the technique is on the verge of extinction in this country, leadership having passed to our European colleagues over a decade ago. With one exception (the neutron source at the National Institute of Standards and Technology), all the neutron sources in the U.S. that support neutron-scattering programs are run by the Department of Energy. Those sources are old and, by modern standards, poorly instrumented. The high-flux reactors at Brookhaven and Oak Ridge national laboratories may reach the end of their useful lives before the end of this decade. A pulsed spallation source at Argonne National Laboratory provides only one-tenth of the intensity of the ISIS facility at the University of Oxford. The LANSCE source has a peak neutron flux that is slightly higher than the ISIS source but has suffered from poor reliability and short annual operating periods. The problems of LANSCE have been exacerbated by constant erosion of the operating budget of LAMPF over the past five years. LANSCE is now threatened with closure because LAMPF is no longer the highest priority of the nuclear-physics community. Without a new neutron source, U.S. researchers will not remain competitive. Fundamental research as well as technology will suffer.

Nor are our competitors standing still. A consortium of European laboratories has proposed a design study for an advanced spallation source (the European Spallation Source, or ESS) that would provide capabilities well beyond those available at the ILL. The proposed ESS—consisting of a high-energy linac and an accumulator ring—looks very much like an upgraded version of LAMPF and the Proton Storage Ring. It is ironic that just as the Laboratory's spallation source faces shutdown, the Europeans are realizing that spallation sources are the way of the future.

Recognizing the national need for new neutron-scattering capabilities, as well as the value of existing infrastructure at LAMPF and the strength of its expertise in neutron scattering, Los Alamos National Laboratory has proposed the construction of a new pulsed spallation source with an initial power of 1 megawatt (the power of the present LANSCE source is 60 kilowatts) and a possible future power of 5 megawatts. On August 19, 1992, Laboratory Director Sig Hecker announced the proposal to a visiting review committee, noting that the Laboratory wants to change the emphasis of research at its 800-MeV linac from nuclear physics to neutron scattering. The committee Hecker addressed was chaired by Walter Kohn of the University of California, Santa Barbara, and had been charged by Will Happer, head of the DOE's Office of Energy Research, to examine the relative merits of reactor and spallation sources for the country's future neutron-scattering program. After several contentious weeks, the committee finally concluded that the country would be best served if two new sources—one of each type—could be built.

Since its beginning in 1987, LANSCE has been operated as a National User Facility, open to scientists from industry, academia, and other national laboratories. Experimental proposals submitted by potential users are peer-reviewed to ensure that the best use is made of the facility. Most of the national laboratories host a user facility of some sort in fulfillment of one of the DOE's most important missions in the area of basic research. The new pulsed source proposed by the Laboratory will remain a user facility, and the community of users will define the facility specifications that best suit its needs. Of course, we have an idea of the facility we would like to build—it is described below—but it is important to recognize that the final design parameters will come from the users rather than from the Laboratory.

A new spallation source at the Laboratory will make use of a number of existing LAMPF assets that would be expensive to reproduce elsewhere and are very appropriate as part of a modern accelerator complex. The 700-meter-long shielded tunnel that contains the present linac will remain, as will buildings, cooling towers, 30 megawatts of site electrical power, and a 600-meter part of the existing accelerator called a coupled-cavity linac. The latter is basically a copper pipe—albeit of somewhat exotic design—that is not expected to wear out.

All of the high-tech parts of the proposed accelerator complex will be new and will take full advantage of accelerator technology developed here and elsewhere as part of the Strategic Defense Initiative. Our present reference design calls for injection of 800-MeV protons from the upgraded linac into an accumulator ring that is similar in concept to the the existing Proton Storage Ring. However, we are studying an option that would increase the proton energy and, perhaps, permit an easier upgrade to 5 megawatts of beam power in the future.

The new accelerator complex will produce 60 proton pulses per second, each of about 0.5 microsecond in duration, and distribute them between two neutron-production targets. One target will receive 40 pulses per second and the other 20 pulses per second. We expect the 40-hertz target to provide about five times the average neutron flux generated by the ISIS source. Coupled cold moderators at the 20-hertz target will give twenty-five times the peak flux of either the ISIS or the present LANSCE source. There will be room for between twelve and sixteen beam lines around each target.

What does all this buy us? How does it compare with the ILL, for example? This question was answered by a group of European and American neutron-instrumentation experts who advised the Kohn panel. That eminent group concluded that a 1-megawatt pulsed spallation source could duplicate the capabilities of the ILL *and* provide facilities that exceed those at the ILL for experiments requiring the intense high-energy neutron beams produced by spallation. In other words, a 1-megawatt spallation source would give the U.S. the same capability as the ILL *plus* the obvious advantages over ISIS. Such a source would also be complementary to the reactor—the so-called Advanced Neutron Source—that the DOE proposes to build at Oak Ridge National Laboratory. Although there is a large area of overlap in the capabilities of these two sources, each has unique strengths.

Any prediction of the scientific impact of a 1-megawatt spallation source is bound to be incomplete, at best. Nevertheless, the impact is fairly obvious in those areas that are extrapolations of current research. For example, many experiments are beyond our current capabilities because samples of sufficient size are not available. Neutron scattering is inherently a signal-limited technique because, as mentioned above, only a small fraction of the neutrons incident on a sample are scattered. If the sample is not large enough, the informative scattered neutrons—the signal—cannot be discriminated from background neutrons that have suffered spurious scattering processes. One way of overcoming this limitation is to increase the flux of incident neutrons as our new source is designed to do. Such a source would make many important experiments possible. For example, we would be able to probe the collective excitations of high-temperature superconductors and fullerenes and perhaps understand the bases for their bizarre properties. Experiments of this sort now require single crystals of a size that cannot be grown. The problem of sample size is even more critical in structural biology, where the structures of only about one in every two hundred interesting proteins are now accessible to neutron scattering—sufficiently large crystals of the others just cannot be produced. Although some improvement in sample size is likely in some instances, there are other areas—the study of interfacial structure by neutron reflection, for example—where the scattering volume is inherently small and will always remain so. For such systems the only way forward is through the use of neutron sources with higher flux.

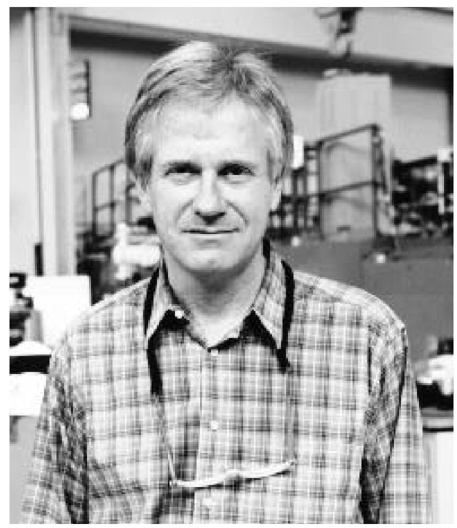
Higher-flux sources will also offer scientists the possibility to study structures as they evolve over time. Examples include changes in the structure of interfaces during corrosion and of electrolytes during battery discharge; phase transformations induced by propagating shock waves; and conformational changes of polymers during extrusion molding. Presently such experiments are restricted to model systems that change relatively slowly with time or to systems that can be arrested or cycled repeatedly. Examining the change in structure of a catalyst during its active phase, for example, is beyond current capabilities because the entire reaction is completed in a time that is much shorter than that needed for a neutron measurement.

The techniques used for neutron scattering at high-flux pulsed sources are well adapted to neutron-scattering experiments in which samples are subjected to high pressures or high magnetic fields. The equipment required to achieve high pressures would fail if it had to have large windows to let neutrons in and out, and high magnetic fields can be maintained only for short periods. Powder-diffraction measurements at

LANSCE have already been made at pressures above 100 kilobars, a pressure that is three or four times higher than has been achieved by similar experiments elsewhere in the country. Achieving still higher pressures, such as those needed to mimic some geological conditions, will require the use of smaller samples and, concomitantly, more powerful neutron sources.

One of the most exciting new capabilities offered by a 1-megawatt pulsed source couples the characteristics of the source with the great progress that has been made in computer science over the past two decades. Twenty-five years ago, neutron spectrometers were deliberately designed to avoid collecting too much data—it would have been just too confusing to the poor scientists! Spectrometers were designed to focus on phenomena that occurred over a small range of length and time scales and to ignore the rest. Although that approach delayed some discoveries a decade or two, it worked reasonably well for simple samples, especially those that could be grown in the form of single crystals. Unfortunately, many of the complex materials of interest today—both in materials science and structural biology—are interesting precisely because they have structure on a wide variety of length scales. Examples range from DNA molecules packaged as chromatin to the fractal structures found in porous media. To study such materials with neutrons requires spectrometers with access to a large range of length scales, a feature provided quite naturally by pulsed sources. However, to find some meaning in the vast quantities of information obtained by such spectrometers requires the speed of modern computers and the wizardry of modern techniques for data manipulation and display. The payoff could be immense, however. One can easily imagine a neutron spectrometer at a new pulsed source with 1 million or 10 million parallel information channels instead of just one.

Although this article has focused on neutron scattering, we expect that a new spallation source at the Laboratory would support other types of research as well. Indeed, it would be indefensible from the taxpayers' point of view *not* to exploit synergistic uses of the facility, some of which might help to resolve important issues in areas such as the management of radioactive waste. And there are exciting experiments in basic physics and nuclear-physics research to be done with neutrons, as well as complementary investigations of materials by muon spin resonance. Experiments with ultracold neutrons—those with velocities of only a few meters per second—can accurately measure the lifetime of the neutron and determine whether it has an electric dipole moment. Both of those properties are important inputs to the standard model used to understand our universe. High-power spallation sources also have practical applications. They can produce neutron-poor radioisotopes, many of which have become indispensable to modern nuclear medicine; they can be used to study radiation damage of materials in regimes that are relevant to fusion-energy systems; and they are the basis for many transmutation schemes that have been proposed to solve problems ranging from the production of tritium to the destruction of long-lived fission products and plutonium from the weapons stockpile. We envisage that those other types of research could be carried out without jeopardizing the primary mission of neutron scattering. The prospects of such a multidisciplinary research facility are indeed exciting and a fitting continuation of Los Alamos expertise in the science of neutrons. ■

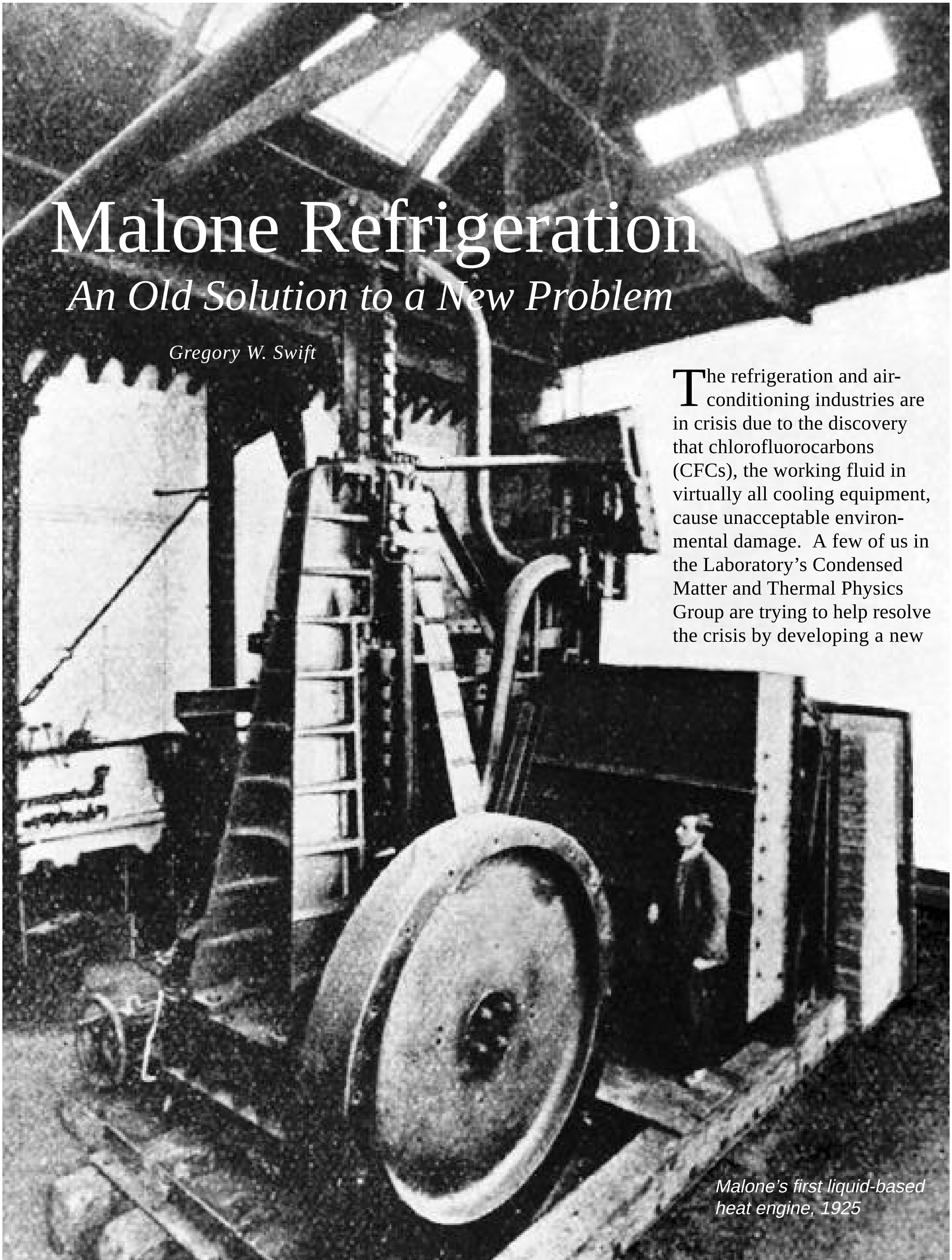


Roger Pynn was born and educated in England. He received his M.A. from the University of Cambridge in 1966 and his Ph.D. in neutron scattering, also from the University of Cambridge, in 1969. He was a Royal Society European Fellow to Sweden in 1970; he did two years of postdoctoral research in Norway; and then he was an associate physicist for two years at Brookhaven National Laboratory. After spending eleven years at the world's leading center for neutron scattering, the Institut Laue Langevin in Grenoble, France, Pynn was appointed director of the Laboratory's Manuel Lujan, Jr. Neutron Scattering Center. He served as a member of the Kohn panel whose deliberations are described in this article.

Malone Refrigeration

An Old Solution to a New Problem

Gregory W. Swift



Malone's first liquid-based heat engine, 1925

The refrigeration and air-conditioning industries are in crisis due to the discovery that chlorofluorocarbons (CFCs), the working fluid in virtually all cooling equipment, cause unacceptable environmental damage. A few of us in the Laboratory's Condensed Matter and Thermal Physics Group are trying to help resolve the crisis by developing a new

CFCs and Cooling Equipment: The Size of the Problem

As recently as ten years ago, only a few experts appreciated the potentially disastrous consequences of releasing chlorofluorocarbons (CFCs) into the environment. Today it is generally believed that chlorine from CFCs is destroying stratospheric ozone, which shields the earth from the sun's ultraviolet radiation, and that the resulting increased ultraviolet radiation will soon lead to millions of fatal and nonfatal skin cancers and cataracts, to other threats to human health, and to severe damage to agriculture and ecosystems. To mitigate those effects, the United States and most other countries have committed themselves, through the 1987 Montreal Protocol on Substances That Deplete the Ozone Layer and its later revisions, to rapid elimination of CFC production.

The rate of CFC production, though being reduced, is on the order of a million tons per year. CFCs are used extensively as working fluids in refrigerators and air conditioners, as cleaning solvents in electronics and sheet-metal fabrication, and as foaming agents in foam insulation and cushions. CFCs have the advantage for many purposes of being almost chemically inert (their behavior in the upper atmosphere is an exception); in particular they present no fire or poison danger. In addition their lack of interaction with the lubricants used in refrigerator compressors improves the efficiency and the lifetime of the refrigerator. Cooling-system efficiency has major economic and environmental impacts. Cooling consumes about 20 percent of the nation's electricity, at a cost of tens of billions of dollars per year. Kitchen refrigerators alone use 8 percent of the nation's electricity. Most of that electricity is produced by burning fossil fuels. Congress is accordingly requiring ever more efficient cooling equipment; for example, kitchen refrigerators built in 1993 must use 30 percent less electricity than those built in 1990. Clearly the elimination of CFCs should not involve making refrigerators that are much less efficient than present models. Therefore the use of CFCs in cooling is the most difficult of their common uses to eliminate.

The cooling industry is enormous. Cooling equipment worth $40 billion is sold in the United States each year, and, since it has a long useful lifetime, the total value of installed equipment is about $200 billion. Thus the appliance industry and other cooling-equipment manufacturers face a daunting challenge. The prospects of millions of fatal cancers, tens of millions of cataracts, and continued enormous energy consumption and attendant greenhouse-gas emissions have led to government regulations that are driving a $40-billion-per-year industry to an unprecedented crisis. The situation dwarfs the problems of the DOE's nuclear-weapons complex, a mere $12-billion-per-year industry responsible for less environmental and human-health damage.

Stopgap solutions to the CFC crisis are required immediately and are, in fact, in progress. Industry has begun extensive recycling of CFCs, especially in air-conditioner repair. Some new appliances will soon use hydrochlorofluorocarbons, which tend to break down and then rain out before carrying their chlorine to the stratosphere, and hydrofluorocarbons, which contain no chlorine at all. One German manufacturer is using ordinary hydrocarbons (a mixture of propane and butane) as working fluids, recognizing that their flammability poses negligible danger since only small quantities are used.

But these new chemicals have disadvantages. They are less compatible with lubricants than are CFCs, so they may cause present compressors to wear out more quickly. HCFCs and HFCs also significantly reduce the efficiency of cooling machinery. They are also greenhouse gases, with roughly 1000 times the global-warming potential of carbon dioxide per molecule, so their use will probably eventually be limited by international agreements. Finally, as they do not occur in nature, their release into the environment in million-ton quantities, like the release of CFCs, will be an experiment in atmospheric chemistry with unpredictable consequences.

Completely different cooling technologies that don't use any of these chemicals are still needed. Many are being developed, including Rankine cycles using CO_2, Stirling cycles using helium, and cooling by the Peltier (thermoelectric) effect. Almost by accident three new cooling technologies—thermoacoustic refrigeration, the related Sonic Sompressor, and Malone refrigeration—have been developed in part here at the Laboratory; each may become part of intermediate or long-term solutions to the CFC problem. They are described in the accompanying articles. □

type of cooling machinery. Its design takes advantage of modern fabrication techniques and environmentally benign materials but is inspired by a turn-of-the-century invention.

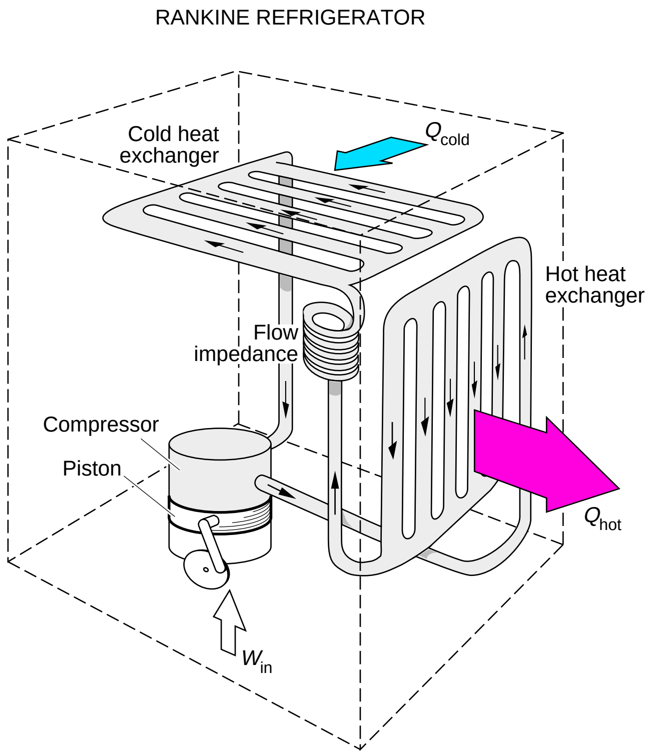
Refrigerators and air conditioners are based on heat pumps, machines that exploit mechanical power to pump heat from low temperature to high temperature. They operate by taking a working fluid through cycles of temperature and pressure changes in which the working fluid absorbs heat at low temperature (from the air inside the refrigerator, say) and loses ("rejects") heat at the higher temperature in the room. Figure 1 shows an example of such a thermodynamic cycle, the Rankine cycle used in all present household refrigerators, in which the cooling (absorption of heat at low temperature) is produced by the evaporation of the CFC working fluid. Reversing the cycle of a heat pump makes a heat engine, which converts heat to mechanical power as heat flows from high temperature to low. For example, the steam turbines that drive electric generators in large power plants use the reverse of the Rankine cycle shown in Figure 1. Water is the working fluid; its expansion on evaporation drives the turbines.

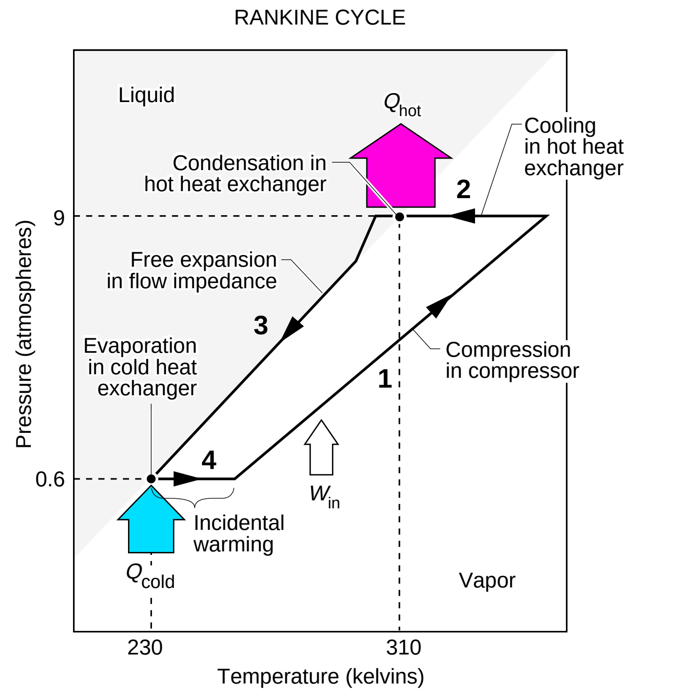
Heat pumps and engines are among the wonderful machines developed during the nineteenth century, when pistons, crankshafts, flywheels, and automatically timed valves began to replace the labor of horses, oxen, and people. Engines removed water from mines and later propelled ships and trains. Refrigerators preserved beef on the two-month voyage from Argentina or New Zealand to England. Today, the internal-combustion engines in cars and the heat pumps in refrigerators, the mature descendants of those nineteenth-century inventions, work unobtrusively and reliably—but not so well that further improvement is impossible.

In the 1970s, John Wheatley, then a physics professor at the University of California, San Diego, and at the height of a distinguished career of academic research into the properties of liquid helium, became interested in improving the efficiency of such ubiquitous heat-engine machinery. But there was trouble at UCSD. Some faculty thought that the work was "not really physics," and Wheatley shared their concern that heat-engine research would not prepare graduate students for a traditional career in academia. Furthermore, the research required more sophisticated fabrication techniques than were available at universities. So (with much encouragement from Jay Keyworth, then leader of the Laboratory's Physics Division) Wheatley moved to Los Alamos in early 1981 and assembled a team that included Al Migliori, Tom Hofler, Heikki Collan, and me, to begin fundamental investigations of old and new concepts in the fields of refrigeration and power generation. One of our research areas is described in "Thermoacoustic Engines and Refrigerators."

This article discusses another area of our investigations: refrigerators and heat engines that use liquids, without change of phase to gas, as the working fluid. We called such devices Malone refrigerators and engines after the first engineer to build such machines (see "John Malone and the Invention of Liquid-Based Engines"). Malone's ideas had been ignored for fifty years, in part because of a common misconception among scientists that liquids have small thermal-expansion coefficients (part of a larger misconception that liquids resemble idealized hydraulic fluid) and therefore do not couple heat to work well enough for use in engines and heat pumps. In heat pumps, working fluids must cool as the fluid is depressurized in order to absorb heat from the area to be cooled; in engines, working fluids must expand on heating in order to do work. The cooling on depressurization and thermal expansion of a material are both proportional to one thermodynamic property: its thermal-expansion coefficient. The thermal-expansion coefficients of gases are large, those of liquid-gas mixtures at the boiling point are essentially infinite (a fact that explains why evaporation and condensation are so useful in the Rankine cycle), but those of liquids are usually small. However, as Figure 2 illustrates, near their critical points liquids do have large thermal-expansion coefficients, larger in fact than that of an ideal gas. (The critical point is the temperature, T_{critical}, and pressure, P_{critical}, above which the liquid and gas phases of a substance are indistinguishable.) Hence liquids can indeed serve as working fluids in engines and refrigerators.

Liquids also have advantages over gases as working fluids. For instance, as shown in Figure 3, liquids are far less compressible than gases; that is, a given volume change causes a larger pressure change in a liquid than in a gas. Large pressure changes are desirable because the heat transferred in the heat exchangers is proportional to the pressure change in the previous step. Low compressibility allows fractional pressure changes to be made large with modest fractional





Conservation of energy: $Q_{hot} = Q_{cold} + W_{in}$. That is, heat rejected (at high temperature) equals heat added (at cold temperature) plus work done by the compressor.

Figure 1. The Rankine Cycle in a Household Refrigerator
The heat pumps in household refrigerators and nearly all other present cooling equipment use the Rankine cycle. This design has been by far the most popular for decades because it is simple and reliable: Many refrigerators run for thirty years with little or no maintenance, and the cost to buy and to run them is low. On the left is a sketch of the heat pump in a kitchen refrigerator. As the working fluid flows through the heat pump in a continuous loop, each element of the working fluid helps cool the refrigerator compartment by going through the thermodynamic cycle shown on the right, the Rankine cycle. The cycle requires work to be done on each fluid element by compressing it in order to raise its temperature above room temperature. The cycle has four steps: (1) The power piston in the compressor does work W_{in} on the element of working fluid, which is in vapor form, by greatly increasing its pressure. As the vapor is compressed, its temperature rises above room temperature. The high pressure is what drives the working fluid around the heat pump. (2) In the hot heat exchanger the hot vapor condenses as it rejects an amount of heat Q_{hot} to the air in the room. The pressure of the fluid remains constant. (3) The working fluid, initially all in liquid form, cools to below the temperature inside the refrigerator compartment by undergoing a free expansion in the flow impedance, a narrow tube that resists the flow of the working fluid so that it emerges at the pressure required for the next step. (4) In the cold heat exchanger the cold liquid absorbs heat from the refrigerator compartment by evaporating. The vapor absorbs a little additional heat on its way to the compressor; the total heat absorbed is Q_{cold}. The pressure of the fluid remains constant.

volume changes of the liquid. Therefore the volume change created by the power piston can be small compared to the volume available in the heat exchangers. Thus either the power piston and other mechanical components involved in volume changes can be smaller and simpler than in a gas engine, or the heat exchangers can be more capacious and consequently more efficient. The low compressibility of liquids also leads to low stored elastic energy per unit volume when they are pressurized, diminishing the hazard of explosions.

Another advantage of liquid working fluids is heat capacities per unit volume that are orders of magnitude larger than those of gases at the pressures typically reached in refrigerators and engines. Therefore when the working fluid is a liquid, the volume of working fluid that must flow through a heat exchanger is orders of magnitude less. As a result, far less mechanical power is required to pump a liquid through the heat exchanger, and heat exchangers for liquids can be far smaller. Because the heat exchanger can be particularly compact if it transfers heat to or from another liquid stream, particularly promising applications for Malone refrigerators and engines include those that both absorb heat from and reject heat to water. Among practical examples are water-cooled water chillers providing air conditioning for large buildings and perhaps ocean thermal energy conversion, in which the temperature difference between the surface and the depths of the ocean is used to produce mechanical power and ultimately electricity.

On the other hand, the high heat capacities of liquids also present a problem. Compressing or depressurizing a liquid without transferring heat changes its temperature relatively little, but refrigeration requires that the working fluid undergo relatively large temperature changes. In particular, as the working fluid flows from the hot heat exchanger to the cold heat exchanger, its temperature must decrease from room temperature to below the temperature inside the refrigerator. Cooling a liquid through such a wide temperature range must be accomplished by removing heat from the liquid in addition to depressurizing it. The most efficient way to remove the heat is to store it and use it later to warm the liquid flowing from the cold heat exchanger to the hot heat exchanger. This process is called regeneration.

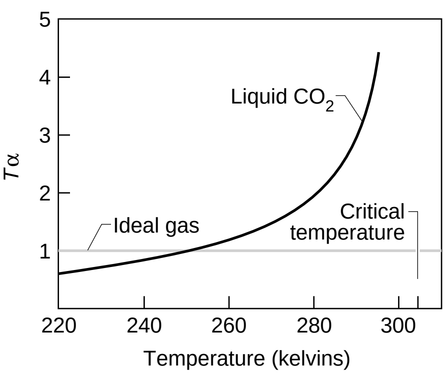


Figure 2. Thermal-Expansion Coefficient of CO_2 near $T_{critical}$
The thermal-expansion coefficient, denoted by Æ , is defined as (@V =@T)P =V , where V is volume, T is temperature, and P is pressure. Here TÆ (the dimensionless thermal-expansion coefficient) of CO_2 is plotted as a function of temperature at the critical pressure of CO_2, 73 atmospheres. Like many thermodynamic derivatives, the thermal-expansion coefficient goes to infinity exactly at the critical point. Plotted in gray is TÆ for an ideal gas, which is unity at any temperature and pressure. Since the gases often used in heat engines and heat pumps are nearly ideal, the fact that TÆ of CO_2 near the critical point exceeds TÆ of an ideal gas implies

The first thermodynamic cycle to use regeneration was the Stirling cycle, invented for engines by the Reverend Robert Stirling in 1816. As shown in Figure 4, a Stirling heat pump or engine differs from a Rankine heat pump or engine in including a regenerator, which consists of walls that bound a number of narrow channels through which the working fluid flows from the hot heat exchanger to the cold heat exchanger and back. (A Stirling machine also differs from a Rankine machine in that the working fluid flows back and forth rather than around a continuous loop.) The temperature of the channel walls decreases steadily from the temperature of the hot heat exchanger at one end of the regenerator to that of the cold heat exchanger at the other end. The walls should have low heat conductance along the channels so that they do not conduct heat from the hot heat exchanger to the cold heat exchanger. The narrowness of the channels creates excellent thermal contact between the working fluid and the walls. Therefore the temperature of each small element of the working fluid is always nearly the same as that of the adjacent part of the channel walls. As the working fluid is displaced through the hot heat exchanger and the regenerator (step 4 in Figure 4), it gradually cools to the temperature of the cold heat exchanger by transferring heat to the channel walls, which store the heat. (The heat capacity of the walls should be sufficiently high that the stored heat does not change their temperature significantly.) The working fluid is then ready to be further cooled by decompression in step 1. When the cold working fluid flows back through the regenerator in step 2, it is warmed by the heat stored in the channel walls. Since the fluid is always at almost the same temperature as the nearby walls, heat transfers to and from the regenerator are nearly the reverses of each other. Therefore each part of the regenerator is restored to its original state at the end of each cycle.

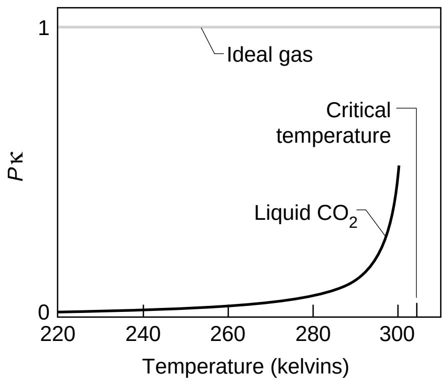


Figure 3. Compressibility of CO_2 near $T_{critical}$
The compressibility, denoted by ∑ , is defined as (@V =@P)T =V . Here, P∑ (the dimensionless compressibility) of CO_2 is plotted versus temperature at the critical pressure of CO_2, 73 atmospheres. Plotted in gray is P∑ for an ideal gas, which, like TÆ , is equal to unity at any temperature and pressure. Since P∑ is lower for liquid CO_2 than for an ideal gas except at temperatures closer to the critical point than are shown, liquid CO_2 has advantages as a working fluid over gases.

In 1981 we chose the Stirling cycle for our research on engines and heat pumps that used liquids as working fluids. As the working fluid in our first Stirling heat pump, we chose liquid propylene (C_3H_6) from more than a dozen hydrocarbon, CFC, and inorganic fluids with critical points just above room temperature. The machine we built,

John Malone and the Invention of Liquid-based Engines

It is hard to imagine using a liquid instead of a gas in a heat engine, but John Malone did—perhaps partly because he was not prejudiced by a proper scientific education. Malone was born in England in 1880. His formal education ended in his eighteenth year, when (probably in part to avoid some trouble with the police) he joined the merchant marine. He remained at sea for nearly all of the next fourteen years; during that time he was wounded seventeen times in Arab and Latin-American wars.

Leaving the merchant marine, Malone founded the Sentinel Instrument Company and, later, the Fox Instrument Company. He began experimenting with liquids as engine working fluids in the 1920s. As part of that project, he measured the compressibilities and thermal-expansion coefficients of many liquids, including hydrocarbons, mercury, carbon dioxide, and sulfur dioxide. In 1925 he completed his first liquid-based engine, shown on the title page of "Malone Refrigeration. It burned coal, used high-pressure liquid water as working medium, and produced 50 horsepower. Malone referred to that first engine as crude and cumbersome, but claimed that with perseverance it would have eventually produced 500 horsepower.

Instead, in 1927 Malone completed a much smaller and more versatile 50-horsepower water engine and began an extensive program of experimentation with it. Malone claimed that his second engine was very efficient. In 1931 he wrote, "Trials by three different independent engineers gave 27% indicated efficiency. Thus, after allowing for furnace and mechanical losses in a commercial engine, 20% overall efficiency between the heat in the coal and the shaft horsepower can be expected." The efficiencies of the steam engines that powered ships at the time were between 9 and 12 percent and those of locomotives were between 5 and 7 percent, much lower than the efficiency of Malone's engine.

Curiously, the "27% indicated efficiency" quoted above is the only quantitative experimental datum in any of Malone's publications and patent disclosures. In a 1939 letter to Selwyn Anderson, Malone wrote about his measurements, "I refused to publish this information because it cost me a lot to learn it and I may yet obtain some reward if it is not known. Also because to my amazement I found my enemies were alleged centers of learning. Universities and the like." Later in the same letter his bitterness is more evident: "A study of liquids as mediums in thermodynamics will teach an engineer more about the art of thermodynamics than all the universities on earth, or the memory men who infest them, and knowledge for knowledge's sake is better than their parasitical life." After Malone's death in 1959, his son Ray wrote, "Now as patent rights have long expired I can see no advantage in publishing any of the information which he accumulated while developing his liquid engine."

We can only guess why Malone's promising work came to an end. The worldwide economic depression of the 1930s must have made venture capital scarce. Some may have dismissed the idea of liquid working fluids because it contradicted conventional wisdom. Large coal-fired steam turbines with 20 percent efficiency were in the ascendancy for applications above 10,000 horsepower. The internal-combustion engine (including what we know today as the diesel engine) was already more advanced than Malone's engine, and its incomparable power-to-weight ratio made it seem the only practical choice for airplanes and automobiles. By the time the Great Depression and then World War II had ended, the steam engine was disappearing, internal-combustion engines and turbines were becoming ubiquitous, and Malone's work had been forgotten. It took another independent thinker, the late John Wheatley, to see the promise in Malone's work fifty years later and resume the study of liquid-based engines. □

Figure 4. A Malone Heat Pump Using the Stirling Cycle

Typically an electric motor (not shown) supplies the work required to compress the liquid working fluid by turning a crankshaft (also not shown) that operates the pistons. (Only the power piston does work; the displacer piston is operated by the same crankshaft for proper phasing.) The graph shows the thermodynamic cycle of pressure and temperature changes undergone by the working fluid; the object is the absorption of heat Q_{cold} at the cold heat exchanger in step 2. The graph of the Stirling cycle has been simplified by assuming that the four steps are separate from each other; that is, that each piston is stationary while the other one moves. In real machines the pistons oscillate sinusoidally in time. The simultaneous motion of the two pistons causes consecutive steps of the cycle to overlap; the overlaps would be reflected in a graph of the Stirling cycle by rounded corners. In addition, because the volume of working fluid displaced by the pistons is small compared to the volume in the heat exchangers and regenerator, different elements of the working fluid are carried through variations of the single thermodynamic cycle shown. Liquids are practical working fluids for the Stirling cycle and others because they possess certain thermodynamic properties. Specifically, their thermal-expansion coefficients (α) are large, their compressibilities (κ) are small, and their heat capacities per unit volume at constant pressure (C_P) are large. The heat Q_{cold} absorbed in step 2 is given by $(\partial Q/\partial P)_T \Delta P$, where ΔP is the pressure difference brought about in step 1. From a thermodynamic identity and the definition of α,

$$\left(\frac{\partial Q}{\partial P}\right)_T = T\left(\frac{\partial S}{\partial P}\right)_T = -T\left(\frac{\partial V}{\partial T}\right)_P = -T\alpha V\,;$$

where S is entropy. It follows that $Q_{cold} \approx T_{cold}\alpha V\Delta P$, where T_{cold} is the temperature of the cold heat exchanger, and V is the volume of liquid displaced through the heat exchangers. Thus a large thermal-expansion coefficient leads to absorption of a large amount of heat. On the other hand, the energy wasted in heat absorption is proportional to the square of the temperature drop ΔT caused by the depressurization in step 1, which is given by $\Delta T = (T\alpha/C_P)\Delta P$. Therefore a working fluid with a small value of α/C_P is desirable. Fortunately, liquids near their critical points typically have thermal-expansion coefficients that are a little larger than that of an ideal gas and heat capacities per unit volume that are orders of magnitude larger than that of an ideal gas, so that the cooling can be both powerful and efficient. Finally, the low compressibilities of liquids are an advantage in steps 1 and 3, as can be seen from the relation $\Delta P = \Delta V/V\kappa$, which follows from the definition of κ. Thus large pressure changes ΔP can be achieved in the entire liquid volume V by power-piston strokes that displace small volumes ΔV.

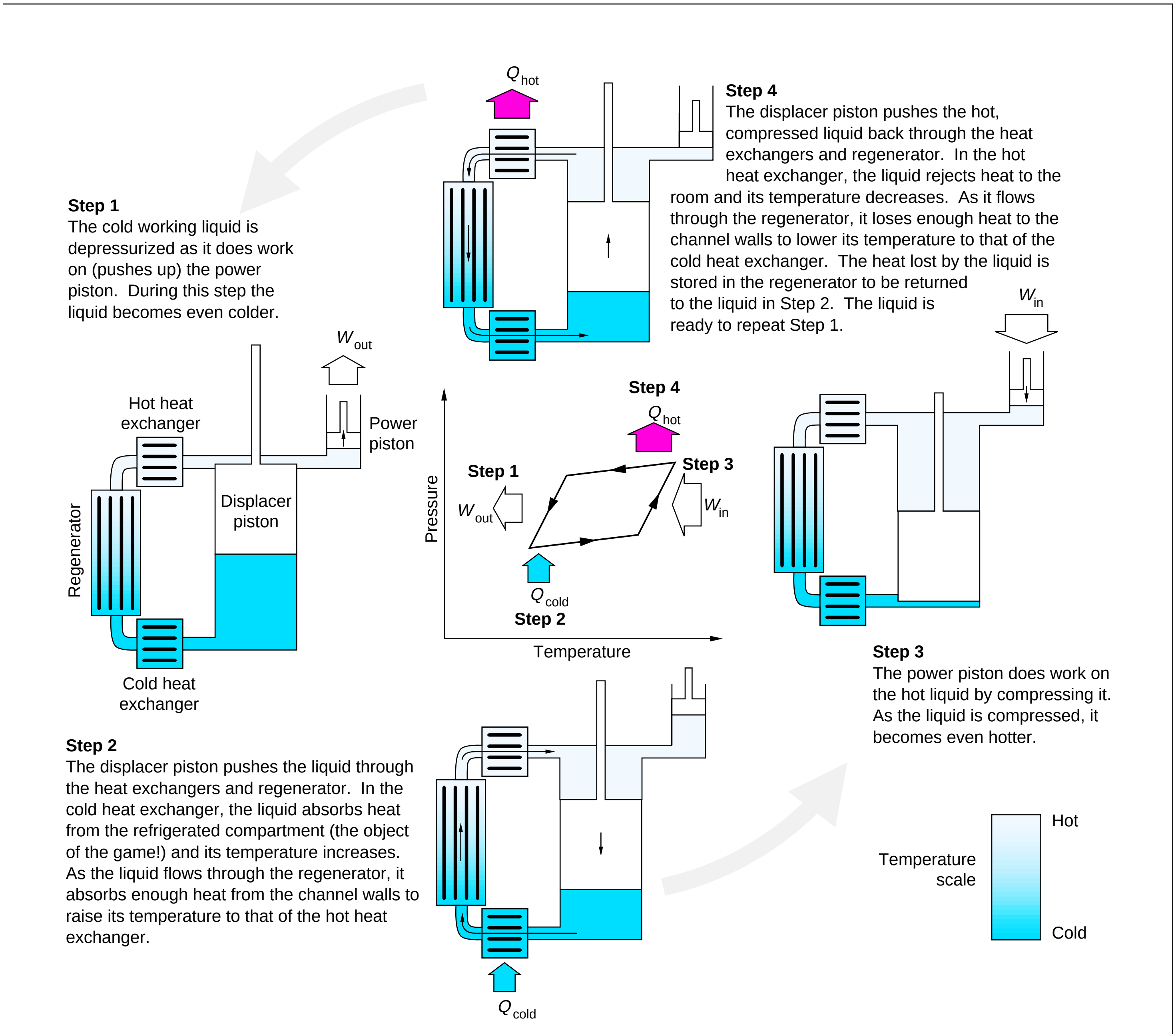
shown in Figure 5, could function as engine or heat pump, and was heavily instrumented to allow simultaneous measurement of mechanical power, heat flow, and temperatures and pressures throughout. Our extensive program of measurement on that machine, coupled with simple theoretical work, taught us how liquid properties, machine geometry, and cycle thermodynamics all work in concert to process each of the countless volume elements of the liquid through its own closed cycle. As a result, we believe we can now predict the power and efficiency of Malone machines with reasonable accuracy.

As this fundamental phase of the work drew to a close in the middle and late 1980s, our situation underwent a number of important changes. Recognition that CFCs cause unacceptable health hazards and environmental damage by destroying atmospheric ozone (see “CFCs and Cooling Equipment: The Size of the Problem”) gave new motivation to our development of cooling technologies based on other working fluids. But the pace of the research did not increase. Wheatley’s sudden death deprived us of our leader, our keenest intellect, and our most successful fundraiser.

Escalating costs of experimental work—especially of fabrication—slowed progress further. The next step in the development of a Malone refrigerator needed much more support than our earlier work, but support was hard to find: the DOE’s Office of Basic Energy Sciences was not interested in increasing our funding for this applied work, and the DOE offices that support conservation and renewable energy seemed interested in funding only those projects that promised to make an impact in the marketplace within two or three years. A prolonged, time-consuming attempt to get Navy support for our research ended with no Navy funds coming to Los Alamos, but with researchers at a Navy laboratory beginning their own development of Malone machines. Fortunately, during this difficult interim phase of basic technology development, the Laboratory’s own research funds have provided a partial bridge.

For our current Malone-refrigerator research, we picked the most environmentally acceptable of the liquids with critical points just above room temperature—carbon dioxide. Liquid carbon dioxide and dilute mixtures of methanol or ethanol in liquid carbon dioxide are efficient and safe working fluids for Malone refrigeration. The amount of carbon dioxide used has negligible environmental impact compared even with the effect of the pounds per capita per day we each exhale; the amount of alcohol also has negligible environmental impact.

Our first Malone machine, though designed for ease of measurement rather than for efficiency, was half as efficient as present-day CFC-based equipment. There is little



doubt that a liquid-CO_2 Malone refrigerator can be built with an efficiency higher than that of existing CFC-based refrigerators. There are two important questions, which we are pursuing simultaneously. Can an efficient Malone refrigerator be built inexpensively enough to enjoy widespread manufacture? What are the environmental costs of the manufacturing process?

The CO_2 Malone refrigerator we are building now should provide partial answers to those questions. We are using modern fabrication techniques when necessary but are avoiding expensive (and sometimes environmentally questionable) "space-age" materials and techniques. For example, the heat-exchanger/regenerator assembly is a furnace-brazed stack of stainless steel sheets; slots in some of the sheets form fluid channels when the sheets are assembled. Although we now fabricate the sheets by photolithography and chemical milling for speed and flexibility, we know that they can ultimately be mass-produced very inexpensively and cleanly by punching. The brazing metal is pure copper, which is cheaper than the more commonly used silver alloys. The copper can be very thin and can be applied to the sheets by electroplating before punching or chemical milling, as another cost-saving measure. And we hope we can eventually save still more money by making some of the parts from ordinary carbon steel instead of stainless steel.

The configuration of the pistons and other moving parts is also influenced by the need to reduce costs. Our original Malone machine had too many high-precision and hence expensive moving parts, including a dozen roller bearings. The cost of the bearings alone was higher than the cost of the compressor in a con-

Thermoacoustic Engines and Refrigerators

Thermoacoustic effects, which convert heat energy to sound, have been known for over a hundred years. They have generally been considered mere curiosities, but in the early 1980s our engine-research group at Los Alamos, led by John Wheatley, began to consider thermoacoustic effects as a practical way to make efficient engines. One serious impediment to rapid progress on our experimental Malone engines was the large number of precision moving parts required. While looking for simpler engine designs, we came across the work of Peter Ceperley at George Mason University, who had realized that the timing between pressure changes and motion in Stirling engines and heat pumps is the same as in a traveling sound wave. Inspired by his work, we eventually invented thermoacoustic heat pumps (and new types of thermoacoustic engines) that had at most one moving part.

As Figure 1 shows, our thermoacoustic heat pumps use standing (rather than traveling) sound waves to take the working fluid (a gas) through a thermodynamic cycle. They rely on the heating and cooling that accompany the compression and expansion of a gas in a sound wave. Although ordinary, conversational-level sound produces only tiny heating and cooling effects, extremely loud sound waves produce heating and cooling effects large enough to be useful. Whereas typical heat pumps have crankshaft-coupled pistons or rotary compressors, thermoacoustic heat pumps have no moving parts or a single flexing moving part, such as a loudspeaker, and have no sliding seals. The lack of moving parts gives thermoacoustic refrigerators the advantages of simplicity, reliability, and low cost. Because the sound waves are confined in sealed cavities, the machines are fairly quiet.

For us thermoacoustic heat pumps had the additional advantages of conceptual elegance and easy, low-cost prototype development. We hoped that those features would lead to near-term successes (which would help keep our research well funded and lend credibility to our longer-term Malone program). The development of thermoacoustic refrigerators has indeed had successes, such as the 1992 flight in a space shuttle of a thermoacoustic refrigerator built at the Naval Postgraduate School and a 1993 test of a thermoacoustic sonar projector (an engine rather than a heat pump) by Bill Ward in the Laboratory's Advanced Engineering Technology Group.

After Tim Lucas, an inventor, noticed an article about our thermoacoustic work in a popular science and technology magazine, he added yet another chapter to the story of novel refrigeration at the Laboratory. Lucas had invented the Sonic Compressor (Figure 2), a device for compressing conventional refrigerant vapors that con-

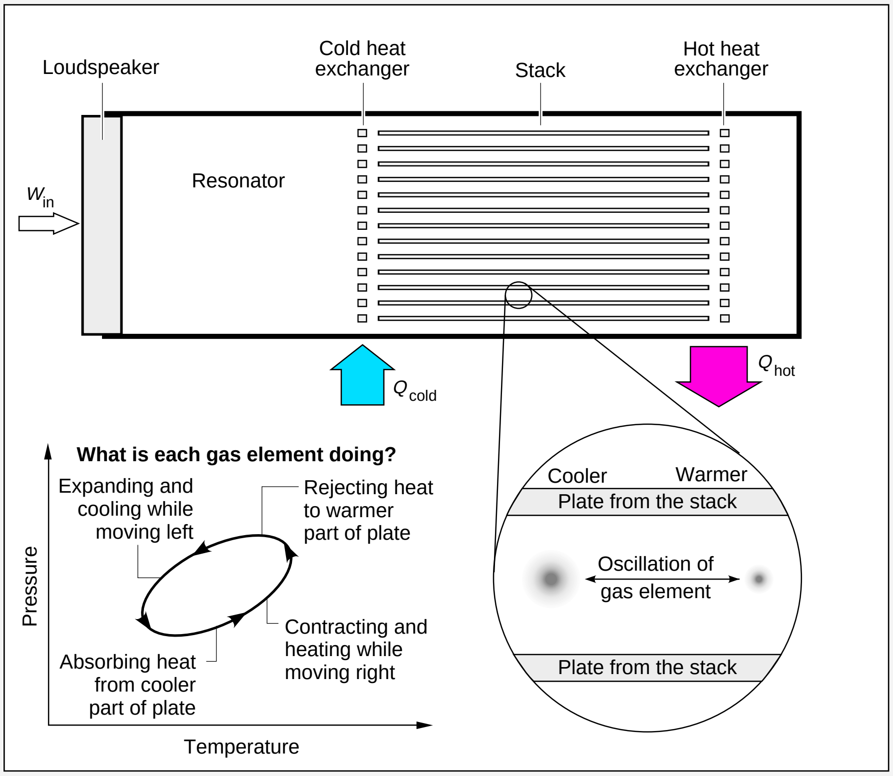


Figure 1. The Thermoacoustic Refrigerator

An electrically driven, radically modified loudspeaker maintains a standing sound wave in an inert gas in a resonator. The sound wave interacts with an array of parallel solid plates called the stack. The resulting refrigeration can be understood by examining a typical small element of gas between the plates of the stack. As the gas oscillates back and forth because of the standing sound wave, it changes in temperature. Much of the temperature change comes from compression and expansion of the gas by the sound pressure (as always in a sound wave), and the rest is a consequence of heat transfer between the gas and the stack. In the example shown the length of the resonator is one-fourth the wavelength of the sound produced by the speaker, so all the elements of gas are compressed and heated as they move to the right and expanded and cooled as they move to the left. Thus each element of gas goes through a thermodynamic cycle in which the element is compressed and heated, rejects heat at the right end of its range of oscillation, is depressurized and cooled, and absorbs heat at the left end. Consequently each element of gas moves a little heat from left to right, from cold to hot, during each cycle of the sound wave. The combination of the cycles of all the elements of gas transports heat from the cold heat exchanger to the hot heat exchanger much as a bucket brigade transports water. The spacing between the plates in the stack is crucial to proper function: If the spacing is too narrow, the good thermal contact between the gas and the stack keeps the gas at nearly the same temperature as the stack, whereas if the spacing is too wide, much of the gas is in poor thermal contact with the stack and does not transfer heat effectively to and from it.

tains no sliding parts. Instead a resonant sound wave in a cavity compresses the vapor and two one-way valves ensure that only low-pressure vapor enters and only high-pressure vapor leaves the compressor. Since the Sonic Compressor needs no lubricating oil, it is attractive for compressing HFC refrigerants, which do not destroy the ozone layer but have the drawback of being less compatible with lubricants than CFCs are (see "CFCs and Cooling Equipment: The Size of the Problem"). The lack of sliding parts should also lead to higher efficiency in small systems. Furthermore, the Sonic Compressor can replace the piston-driven compressor in present refrigerators without requiring any changes in other parts.

Lucas needed to suppress the production of shock waves in his compressor by the high-amplitude sound because the shock waves wasted energy by turning it into heat. He sought help from us because of our experience with high-amplitude sound in thermoacoustics. Working together we found that the shock waves resulted from nonlinear self-interactions in the desired fundamental resonance in the cavity and from unwanted resonances at frequencies that were exact integral multiples of the fundamental frequency. When we changed the shape of the cavity to that shown in Figure 2, the frequencies of the extra resonances changed so that they were no longer significantly excited by nonlinear self-interaction in the fundamental.

Lucas's collaboration with us was an example of totally successful "tech transfer." During

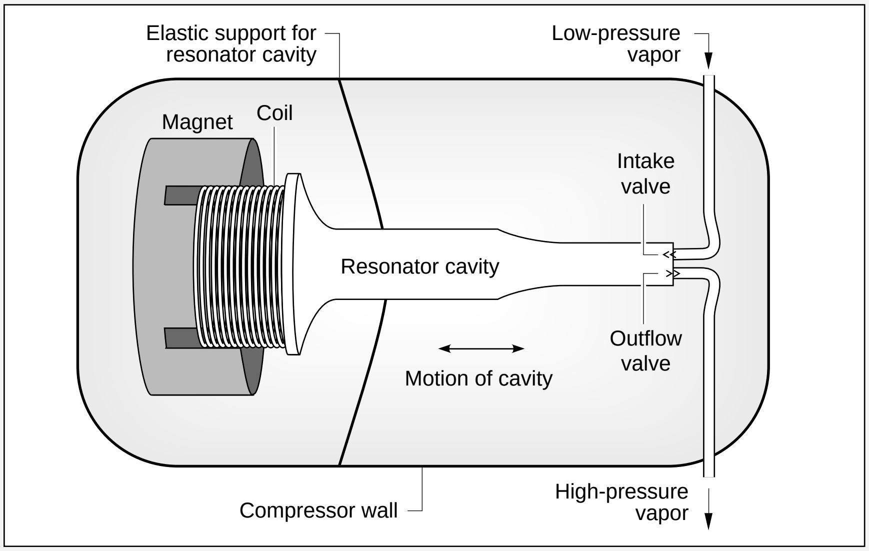


Figure 2. The Sonic Compressor
The Sonic Compressor uses electric power to compress a conventional refrigerant vapor by means of a high-amplitude sound wave; the model depicted can replace the piston compressor in a conventional cooling system such as the household refrigerator shown in Figure 1 of "Malone Refrigeration." The electricity drives a radically modified loudspeaker that shakes a cavity back and forth at a resonance frequency of the working-fluid vapor inside (300 hertz). In the figure the cavity is shown at the rightmost point of the vibration. The motion of the cavity causes the vapor to slosh back and forth—in other words, the motion generates a standing sound wave. The shape of the cavity is designed to prevent the formation of shock waves. The standing sound wave compresses and expands the gas; at the end of the tube farther from the loudspeaker, the range of pressure is 8 atmospheres. A pair of one-way valves at that end, which are opened and closed at the operating frequency by the pressure itself, admits low-pressure vapor from the intake pipe and ejects high-pressure vapor into the outflow pipe.

the year he spent here, we solved his shock problem. Of equal importance to him, we did not jointly invent anything patentable, so the business aspects of his project were not complicated by the involvement of intellectual-property rights belonging to the Laboratory. As Lucas's visit was successful, the Sonic Compressor could come into production in a few years. Thermoacoustic refrigeration will not be ready for the market until a few years later. Malone refrigeration will take still more time to develop but appears to be the most efficient option of the three to which we at Los Alamos are contributing. ▫

ventional CFC refrigerator! A totally different design was clearly required if Malone technology was ever to enjoy widespread use. So in our present CO_2 machine we are using a linear free-piston configuration, which was invented only recently and is being employed in gas-based Stirling engines intended for solar power or for use in space. The pistons in a linear free-piston machine are driven, not by a rotating motor connected to a crankshaft, but by a "linear" electric motor that provides reciprocating force and motion directly in the same way as a loudspeaker. This configuration minimizes the number of moving parts and eliminates the need for high-force bearing surfaces. Careful design can even eliminate the need for any mechanical connection to the displacer piston—the piston moves with the correct amplitude and phase simply in response to the fluid pressures acting on it.

The present crisis in the cooling industry is a unique opportunity for a new, potentially more efficient technology to break the monopoly of a technology that has enjoyed decades of incremental improvement. The primary challenge of the next year or two is to keep this difficult experimental project moving ahead, though the funding only pays for a third of the time of one researcher, while trying to attract the interest of an industrial collaborator. At best, years of further work costing millions of dollars will be required to bring Malone refrigeration to the threshold of possible widespread application. Meanwhile, as described in "CFCs and Cooling Equipment: The Size of the Problem," industry is proceeding promptly with more straightforward interim measures. In the intermediate time scale, mature new technologies such as the Sonic Compressor and perhaps thermoacoustic re-

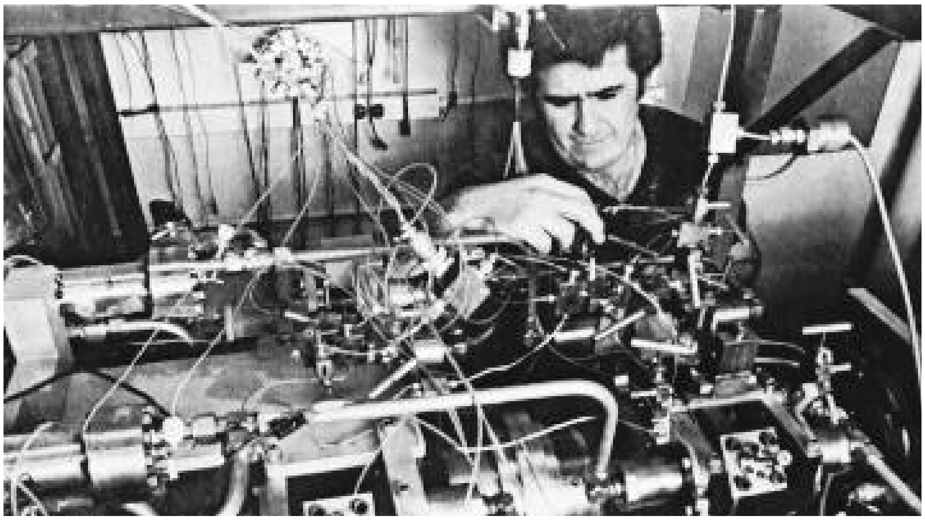


Figure 5. Chris Espinoza Adjusts Our First Malone Refrigerator (1989)

frigeration may come into use.

If we succeed in developing a widely used refrigerator, it would be the first Laboratory product since the implementation of the Atmospheric Test Ban Treaty to find its way into homes and businesses throughout the world. If we fail for unforeseen technical reasons, we will not regret having tried. But if we fail because of inadequate support, an opportunity to improve the world environment and reduce its energy consumption will have been lost.

It is a pleasure to acknowledge the contributions of Al Migliori, Sonia Balcer, Chris Espinoza, Frank Murray, and Alex Brown to the development of Malone refrigeration at the Laboratory and to thank the Department of Energy's Office of Basic Energy Sciences for steady support of fundamental engine and refrigerator research at Los Alamos National Laboratory.

Further Reading

R. E. Benedick. 1991. *Ozone Diplomacy*. Cambridge, Massachusetts: Harvard University Press, 1991.

Energy Information Administration. 1989. Potential costs of restricting chlorofluorocarbon use. EIA Service Report SR/ESD/89-01. Washington, DC: U. S. Department of Energy.

Gregory W. Swift. 1990. Malone Refrigeration. *ASHRAE Journal*, November: 28–34.

Gregory W. Swift. 1989. A Stirling engine with a liquid working substance. *Journal of Applied Physics* 65: 4157–4172.

Steven L. Garrett and Thomas J. Hofler. 1992. Thermoacoustic refrigeration. In *Technology 2001: Proceedings of the Second National Technology Transfer Conference and Exposition*. NASA Conference Publication 3136, 2: 397–406. Washington, DC: NASA.

Gregory W. Swift. 1988. Thermoacoustic engines. *Journal of the Acoustical Society of America* 84: 1145–1180.

Information about the Sonic Compressor can be obtained by calling Timothy S. Lucas (president of Sonic Compressor Systems in Glen Allen, Virginia) at (804) 262-3700.

Gregory W. Swift is a staff member in the Condensed Matter and Thermal Physics Group, where he has been working on novel heat engines and refrigerators since 1981. He received his B.S. in physics and mathematics from the University of Nebraska and his Ph.D. in physics from the University of California, Berkeley. From 1983 to 1985 he held an Oppenheimer Fellowship at Los Alamos. He is a fellow of the Acoustical Society of America.

CFCs and Cooling Equipment: The Size of the Problem

As recently as ten years ago, only a few experts appreciated the potentially disastrous consequences of releasing chlorofluorocarbons (CFCs) into the environment. Today it is generally believed that chlorine from CFCs is destroying stratospheric ozone, which shields the earth from the sun's ultraviolet radiation, and that the resulting increased ultraviolet radiation will soon lead to millions of fatal and nonfatal skin cancers and cataracts, to other threats to human health, and to severe damage to agriculture and ecosystems. To mitigate those effects, the United States and most other countries have committed themselves, through the 1987 Montreal Protocol on Substances That Deplete the Ozone Layer and its later revisions, to rapid elimination of CFC production.

The rate of CFC production, though being reduced, is on the order of a million tons per year. CFCs are used extensively as working fluids in refrigerators and air conditioners, as cleaning solvents in electronics and sheet-metal fabrication, and as foaming agents in foam insulation and cushions. CFCs have the advantage for many purposes of being almost chemically inert (their behavior in the upper atmosphere is an exception); in particular they present no fire or poison danger. In addition their lack of interaction with the lubricants used in refrigerator compressors improves the efficiency and the lifetime of the refrigerator. Cooling-system efficiency has major economic and environmental impacts. Cooling consumes about 20 percent of the nation's electricity, at a cost of tens of billions of dollars per year. Kitchen refrigerators alone use 8 percent of the nation's electricity. Most of that electricity is produced by burning fossil fuels. Congress is accordingly requiring ever more efficient cooling equipment; for example, kitchen refrigerators built in 1993 must use 30 percent less electricity than those built in 1990. Clearly the elimination of CFCs should not involve making refrigerators that are much less efficient than present models. Therefore the use of CFCs in cooling is the most difficult of their common uses to eliminate.

The cooling industry is enormous. Cooling equipment worth $40 billion is sold in the United States each year, and, since it has a long useful lifetime, the total value of installed equipment is about $200 billion. Thus the appliance industry and other cooling-equipment manufacturers face a daunting challenge. The prospects of millions of fatal cancers, tens of millions of cataracts, and continued enormous energy consumption and attendant greenhouse-gas emissions have led to government regulations that are driving a $40-billion-per-year industry to an unprecedented crisis. The situation dwarfs the problems of the DOE's nuclear-weapons complex, a mere $12-billion-per-year industry responsible for less environmental and human-health damage.

Stopgap solutions to the CFC crisis are required immediately and are, in fact, in progress. Industry has begun extensive recycling of CFCs, especially in air-conditioner repair. Some new appliances will soon use hydrochlorofluorocarbons, which tend to break down and then rain out before carrying their chlorine to the stratosphere, and hydrofluorocarbons, which contain no chlorine at all. One German manufacturer is using ordinary hydrocarbons (a mixture of propane and butane) as working fluids, recognizing that their flammability poses negligible danger since only small quantities are used.

But these new chemicals have disadvantages. They are less compatible with lubricants than are CFCs, so they may cause present compressors to wear out more quickly. HCFCs and HFCs also significantly reduce the efficiency of cooling machinery. They are also greenhouse gases, with roughly 1000 times the global-warming potential of carbon dioxide per molecule, so their use will probably eventually be limited by international agreements. Finally, as they do not occur in nature, their release into the environment in million-ton quantities, like the release of CFCs, will be an experiment in atmospheric chemistry with unpredictable consequences.

Completely different cooling technologies that don't use any of these chemicals are still needed. Many are being developed, including Rankine cycles using CO_2, Stirling cycles using helium, and cooling by the Peltier (thermoelectric) effect. Almost by accident three new cooling technologies—thermoacoustic refrigeration, the related Sonic Sompressor, and Malone refrigeration—have been developed in part here at the Laboratory; each may become part of intermediate or long-term solutions to the CFC problem. They are described in the accompanying articles. □

John Malone and the Invention of Liquid-based Engines

It is hard to imagine using a liquid instead of a gas in a heat engine, but John Malone did—perhaps partly because he was not prejudiced by a proper scientific education. Malone was born in England in 1880. His formal education ended in his eighteenth year, when (probably in part to avoid some trouble with the police) he joined the merchant marine. He remained at sea for nearly all of the next fourteen years; during that time he was wounded seventeen times in Arab and Latin-American wars.

Leaving the merchant marine, Malone founded the Sentinel Instrument Company and, later, the Fox Instrument Company. He began experimenting with liquids as engine working fluids in the 1920s. As part of that project, he measured the compressibilities and thermal-expansion coefficients of many liquids, including hydrocarbons, mercury, carbon dioxide, and sulfur dioxide. In 1925 he completed his first liquid-based engine, shown on the title page of "Malone Refrigeration. It burned coal, used high-pressure liquid water as working medium, and produced 50 horsepower. Malone referred to that first engine as crude and cumbersome, but claimed that with perseverance it would have eventually produced 500 horsepower.

Instead, in 1927 Malone completed a much smaller and more versatile 50-horsepower water engine and began an extensive program of experimentation with it. Malone claimed that his second engine was very efficient. In 1931 he wrote, "Trials by three different independent engineers gave 27% indicated efficiency. Thus, after allowing for furnace and mechanical losses in a commercial engine, 20% overall efficiency between the heat in the coal and the shaft horsepower can be expected." The efficiencies of the steam engines that powered ships at the time were between 9 and 12 percent and those of locomotives were between 5 and 7 percent, much lower than the efficiency of Malone's engine.

Curiously, the "27% indicated efficiency" quoted above is the only quantitative experimental datum in any of Malone's publications and patent disclosures. In a 1939 letter to Selwyn Anderson, Malone wrote about his measurements, "I refused to publish this information because it cost me a lot to learn it and I may yet obtain some reward if it is not known. Also because to my amazement I found my enemies were alleged centers of learning. Universities and the like." Later in the same letter his bitterness is more evident: "A study of liquids as mediums in thermodynamics will teach an engineer more about the art of thermodynamics than all the universities on earth, or the memory men who infest them, and knowledge for knowledge's sake is better than their parasitical life." After Malone's death in 1959, his son Ray wrote, "Now as patent rights have long expired I can see no advantage in publishing any of the information which he accumulated while developing his liquid engine."

We can only guess why Malone's promising work came to an end. The worldwide economic depression of the 1930s must have made venture capital scarce. Some may have dismissed the idea of liquid working fluids because it contradicted conventional wisdom. Large coal-fired steam turbines with 20 percent efficiency were in the ascendancy for applications above 10,000 horsepower. The internal-combustion engine (including what we know today as the diesel engine) was already more advanced than Malone's engine, and its incomparable power-to-weight ratio made it seem the only practical choice for airplanes and automobiles. By the time the Great Depression and then World War II had ended, the steam engine was disappearing, internal-combustion engines and turbines were becoming ubiquitous, and Malone's work had been forgotten. It took another independent thinker, the late John Wheatley, to see the promise in Malone's work fifty years later and resume the study of liquid-based engines. □

Thermoacoustic Engines and Refrigerators

Thermoacoustic effects, which convert heat energy to sound, have been known for over a hundred years. They have generally been considered mere curiosities, but in the early 1980s our engine-research group at Los Alamos, led by John Wheatley, began to consider thermoacoustic effects as a practical way to make efficient engines. One serious impediment to rapid progress on our experimental Malone engines was the large number of precision moving parts required. While looking for simpler engine designs, we came across the work of Peter Ceperley at George Mason University, who had realized that the timing between pressure changes and motion in Stirling engines and heat pumps is the same as in a traveling sound wave. Inspired by his work, we eventually invented thermoacoustic heat pumps (and new types of thermoacoustic engines) that had at most one moving part.

As Figure 1 shows, our thermoacoustic heat pumps use standing (rather than traveling) sound waves to take the working fluid (a gas) through a thermodynamic cycle. They rely on the heating and cooling that accompany the compression and expansion of a gas in a sound wave. Although ordinary, conversational-level sound produces only tiny heating and cooling effects, extremely loud sound waves produce heating and cooling effects large enough to be useful. Whereas typical heat pumps have crankshaft-coupled pistons or rotary compressors, thermoacoustic heat pumps have no moving parts or a single flexing moving part, such as a loudspeaker, and have no sliding seals. The lack of moving parts gives thermoacoustic refrigerators the advantages of simplicity, reliability, and low cost. Because the sound waves are confined in sealed cavities, the machines are fairly quiet.

For us thermoacoustic heat pumps had the additional advantages of conceptual elegance and easy, low-cost prototype development. We hoped that those features would lead to near-term successes (which would help keep our research well funded and lend credibility to our longer-term Malone program). The development of thermoacoustic refrigerators has indeed had successes, such as the 1992 flight in a space shuttle of a thermoacoustic refrigerator built at the Naval Postgraduate School and a 1993 test of a thermoacoustic sonar projector (an engine rather than a heat pump) by Bill Ward in the Laboratory's Advanced Engineering Technology Group.

After Tim Lucas, an inventor, noticed an article about our thermoacoustic work in a popular science and technology magazine, he added yet another chapter to the story of novel refrigeration at the Laboratory. Lucas had invented the Sonic Compressor (Figure 2), a device for compressing conventional refrigerant vapors that con-

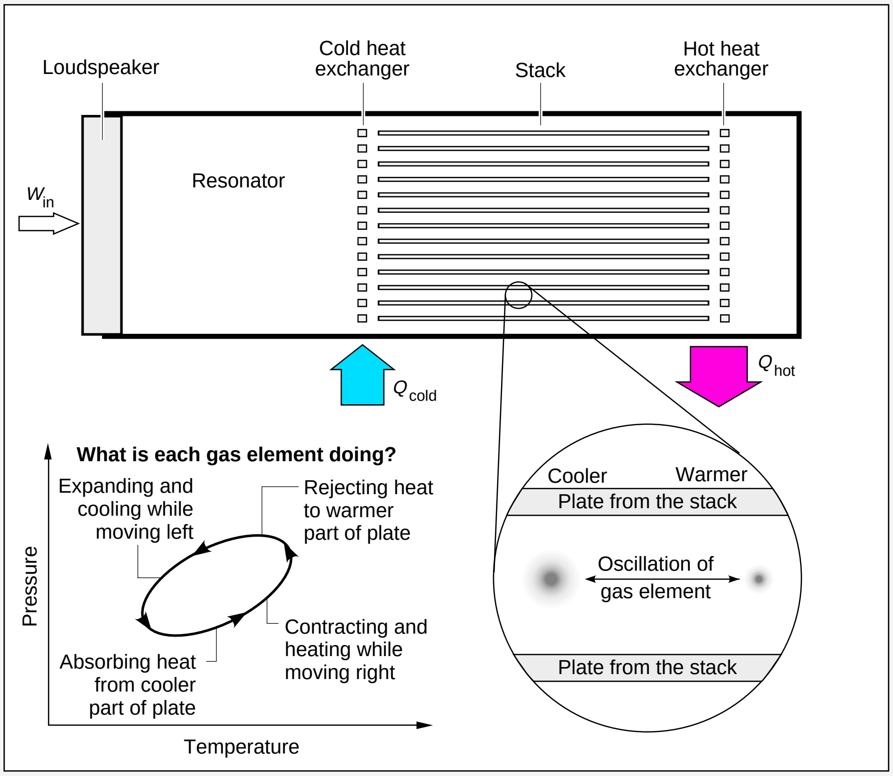


Figure 1. The Thermoacoustic Refrigerator

An electrically driven, radically modified loudspeaker maintains a standing sound wave in an inert gas in a resonator. The sound wave interacts with an array of parallel solid plates called the stack. The resulting refrigeration can be understood by examining a typical small element of gas between the plates of the stack. As the gas oscillates back and forth because of the standing sound wave, it changes in temperature. Much of the temperature change comes from compression and expansion of the gas by the sound pressure (as always in a sound wave), and the rest is a consequence of heat transfer between the gas and the stack. In the example shown the length of the resonator is one-fourth the wavelength of the sound produced by the speaker, so all the elements of gas are compressed and heated as they move to the right and expanded and cooled as they move to the left. Thus each element of gas goes through a thermodynamic cycle in which the element is compressed and heated, rejects heat at the right end of its range of oscillation, is depressurized and cooled, and absorbs heat at the left end. Consequently each element of gas moves a little heat from left to right, from cold to hot, during each cycle of the sound wave. The combination of the cycles of all the elements of gas transports heat from the cold heat exchanger to the hot heat exchanger much as a bucket brigade transports water. The spacing between the plates in the stack is crucial to proper function: If the spacing is too narrow, the good thermal contact between the gas and the stack keeps the gas at nearly the same temperature as the stack, whereas if the spacing is too wide, much of the gas is in poor thermal contact with the stack and does not transfer heat effectively to and from it.

tains no sliding parts. Instead a resonant sound wave in a cavity compresses the vapor and two one-way valves ensure that only low-pressure vapor enters and only high-pressure vapor leaves the compressor. Since the Sonic Compressor needs no lubricating oil, it is attractive for compressing HFC refrigerants, which do not destroy the ozone layer but have the drawback of being less compatible with lubricants than CFCs are (see "CFCs and Cooling Equipment: The Size of the Problem"). The lack of sliding parts should also lead to higher efficiency in small systems. Furthermore, the Sonic Compressor can replace the piston-driven compressor in present refrigerators without requiring any changes in other parts.

Lucas needed to suppress the production of shock waves in his compressor by the high-amplitude sound because the shock waves wasted energy by turning it into heat. He sought help from us because of our experience with high-amplitude sound in thermoacoustics. Working together we found that the shock waves resulted from nonlinear self-interactions in the desired fundamental resonance in the cavity and from unwanted resonances at frequencies that were exact integral multiples of the fundamental frequency. When we changed the shape of the cavity to that shown in Figure 2, the frequencies of the extra resonances changed so that they were no longer significantly excited by nonlinear self-interaction in the fundamental.

Lucas's collaboration with us was an example of totally successful "tech transfer." During

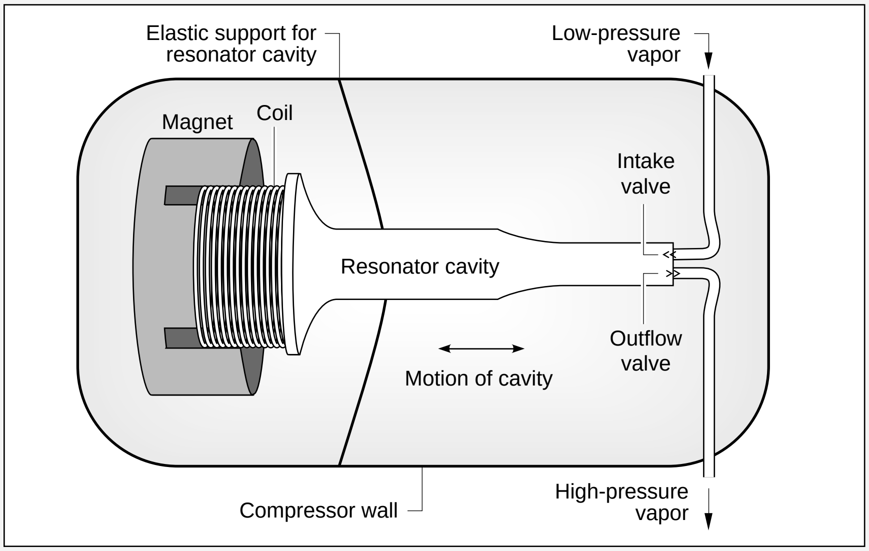


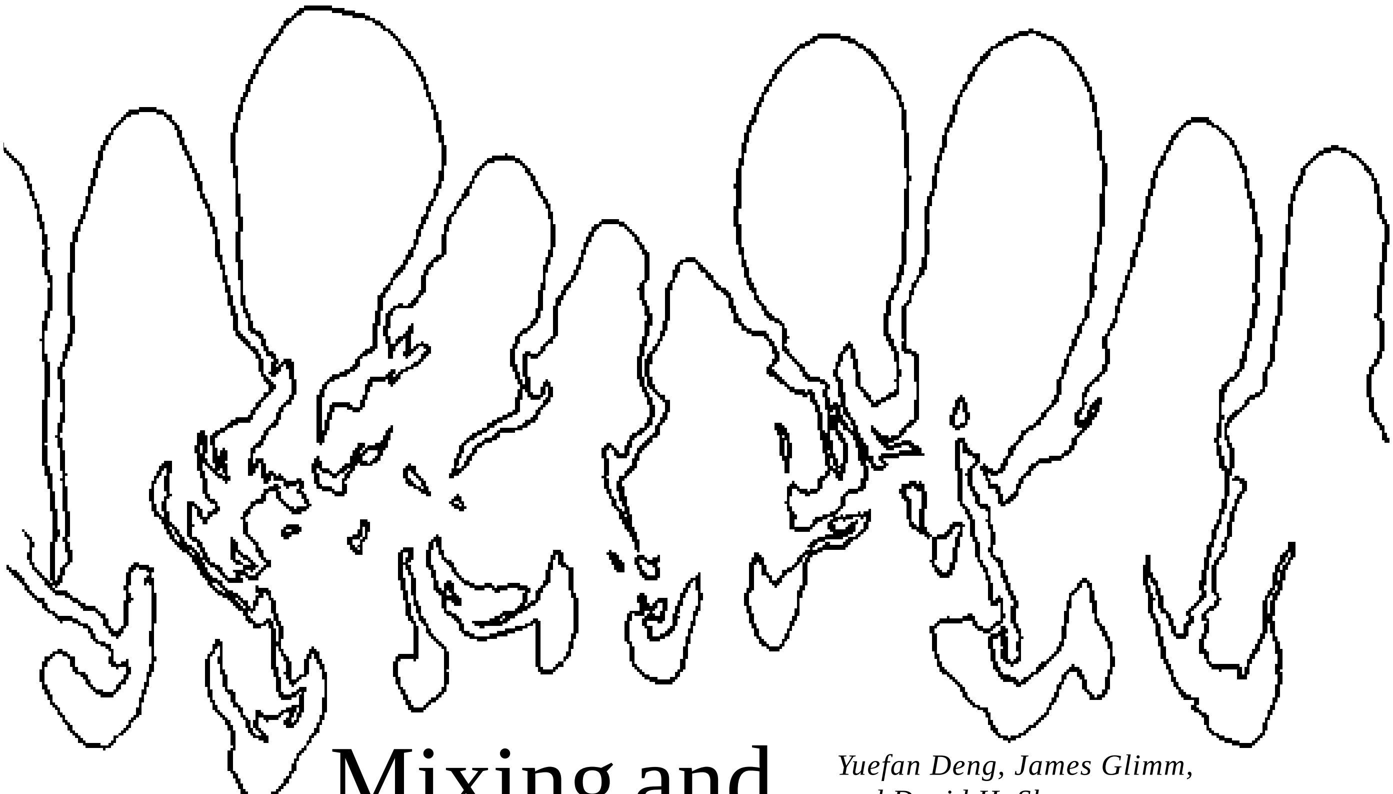
Figure 2. The Sonic Compressor

The Sonic Compressor uses electric power to compress a conventional refrigerant vapor by means of a high-amplitude sound wave; the model depicted can replace the piston compressor in a conventional cooling system such as the household refrigerator shown in Figure 1 of "Malone Refrigeration." The electricity drives a radically modified loudspeaker that shakes a cavity back and forth at a resonance frequency of the working-fluid vapor inside (300 hertz). In the figure the cavity is shown at the rightmost point of the vibration. The motion of the cavity causes the vapor to slosh back and forth—in other words, the motion generates a standing sound wave. The shape of the cavity is designed to prevent the formation of shock waves. The standing sound wave compresses and expands the gas; at the end of the tube farther from the loudspeaker, the range of pressure is 8 atmospheres. A pair of one-way valves at that end, which are opened and closed at the operating frequency by the pressure itself, admits low-pressure vapor from the intake pipe and ejects high-pressure vapor into the outflow pipe.

the year he spent here, we solved his shock problem. Of equal importance to him, we did not jointly invent anything patentable, so the business aspects of his project were not complicated by the involvement of intellectual-property rights belonging to the Laboratory. As Lucas's visit was successful, the Sonic Compressor could come into production in a few years. Thermoacoustic refrigeration will not be ready for the market until a few years later. Malone refrigeration will take still more time to develop but appears to be the most efficient option of the three to which we at Los Alamos are contributing. ◻

ventional CFC refrigerator! A totally different design was clearly required if Malone technology was ever to enjoy widespread use. So in our present CO_2 machine we are using a linear free-piston configuration, which was invented only recently and is being employed in gas-based Stirling engines intended for solar power or for use in space. The pistons in a linear free-piston machine are driven, not by a rotating motor connected to a crankshaft, but by a "linear" electric motor that provides reciprocating force and motion directly in the same way as a loudspeaker. This configuration minimizes the number of moving parts and eliminates the need for high-force bearing surfaces. Careful design can even eliminate the need for any mechanical connection to the displacer piston—the piston moves with the correct amplitude and phase simply in response to the fluid pressures acting on it.

The present crisis in the cooling industry is a unique opportunity for a new, potentially more efficient technology to break the monopoly of a technology that has enjoyed decades of incremental improvement. The primary challenge of the next year or two is to keep this difficult experimental project moving ahead, though the funding only pays for a third of the time of one researcher, while trying to attract the interest of an industrial collaborator. At best, years of further work costing millions of dollars will be required to bring Malone refrigeration to the threshold of possible widespread application. Meanwhile, as described in "CFCs and Cooling Equipment: The Size of the Problem," industry is proceeding promptly with more straightforward interim measures. In the intermediate time scale, mature new technologies such as the Sonic Compressor and perhaps thermoacoustic re-



Mixing and Chaotic Microstructure

Yuefan Deng, James Glimm, and David H. Sharp

Mixing is a process in which distinct fluids intermingle in a complex way. Chaotic microstructure refers, more broadly, to all small-scale stochastic or chaotic behavior that affects dynamics at large length scales. Problems that involve mixing and chaotic microstructure are of fundamental importance in both basic science and engineering: They are the central issue in turbulence, pipeline flow, and the dynamics of supernovae. The microstructure of porous rocks, which is stochastic on all length scales, is a dominant aspect of the geology of aquifers and oil reservoirs. Chaotic microstructure also plays a key role in determining the properties of common materials such as metals and plastics. Mixing is a dominant phenomenon limiting the performance of pellets in laser-fusion experiments.

Without doubt the problem of chaotic microstructure, in its many ramifications, is the most fundamental in classical continuum physics. Why is this? First, because the problem is pervasive, and second because the only systematic solution occurs at the engineering level through the method of conservative overdesign, or experimental trial and error. The classical methods of science—theory, computation, and experiment—have so far delivered much less than is needed either for engineering purposes or for scientific understanding.

These problems are difficult for three reasons. They possess an exceedingly large number of active degrees of freedom, they are usually nonlinear, and they typically do not admit a small parameter for useful expansions. Together these features lead to an extremely complex phenomenology.

We propose here a systematic program for addressing this class of problems. This program follows a line of development that will be familiar to many-body theorists. The first step is the identification and analysis of elementary modes, or coherent structures. For example, in turbulence the elementary modes are vortices; in phase transitions, dendrites; in Hele-Shaw cells, viscous fingers; in solid materials, lattice dislocations. The second step is to understand the pairwise interaction of elementary modes. This step is followed, not by an analysis of multi-mode interactions, but by the statistical analysis of an ensemble of modes governed by pairwise interactions. From this analysis we want to derive continuum-level constitutive laws (such as equations of state), which are then used (fifth step) in macroscopically averaged flow equations such as the Navier-Stokes equation. Thus we have a five-step procedure relating microscopic dynamics to macroscopic observables. Although each of the steps may seem rather obvious, their integration is less commonly discussed. A large fraction of work in classical physics is related to this program in that it addresses one or another of its steps.

In this paper we will illustrate this program by discussing its implementation in examples drawn from our work and that of our collaborators. This work as well as collateral work of others can be traced from the further reading given at the end of the article.

Rayleigh-Taylor Instability

Our first example concerns Rayleigh-Taylor instability. The occurrence of this instability can be understood in the following simple way. Imagine the ceiling of a room plastered uniformly with water to a depth of 1 meter (Figure 1). The layer of water will fall. However, it is not through lack of support from the air that the water falls. The pressure of the atmosphere is equivalent to that of a layer of water 10 meters thick, quite sufficient to hold the water against the ceiling. But in one respect the atmosphere fails as a supporting medium. It fails to constrain the air-water interface to flatness. No matter how carefully the water layer was prepared, small irregularities are inevitably present at the interface. Those portions of the fluid at the interface that lie higher than the average experience more pressure than is necessary for their support. They begin to rise, pushing aside water as they do so. Neighboring portions of the fluid, where the surface hangs a little lower than average, require more than average pressure for their support. They begin to fall. The air cannot supply the specific variations in pressure from place to place necessary to prevent the interface irregularities from growing. The initial irregularities therefore increase in magnitude, exponentially in time at the beginning. The water falls to the floor.

The same layer of water lying on the floor would have been perfectly stable. Irregularities die out. Thus we can infer a simple criterion for the onset of Rayleigh-Taylor instability at the interface between two fluids of different densities: *If the heavy fluid pushes the light fluid, the interface is stable. If the light fluid pushes the heavy fluid, the interface is unstable.*

Rayleigh-Taylor instability occurs in diverse situations. Let us take a quick look at one example from research into inertial-confinement fusion. Figure 2 shows a highly schematic picture of the implosion of a deuterium-tritium (DT) pellet. A spherical glass or metal tamper is filled with DT gas. The tamper is irradiated with intense laser light, which causes it to accelerate inward, compressing the DT gas inside the cavity in order to bring about nuclear fusion. During irradiation the outer surface of the tamper is the interface between a heavy fluid (glass or metal) and a light fluid (vaporized glass or metal) and is unstable during the initial phase of the implosion, when the light fluid pushes the heavy fluid. As the pellet is compressed to perhaps 1000 times its normal density, the pressure in the cavity increases until it is sufficient

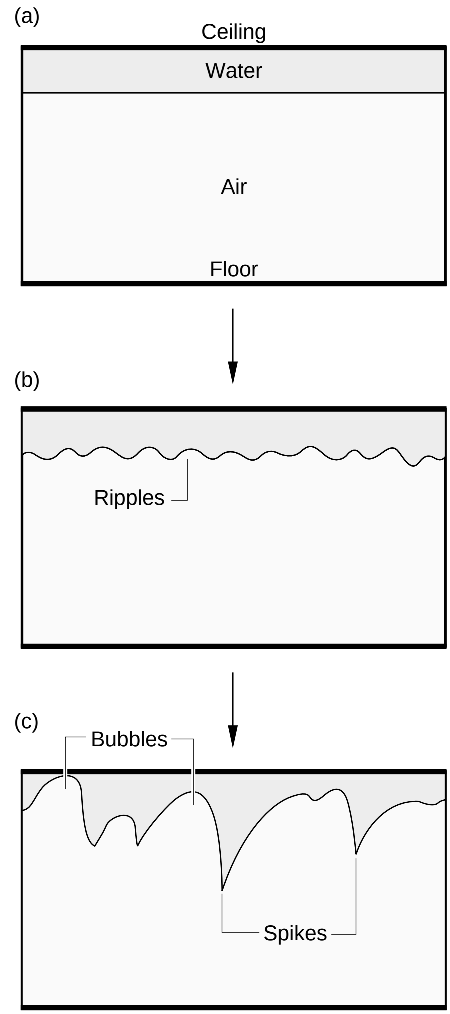


Figure 1. An Example of Rayleigh-Taylor Instability
(a) The pressure of the air is quite sufficient to support a perfectly uniform layer of water 1 meter thick against the ceiling. (b) But the air pressure cannot constrain the air-water interface to flatness. Ripples or irregularites will inevitably be present at the interface. (c) The irregularities grow, forming “bubbles” and “spikes.” The water falls to the floor.

to slow the inward motion of the tamper. This phase of the implosion is also Rayleigh-Taylor unstable; here the high-pressure light gas in the cavity is pushing the tamper. Although this picture of the implosion of a DT pellet is oversimplified in a number of ways, it nevertheless suggests quite clearly that Rayleigh-Taylor instability can have an important effect on the compression of the pellet.

A closer look at the time evolution of a Rayleigh-Taylor unstable interface reveals a complex phenomenology. As the instabilities develop, bubbles of light fluid and spikes of heavy fluid form, each penetrating into the other phase. Complex interactions among bubbles and spikes lead to the formation of a chaotic mixing layer at the interface. Understanding the growth rate and structure of this mixing layer is the central problem in the study of Rayleigh-Taylor instability.

This question has been investigated experimentally by Read and Youngs. Their important findings bring the questions to be understood into sharp focus and can be summarized as follows. The mixing layer has three principal regions: the edge where bubbles of light fluid are penetrating into the heavy fluid (bubble regime), the edge where spikes of heavy fluid are penetrating into the light fluid (spike regime), and the connecting interior region (mixing layer). The bubble regime has been the most carefully studied and has the simplest properties. It is found in the Read and Youngs experiments, as well as in numerical simulations (Figure 3), that the edge of the mixing layer in the bubble regime is dominated by a collection of bubbles.

The average height of the bubbles relative to the mean position of the interface, $h(t)$, grows with time according to the scaling law

$$h(t) = \text{Æ}\, Agt^2,$$

where A, the Atwood number, is a modification due to buoyancy of the acceleration force (gravity) g and is defined as $A = (\Omega_2 \circ \Omega_1)/(\Omega_2 + \Omega_1)$ in which Ω_i represents the density of fluid i. It is a remarkable experimental finding of Read and Youngs that for incompressible fluids the acceleration constant Æ in this equation is an approximately *universal* constant, in that it is nearly independent of initial conditions and of fluid properties such as density, viscosity, and surface tension. Finally, it is observed that the average number of bubbles at the interface decreases with time, and that their average radius increases.

The main objective of our work has been to understand the physical properties of the Rayleigh-Taylor mixing layer. The fluid in the mixing layer is in a chaotic state, and it is necessary to have a strategy to guide one through the complexities of this problem. The fundamental challenge in modeling fluid chaos is to provide a simple macroscopic description of a chaotic fluid state that expresses a statistical average of information describing its chaotic microstructure. For the case of chaotic microstruc-

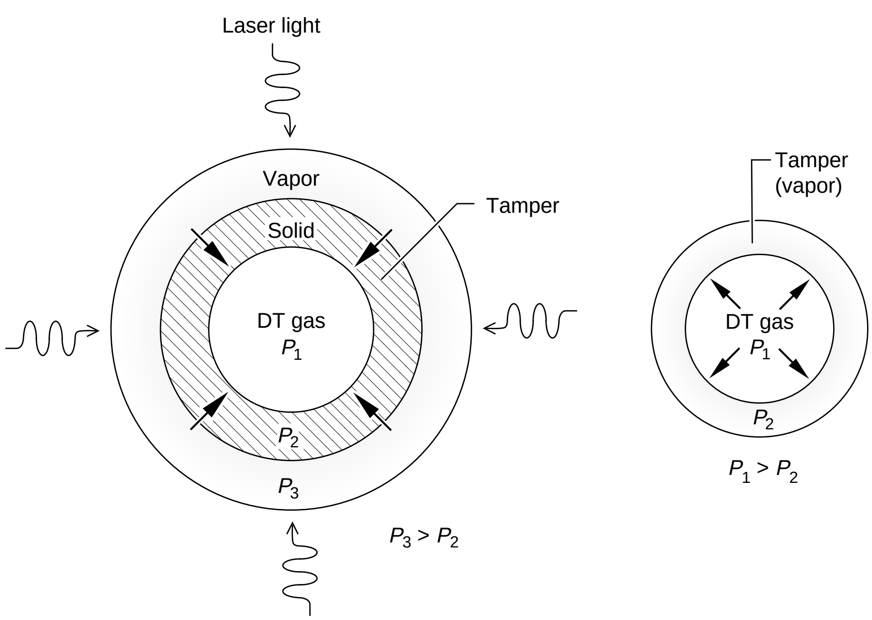


Figure 2. Rayleigh-Taylor Instabilities in Laser Fusion
On the left appears an extremely simplified view of a laser-fusion pellet at an early stage of a fusion experiment. Laser irradiation of the solid tamper vaporizes the outer layer of the tamper and pushes the tamper inward. Because the light vapor is pushing the heavier solid, this stage of the process is Rayleigh-Taylor unstable. On the right is sketched the same experiment at a later stage. Here the DT gas, now at high pressure (P_1), slows the inward motion of the tamper. Since the DT gas is less dense than the tamper, this stage of the process is also Rayleigh-Taylor unstable.

ture at the molecular level, the solution to this problem is highly developed and has given rise to the subjects of thermodynamics and statistical physics to describe the macro- and microphysics respectively.

Adapting the statistical-physics viewpoint to the analysis of fluid chaos,we proceed to study the one- and two-body problems and follow this with an analysis of a statistical ensemble. A renormalization-group fixed point emerges as a simple description of this ensemble.

In molecular physics the elementary modes, or units of analysis, are atoms or molecules. We identify bubbles as the fundamental modes governing the microphysics at one edge of the Rayleigh-Taylor mixing zone. The one-body problem concerns the dynamics of a single bubble. A relatively complete theory of the motion of a single bubble has been developed, which is valid for both compressible and incompressible fluids and includes a formula for the bubble velocity as a function of mode amplitude. This formula contains a small number of parameters, which have been fixed on the basis of analytic formulae, numerical computations, and experiments on periodic arrays of bubbles.

When more than one bubble is present, the bubbles interact. The interactions have a pronounced effect on the behavior of bubbles. The first effect, observed in both experiments and simulations, is that the velocity of a single bubble in a chaotic array of bubbles is typically a factor of 2 greater than that predicted by the single-bubble theory. This is a rather surprising result, which can be understood as follows. In a chaotic array of bubbles, a given bubble has left and right neighbors which generally differ in height and radius. We draw an envelope through the tips of adjacent bubbles. The envelope defines a long-wavelength collective excitation of the fluid and has the appearance of a set of broader bubbles and spikes. These broader bubbles can be viewed as additional elementary modes, which themselves obey the single-bubble theory. Both experiment and simulation led us to the assumption that, although each fundamental bubble is in a deeply nonlinear regime, the interaction of such a bubble with the collective mode is linear, so that the net velocity of a bubble in a chaotic array is given by a simple superposition:

$$v = v_b + v_e,$$

where v_b is the velocity of a bubble of the same amplitude as given by the single-bubble theory, and v_e is the velocity of a bubble in the envelope. The resulting formula agrees very well with experiment; note that it contains *no* free parameters.

The superposition hypothesis captures only one aspect of bubble interactions. The other is bubble merger. Merger refers to the behavior of a pair of neighboring bubbles: a larger, advanced, faster-moving bubble and a smaller, less advanced, slower-moving bubble. (Since larger bubbles move faster, eventually they are always more advanced than smaller bubbles.) It is observed that the smaller bubble is rapidly washed downstream, while the larger bubble increases in size to occupy the portion of the interface vacated by the smaller bubble. Thus, the dynamics of the fundamental modes comprises the dynamics of a single nonlinear

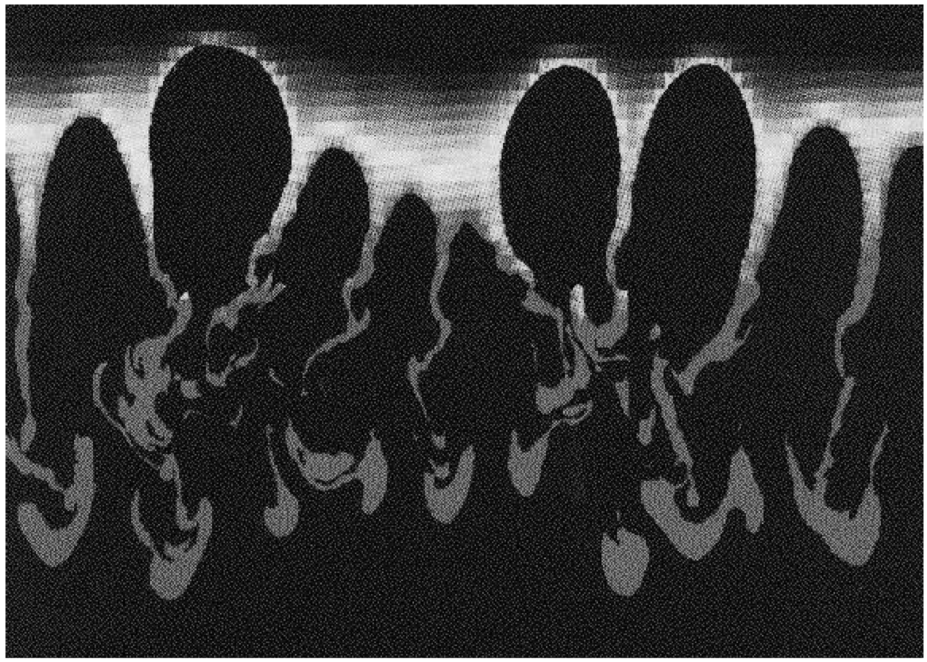


Figure 3. A Simulation of Rayleigh-Taylor Mixing
The figure shows the late-time behavior of Rayleigh-Taylor instability. The dimensionless compressibility $M^2 = 0.5$. The simulation was performed on a 32-node parallel iPSC/860 computer.

mode (single bubble) plus an approximately linear mode-mode velocity coupling plus a highly nonlinear mode-merger process.

We next consider a statistical ensemble of interacting modes (bubbles). We suppose there is a probability distribution for an infinite collection of bubbles, which defines the single-particle distribution (the probability of picking a bubble of height *h* from the ensemble), the pair correlations, and so forth. The bubble dynamics outlined above defines an evolution in this statistical ensemble of modes. Although the full statistical model could be solved numerically, we instead introduce simplifying assumptions to arrive at a more tractable form of the model. As when the Boltzmann equation is derived in statistical mechanics, we neglect pair correlations; then the probability measure defining the statistical ensemble is determined by the single-particle distributions. Those distributions in turn are approximated in terms of just three time-dependent parameters: the average height of the bubbles, $h(t)$, the variance in bubble height, æ (t), and the bubble radius, $r(t)$. The bubble dynamics leads to a set of ordinary differential equations for these quantities.

The bubble dynamics can be understood from the perspective of the renomalization-group method. This method has two key steps: "coarse-graining" and "rescaling." The bubble-merger process accomplishes a dynamical coarse-graining of the dynamics, because sets of smaller bubbles are replaced by larger bubbles, without changing the physics. The differential equations for the bubble dynamics thus reflect the coarse-graining of the merger process, as well as a dynamic increase in the length scale of the instability. The next step is to introduce rescaled variables, which subtract out these changing length scales. Once this is done, the result is a renormalization-group equation.

Analysis of this equation revealed the existence of a non-trivial renormalization-group fixed point. This means that the equations have the property that, no matter what the initial values of the variables are, they approach the same asymptotic values, or fixed point. At this fixed point, the width of the mixing layer grows at a constant acceleration, Æ . The theory provides a zero-parameter determination of Æ , which turns out to be in remarkable agreement with experiment. Since Æ is determined by a fixed point, its insensitivity to initial conditions is also explained. Thus our analysis has shown that the main properties of the bubble-dominated part of the chaotic mixing layer are direct consequences of scaling and the existence of a renormalization-group fixed point.

We emphasize that our program has depended in an absolutely essential way on numerical simulation of the solutions of the full two-dimensional, two-fluid Euler equations. These calculations were used in a variety of ways. Calculations of the evolution of Rayleigh-Taylor instability at a random interface were used as a model-independent way to calculate the magnitude of Æ , as well as to establish the insensitivity of Æ to a wider range of initial conditions than were explored experimentally and in a way that was not tied to possible idiosyncrasies of the experimental apparatus. Numerical simulations were used to define parameters in our single-bubble model. They were also used to establish the approximate validity of the modeling assumptions employed to understand bubble interactions and to identify conditions where those assumptions were not valid.

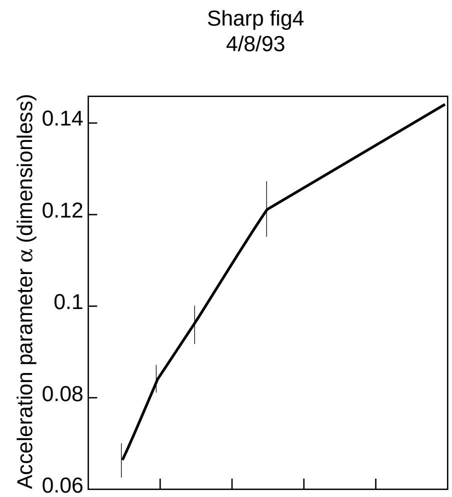


Figure 4. Dependence of the Mixing-Layer Growth Rate on Compressibility

Plot of the growth rate Æ versus the compressibility, M^2. The vertical bars indicate the variance associated with the choice of random-number seed.

Our numerical simulations of Rayleigh-Taylor mixing employed a front-tracking method. Front-tracking works by "hard-wiring" into the code maximal information about the analytically known behavior of the solution near a discontinuity or front (such as the jump in density at the heavy/light interface in the Rayleigh-Taylor problem or at a shock front in gas-dynamics problems). This approach leads to two benefits as well as a cost. The first benefit is the scientific understanding about the behavior of the solution near a discontinuity, generated in the course of implementing the method. The second benefit is that highly accurate solutions can be achieved without recourse to very fine computational grids. In fact, the front-tracking method typically allows an increase in computational resolution by a factor of about 3.5 in each spatial dimension and in time, that is, by a factor of about 40 in a two-dimensional, time-dependent problem and 150 in a three-dimensional, time-dependent problem. This dramatic increase in efficiency allowed a corresponding increase in the detail and scope of the computations attempted and was instrumental to the success of the program. The cost is that code development for front-tracking is more demanding because more complex algorithms are required to express the information about solution discontinuities. This is basically a software-complexity issue, which has been dealt with by the construction of highly modular code, the use of more powerful programming languages, and other modern code-development methods.

The Read-Youngs experiments were carried out with nearly incompressible fluids. Both our simulation methods and our model apply also to more compressible fluids. It is thus possible to explore, and indeed for the first time to predict, the properties of a chaotic mixing layer in a parameter regime outside the range of existing experiments. Our analysis shows that the mixing-layer growth rate in a compressible fluid can be a factor of 2 larger than in incompressible fluids (see Figure 4), and that α develops a dependence on initial conditions. The importance of these results in various applications seems clear.

The final step in our program for the study of high-dimensional chaos is to study continuum-level constitutive laws and equations. Here the flow is regarded as stochastic and flow variables $\rho = \langle \rho \rangle + \delta\rho$, $v = \langle v \rangle + \delta v$, etc., are expressed as sums of mean quantities $\langle \rho \rangle$, $\langle v \rangle$, etc., and fluctuating quantities $\delta\rho$, δv, etc. An effort to write governing equations directly for mean quantities such as $\langle \rho \rangle$ fails due to the nonlinearity of the governing equations and the fact that the average of a product is not equal to the product of the averages. For this reason, the averaged conservation laws contain new quantities, such as the Reynolds

$$R_{ij} = \langle \rho v_i v_j \rangle - \frac{\langle \rho v_i \rangle \langle \rho v_j \rangle}{\langle \rho \rangle}.$$

stress:

The continuum equations, even after quantities such as R_{ij} are introduced as new unknowns with their own dynamical equations, require the introduction of further new quantities (do not close). This means that further modeling assumptions must be made in order to arrive at a complete system of equations. We view these modeling relations as an extension of thermodynamics. Our work in progress in the area is to evaluate the correctness of different modeling approaches. One step in this program is to understand the structure of the fluctuating flow moments, such as R_{ij}, in a Rayleigh-Taylor unstable flow.

Flow in Porous Media

In the example of Rayleigh-Taylor instability discussed above, multiple-length-scale chaos arises spontaneously in a nonlinear dynamical problem. We now turn to an example in which chaos is forced, through the influence of multiple-length-scale data in the definition of the problem. This example concerns the transport and dispersion of pollutants (solutes) in groundwater. There is a nontrivial linear version of this problem, and in this sense it is simpler than the Rayleigh-Taylor problem. The fundamental methodology again consists of the integration of field observations, theory, and computation, but the linearity permits the specialization of the full arsenal of techniques described above and a more complete analysis of the problem. In particular, since the modes in a linear problem do not interact with each other, the study of mode-mode interactions is unnecessary.

As Einstein recognized, the phenomenon of dispersion (the spread of one medium through another) is the macroscopic consequence of random microscopic events. The classical theory of dispersion, based on the assumption that the random events occur on a much shorter scale than the macroscopic observations, leads to the familiar relation

$$l(t) = O(t^{1/2}),$$

where l is the position of the dispersion front and t is time. Field data for transport of fluids through porous media do not satisfy this simple relationship. Porous media of practical interest are heterogeneous on all length scales owing to random variations in the geology. Thus the scale on which the randomness operates is not small relative to the observations, and the assumptions underlying the classical theory of dispersion are not valid. This creates a serious problem for practical engineering in hydrology and in environmental studies of contaminants dispersing in groundwater, for example. In fact there is no known body of knowledge to connect laboratory and field-scale values of dispersivity. Therefore in practice, dispersivity

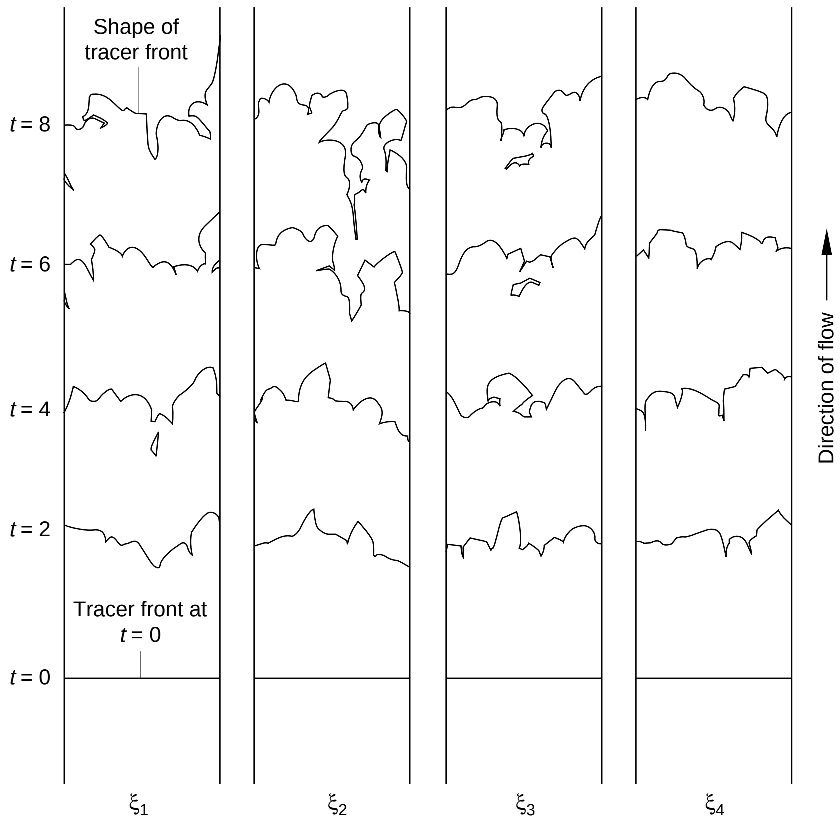


Figure 5. Simulation of Flow through a Porous Medium
Depicted are positions of a tracer front in a fluid flowing through a porous medium. The front is shown at five times ($t = 0, 2, 4, 6, 8$) for each of four independent simulations. Each simulation uses a different realization of the ª field drawn from a statistical ensemble of geologies.

must be estimated conservatively from carefully instrumented field studies conducted at some other site but at comparable length scales.

Flow through porous media is controlled by unknown (and, at necessary fine scales of detail, unknowable) features of the geological medium such as its porosity. In our studies, we treat those features as random fields, that is, random variables defined at every point of the media. The transmissibility ∏ is the most important variable relating the geology to flow properties. In analogy to electrical conductivity, it is an inverse resistivity for the flow in the sense that the flow rate equals ∏ times the negative pressure gradient. We make the common assumption that ∏ = exp ª, where ª is a normally distributed, random field that depends on the position **x** in space and represents the random variations in the structure of the porous medium.

Assuming stationary statistics for simplicity and denoting the ensemble average with angle brackets, we note that hª(**x**)i is independent of **x** and can be normalized to be zero, so that the entire statistical variation of the flow is determined by the two-point correlation function hª(**x**)ª(**y**)i, which is then a function of the difference variable **x**°**y**; that is, hª(**x**)ª(**y**)i= *f*(**x**°**y**). Any choice of *f* defines a model of the stochastic structure of the rock. A common choice, *f* = *b* exp(°|**x**°**y**|/*l*), has a single correlation length scale, *l*, and thus does not represent variability on all length scales. The simplest form of *f* that does allow variability on all length scales is the scale-invariant function

$$f = b|\mathbf{x}°\mathbf{y}|^{°Ø}$$

with Ø > 0. Such a scale-invariant, or self-similar, function is known as a fractal. Somewhat greater generality is obtained by allowing Ø to be a slowly varying function of |**x**°**y**|.

Randomly varying porous structure (that is, the specification of a ª field) introduces a statistical aspect to the flow. In Figure 5 we plot successive positions of a tracer front (an interface between one part of the fluid, which has been marked with a dye or in some other way, and the rest of the fluid) in four simulations. The simulations used the same initial configuration of the tracer front but different realizations of the ª field, each drawn from the same statistical ensemble of geologies. The statistical variation between the simulations arises, not from a lack of determinism of the flow, but from unpredictability of the flow due to unknown details of the geology. To illustrate this idea, we consider the simplest flow problem: transport of a passive concentration sample of an impurity by a background flow through porous rock. The flow uncertainty would be observable partly as an uncertainty of arrival times of a concentration sample and partly in the spreading of the concentration gradient due to differing arrival times of the individual particles that compose it. Apart from the rather small effect of molecular mixing on length scales typically of interest, these two aspects of flow uncertainty are identical because the structure of the rock is heterogeneous on all length scales and consequently affects the flow on all scales. The distinction between the two types of uncertainty lies in the resolving ability of the measuring instrument, the size or volume of the initial concentration profile, its initial concentration gradient, and the distance of travel before measurement occurs.

We use dispersion to represent both aspects of uncertainty described above. We can study this dispersion mathematically in terms of the expected values $\langle c(\mathbf{x},t)\rangle$ of a concentration c. The fluctuations, $\delta c = c - \langle c\rangle$, and higher moments such as the covariance, $\langle \delta c \delta c\rangle$, are of interest as well. Similarly, the flow velocity v can be expressed as the sum of its mean and its fluctuating parts: $v = \langle v\rangle + \delta v$. At this point, the flow problem is similar to the transport of a concentration by a turbulent velocity field, such as a pollutant in the atmosphere.

The following methods are available for this class of problems: numerical simulation of the transport process, followed by an ensemble average, to determine $\langle c\rangle$ directly; perturbative solutions of the transport equations in powers of δv, including resummation methods such as renormalized perturbation theory; analysis of field data; and exact solution of simplified model problems. These methods lead to consistent results and show anomalous (that is, nonclassical) dispersion. This

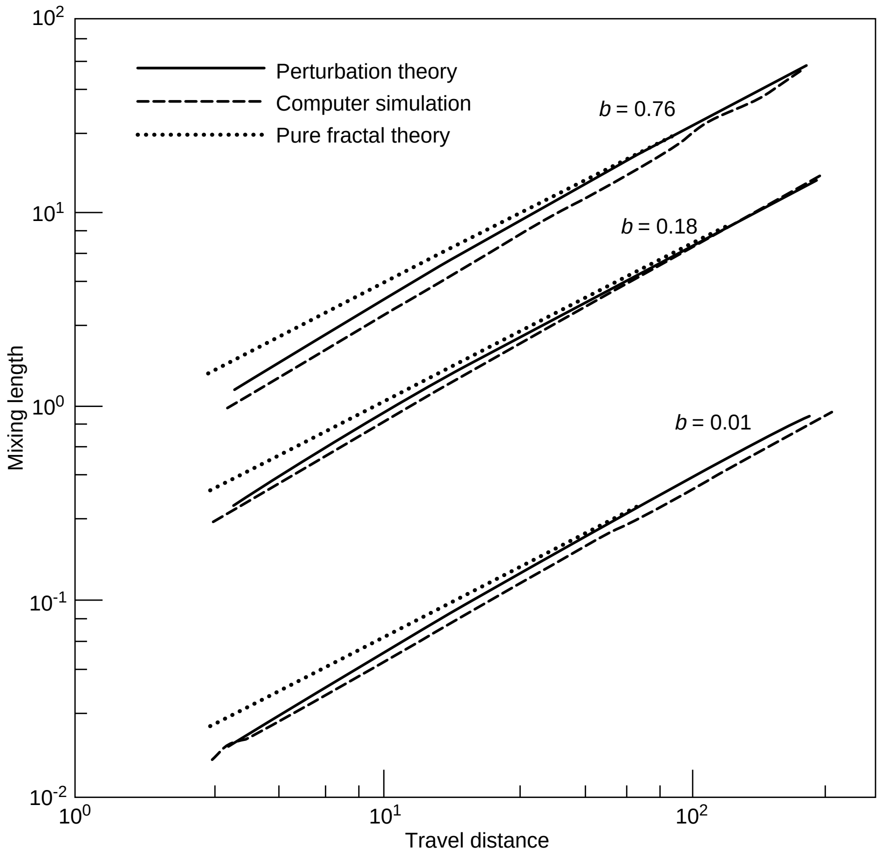


Figure 6. Three Calculations of the Mixing Length in a Porous Medium
The figure is a log-log plot of mixing length, *l*, as a function of travel distance in a medium having $\beta = 0.5$, where β is the fractal exponent characterizing the stochastic structure of the rock. Each triple of curves shows a comparison between pure fractal theory, second-order transport perturbation theory including transients, and a numerical simulation. Each triple of curves corresponds to a different value of the permeability field coefficient variation, as labeled.

means that the mixing length $l(t)$ associated with the concentration profile $\langle c(\mathbf{x},t)\rangle$ grows in time more rapidly than $t^{1/2}$. When the geology is fractal, $l(t)$ has a fractal behavior as well in the limit of long time, namely,

$$l(t) \sim O(t^{\gamma}),$$

and γ can be related to the fractal exponent β of the geology:

$$\gamma = \max\left(\frac{1}{2}; 1 - \frac{\beta}{2}\right)$$

for $\beta > 0$. However, $l(t)$ at short and intermediate times deviates from this law and in fact at short times is fractal with a different exponent.

In contrast to many problems of turbulent mixing, the different methods give consistent results. As Figure 6 shows, perturbative and numerical solutions agree quantitatively, simple soluble models agree asymptotically at long times, and each is consistent with geological field data. However, each of these methods is at variance with the results of classical diffusion theory, with a fixed diffusion constant.

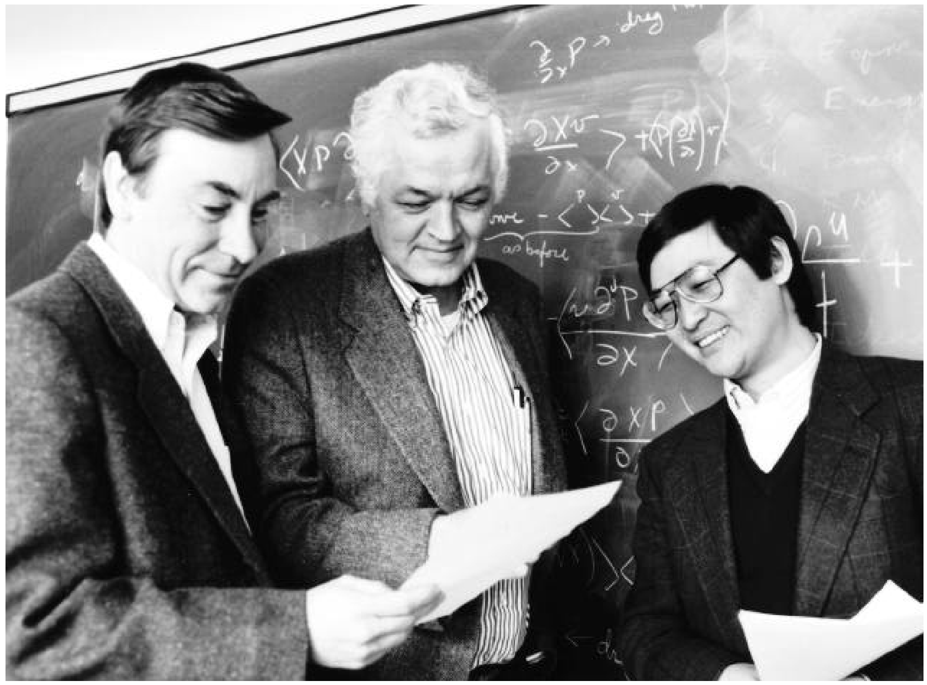
Our conclusion is that for geology with unknown fine-scale heterogeneity, dispersion of contaminants in groundwater should depend on time or travel distance, with transient corrections to fractal asymptotics. The understanding of dispersivity and its relation to multiscale geological heterogeneities is important because dispersivity is an essential input into large-scale computational modeling of groundwater-remediation projects. ■

Further Reading

Y. Chen, Y., Deng, J. Glimm, G. Li, D. H. Sharp, and Q. Zhang. A renormalization group scaling analysis for compressible two-phase flow. University at Stony Brook Preprint SUNYSB-AMS-92-02. Submitted to *Physics of Fluids A*

F. Furtado, J. Glimm, W. B. Lindquist, F. Pereira, and Q. Zhang. 1992. Time dependent anomalous diffusion for flow in multi-fractal porous media. In *Proceedings of the Workshop on Numverical Methods for the Simulation of Multiphase and Complex Flow*, edited by T. M. M. Verheggan. New York: Springer Verlag.

J. Glimm, Q. Zhang, and D. H. Sharp. 1991. The renormalization group dynamics of chaotic mixing of unstable interfaces. *Physics of Fluids A* 3: 1333–1335.



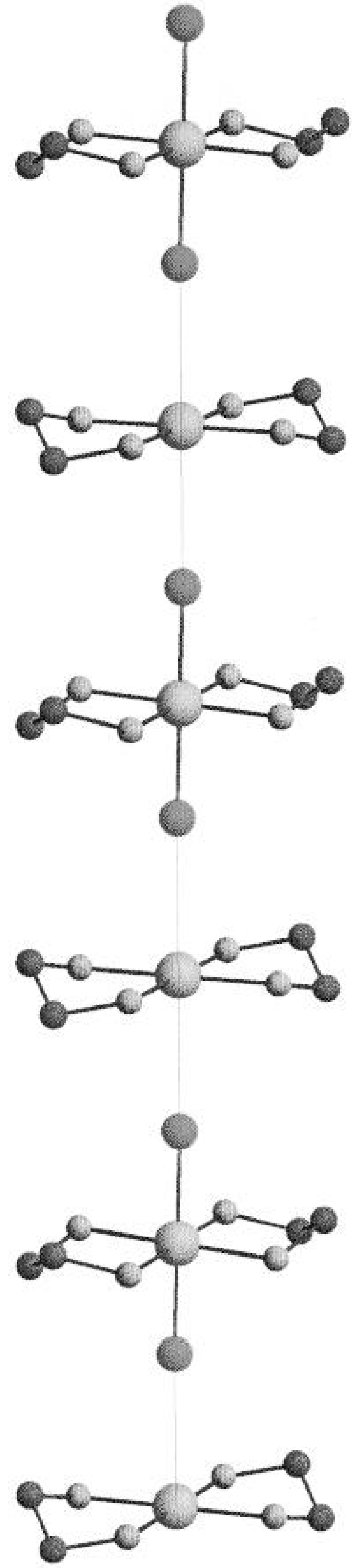
Yuefan Deng (right) was admitted to Columbia University after passing a China-U.S. joint examination and received a Ph. D. in theoretical physics in 1989. He is now an assistant professor in the Department of Applied Mathematics and Statistics at the University at Stony Brook. His main interests include parallel processing and its application to large-scale scientific and engineering problems such as the modeling of multiphase flows, electromagnetic scattering, lattice QCD, and protein folding. His "extra-curricular" activities include writing in Chinese and fixing broken electronic devices.

James Glimm (center) is Chair of the Department of Applied Mathematics and Statistics and Director of the Center for Advanced Manufacturing, both of the University at Stony Brook. He previously held faculty positions at New York University, Rockefeller University, and the Massachusetts Institute of Technology. He is a member of the National Academy of Sciences and recently received the Steele prize of the American Mathematical Society. He has long been interested in DOE problems and visits the Laboratory periodically. His research interests include computation and modeling for turbulent flows, the mathematical theory of conservation laws, and stochastic methods. He enjoys listening to operas and hiking in the mountains.

David H. Sharp (left) received an A.B. from Princeton University in 1960 and a Ph.D. in theoretical physics from the California Institute of Technology. He joined Los Alamos National Laboratory in 1974 and currently holds the position of Laboratory Fellow. Sharp's current research interests include the modeling of complex fluid flows and the formulation and analysis of models of gene regulation. He is a Fellow of the American Association for the Advancement of Science and the American Physical Society.

Novel Electronic Materials
the MX family

Alan R. Bishop and Basil I. Swanson



Side view of a PtCl chain in an MX solid. The Pt ions are bonded to ethylenediamine ligands.

The Laboratory and the nation are increasingly faced with demands for novel electronic materials needed to advance special-purpose technologies. Those demands have already led to extraordinary discoveries and applications of novel materials. Recent examples include self-assembling multilayer thin films for use in sensors and nonlinear optical devices and electrically conducting organic polymers that produce blue light by electroluminescence (light in that spectral region is unavailable from conventional light-emitting diodes).

Standard, or "textbook," electronic materials (such as copper and silicon crystals) are ordered materials made up of simple unit cells (the repeating units of a crystal). The theories to describe those materials are relatively simple: the electrons are treated as nearly noninteracting and the theories are applied only in situations where the displacements of the atoms from their equilibrium positions in the crystal lattice are small. However, the last fifteen years or so have been marked by the synthesis, microscopic characterization, and increasingly predictive modeling of rich families of materials that are very different from traditional materials. Research on these families of materials has led to major awards (including several Nobel Prizes) and new technologies. The families include amorphous semiconductors, quantum-Hall-effect materials, materials with artificial structures on mesoscopic scales (a few hundred or thousand atoms), ultrathin films, synthetic metals (such as conducting polymers), quasi-crystals, organic superconductors, high-temperature superconductors, and buckminsterfullerenes, to name a few.

These "novel electronic materials" generally display significant differences from the former textbook materials. In particular, many are anisotropic (composed of chain-like or plane-like configurations of atoms), artificially structured into stable or long-lived metastable forms, inhomogeneous, or composed of complex, low-density unit cells, and many have various combinations of those properties. Indeed, the subject of novel electronic materials lies between textbook chemistry, physics, and biology, and we speak of "soft condensed matter" or "condensed-phase chemistry" in an attempt to capture the spirit of the new interdisciplinary ground.

The essential and newly emerging research strategy for all of these fascinating materials is a very close team approach involving synthesis, characterization, and theoretical modeling. This approach has been adopted vigorously at the Laboratory, since it is well suited to the Laboratory's extreme breadth of capabilities and tradition of tackling complex issues from a multidisciplinary perspective. Those strengths have proven very attractive to industrial collaborators as the Laboratory develops closer relations with industry. We describe here some notable successes of the application of our strategy to one newly emerging family of materials—metal-halogen (MX) linear-chain complexes. MX materials are discussed in detail below, but first we should explain the general challenges posed by the development of novel electronic materials.

It is now abundantly clear that the method of synthesis, the resulting microstructure (local atomic patterns), and the macroscopic (bulk) properties of novel electronic materials are inextricably linked. Desirable optical, magnetic, and other macroscopic properties are found to depend crucially on details of microstructure, which in turn are controlled by the details of the synthesis. Understanding those relationships is a goal of synthetic chemists and materials scientists as well as of experimentalists developing characterization techniques with ever-improving spatial and temporal resolution—such as ultrafast optical spectroscopy (which can resolve times approaching 10^{-14} second) and scanning tunneling microscopy.

Therefore, the ultimate aim of theoretical modeling must be to short-circuit trial and error in the synthesis-characterization-modeling-application program by predicting which modified or entirely new materials will have optimal properties and how those materials can be synthesized. However, that aim is extremely difficult to achieve using current theoretical techniques. The materials in which we are interested have strong electron-lattice and electron-electron interactions that compete on different length and time scales, produce strongly nonlinear and nonequilibrium features, and make the properties of the material very sensitive to its microstructure. In fact, the existence of and competition between these complex properties are directly responsible for the very novelty of the materials and give rise to their potential for meeting the rigorous demands of device engineering. At the Laboratory we are trying to develop new techniques for modeling those materials by bringing together traditional techniques from quantum chemistry, the study of electronic band structure, and many-body physics. We are validating theory by direct experimental tests as part of our synthesis-characterization-theory-application team approach.

MX Materials

The MX family of compounds is particularly rewarding to study for a number of reasons. They are fine subjects for working out an integrated development strategy. They are complex enough to exhibit much of the sensitivity discussed above but just simple enough to allow substantial microscopic characterization. The various compounds of the family exhibit a remarkable range of strengths of competing forces and consequent physical properties. They share fundamental features with other novel electronic materials, such as conducting polymers and organic and high-temperature-superconductors, and therefore test our ability to model an extremely wide range of physical properties. Finally, in learning how to synthesize, model, and fabricate them, we have discovered that they have substantial technological potential in their own right.

Although the MX family now represents one of the very best examples of international success in the synthesis-characterization-modeling-application strategy, it is important to emphasize that the research challenge is far from completely met—at the Laboratory or anywhere else. For the MX class, we have had success in controlling microstructure by varying the method of synthesis. We can make single crystals, thin films, and junctions between different members of the MX family. Likewise by applying many-body modeling techniques to known specific microstructures, we have made precise predictions of optical, electronic, and magnetic properties, which in turn we measured with precise microscopic characterization tools. Some of those successes are outlined below. Yet the "holy grail," the ability to predict what synthesis method will lead to desired macroscopic properties, remains unachieved for the MX materials, as indeed for all such complex materials.

Figure 1 depicts the basic crystal structure of MX materials. Most importantly they contain parallel chains of metal ions (M), such as platinum, palladium or nickel, alternating with halide ions (X). The metal ions are attached to molecular complexes (ligands) that act to hold the chains in place in the three-dimensional crystal structure. However, electronic, optical, and magnetic properties of the MX solids are determined mostly by the electrons and ions along the metal-halide chains, so the electronic character of the solids is strongly *one*-dimensional.

Two properties of MX materials are crucial in producing their novel electronic and optical properties: first, they have broken-symmetry ground states whose crystal struc-

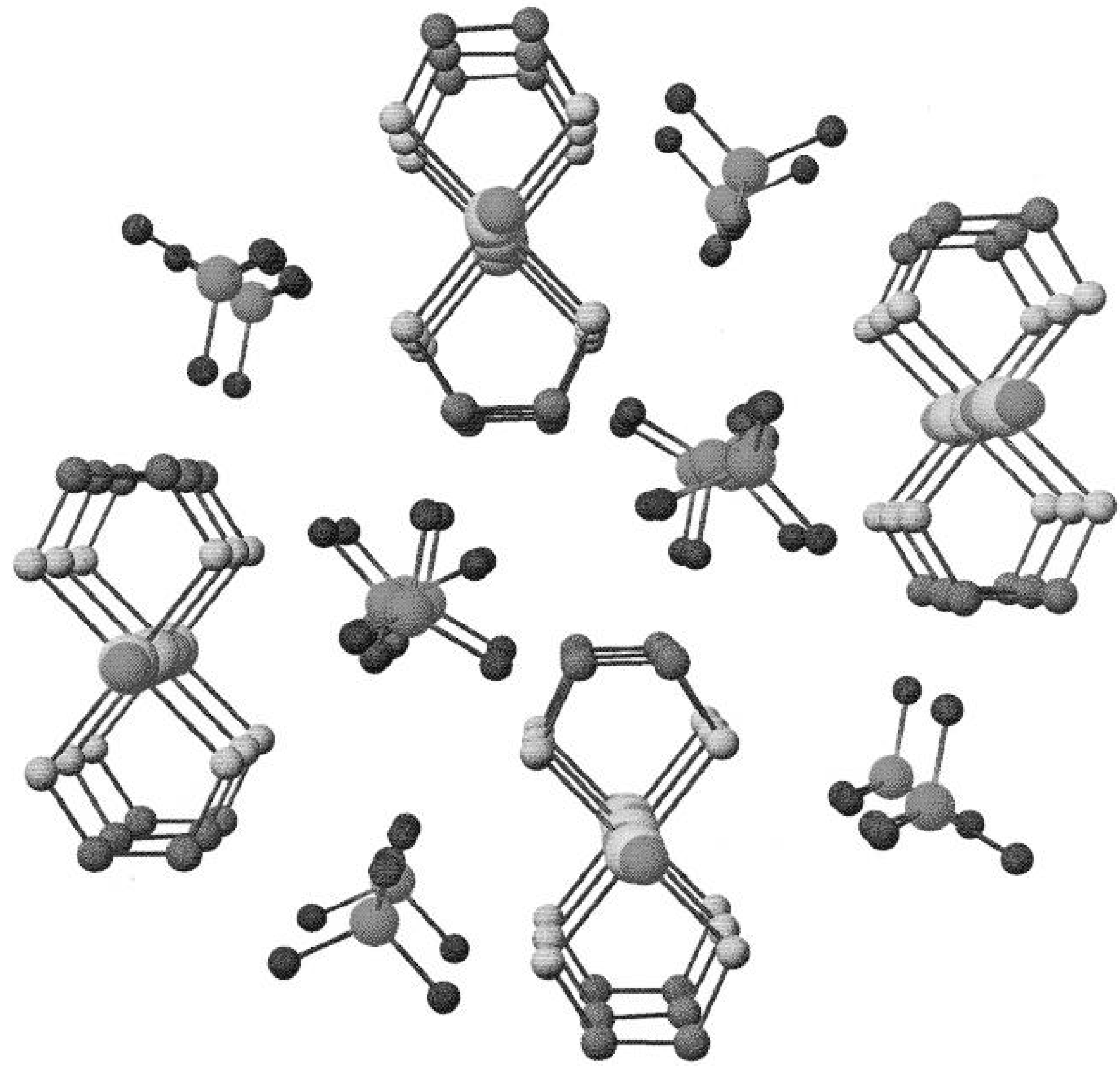


Figure 1. The Structure of a Typical MX Material
The crystal structure sketched here is that of $[Pt(en)_2][Pt(en)_2Cl_2](ClO_4)_4$, where en stands for the molecule ethylenediamine, $(CH_2NH_2)_2$. The view is approximately down the Pt-Cl chains, the most important axis for electronic activity. Pt atoms are represented in blue and Cl in green. Each ethylenediamine ligand binds to a Pt ion on a chain. The spaces between the chains are filled with rows of ClO_4^- ions, called counterions because they provide the charges needed for overall electric neutrality.

tures are nearly perfect over large domains, and second, they have gap states, electronic states above the ground state that are created by defects, impurities, structural inhomogeneities, and photoexcitation.

Broken-Symmetry Ground States

Figure 2 illustrates two examples of broken-symmetry ground states in MX materials. In both cases the metal-halide chains in their ground states are less symmetrical than a regular alternation of metal and halide ions. Specifically, each unit cell of the chain contains two metal and two halide ions rather than just one of each.

In the ground state of the PtCl compound the unit cell is doubled because the electrons and ions are shifted from their symmetric positions. The structural distortion creates a periodic variation in charge density called a "charge-density wave" (CDW). The distortion is a direct result of electron-electron and electron-ion interactions. Structural distortions that double the unit cell were predicted to occur in low-dimensional electronic materials by Peierls in the 1950s and have since been found in many such materials, including most MX compounds. (Incidentally, Peierls was an important member of the Laboratory's Theoretical Division during the Manhattan Project.) The existence of the CDW ground state of the PtCl solid was successfully predicted by our theoretical modeling.

Figure 2 also illustrates a second type of broken-symmetry ground state

(a) Hypothetical Symmetric Ground State

Metal-halogen distances are equal; charge and electron-spin configurations are symmetric. Spin of unpaired electron on metal ion has equal probability of pointing up or down.

(b) Charge-Density Wave in PtCl Chain

Electron-spin configuration is symmetric but charges on metal ions alternate between +4 and +2 and halogens move closer to Pt^{+4} ions, breaking the charge and lattice symmetry in (a).

(c) Spin-Density Wave in NiCl Chain

Metal-halogen distances are equal; charge configuration is symmetric but unpaired electrons on metal ions alternate in direction, breaking the spin symmetry in (a).

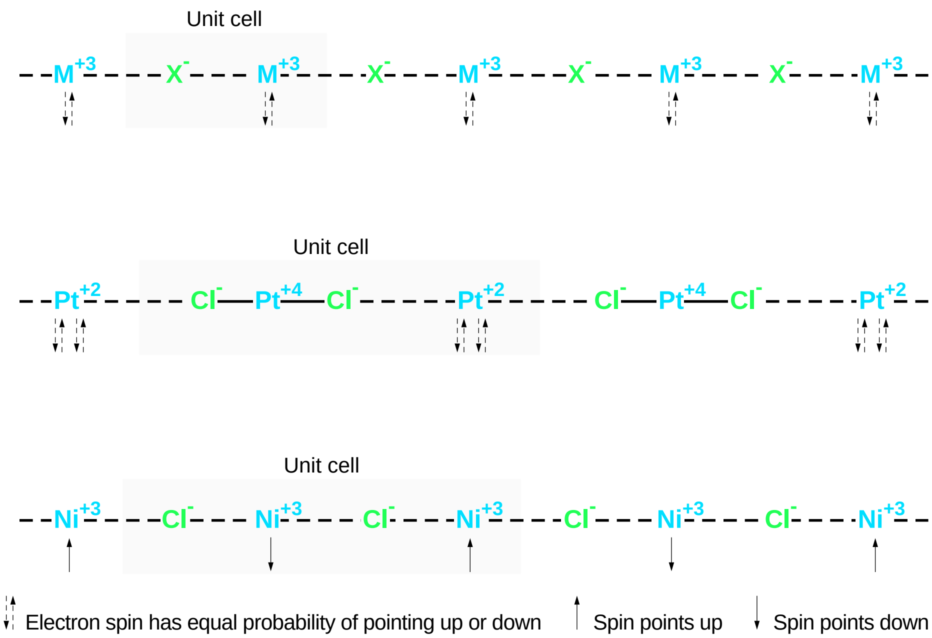


Figure 2. Peierls Transitions to Broken-Symmetry Ground States
The MX chains exhibit Peierls transitions to broken-symmetry states. A Peierls transition is a spontaneous transition of a one-dimensional chain from a symmetric configuration like that shown in (a) to a less symmetric but lower-energy configuration like those shown in (b) and (c). The symmetry breaking causes the unit cell to contain twice as many atoms and thus to double in size. Two types of broken-symmetry ground states are found in MX solids, depending on the metal (M) and halide (X). (b) The charge-density-wave (CDW) ground state in a PtCl chain. The charges on the metal ions (Pt) alternate between +2 and +4. The lattice is distorted from the symmetric configuration so that the halide ions (Cl°) are bound more closely to the more highly charged neighboring metal ion than to the other. (c) The spin-density-wave (SDW) ground state in a NiCl chain. The lattice is symmetric because all bond lengths are equal, but the spins of the unpaired electrons on the metal ions (Ni) alternate in direction. The symmetry breaking of the CDW and SDW ground states results in electronic configurations with an energy gap like the energy gaps characteristic of semiconductors (illustrated in Figure 3).

in which the spins of the unpaired electrons on the metal ions alternate in direction to form a “spin-density wave” (SDW). Spin-density waves occur in the ground states of only a few known MX materials, such as the NiBr and NiCl compounds. They are similar to the insulating states of high-temperature superconductors.

In both CDWs and SDWs, the doubling of the unit cell creates a gap in the band of allowed electronic energy levels. The gap forms between the occupied and unoccupied levels of the ground state and thus plays a role similar to that of the band gap in semiconductors (see Figure 3), even though the mechanisms producing the gaps are quite different. Further, by “tuning” MX materials (that is, by varying the metal, halogen, ligand, pressure, and so forth) we can vary the size of the gap to values between approximately 1 and 3.5 electron volts, an unusually wide and technologically important range.

Gap States

In both MX materials and semiconductors, local lattice distortions caused by inhomogeneities or local defects (including impurities) can produce spatially localized electronic wave functions and cause the electronic energy levels near the band gap to move into the gap. The resulting gap states may involve zero, one, or two electrons (or holes, which are missing electrons), depending on the charge of the inhomo-geneity. As shown in Figure 4, MX materials, like semiconductors, can be deliberately doped with a low density of impurity atoms that, by donating electrons to or accepting electrons from the host material, create charged gap states such as polarons and bipolarons. Those excitations are local lattice distortions containing extra electrons or holes. Also illustrated in Figure 4 is the use of "photodoping" to create neutral gap states such as excitons. (Structural imperfections also cause electronic energy levels to move into the gap. Like photodoping, structural defects do not change the overall electronic charge in the material, so the resulting gap states are neutral.)

Much semiconductor device physics is based on modifying material properties to control the energies and occupation densities of the gap states. For example, if excitons are created, their decay can result in electroluminescence, a phenomenon exploited in light-emitting diodes. The presence of gap states also allows the absorption of light at energies (and therefore colors) different from the energy across the band gap (see Figure 5). Finally, conductivity, for either direct current or alternating current, is strongly influenced by the occupation density and nature of the gap states, especially charge-carrying states such as polarons or bipolarons. The gap states may diffuse along chains or quantum-mechanically hop along or between chains, as they are believed to do in conducting organic polymers.

Semiconductor research and development is therefore largely devoted to "engineering" the size of the band gaps and the nature and occupation density of the gap states. That type of engineering is similarly important in the MX family, but easier for two reasons. First, because MX materials are low-dimensional and have strong electron-ion interactions, their lattice distortions are large, leading to gap states whose energies are relatively far from the valence- and conduction-band edges.

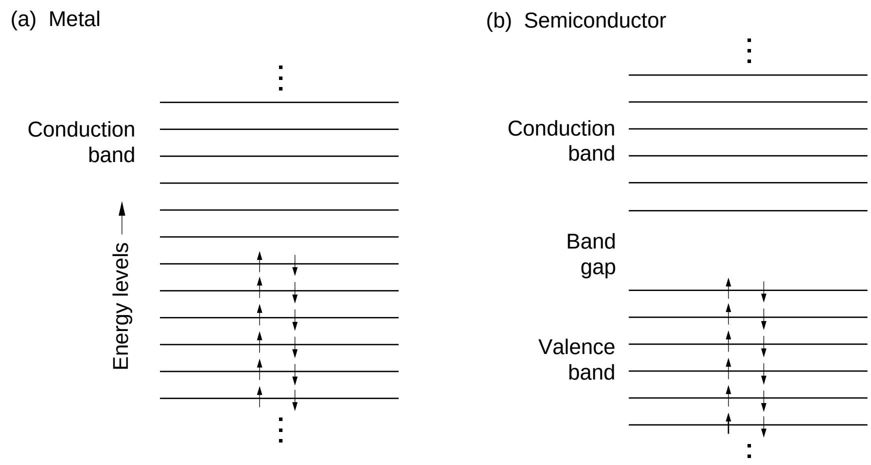


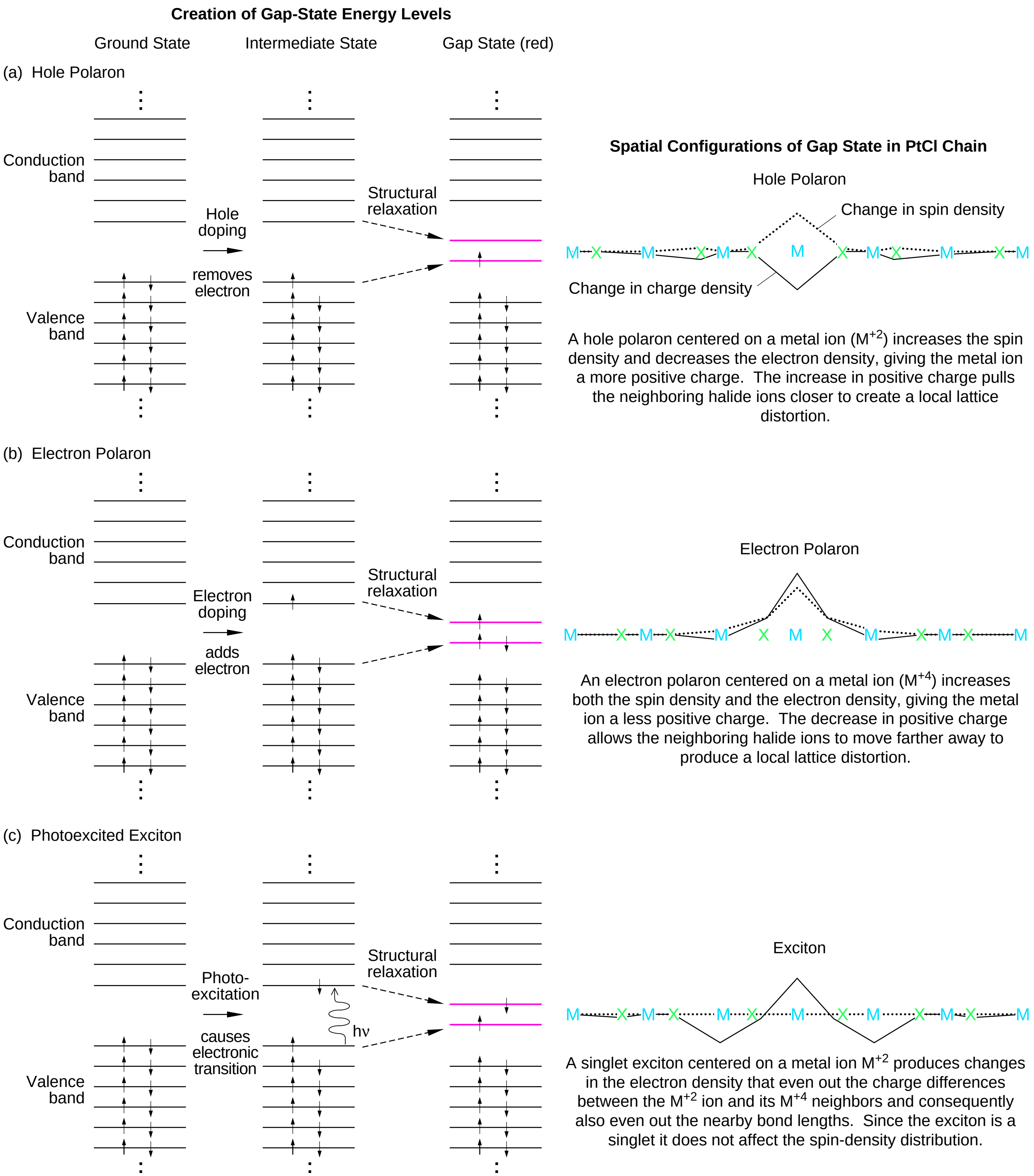
Figure 3. Electronic Energy Levels of a Metal and a Semiconductor
In solids, the electronic energy levels are so closely spaced in energy that they form bands. For clarity only a few of the levels are shown here. Each level can be occupied by an up-spin electron (↑) and a down-spin electron (↓). Additional occupation is forbidden by the Pauli principle. (a) In a metal, the band of energy levels is partially filled. Transitions to the higher, unoccupied levels are easily induced, for example by applying an electric field. Once such transitions have occurred, electric charge can flow. (b) A semiconductor contrasts with a metal in having an energy gap between the top occupied level—the top of the valence band—and the lowest unoccupied level—the bottom of the conduction band. Therefore free flow of charge is inhibited unless a sufficiently strong electric field or light of sufficient energy transfers an electron across the gap.

(In traditional semiconductors, the lattice distortions are much smaller and gap states remain near the edges.) Second, because of the *tunability* of the MX family, the size of the energy gaps can be varied over a strikingly large range to produce a range in the size of local lattice distortions from one or two unit cells (in PtCl) to more than thirty unit cells (in PtI) , a potentially useful property for photonic and other technologies.

Gap states may also be created in MX materials by controlled inhomogeneities such as interfaces between different compounds in the family. The variety of available gap states allows for stringent tests of theory and modeling and opens up a wide range of technological possibilities.

Pure MX Solids

The optical and electronic properties of MX materials are most easily studied in pure MX compounds like that shown in Figure 1. At the beginning of our research, we discovered that all the earlier studies of the PtBr solid had actually employed samples in which as many as 12 percent of the halide ions were chloride rather than bromide. The assumption that those samples were pure had made it impossible to properly interpret experimental results, and the resulting misinformation further hampered the development of theoretical models. This episode illustrates the essential role of synthesis coupled with detailed characterization in materials research.



Once we had pure MX materials, we determined their crystal structures more precisely than had previously been done (particularly by finding structural phase transitions involving slight shifts of atomic positions within a CDW configuration). We then took advantage of the MX family's variable chemical make-up by studying the dependence of ground states on composition. It has long been known that the choices of both metal and halogen affect the properties of an MX material. For instance, using chlorine as the halogen yields a strong CDW ground state; that is, the ions shift considerably from their symmetric positions, so the periodic

Figure 4. Gap States

Three types of gap states—a hole polaron, an electron polaron, and a singlet exciton—are illustrated as they would appear in a PtCl chain. Also shown are the mechanisms that typically create them. (a) An acceptor dopant (an ion that readily binds to an electron) is close to a PtCl chain. The acceptor dopant withdraws an electron from the chain, becoming negatively charged and leaving behind a vacancy or hole in the highest level of the valence band. The hole behaves as if it had a positive charge. Once the hole has been created, a lattice distortion takes place that causes the highest energy level of the valence band and the lowest energy level of the conduction band to move into the gap. The resulting gap levels are occupied by any electrons that occupied those levels before the lattice distortion. The combination of the missing electron and the lattice distortion is called a hole polaron. As shown at right, the lattice distortion traps both the charge density and the spin density associated with the hole. The lattice distortion shown is in the strong CDW of a PtCl chain and therefore is highly localized, extending over only two unit cells. Polarons in chains with weaker CDWs, such as PtBr and PtI, are less localized. (b) A donor dopant (an ion that readily loses an electron) donates an electron to the lowest level of the conduction band. The resulting electron polaron is very similar to a hole polaron except that it contains an electron and therefore has a negative rather than a positive charge. (c) Here no extra charge has been added to the MX chain. Instead, light of an energy greater than the band gap has excited an electron from the top of the valence band to the bottom of the conduction band. As in (a) and (b), a lattice distortion follows, forming the gap levels illustrated. One level is occupied by a spin-up electron and the other by a spin-down electron to form a singlet exciton (one with zero total spin) that does not alter the spin density. The charge-density distribution of the exciton follows a "multipole" pattern. Other kinds of excitons are possible, including a triplet exciton in which the electrons occupying the two gap levels are pointing in the same direction.

variation in charge density caused by those shifts has a large amplitude. On the other hand, using iodine yields a weak CDW (the ions are close to their symmetric positions and the amplitude of the CDW is small). Materials with strong CDWs have large band gaps whereas those with weak CDWs have small band gaps.

Our group and those led by Yamashita (Japan) and Clark (Britain) have recently recognized that changing the ligands or counterions changes the spacing along the metal-halide chains and thereby provides an alternate method for tuning the material's electronic structure. We call that control the "template effect" because the ligands and counterions act as a template to which the chain geometry must adjust. The template effect allows us to vary electronic properties of MX solids over a wide range with high precision and to simplify both experimental characterization and theoretical modeling. We hope soon to use the template effect to create materials that have a metal-like phase.

Excitations in Pure MX Solids

Our study of gap states in MX materials is an example of strong synergy between theory and experiment. We used several experimental techniques to produce and characterize a variety of gap states. We also developed models of the different possible gap states and solved them to predict what experimental results each type of gap state would lead to. We then used the experimental data to refine the models and theoretical interpretations, which, in turn, improved our understanding of the data.

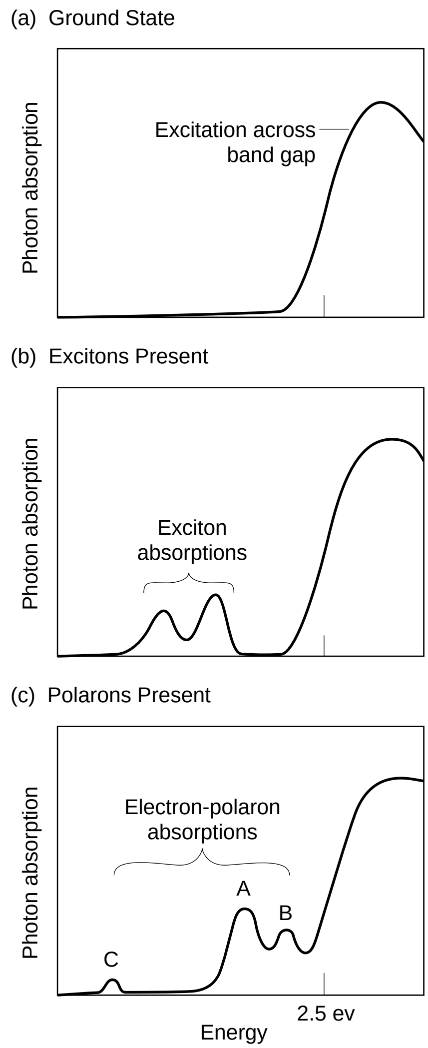
Our basic model is a Schrödinger equation for electrons on a single chain. The Hamiltonian contains the Hubbard-model Hamiltonian, which describes interactions between spins of electrons on the same or nearby ions as well as quantum-mechanical "hopping" of an electron from one ion to an adjacent ion. Our Hamiltonian also contains "Peierls" terms, which describe the interaction between electrons and lattice distortions (so that as the lattice distorts, both the local and hopping electronic energies change self-consistently) and terms for the elastic energy of lattice distortions.

Since different gap states are at different energy levels, they absorb light of different energies. A typical method for detecting and characterizing them is to measure their absorption spectra (see Figure 5). However, gap states in MX materials are rare. To enhance spectral signatures from them and to learn about the associated lattice distortions, we also used resonance Raman spectroscopy, which measures the energies of the characteristic lattice vibrations associated with the local distortions of the gap states.

In resonance Raman spectroscopy, material is illuminated with light that is in resonance with a specific electronic transition. After interacting with that light, the material scatters light not only at the incident energy (Rayleigh scattering) but at higher and lower energies as well that differ by the energies of the

Figure 5. Optical Absorption Spectra of Gap States

For light to be absorbed by an MX solid, the light must have enough energy to excite an electron from an occupied electronic level to a level that is not completely filled. (a) For an MX material in its ground state to absorb light, the energy of the light must be at least as great as the band gap. (b) If excitons have been created, for instance by photoexcitation, two energy levels move into the gap, so light can be absorbed at energies below the band gap indicated by peaks in the absorption spectrum. (c) Polarons can also absorb light at energies below the band gap. The energies are different from those absorbed by excitons as indicated by the absorption spectrum of electron polarons.



material's lattice-vibrational modes. The signals from vibrational modes that are coupled to the electronic transition are greatly enhanced, by as much as a factor of a million in some instances, compared with the signals that would be obtained by ordinary Raman spectroscopy. (In ordinary Raman scattering, as opposed to resonance Raman scattering, the incident light is not in resonance with any transition and so has a much smaller probability of interacting.)

To perform resonance Raman spectroscopy on the gap states in MX materials, we first cooled the material to below 10 kelvins and created gap states by photoexcitation. As illustrated in Figure 6, photoexcitation typically creates excitons which can decay to polarons. Gap states in MX materials are short-lived, but at temperatures below 10 kelvins a small fraction of them live long enough for the resonance Raman measurement to be made. We tuned the incident light to an absorption peak below the band gap that had been measured previously. From the frequencies of the emitted light we determined the vibrational energies of the local lattice distortions associated with the gap states. The measured energies of gap-state vibrations could then be used to estimate the number and lengths of bonds affected by the gap states' local lattice distortions because the energies of lattice vibrations are directly related to such structural features as the lengths and strengths of interatomic bonds.

We found that the resonance Raman and optical-absorption spectra of the gap states in the PtCl solid matched our theoretical prediction for the vibrational spectra of polarons. Studies with electron paramagnetic resonance (EPR) spectroscopy, which measures electron spins and their interactions with the lattice, corroborated our identification of the gap states as polarons. Using the same technique we found that at low temperatures the excitations in the PtBr solid (which is a weaker CDW than the PtCl solid) are bipolarons. Bipolarons are similar to polarons but contain pairs of electrons or holes whose spins point in opposite directions. Above about 60 kelvins the bipolarons are observed to dissociate into polarons.

Since polarons and bipolarons have electric charge and can move along the chains (by diffusion or hopping), they play a critical role in the electrical conductivity of MX solids. Their mobility is low because the motion of polarons requires the motion of ions in the lattice, which are much heavier than electrons. Hence most MX materials are poor conductors. The mobility of the polarons increases, however, as the strength of the CDW decreases, because in a weak CDW the local lattice distortions of the polarons are less localized (as measured by EPR). We are hoping that if we can make an MX material in the transition regime between CDW and SDW ground states, its conductivity will be high.

Several groups had predicted that electron and hole polarons would have exactly the same absorption spectra. Thus, it was a surprise to find that, when we measured the absorption spectra using resonance Raman spectroscopy, the electron-polaron spectra were slightly different from the hole-polaron spectra. The difference is consistent with the involvement of the halide-ion orbitals as well as the metal-ion orbitals in the polarons. When we incorporated the halide-ion orbitals into our theory, we were able to reproduce the spectra of both electron and hole polarons. The discovery of the unexpected role of the halides has been useful to Los Alamos theorists studying high-temperature superconductors. Those materials share many general features with MX solids; the oxide and copper ions they contain are analogous to the halide and metal ions in MX solids. There had been considerable debate about whether the oxide ions had any important effect on superconductivity in high-temperature superconductors. Guided by our experience with MX solids, Laboratory theorists are taking the oxide ions into account.

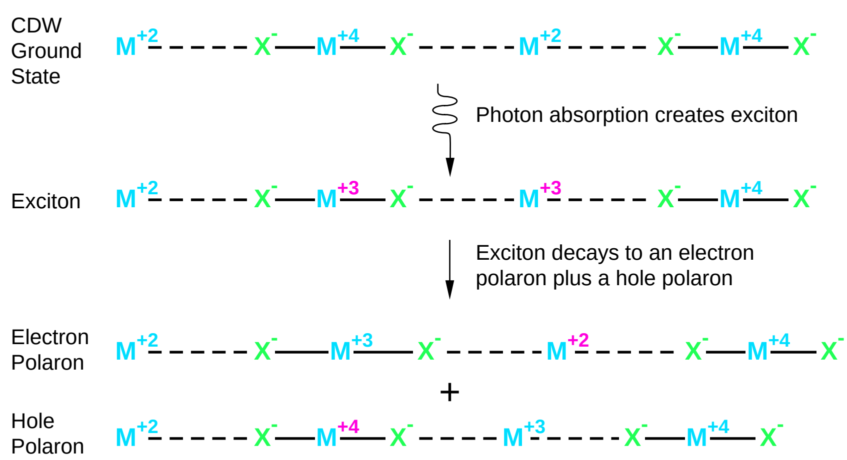


Figure 6. Photoexcitation of Excitons and Polaron Pairs
The figure illustrates the creation of an exciton through photoexcitation and its spontaneous decay to an electron polaron and a hole polaron. The first step is the same as that shown in Figure 4 c. In the second step the exciton, which is neutral, breaks up into a hole polaron (positive) and an electron polaron (negative). Here, instead of displaying the charge distributions of the excitons and polarons as in Figure 4, we schematically represent changes in the charge distributions by changes in the valences of the metal ions. The changes in valence are highlighted in red. Note that no charge is added in either step.

Mixed-Halide MX Solids and Quantum Confinement

Platinum-halide solids made with significant quantities of two halogens can form high-quality single crystals. This result is another surprise because the metal-metal spacing of the PtI solid is quite different from those of the PtCl and PtBr solids, and normally solids of different structures form separate crystals. We attribute the unusual miscibility to the template effect. The chains in mixed crystals contain segments 1 to 10 nanometers long in which only one halide occurs, and the metal-metal spacing of the minority segments is forced to approach the spacing of the majority of the crystal. Thus mixed crystals contain metal-halide chains with metal-metal spacings, and consequently CDW strengths, quite different from those found in pure crystals made with the same halide. The most striking example is a mixed PtI/PtCl solid in which there is much more Cl than I. In that case the Pt-Pt spacings in both the PtI and PtCl segments are nearly as small as they are in a pure PtCl material. Consequently, in the PtI segments the iodide ions appear to be halfway between the Pt ions, as the halide ions are in an SDW, rather than closer to one Pt ion as they are in the CDW ground state of a pure PtI solid (see Figure 2). Evidently the PtI segments have been forced past the boundary of the CDW phase into some other phase, possibly an metallic phase or an SDW.

Experimental evidence suggests that in mixed-halide MX chains, electrons are confined to segments of one composition. The confinement is particularly clear-cut in PtCl/PtBr solids. As mentioned earlier, the lattice and charge-density distortions of polarons in the strong CDW of PtCl compounds are highly localized. In other words, there is very little electronic communication along PtCl chains. When PtBr segments are embedded in a PtCl chain, the chloride ions at either end of a PtBr segment act as hard “electronic rocks” that isolate the electrons in the PtBr segment. The “quantum confined” behavior of electrons in such sharply bounded segments, or quantum rods, is amenable to theoretical modeling. For instance, electron configurations affect vibrational frequencies in such a way that short confined segments have higher vibrational frequencies than longer segments. We used resonance Raman spectroscopy to measure this effect in PtBr segments. We tuned the incident light to the band gaps in the PtBr segments, which are

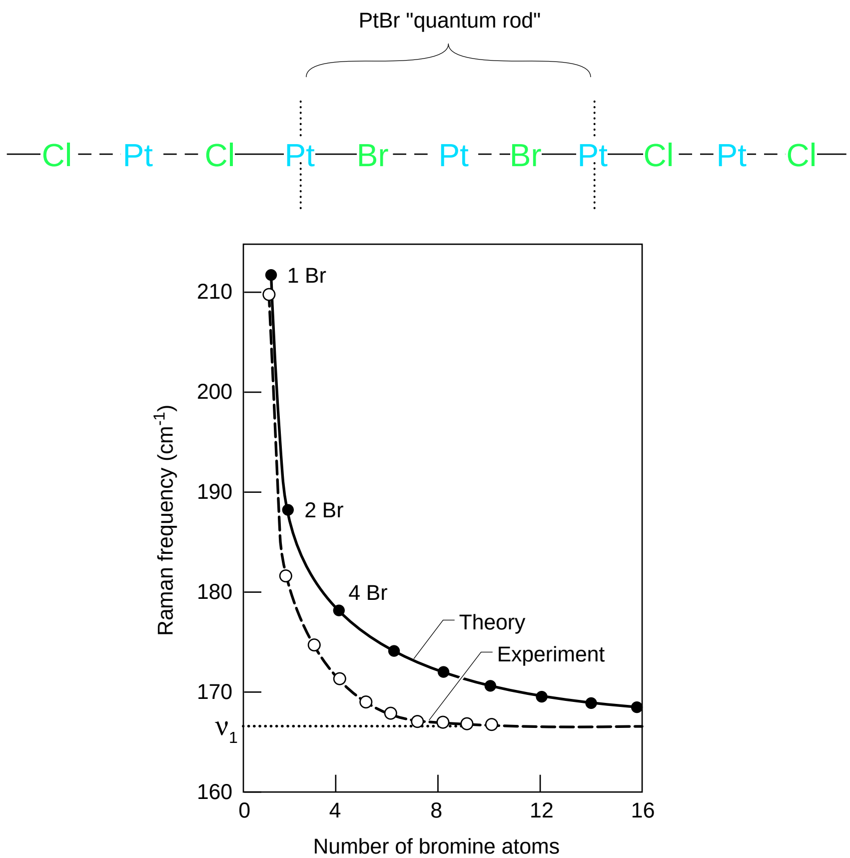


Figure 7. Segment-Length Dependence of Raman Frequency
The molecule depicted is a PtCl chain containing a short PtBr segment. The theoretical curve (solid circles) gives the results of our calculations for the lowest vibrational (Raman) frequency of such a segment as a function of the number of Br ions in the segment. Frequencies measured by resonance Raman spectroscopy are also shown (open circles). The many measured frequencies close to the value $\int_1$ cannot be distinguished, so no data points are given beyond 10 bromine atoms. The frequencies have been assigned to numbers of Br ions based on the principle, thoroughly established experimentally and theoretically, that the vibrational frequency of a segment decreases with increasing segment size. The uncertainties in the measured frequencies are approximately the same size as the circles representing the data points.

lower than the band gap in the PtCl chain. Since the band gaps of the Pt-Br segments increase with decreasing segment length, we knew that if we gradually increased the energy of the incident light, starting from the band-gap energy of pure PtBr, we could excite PtBr segments in order of decreasing length. We measured the resonance Raman spectra of the PtBr segments by that method. Figure 7 shows that the Raman (vibrational) frequencies are close to those predicted by our many-body model for PtBr segments of various lengths.

The properties we have discussed so far are of interest mostly from the standpoint of basic science. Potentially more useful in devices are the properties of the junctions between segments. As shown in Figure 6, excitation of an MX solid with a photon whose energy is greater than or equal to the band gap can create an electron polaron and a hole polaron. Thus, exposing an MX solid to light creates equal numbers of electron and hole polarons throughout the material. Surprisingly, though, when a mixed PtCl/PtBr material is photoexcited, the majority of the electron polarons have resonance Raman spectra typical of polarons in PtBr whereas the majority of the hole polarons have spectra typical of polarons in PtCl. That result implies that, once created, the electron polarons tend to migrate to the bromide segments while the hole polarons tend to migrate to the chloride segments. Such separation of charge carriers also occurs in doped semiconductors and is the basis of many of their useful properties. We hope eventually to synthesize MX samples containing single junctions, in the same place in every chain, between regions of different composition. Such samples could be incorporated into photovoltaic devices. We expect that samples with single junctions could also efficiently generate coherent light at double the frequency of incident laser light (second-harmonic generation).

MX solids also appear promising for nanotechnology because they can be made in high-quality single crystals in which the electronic and optical properties vary on a scale of nanometers and because those properties can be controlled chemically through, for example, synthesis of mixed-halide crystals. If we could produce conducting phases, mixed MX solids could have junctions between conducting and semiconducting segments; such junctions are needed to create nanoscale computers. A further requirement is the fabrication of superlattices: samples in which junctions are spaced at controlled intervals along chains and the junctions of different chains are lined up in parallel. Superlattices could also serve in solid-state lasers. We hope to prepare nanoscale superlattices using electrochemical growth methods that have already been used successfully with other materials. At present, superlattices and other nanoscale structures in materials for electronics can be created only through expensive vapor-deposition methods.

Another direction of our current work is the use of picosecond and femtosecond spectroscopy. At present, we can study only the longest-lived excitations: polarons and bipolarons at low temperatures. Using the laser apparatus that has been developed for ultrafast measurements, we hope to study short-lived excitations such as polarons at high temperatures and the still shorter-lived excitons (see Figure 6) as well. We will also look for such exotic states as "breathers" (stable, nonlinear, localized vibrational states) that have been predicted to occur in MX and related materials. Another goal that we and others are pursuing is making MX materials with different compositions. A particularly attractive possibility is replacing the metal ions with "binuclear" ions consisting of two metal ions (MM) bound together by ligands. Such ions are the subject of great excitement in the field of inorganic chemistry and the synthesis of MMX solids might enormously expand the potentials of MX materials. We hope that through continued collaborations involving synthesis, experiment, and theory, we will learn more about the fascinating properties of MX solids and make available new materials that may be the basis of future advanced technology. ■

Acknowledgements

We are grateful to many colleagues with whom we have developed our understanding of MX solids, including Bob Albers, Meb Alouani, Tony Arrington, Dionys Baeriswyl, Alain Bulou, Robert Donahoe, Tinka Gammel, Joe Huang, Sara Huckett, Susumu Kurita, Steve Love, Richard Martin, Avadh Saxena, Brian Scott, and Laura Worl.

Further Reading

D. Baeriswyl and A. R. Bishop. 1987. Halogen-bridged metal complexes: Model compounds for high-T_c superconductors? *Physica Scripta* T19: 239–245.

Sara C. Huckett, Robert J. Donohoe, Laura A. Worl, Alain D. F. Bulou, Carol J. Burns, Joseph R. Laia, David Carroll, and Basil I. Swanson. 1991. Mixed-halide MX chain solids: Effect of chloride doping on the crystal structure of resonance Raman spectra of $[Pt(en)_2Br_2][Pt(en)_2](ClO_4)_4$. *Chemistry of Materials* 3: 123–127.

J. Tinka Gammel, A. Saxena, I. Batistić, A. R. Bishop, and S. R. Phillpot. 1992. Two-band model for halogen-bridged mixed-valence transition-metal complexes. I. Ground state and excitation spectrum. *Physical Review B* 45: 6408–6434.

S. M. Weber-Milbrodt, J. Tinka Gammel, A. R. Bishop, and E. Y. Loh, Jr. 1992. Two-band model for halogen-bridged mixed-valence transition-metal complexes. II. Electron-electron correlations and quantum phonons. *Physical Review B* 45: 6435–6458.

S. P. Love, L. A. Worl, R. J. Donohoe, S. C. Huckett, and B. I. Swanson. Origin of the fine structure in the vibrational spectrum of $[Pt(C_2H_8N_2)_2][Pt(C_2H_8N_2)_2Cl_2](ClO_4)_4$: Vibrational localization in a quasi-one-dimension system. *Physical Review B* 46: 813–816.

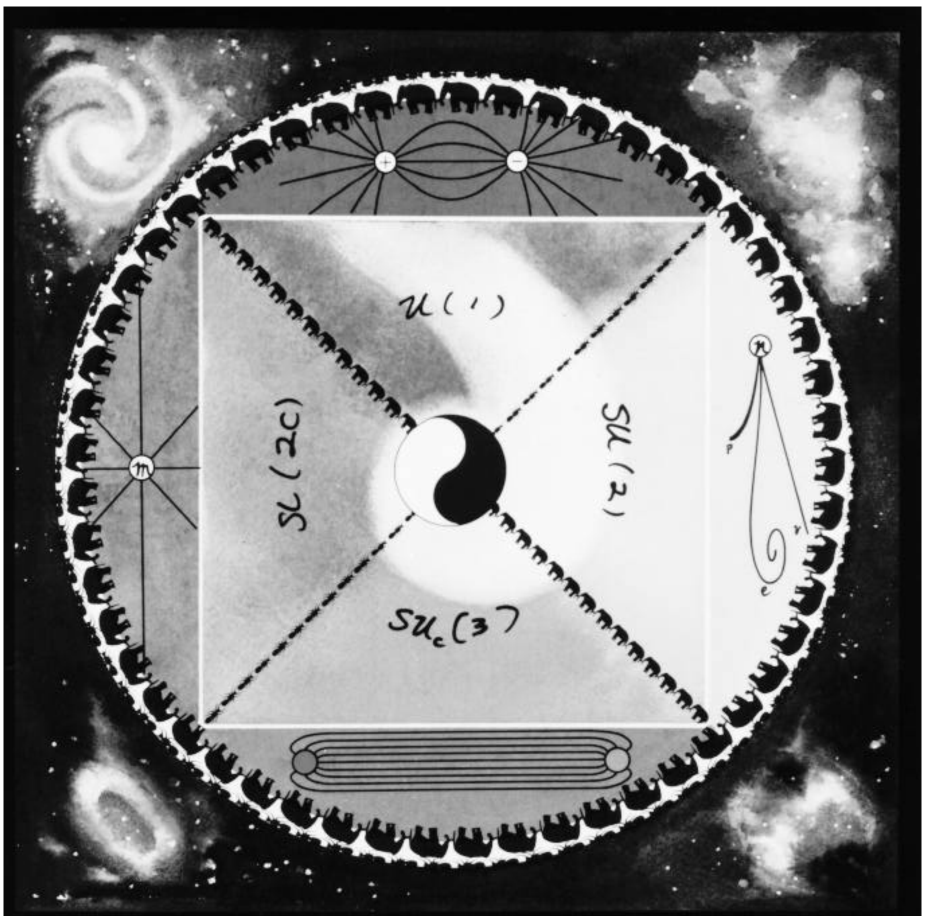
Alan R. Bishop (left) joined the Laboratory in 1979. He is currently leader of the Condensed Matter and Statistical Physics Group in the Laboratory's Theoretical Division. His interests include electronic and magnetic materials and nonlinear phenomena in er and materials science. He is actively involved in the Laboratory's Center for Materials Science and Center for Nonlinear Studies.

Basil I. Swanson (right) received a Ph.D. in physical-inorganic chemistry from Northwestern University in 1970. After a postdoctoral fellowship at the Los Alamos Scientific Laboratory and two university professorships, he returned to the Laboratory in 1980 as a staff member in the Inorganic Chemistry group. He served as the leader of that group from 1984 until 1988. Swanson is now in the Spectroscopy and Biochemistry group and was named a Laboratory Fellow in 1991. His areas of research include the structure and dynamics of solids, materials chemistry, and advanced electronic materials.

Unification of Nature's Fundamental Forces

a continuing search

Geoffrey B. West
Fredrick M. Cooper
Emil Mottola
Michael P. Mattis



It is a well-known, and much-overworked, adage that the group of scientists brought to Los Alamos to work on the Manhattan Project constituted the greatest assemblage of scientific talent ever put together. Many of those scientists had made (or in some cases were destined to make) highly significant contributions to the development of our understanding of the basic laws of physics. They had grappled with deep questions concerning the consequences of quantum mechanics, the structure of the atom and its nucleus, and the development of quantum electrodynamics (QED, the relativistic quantum field theory describing the interaction of charged particles with radiation). Although those questions had to be put on the back burner for the duration of the project, the pot, so to speak, continued to simmer. Indeed, it was explicitly recognized at the time that basic research had an important and seminal role to play even in the highly programmatic environment of the Manhattan Project. Not surprisingly this mode of operation evolved into the remarkable and unique admixture of pure, applied, programmatic, and technological research that is the hallmark of the present Laboratory structure. Nowhere in the world today can one find under one roof such diversity of talent dealing with such a broad range of scientific and technological challenges—from questions concerning the evolution of the universe and the nature of elementary particles to the structure of new materials, the design and control of weapons, the mysteries of the gene, and the nature of AIDS!

Many of the original scientists would have, in today's parlance, identified themselves as nuclear or particle physicists. They explored the most basic laws of physics and continued the search for and understanding of the "fundamental building blocks of nature'' and the principles that govern their interactions. It is therefore fitting that this area of science has remained a highly visible and active component of the basic research activity at Los Alamos. Furthermore, in keeping with the historical development of the Laboratory, there continues to be a vigorous and productive interplay with other more applied and programmatic parts of the Laboratory (see "Testing the Standard Model of

Particle Interactions using State-of-the-Art Computers").

This article will give a brief overview of some of the exciting developments in this area of science that have taken place in recent years and in which scientists at Los Alamos have played significant roles.

From the end of the Manhattan Project until approximately 1970, the general taxonomy and phenomenology of the elementary particles was intensely studied. When the various mesons (such as the pions and the kaons) and baryons (such as the proton, neutron, §, ß, and •) were classified, it was discovered that there were four apparently quite distinct ways in which they interacted among themselves: (1) gravitationally, (2) electromagnetically, (3) weakly and (4) strongly. The first two interactions, or forces, are the most familiar since they are long-range and so manifest themselves macroscopically even over astronomical distances. The second two have very short ranges (less than 1 fermi, or 10^{-13} centimeter) and so are only important at the nuclear and subnuclear level. The weak interaction is the one responsible for beta decay (such as the decay of the neutron to a proton, an electron, and a neutrino), and the strong interaction is the one responsible for the binding of protons and neutrons to form the nuclei of atoms.

One of the great intellectual achievements of the twentieth century has been the realization that these four wildly different "fundamental'' forces may, in fact, be viewed as manifestations of a single unified force. Although this paradigm was originally the dream of Einstein, not until the 1970s was any serious progress made toward developing the theoretical framework for unification.

The initial major breakthrough occurred around 1970 when a theory was proposed that unified the weak force with the electromagnetc force in a mathematically consistent fashion. This unification came almost a century after Maxwell had shown that electric and magnetic phenomena could themselves be unified in a single force carried by the electromagnetic field. In the modern language of QED, *all* electromagnetic phenomena can be understood in terms of a force-carrier that is exchanged between electrically charged particles such as electrons. This force-carrier, called the photon, is the quantum of the electromagnetic field and can be thought of as a massless elementary particle with no electric charge and spin angular momentum equal to one (in units of $\hbar$). The masslessness of the photon is a consequence of the so-called gauge (or phase) symmetry of the electromagnetic field and gives rise to the long-range nature of the electrostatic force between charged particles. Furthermore, it guarantees that electric charge is conserved. Technically, the gauge symmetry responsible for the masslessness of the photon is a reflection of the invariance of Maxwell's equations to arbitrary phase changes in the quantum fields associated with the electron and the photon. (Specifically, an arbitrary phase change of the electron field can be compensated for by a redefinition of the photon field, which leaves the form and structure of the equations of motion for electromagnetic interactions unchanged.) Because of its deep and seminal role, this local phase, or gauge, invariance of the electromagnetic fields is now viewed as a fundamental principle, or constraint, used to derive the form of the basic interactions among the quantum fields describing all elementary particles.

The unification of electromagnetism with the weak interactions was accomplished in the context of quantum field theory by extending the principle of gauge symmetry and the idea of a force-carrier to the case where the force-carrier can itself carry electric charge and therefore interact with itself. (In QED the photon is neutral and therefore cannot interact with itself.) A force-carrier of the weak interactions must have electric charge because some weak interactions cause the charge of a particle to change. For example, the neutron, which is neutral, decays through the weak interactions to a proton, which has one unit of positive charge. Further, the force-carrier of the weak interaction must be massive rather than massless because the range of the weak force is so short, less than 10^{-14} centimeter. (This follows from the original observation of Yukawa that the range of a force varies inversely with the mass of the force-carrier.)

The remarkable accomplishment of Glashow, Weinberg and Salam was to fashion a well-defined quantum field theory like QED for the weak interaction. It is based on an extended notion of gauge symmetry that Yang and Mills had invented earlier in an attempt to describe the strong interactions, but, in addition, is ideally suited for describing the charge-changing and charge-conserving interactions of the weak force. To obtain massive force-carriers, called the W^+, W^-, and Z^0, the gauge (or phase) symmetry of the quantum fields had to be dynamically broken in a rather subtle way analogous to the way the spontaneous symmetry breaking in a superconductor produces the Meissner effect. This dynamical symmetry breaking is referred to as the Higgs mechanism. It was designed to yield massive mediators of the weak force while preserving the masslessness of the photon. As mentioned above, the masslessness of the photon is the origin of both the long-range nature of the electrostatic force and the conservation of electric charge, so these properties of electromagnetism are sacrosanct and had to be maintained in any attempt at unifying electromagnetic and weak forces into a unified force law. These constraints turn out to be so tight that they lead to precise predictions for the masses of the W^+, W^-, and Z^0, which were brilliantly confirmed by experiment. The prediction and discovery of the mediators of the weak force can be likened in their profundity to the prediction and discovery of electromagnetic waves that followed from Maxwell’s formulation of electromagnetism.

The Higgs mechanism of dynamical symmetry breaking can be thought of as a spontaneous alignment of the vacuum (the state of lowest energy). The alignment is roughly analogous to the spontaneous magnetization encountered in ferromagnetic materials such as ordinary bar magnets. As particles such as the W^+, W^-, and Z^0 propagate through this aligned vacuum, they become massive as a result of their coherent interaction with the Higgs field (analogous to a magnetic field). It is believed at present that the dynamical interaction of particles with the vacuum is the origin of all the mass in the universe (including us!). The precise nature of the Higgs mechanism is not, however, well understood and remains the subject of intense study. In its simplest version small fluctuations of the Higgs alignment field manifest themselves as an elementary massive particle, dubbed the Higgs particle. The mass of the Higgs particle is not certainly known, but it must be less than approximately 1 TeV = 10^{12} eV.

The existence of the Higgs particle and the origin of electroweak symmetry breaking are therefore intimately related to the ancient and vexing question of the origin of mass itself. Indeed the hope of verifying this line of thinking has been one of the main justifications for constructing the $10 billion Superconducting Super Collider (SSC) just outside of Dallas. The SSC will collide two proton beams, each composed of protons accelerated to energies of 20 TeV. Compelling arguments suggest the collision energies are ample to either create and discover the elusive Higgs particle or, even if it doesn’t exist, to unravel the deep mystery as to the “origin of mass.”

Scientists at Los Alamos have been actively involved in several aspects of SSC physics. These have included theoretical studies concerning the nature and phenomenological implications of the Higgs mechanism as well as vigorous involvement in designing the huge (and very expensive) detectors needed to investigate these questions experimentally. Indeed the Los Alamos contribution to the SSC nicely exemplifies the breadth of the Laboratory’s unique and special characteristics: fundamental theoretical studies, detailed physics involvement with the design of the detectors, and superb engineering capability for its actual construction. This year over $7 million of SSC engineering funds will be spent at Los Alamos.

Baryon-Number Nonconservation and the Stability of Matter

One of the most intriguing possibilities for new physics at the SSC and an area in which Los Alamos has been heavily involved is the possibility of experimentally observing the violation of baryon-number (B) conservation. Baryon number appears to be an exactly conserved quantum number in nature. Its conservation, for example, prevents a proton ($B = 1$) from decaying into a state with $B = 0$ such as an electron and a pion or an electron and a photon, even though these decays are energetically very favorable. Indeed, it is the conservation of baryon number that keeps ordinary matter stable against radioactive decay (“diamonds are forever’’).

Notwithstanding a complete lack of experimental evidence for the violation of baryon-number conservation, the absolute validity of this law has always been viewed with some suspicion by theorists. Indeed, it has always appeared as a somewhat ad hoc phenomenological law with no fundamental basis for its understanding. There are two basic reasons for this skepticism. (1) Unlike the conservation of electric charge (which *is* believed to be exactly conserved) there is no known long-range force between baryons or local gauge symmetries associated with baryon number. Thus there is no fundamental reason for the exact conservation of baryon number; if exact, it would have to be viewed as an “accident’’ from our present viewpoint. (2) The second reason for skepticism is the well-known and remarkable fact that the universe appears to be overwhelmingly dominated by matter, which has positive baryon number, rather than antimatter, which has negative baryon number. This dominance is particularly hard to understand if baryon number is exactly conserved; in that case one would have to postulate that this gross asymmetry in the universe was put in "by hand'' as an initial boundary condition at the moment of creation!

Suggestively, the unified electroweak theory has, as a dynamical consequence, the breakdown of baryon-number conservation. As with the Higgs mechanism and the generation of mass, so here, the origin of the violation has its roots in the terribly complicated vacuum structure of the theory. It turns out that the field configuration of the vacuum has a periodic structure in which the different minima are separated by a potential barrier whose height is $M_w/\alpha \approx 10$ TeV (where M_w is the mass of the $W^{+,-}$ and α is the strength of the electromagnetic interaction). Furthermore these different minima correspond to different values of baryon number. Normally, an isolated system (such as "the universe'') sits at the bottom of a specific minimum with a given value of B. Now, B can change if either the system quantum mechanically tunnels through the barrier separating it from its neighboring vacuum (analogous to the way in which α particles tunnel through the nuclear potential barriers in the radioactive α decay of a nucleus), or the system has sufficient energy (nominally greater than about 10 TeV) to jump over the barrier (see Figure 1). It turns out that the probability for quantum-mechanical tunneling is ridiculously small, on the order of 10^{-170}. So even though baryon number can indeed change in the electroweak theory, the universe is still waiting for the first event of this kind to happen; diamonds are not going to spontaneously decay into radiation any time soon!

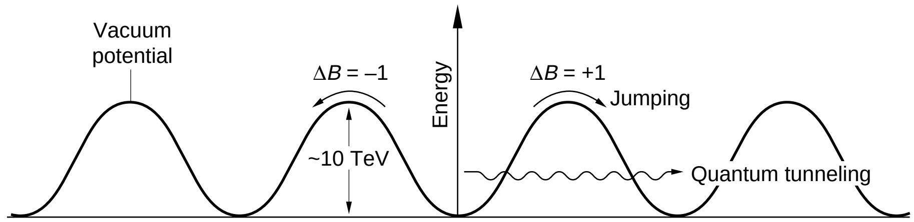


Figure 1. Structure of the Vacuum for the Electroweak Theory and the Possibility of Baryon-Number Nonconservation
This figure illustrates the periodic structure of the vacuum in the electroweak sector of the standard model. A system usually sits in the minimum of ar well of the vacuum potential where it is separated from adjacent minima by a very high energy barrier of about 10 TeV. Baryon number is conserved as long as the system remains in a single well of the potential. In principle, baryon-number conservation can be violated by two mechanisms: Either a system acquires enough energy, through heating or high-energy collisions, to jump over the barrier, or a system leaks through the barrier by quantum tunneling. A system will change its baryon number by one unit each time it jumps over or tunnels through the barrier. Collision energies at the SSC will be high enough to test the former mechanism of baryon-number violation.

However, the second possibility, namely jumping over the barrier, brings up two intriguing alternatives. One can, in principle, pump sufficient energy into a system either (1) by heating it up or (2) by performing some high-energy collision. In the former case one needs to heat the system to a temperature greater than roughly 10^{15} kelvins (100 million times hotter than the center of the sun!). Although this is clearly not feasible in the laboratory, the universe itself was at those extreme temperatures during its evolution following the big bang. The theory predicts that during that period baryon-number conservation was significantly violated by the system's jumping over the 10-TeV potential barrier. It is therefore conceivable, though not yet proven, that the present overwhelming domination of matter over antimatter was generated at that time. As the universe cooled down (its present mean temperature is only 3 kelvins), this asymmetry of matter over antimatter was frozen in since tunneling through the potential barrier is so strongly suppressed. So, we are led to the satisfying possibility that the unified electroweak theory can potentially explain both the dominance of matter over anti-matter, which requires the violation of baryon-number conservation at an earlier epoch and, at the same time, explain the apparent exact conservation of baryon number today (and, consequently, the stability of matter).

The Elementary Particles and Field Theory Group at the Laboratory began to study the breakdown of baryon-number conservation in electroweak theory at high temperatures in 1987, and soon after, it was realized that the baryon-number-violating processes described above could be important in the evolution of the early universe. Los Alamos research played a major role in clarifying the mechanism of the process and convincing a skeptical scientific establishment of its importance. Today we are investigating the very exciting possibility that violations of baryon-number conservation could be directly observed in the very high-energy collisions at the SSC. This research is being carried out in collaboration with the Institute of Nuclear Research in Moscow. By combining Los Alamos computer resources and expertise with that of the Russians, we hope to make reliable predictions long before the SSC turns on in the next century.

The possibility of performing a controlled experiment at the SSC that would detect baryon-number violation experimentally is truly fantastic. Thus far, all experimental attempts to observe a violation of baryon-number conservation by detecting the decay of the proton have failed, implying that the proton lifetime must be greater than approximately 10^{32} years! (Recall that the age of the universe is only (sic!) on the order of 10^{10} years). Further, observing violation of baryon-number conservation in a proton-proton collision at energies below 10 TeV (in the center-of-mass system) is impossible because quantum tunneling through the 10-TeV barrier of the periodic vacuum is exponentially suppressed (just as the universe is suppressed from doing so at temperatures below 10 TeV).

Thus baryon number will always be exactly conserved in a high-energy proton-proton collision below 10 TeV. Recall, however, that the total center-of-mass energy available in a collision at the SSC will be 40 TeV, which is fortuitously in excess of the barrier height. A naive calculation indicates that at these energies the colliding protons could jump over the barrier and baryon-number conservation would be violated. In such a case a very dramatic effect would occur: the copious production of *W*'s and *Z*'s (and Higgs particles) (see Figure 2). Recall that ordinarily the probability of producing just one of these particles is, as with photons, at best on the order of the electromagnetic interaction strength, or less than 1 percent! Unfortunately many subtleties need to be carefully considered in these calculations, and it is not at all clear that this naive prediction will, in fact, be borne out. As stated above, this subject is under intense investigation at Los Alamos and elsewhere at the present time. If the predictions are true, the results would be quite spectacular.

Quantum Chromodynamics—The Theory of the Strong Interactions

Even as gauge symmetry was being recognized as essential to unification of the electromagnetic and weak forces, it was simultaneously playing a no less central role in understanding the strong force that binds protons and neutrons together in atomic nuclei. Our present understanding of the strong force is that it too arises from a force-carrier with spin angular momentum equal to one that is analogous to the photon of QED. The carrier of the strong force is called a "gluon.''

The present theory of the strong interactions was developed during the 1970s and is based on the idea that all baryons and mesons are made of spin-1/2 particles called quarks that interact via the exchange of gluons. Quarks themselves were first postulated in the early 1960s independently by Murray Gell-Mann and George Zweig. They are indeed the "fundamental building blocks" and have the curious property of carrying an electric charge that comes in units of $e/3$, where e is the magnitude of the charge carried by electrons and protons; previously, e was thought to be the smallest unit of electric charge carried by any particle. Interestingly, Zweig is now a theoretical biophysicist at Los Alamos doing pioneering work on the physics of the ear and Gell-Mann is spending his sabbatical year at the Laboratory mostly involved in exploring new ideas concerning the concept of complexity.

At first the existence of these "quarks'' was not taken very seriously, since no one had ever observed a free quark (or fractional charge) in a particle detector of any kind. However, the unmistakable presence of point scattering centers within protons discovered by the now famous deep-inelastic-scattering experiments at the Stanford Linear Accelerator Center (SLAC) in the early 1970s, as well as other indirect checks of the theory, finally led to the acceptance of quarks as the genuine building blocks of all the strongly interacting particles (known collectively as hadrons). Thus, in analogy with QED, quarks play the role of electrons and gluons that of photons. The analog to the electric charge—the "strong charge,'' so to speak—was whimsically dubbed color, and so the theory of the strong interactions became known as quantum chromodynamics' (QCD). A crucially new feature of QCD is that color comes in three varieties and that the gluon itself carries this color. In that respect the gluon resembles more the *W*s and *Z* of the weak interactions than the photon of QED in that it can now self-interact. On the other hand, the gauge symmetry of the vacuum of

QCD is *not* spontaneously broken, so the gluons, at least naively, remain massless like photons. The strength of the gluon self-interaction is, however, so large that it is believed to forbid color and, in particular, quarks from ever being liberated and observed directly in an experiment. Nevertheless, the theory dictates that at high energies quarks and gluons inside the proton should behave essentially as if they were free particles. In other words, in a high-energy collision between two protons, the quarks that make up the protons should briefly act as if they were not bound. The precise predictions of QCD were, in fact, brilliantly confirmed by the SLAC experiments for which the originators received the Nobel Prize. It should be recognized that truly free quarks and gluons have never been observed *directly* in *any* experiment. The situation recalls one encountered in an earlier period of the history of science. By the end of the last century, most working physicists and chemists tacitly assumed the existence of atoms in order to interpret a wide variety of experimental data. But not one of those scientists had ever seen a single atom, so that as late as 1916 no less a figure than Ernst Mach still doubted their existence.

The theoretical aspects of QCD have been studied most intensively in recent years by the technique of "lattice gauge theory" calculations, performed on state-of-the-art supercomputers. Rajan Gupta heads the Los Alamos effort in this area, which is recognized as one of the leading groups worldwide. In this decade the appearance of the Teraflop machine (capable of executing 1 *trillion* basic mathematical operations per second) is expected to make it possible for Gupta and his coworkers to obtain a "solution'' to full QCD at an accuracy of about 10 to 20 percent (see "Testing the Standard Model Using State-of-the-Art Supercomputers). These calculations are based on Monte Carlo algorithms invented at Los Alamos forty years ago by Nick Metropolis and first used by Stanislaw Ulam and Enrico Fermi in now classic work.

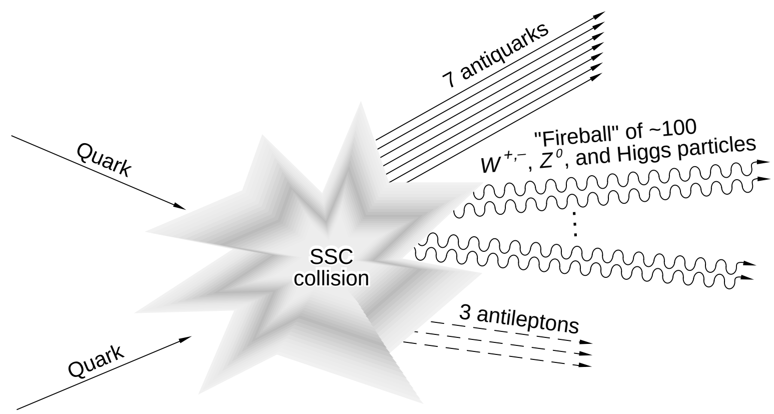


Figure 2. Baryon-Number Violation in a Collision at the SSC
The figure illustrates a collision at the SSC that violates baryon-number (*B*) conservation and lepton-number (*L*) conservation. (Muons, electrons, and neutrinos are leptons and have *L* = 1; their antiparticles have *L* = °1). In this collision *B* and *L* each change by three units. The prominent signature of such an event would be the production of an extraordinarily large number of *W*s, *Z*s, and Higgs particles. Currently theorists are sharply divided on whether *B*-violating events can occur at the SSC at an observable rate. Estimates of the rate published by prominent scientists differ by as much as a hundred orders of magnitude! For a review of the controversy see the article by Michael Mattis listed in Further Reading.

Time-dependent Calculations of Heavy-Ion Collisions and the Quark-Gluon Plasma

Present lattice Monte Carlo simulations of QCD at finite temperature predict that although quarks and gluons can never get completely free of each other, when the temperature gets high enough (around 150 MeV, or more than 100,000 times hotter than the center of the sun), the behavior of matter changes and a phase transition takes place (like the transition from water to steam). In this new high-temperature phase quarks and gluons interact more weakly than they do in normal nuclear matter, somewhat like electrically charged particles in a plasma. For that reason this new phase of matter is called the quark-gluon-plasma phase.

The question naturally arises of how one could detect the presence of such a phase. The requisite high temperatures and densities occurred in the early universe, but it was soon discovered that even if this phase existed in the early universe, it would not have affected very much the standard nucleosynthesis processes that create helium from hydrogen because those processes mainly take place later in the evolution of the big bang. Thus, we would not be able to tell from looking at the relative abundances of hydrogen and helium in the present universe whether or not the quark-gluon plasma ever actually existed.

If the universe won't cooperate and provide us with a suitable laboratory for the quark-gluon plasma, then we have to build one ourselves. About ten years ago, it was realized that relativistic heavy-ion collisions—those in which each nucleon in the heavy-ion projectiles has an energy between 10 and 200 GeV (1 GeV = 10^9 eV)—could conceivably produce energy densities as high as 2–20 Gev/fermi^3 and temperatures on the order of a few Gev (or ten times larger than the quark-gluon-plasma phase-transition temperature). At CERN collisions of light-ion beams (typically consisting of carbon or sulphur nuclei) with other nuclei have produced conditions that approach these extreme temperatures and densities, but as yet they have not provided unmistakable evidence of a quark-gluon plasma. By 1996 the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory will be smashing gold and even heavier nuclei together at even higher energies, well above those required to produce the new phase of quark-gluon plasma in the laboratory. Then the real problem will be to sift through the particle debris of such collisions and to recognize the "smoking gun" of the quark-gluon-plasma phase; that is, a signal that could be produced only by such a phase.

It will not be easy. One of the fundamental and most tantalizing problems in this field is, in fact, what constitutes a "smoking gun." Thus far no one has devised a clear unambiguous methodology for isolating the signal for the elusive quark-gluon plasma. Although lattice calculations predict the existence of the quark-gluon phase, those calculations predict only static equilibrium properties of the plasma. The situation in a high-energy heavy-ion collision is very far from equilibrium, at least initially, and the quark-gluon plasma is expected to persist for no more than 10^{-22} second before the nuclear fireball cools and begins to fragment into ordinary pions, protons, neutrons, and other hadrons. Hence any signal of the quark-gluon plasma could easily be masked by strong interactions taking place in the ordinary hadronic-matter phase. Thus it becomes imperative to understand in great detail the non-equilibrium processes taking place in the plasma.

In particular, we want to calculate the rate at which lepton particle-antiparticle pairs (such as muon-antimuon pairs) would be produced by electromagnetic interactions in the plasma. Leptons interact only through the electroweak interaction, and so, once produced, they would be unaffected by the strong forces and propagate straight through the plasma in the collision region to the particle detectors. In order to determine whether the rate and the energy spectrum of lepton-pair production in the quark-gluon-plasma phase would be different from those in ordinary matter, one needs to know the time evolution of the quark-gluon plasma. In particular, one needs to calculate the distribution of quarks and antiquarks as a function of time since quark-antiquark pairs would be the source of the lepton pairs. Although one can make estimates of these pair-production rates based on equilibrium QCD or perturbation theory, such estimates are neither reliable nor accurate enough to serve as a quantitative test of the presence of quark-gluon plasma.

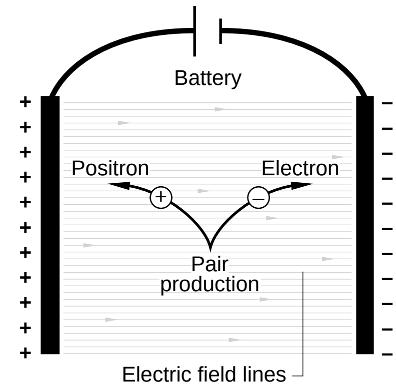
This problem presents a new challenge for quantum field theorists. Previously theorists treated the interaction region of a collision as a "black box.'' They would calculate only the final state resulting from a collision based on the initial state preceding the collision. For example, they might predict how many particles of a certain type and with a certain amount of energy should reach a particular detector following a collision between two protons whose initial energies and momenta are known. This prediction does not require a calculation of the space-time evolution of the interacting particles: one doesn't have to solve the full time-dependent Schrödinger equation to answer questions about the asymptotic, or final, states of particles reaching the detectors. For the quark-gluon plasma, however, one needs to track the space-time evolution of the system *in detail* in order to make quantitative predictions that can be tested by experimenters. Remarkably, the need for a new formulation of quantum field theory had already arisen in connection with the time evolution of the early universe. In that case one needed to follow in detail the time evolution of quantum fields in a time-evolving gravitational field. The formulation required for that problem in early-universe cosmology turns out to be directly applicable to the time evolution of the quark-gluon plasma in a relativistic heavy-ion collision—a beautiful illustration of the underlying unity and interconnectedness of physics.

What then does a relativistic heavy-ion collision look like? The strong forces between hadrons in the colliding nuclei are mediated by the exchange of colored gluons. As a result of these exchanges, the initially colorless hadrons become charged with opposite colors and quickly separate since the hadrons are moving away from each other at velocities near the speed of light. As shown in Figure 3, this situation is analogous to two color-charged plates moving away from each other with a constant color electric field between them. When the energy density in the color field is large enough to produce quark-antiquark pairs, then the "vacuum'' between the hadrons becomes unstable and quark-antiquark pairs begin to pop out of the vacuum, just as electron-positron pairs can be created in a very strong electric field. These quarks then produce more color field that tends to reduce the original one. Thus the resulting quarks and gluons move through and interact with the original time-varying color electric field. The quarks in the quark-gluon plasma also emit pairs of leptons via electromagnetic processes. The lepton pairs leave the collision region and can be seen in standard particle detectors. As the system expands, the energy density decreases and the hadronic phase become energetically favorable so that the quark-gluon-plasma phase changes back into the ordinary hadronic phase.

Even five years ago, calculating the time evolution of this semi-classical model of heavy-ion collisions would not have been possible. Only now that supercomputers have be-

Figure 3. Creation of Quark-Gluon Plasma in Relativistic Heavy-Ion Collisions

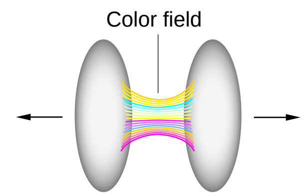
(a) Electron-Positron Pair Production by an Electromagnetic Field



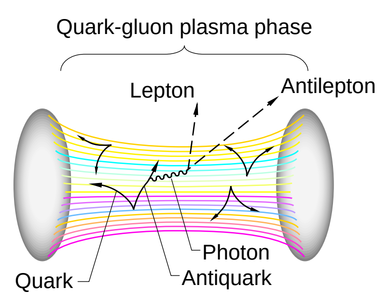
A battery maintains a high voltage difference between two parallel conducting plates. The electric field between the plates is sufficiently high that electron-positron (e^-e^+) pairs are created spontaneously from the vacuum.

(b) Formation of Quark-Gluon Plasma in Relativistic Heavy-Ion Collision

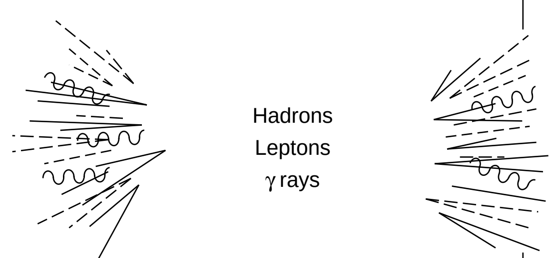
Initial State: Two heavy nuclei approach each other at velocities near the speed of light. At these speeds they appear flattened because of Lorentz contraction.



Immediately Following Collision: The two nuclei have collided and produced a hot region between them.



Creation of Quark-Gluon Plasma: Blobs of nuclear material move away from each other. In the color field between them, pair-production processes analogous to those shown in (a) produce quark-antiquark pairs. The quarks and antiquarks can interact with the electroweak field to produce lepton pairs such as a muon-antimuon ($\mu^+\mu^-$) pair. The leptons are not affected by the strong force as they escape from the quark-gluon plasma.



Final State: Following the collision the quark-gluon-plasma phase has changed back to the hadronic phase and a large number of hadrons (baryons and mesons), leptons, and γ rays emerge from the collision.

come readily available, particularly at the Advanced Computing Laboratory at Los Alamos, is the problem beginning to become tractable numerically. Our calculations of this semi-classical model have already led to some interesting results. We have been able to check many assumptions of more phenomenological pictures of the time evolution of the quark-gluon plasma, such as the classical transport, or hydrodynamic, models. Such models typically do not start from the basic equations of QCD, but rather postulate some effective, essentially classical model. The importance of avoiding such ad hoc assumptions by calculating the evolution of the plasma directly from first principles of QCD (even in simplified approximations) cannot be overestimated. To our knowledge, no other group of researchers anywhere in the world has attempted to carry out this program. Yet, as we have already discussed, any hope of recognizing the quark-gluon-plasma phase in heavy-ion collisions at RHIC depends ultimately upon the detailed theoretical predictions of the signature of the plasma. These, in turn, can come only from such a first-principles QCD attack on the problem with the formidable computational resources available almost uniquely at Los Alamos. As in the SSC effort, our theoretical group is keeping in close contact with the experimenters in the Physics Division who are major participants in planning the search at RHIC for the quark-gluon plasma.

In conclusion, the long quest for a deep and unified understanding of the most basic laws of nature that govern fundamental processes has seen remarkable progress in the last twenty years. These basic laws fashion the incredible structure we see around us; they underlie not only the complex structures that constitute our macroscopic world but also the beginnings of the universe and the evolution of the heavens. From its earliest history Los Alamos has played its part in elucidating this structure. It continues to do so and in so doing provides a unique link between some of the most fundamental questions in science and their integration into a broader spectrum of problems facing society as we approach the twenty-first century. ■

Left to right: Rajan Gupta, Emil Mottola, Geoffrey West, Fred Cooper, and Michael Mattis

Further Reading

Los Alamos Science, Number 11. 1984. The entire issue is devoted to particle physics.

Michael P. Mattis. 1992. The riddle of high-energy baryon number violation. *Physics Reports* 214: 159–221.

Michael P. Mattis and Emil Mottola, editors. 1990. *Baryon Number Violation at the SSC? Proceedings of the Santa Fe Workshop.* Singapore: World Scientific.

Y. Kluger, J. M. Eisenberg, B. Svetitsky, F. Cooper, and E. Mottola. 1992. Fermion pair production in a strong electric field. *Physical Review D* 45: 4659.

Geoffrey B. West received his B.A. in natural sciences (physics) from the University of Cambridge in 1961 and his Ph.D. in Physics at Stanford University in 1966. He came to the Laboratory in 1974 as the first leader of what is now the Elementary Particles and Field Theory Group (T-8). In 1981 he was made a Laboratory Fellow and in 1988 resumed the leadership of T-8. His present interests revolve around the structure and consistency of quantum field theory and, in particular, its relevance to quantum chromodynamics and unified field theory.

Frederick M. Cooper received his Ph.D. from Harvard University in 1968 and came to the Laboratory seven years later. His research interests include hydrodynamical and transport-theory approaches to problems of multiparticle production, analytical methods for studying field theories and nonlinear dynamical systems, and understanding the masses of the heavy quarks. His long-standing interest in the time evolution of relativistic quantum systems led to the work presented here on the evolution of the quark-gluon plasma.

Emil Mottola received his undergraduate degree in astrophysics and, in 1979, his Ph.D. in physics, both from Columbia University. He joined the staff of the Elementary Particles and Field Theory Group at the Laboratory in 1986. His research focuses on quantum fields, quantum gravity, and the early universe and includes studies of baryon-number nonconservation and the evolution of the quark-gluon plasma.

Michael P. Mattis received his A.B. in physics from Harvard University in 1981 and his Ph.D. from Stanford University in 1986. After three years as Enrico Fermi Fellow at the University of Chicago, he came to the Laboratory first as a J. Robert Oppenheimer Fellow and then as a Staff Member. His recent work on baryon-number violation in electroweak theory has earned him a Superconducting Super Collider Fellowship.

Testing the Standard Model of Particle Interactions Using State-of-the-Art Supercomputers

Rajan Gupta

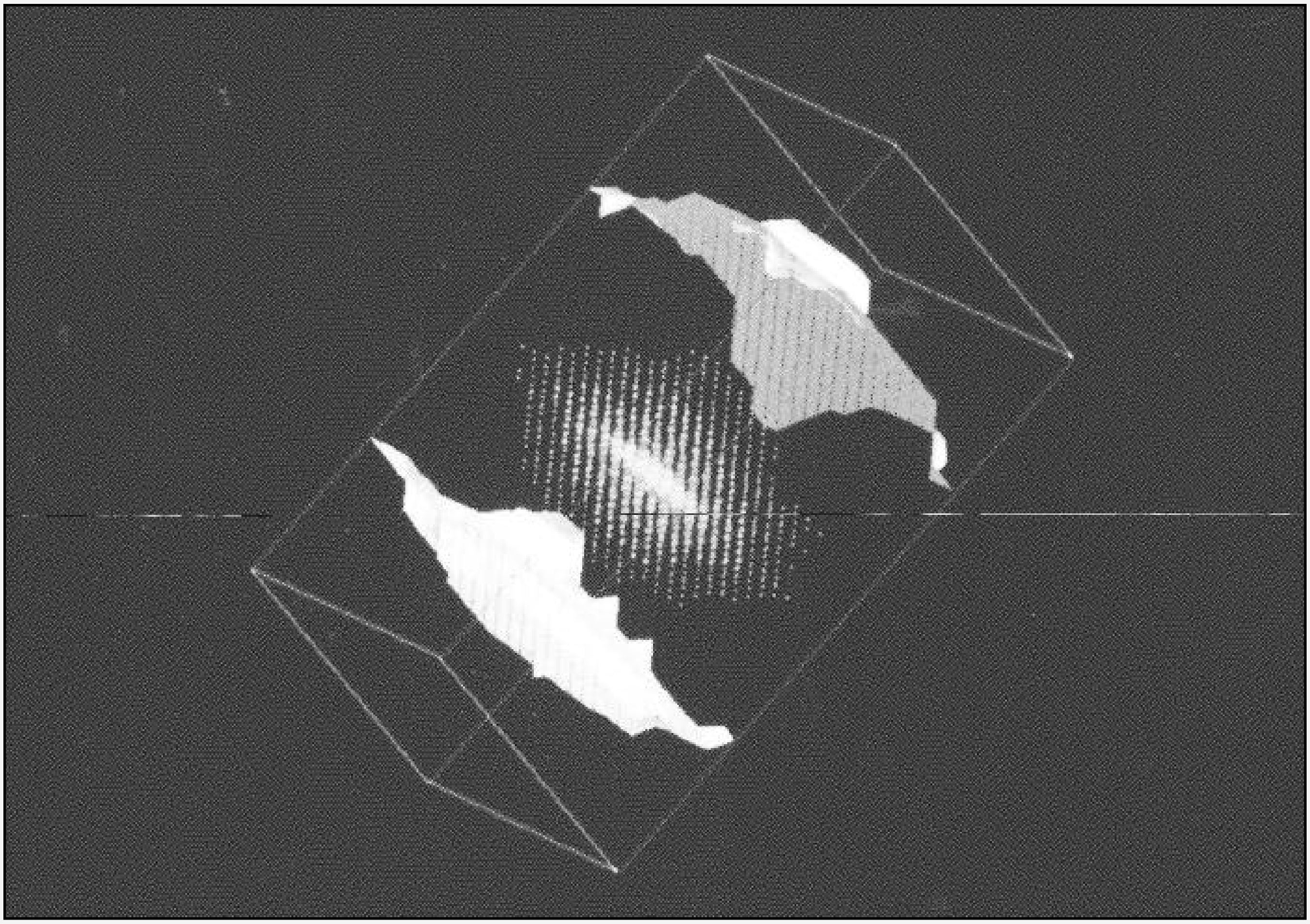
The successes of the standard model of electromagnetic, weak, and strong interactions have been remarkable. Nevertheless, this model contains many assumptions and undetermined parameters that are displeasing aesthetically. Also, the model has not yet produced a satisfying route to unifying gravity with the other three fundamental forces.

The goal of particle physicists now is to discover where the standard model fails and so to find clues to a better theory, perhaps ultimately achieving Einstein's dream of unifying all forces including gravity. It is hoped that the clues will emerge from highly sophisticated experiments in which particles are accelerated and smashed together at very high energies. To date, however, no deviations from the standard model have been found, and so the search for new clues must be performed at even higher energies—such as those that will be achieved at the Superconducting Super Collider (SSC) now under construction in Texas.

Much theoretical analysis is still needed to interpret the results of these extraordinary and expensive experiments. The strong-interaction part of the standard model—quantum chromodynamics, or QCD—presents the major computational stumbling block. Qualitatively, QCD has all the right properties, but so far theoretical physicists have not been able to extract accurate predictions from this precise mathematical model with the traditional tools of the theoretical physicist—pencil and paper. To obtain reliable self-consistent results when dealing with the strong force between, say, two protons requires calculating many subprocesses involving quarks and gluons. In fact, the number of subprocesses is so large that the calculation far exceeds the scope of analytical techniques.

The solution is to turn to a new tool: the supercomputer. Large-scale numerical simulations of QCD are the most promising technique for analyzing the strong interactions. In order to solve QCD on a computer one has to approximate space and time by a four-dimensional grid, or lattice, of points. The discretized version of the theory is called lattice QCD. Experimentally measurable quantities (such as the particle masses and the probabilities of specific transitions) are determined from a statistical average over quantum fluctuations in the quark and gluon fields. The fluctuations at each position in the lattice are simulated by a Monte Carlo procedure, so each Monte Carlo calculation determines one state in a statistical sample of possible states of a system. Monte Carlo methods are an efficient way of sampling the important states, that is, states that give the dominant contributions to the process. The best Monte Carlo calculations to date have used lattices of size up to $32^3 \times 48$ and generated only a small statistical sample (twenty to fifty of the possible states). The three sources of errors in such simulations are the lattice size, the lattice spacing, and the limited statistical sample. These errors can be systematically reduced by making the lattice size larger, the statistical sample larger, and the lattice spacing smaller.

To reduce statistical and systematic errors to the level of a few percent requires a computer with a very large memory and a very high operating speed, over 1000 billions of arithmetic operations per second. For comparison, a typical state-of-the-art home computer has a few million bytes of memory and runs at a few million operations per second. The required technology is just beginning to appear in the form of the parallel supercomputer. In fact, scientists interested in solving the riddle of QCD have played a significant role in the development of parallel supercomputers. The basic principle of these new machines is simple—thousands of



The Pion Propagator

The event depicted here was generated using Monte Carlo simulations of lattice QCD and shows the creation of a pion near the center of the box and its propagation through space both forward and backward in time. The pion propagator, which describes this event, is the correlation function defining the probability amplitude for finding a pion at **x** at time *t* given that it started at the origin at time *t*=0. The propagator is a function of the three spatial coordinates and the time coordinate, but for the purposes of visualization, the data have been averaged over the *z* coordinate and displayed as a function of x and y (the short axes of the box) and time (the long axis). The size of the green "bubbles" at each point represents the magnitude of the propagator at that point in space-time. Note that the "bubbles" decrease in size with distance from the origin, indicating that the magnitude of the propagator decreases. The white surfaces at the ends of the box represent surfaces on which the probability amplitude of the propagator is a constant. The mass of the pion can be calculated from the rate at which the propagator dies out as a function of time. Since the mass of the pion is known from experiment, such calculations of the pion mass can be used to calibrate Monte Carlo simulations of more complicated processes. The data for this Monte Carlo lattice-QCD event were generated on the CM-200, a Connection Machine with 16,000 processors. Many such Monte Carlo events must be calculated to generate a statistically reliable sample for estimating the pion mass. This numerical calculation of the pion propagator exemplifies the first-principles approach to solving strong interactions using lattice QCD.

small but powerful computers work simultaneously to solve one big problem.

The first large-scale simulations of lattice QCD were performed around 1980. Because in those days the fastest computer generally available had the same power as today's desktop workstation, the simulations involved so many approximations that the results were not realistic. To overcome this limitation, physicists turned to parallel computers, often building them themselves. Though the capabilities of the earlier versions of such computers were quite limited, a start had been made, and scientists in other fields became excited by the potential of parallel computation.

A watershed for parallel computing came in 1988. In that year Thinking Machines Corporation introduced the first commercial parallel supercomputer, the Connection Machine 2 (CM-2), and DOE announced its first "Grand Challenges" program, which allocated large grants of supercomputer time to scientists working on key computationally intensive problems. The Los Alamos QCD collaboration was one such recipient. Building on the cruder calculations of 1980–88, we demonstrated, using the most powerful Crays and the CM-2, the viability of numerical methods for QCD. We obtained many new results with an accuracy comparing favorably with the best analytical estimates, though the computational power was still too limited to make definitive predictions. The Los Alamos QCD collaboration also played a key role in technology transfer by pioneering the use of the CM-2 and showing other scientists at the Laboratory and around the world how to use this machine as a production supercomputer.

The revolutionary nature of parallel computing poses new challenges beyond the design of faster hardware: We also need to develop software paradigms for parallel supercomputers that simplify their use for a wider variety of problems and fully exploit the advantages of parallelism. Progress requires close collaboration between scientists, computer engineers, and applied mathematicians. A successful collaboration of this type has begun between our Los Alamos QCD group, the Advanced Computing Laboratory (a DOE High Performance Computing and Research Center at Los Alamos), and Thinking Machines Corporation. This collaboration has received a Grand Challenge grant from the DOE to perform the next generation of QCD calculations on the 1024-node CM-5 located at Los Alamos National Laboratory.

Compared with the computing power of the Cray computers and the CM-2 used in our previous Grand Challenge calculation, the available memory in the CM-5 will be more than an order of magnitude greater, and the available computer speed will be at least two orders of magnitude greater. We plan to use these improvements in computer hardware in three ways. First, we will exploit the increased speed to generate much larger statistical samples for the processes we have already calculated. These simulations will be done on lattices of the size we used previously, $32^3 \times 64$. Second, once the new computers become stable and the proposed hardware upgrades are in place, we will begin simulations on much larger lattices of size $64^3 \times 128$. Finally, we will exploit the larger memory to undertake more complex calculations than have been feasible so far.

Our calculations will yield predictions for a large number of observables affected by strong interactions. The theoretical uncertainties in these quantities will be as low as a few percent. An illustrative and important example concerns the violation of "CP" symmetry. All interactions except the weak are unaffected if one simultaneously interchanges particles and antiparticles (charge conjugation, or C) and takes the mirror image of space (parity transformation, or P). The only experimentally established violation of CP symmetry occurs in the transmutation of a neutral strange meson (K^0) into its antiparticle ($\bar{K}^0$). This discovery was made more than twenty-five years ago and was rewarded with the 1980 Nobel Prize. To test the standard model against this long-standing experimental result requires calculating a parameter called B_K. The more accurately we can determine this quantity, the more precisely we can probe the validity of the standard model. Prior to numerical calculations, the uncertainty in B_K was at least 50 percent. Our present Grand Challenge calculation has reduced this error to about 15 percent. We expect that the planned calculation will further reduce the level of uncertainty to a few percent.

To search for physics beyond or in contradiction to the standard model, one must combine experimental and theoretical knowledge of many quantities. Our calculations will provide theoretical results for a number of the standard-model observables with an accuracy of 10 to 25 percent, small enough to put meaningful constraints on the unknown parameters of the model. Thus, over the next few years, theoretical predictions combined with new experimental results will provide stringent tests of the standard model. ◻

Further Reading

Michael Creutz, editor. 1992. *Quantum Fields on the Computer*. Singapore: World Scientific.

M. Fukugita, Y. Iwasaki, M. Okawa, and A. Ukawa, editors. 1992. *Lattice 91: Proceedings of the International Symposium on Lattice Field Theory, Tsukuba, Japan, 5–9 November 1991.* Amsterdam: North-Holland.

Rajan Gupta was awarded an M.S. by the University of Delhi, India, and a Ph.D. by the California Institute of Technology in 1982. In 1985 he became an Oppenheimer Fellow at the Laboratory and three years later joined the staff of the Elementary Particles and Field Theory Group. Among his scientific interests are the development of non-perturbative methods to solve QCD, using lattice QCD to constrain the standard model, and studying the critical behavior of statistical-mechanics models by means of Monte Carlo renormalization-group methods.

Testing the Standard Model of Particle Interactions Using State-of-the-Art Supercomputers

Rajan Gupta

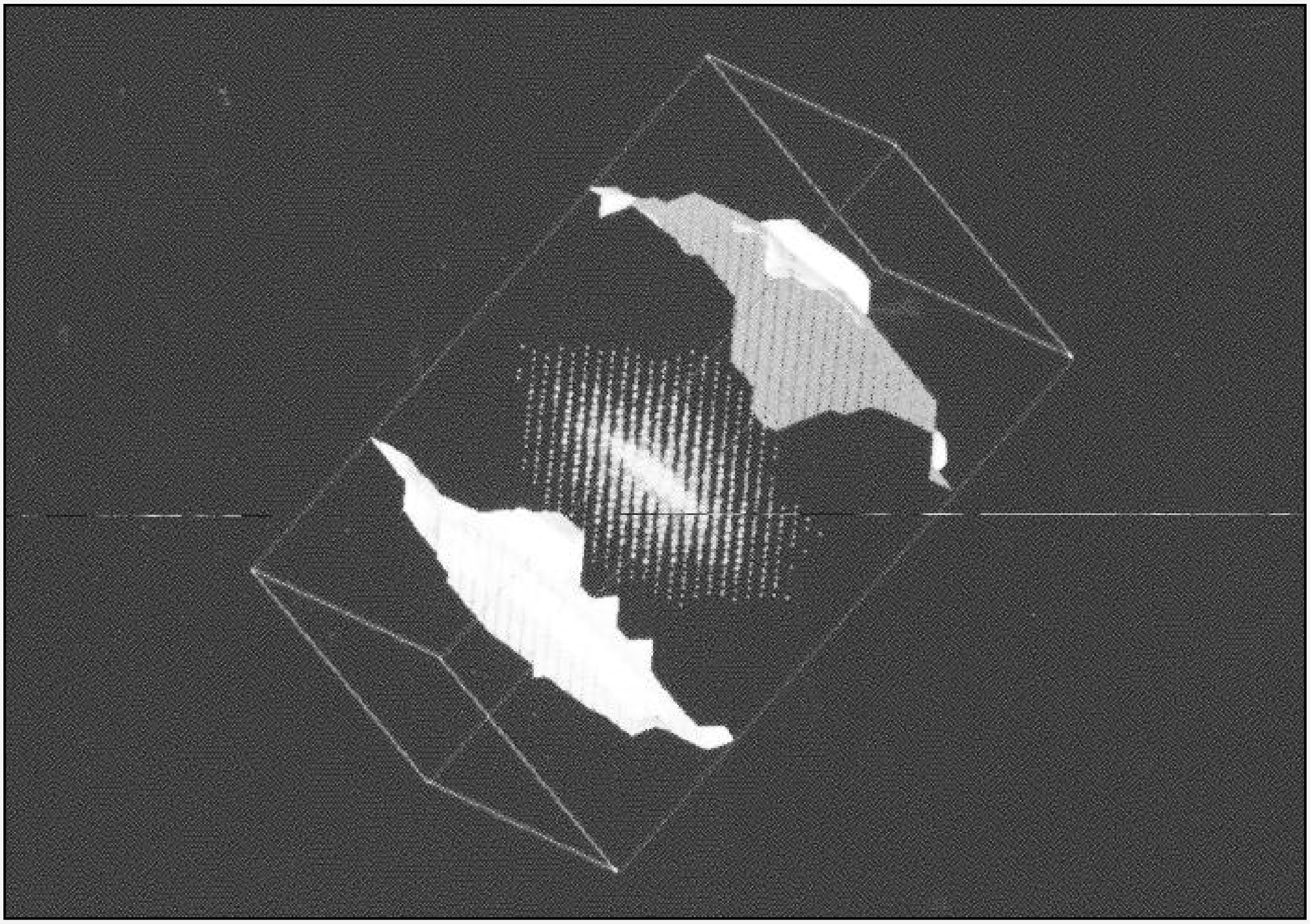
The successes of the standard model of electromagnetic, weak, and strong interactions have been remarkable. Nevertheless, this model contains many assumptions and undetermined parameters that are displeasing aesthetically. Also, the model has not yet produced a satisfying route to unifying gravity with the other three fundamental forces.

The goal of particle physicists now is to discover where the standard model fails and so to find clues to a better theory, perhaps ultimately achieving Einstein's dream of unifying all forces including gravity. It is hoped that the clues will emerge from highly sophisticated experiments in which particles are accelerated and smashed together at very high energies. To date, however, no deviations from the standard model have been found, and so the search for new clues must be performed at even higher energies—such as those that will be achieved at the Superconducting Super Collider (SSC) now under construction in Texas.

Much theoretical analysis is still needed to interpret the results of these extraordinary and expensive experiments. The strong-interaction part of the standard model—quantum chromodynamics, or QCD—presents the major computational stumbling block. Qualitatively, QCD has all the right properties, but so far theoretical physicists have not been able to extract accurate predictions from this precise mathematical model with the traditional tools of the theoretical physicist—pencil and paper. To obtain reliable self-consistent results when dealing with the strong force between, say, two protons requires calculating many subprocesses involving quarks and gluons. In fact, the number of subprocesses is so large that the calculation far exceeds the scope of analytical techniques.

The solution is to turn to a new tool: the supercomputer. Large-scale numerical simulations of QCD are the most promising technique for analyzing the strong interactions. In order to solve QCD on a computer one has to approximate space and time by a four-dimensional grid, or lattice, of points. The discretized version of the theory is called lattice QCD. Experimentally measurable quantities (such as the particle masses and the probabilities of specific transitions) are determined from a statistical average over quantum fluctuations in the quark and gluon fields. The fluctuations at each position in the lattice are simulated by a Monte Carlo procedure, so each Monte Carlo calculation determines one state in a statistical sample of possible states of a system. Monte Carlo methods are an efficient way of sampling the important states, that is, states that give the dominant contributions to the process. The best Monte Carlo calculations to date have used lattices of size up to $32^3 \times 48$ and generated only a small statistical sample (twenty to fifty of the possible states). The three sources of errors in such simulations are the lattice size, the lattice spacing, and the limited statistical sample. These errors can be systematically reduced by making the lattice size larger, the statistical sample larger, and the lattice spacing smaller.

To reduce statistical and systematic errors to the level of a few percent requires a computer with a very large memory and a very high operating speed, over 1000 billions of arithmetic operations per second. For comparison, a typical state-of-the-art home computer has a few million bytes of memory and runs at a few million operations per second. The required technology is just beginning to appear in the form of the parallel supercomputer. In fact, scientists interested in solving the riddle of QCD have played a significant role in the development of parallel supercomputers. The basic principle of these new machines is simple—thousands of



The Pion Propagator

The event depicted here was generated using Monte Carlo simulations of lattice QCD and shows the creation of a pion near the center of the box and its propagation through space both forward and backward in time. The pion propagator, which describes this event, is the correlation function defining the probability amplitude for finding a pion at **x** at time *t* given that it started at the origin at time *t*=0. The propagator is a function of the three spatial coordinates and the time coordinate, but for the purposes of visualization, the data have been averaged over the *z* coordinate and displayed as a function of x and y (the short axes of the box) and time (the long axis). The size of the green "bubbles" at each point represents the magnitude of the propagator at that point in space-time. Note that the "bubbles" decrease in size with distance from the origin, indicating that the magnitude of the propagator decreases. The white surfaces at the ends of the box represent surfaces on which the probability amplitude of the propagator is a constant. The mass of the pion can be calculated from the rate at which the propagator dies out as a function of time. Since the mass of the pion is known from experiment, such calculations of the pion mass can be used to calibrate Monte Carlo simulations of more complicated processes. The data for this Monte Carlo lattice-QCD event were generated on the CM-200, a Connection Machine with 16,000 processors. Many such Monte Carlo events must be calculated to generate a statistically reliable sample for estimating the pion mass. This numerical calculation of the pion propagator exemplifies the first-principles approach to solving strong interactions using lattice QCD.

small but powerful computers work simultaneously to solve one big problem.

The first large-scale simulations of lattice QCD were performed around 1980. Because in those days the fastest computer generally available had the same power as today's desktop workstation, the simulations involved so many approximations that the results were not realistic. To overcome this limitation, physicists turned to parallel computers, often building them themselves. Though the capabilities of the earlier versions of such computers were quite limited, a start had been made, and scientists in other fields became excited by the potential of parallel computation.

A watershed for parallel computing came in 1988. In that year Thinking Machines Corporation introduced the first commercial parallel supercomputer, the Connection Machine 2 (CM-2), and DOE announced its first "Grand Challenges" program, which allocated large grants of supercomputer time to scientists working on key computationally intensive problems. The Los Alamos QCD collaboration was one such recipient. Building on the cruder calculations of 1980–88, we demonstrated, using the most powerful Crays and the CM-2, the viability of numerical methods for QCD. We obtained many new results with an accuracy comparing favorably with the best analytical estimates, though the computational power was still too limited to make definitive predictions. The Los Alamos QCD collaboration also played a key role in technology transfer by pioneering the use of the CM-2 and showing other scientists at the Laboratory and around the world how to use this machine as a production supercomputer.

The revolutionary nature of parallel computing poses new challenges beyond the design of faster hardware: We also need to develop software paradigms for parallel supercomputers that simplify their use for a wider variety of problems and fully exploit the advantages of parallelism. Progress requires close collaboration between scientists, computer engineers, and applied mathematicians. A successful collaboration of this type has begun between our Los Alamos QCD group, the Advanced Computing Laboratory (a DOE High Performance Computing and Research Center at Los Alamos), and Thinking Machines Corporation. This collaboration has received a Grand Challenge grant from the DOE to perform the next generation of QCD calculations on the 1024-node CM-5 located at Los Alamos National Laboratory.

Compared with the computing power of the Cray computers and the CM-2 used in our previous Grand Challenge calculation, the available memory in the CM-5 will be more than an order of magnitude greater, and the available computer speed will be at least two orders of magnitude greater. We plan to use these improvements in computer hardware in three ways. First, we will exploit the increased speed to generate much larger statistical samples for the processes we have already calculated. These simulations will be done on lattices of the size we used previously, $32^3 \times 64$. Second, once the new computers become stable and the proposed hardware upgrades are in place, we will begin simulations on much larger lattices of size $64^3 \times 128$. Finally, we will exploit the larger memory to undertake more complex calculations than have been feasible so far.

Our calculations will yield predictions for a large number of observables affected by strong interactions. The theoretical uncertainties in these quantities will be as low as a few percent. An illustrative and important example concerns the violation of "CP" symmetry. All interactions except the weak are unaffected if one simultaneously interchanges particles and antiparticles (charge conjugation, or C) and takes the mirror image of space (parity transformation, or P). The only experimentally established violation of CP symmetry occurs in the transmutation of a neutral strange meson (K^0) into its antiparticle ($\bar{K}^0$). This discovery was made more than twenty-five years ago and was rewarded with the 1980 Nobel Prize. To test the standard model against this long-standing experimental result requires calculating a parameter called B_K. The more accurately we can determine this quantity, the more precisely we can probe the validity of the standard model. Prior to numerical calculations, the uncertainty in B_K was at least 50 percent. Our present Grand Challenge calculation has reduced this error to about 15 percent. We expect that the planned calculation will further reduce the level of uncertainty to a few percent.

To search for physics beyond or in contradiction to the standard model, one must combine experimental and theoretical knowledge of many quantities. Our calculations will provide theoretical results for a number of the standard-model observables with an accuracy of 10 to 25 percent, small enough to put meaningful constraints on the unknown parameters of the model. Thus, over the next few years, theoretical predictions combined with new experimental results will provide stringent tests of the standard model. ◻

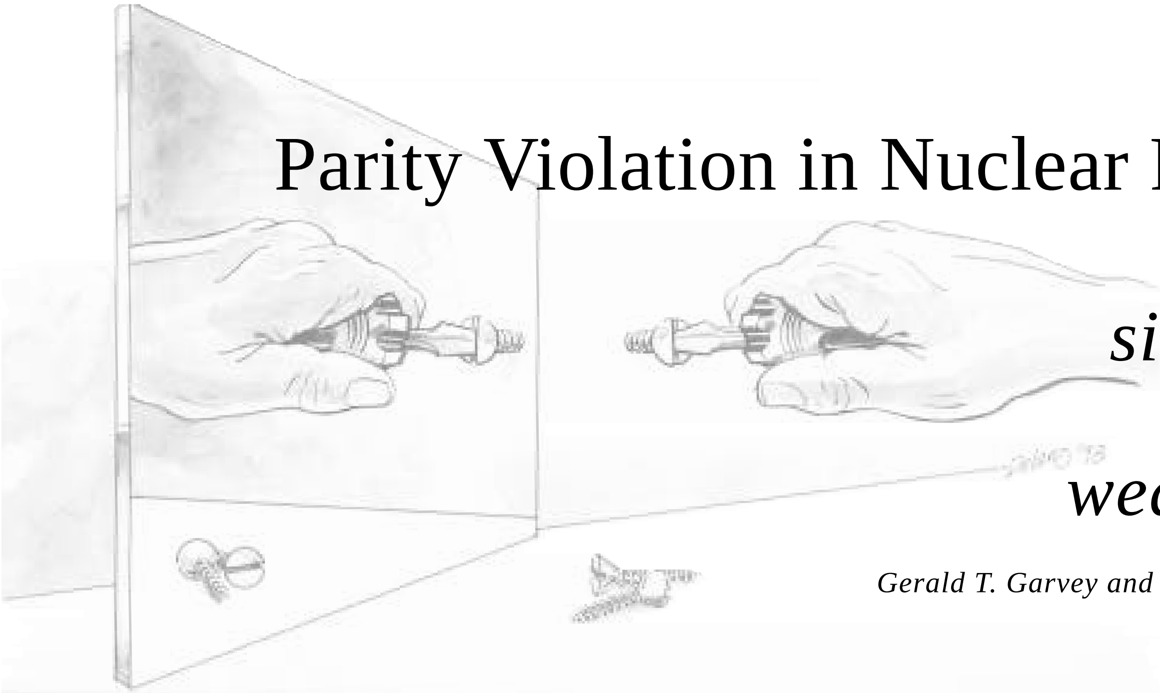
Further Reading

Michael Creutz, editor. 1992. *Quantum Fields on the Computer*. Singapore: World Scientific.

M. Fukugita, Y. Iwasaki, M. Okawa, and A. Ukawa, editors. 1992. *Lattice 91: Proceedings of the International Symposium on Lattice Field Theory, Tsukuba, Japan, 5–9 November 1991.* Amsterdam: North-Holland.

Rajan Gupta was awarded an M.S. by the University of Delhi, India, and a Ph.D. by the California Institute of Technology in 1982. In 1985 he became an Oppenheimer Fellow at the Laboratory and three years later joined the staff of the Elementary Particles and Field Theory Group. Among his scientific interests are the development of non-perturbative methods to solve QCD, using lattice QCD to constrain the standard model, and studying the critical behavior of statistical-mechanics models by means of Monte Carlo renormalization-group methods.

Parity Violation in Nuclear Physics



signature of the weak force

Gerald T. Garvey and Susan J. Seestrom

The elegant and concise description of the physical world contained in the standard model of electromagnetic, weak, and strong interactions (see "Unification of Nature's Fundamental Forces") is supported by a vast body of experimental data. Often only the first or the most precise experiments are cited as providing the requisite body of supporting data. However, the information and techniques developed by the world community of experimental particle and nuclear physicists has provided the broad base on which this powerful model has been constructed.

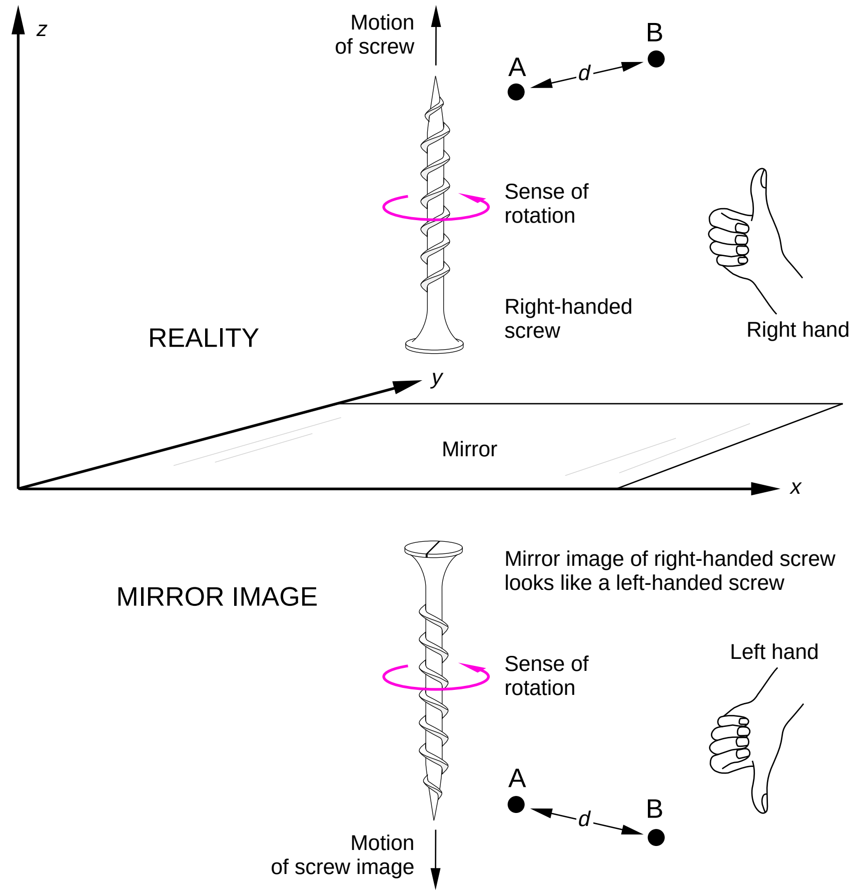
Los Alamos experimentalists, particularly in the Physics and Medium-Energy Physics Divisions, along with university users of the Los Alamos Meson Physics Facility (LAMPF), have played a role in building that base. Many of the major contributions by Laboratory scientists have been in the realm of neutrino physics. For example, they have established the most accurate upper limit on the mass of the electron antineutrino from very careful measurements of tritium beta decay. They also made the first measurements of the scattering of electron neutrinos from electrons, which showed that the interference between the charged and neutral currents of the weak interaction has a negative sign. This result has implications for the well-known solar-neutrino puzzle. The findings of Laboratory scientists support the standard model and have been among the most sensitive tests of its validity. However, since aspects of this research have been reported in previous editions of *Los Alamos Science*, it seems opportune to discuss another area of fundamental research in nuclear and particle physics at the Laboratory. This tale involves the measurement of the strength of the parity-violating interactions between strongly interacting particles—for example, between two neutrons or a neutron and a proton.

Before 1956 physicists believed that all the fundamental interactions in nature would be unchanged by a mirror reflection (or parity inversion). Imagine a basic interaction between two particles described in the orthogonal coordinate system (x, y, and z) as shown in Figure 1. A mirror reflection that inverts the z-axis (z goes to $-z$) results in the configuration shown in the mirror image. The familiar interactions, such as gravity and the interaction between electric charges, depend only on the distance d between the interacting particles, and so a description of those forces is completely unchanged by a parity transformation in which any one of the coordinate axes is inverted, such as in the mirror reflection shown in Figure 1. In the second type of parity transformation, all three axes are simultaneously inverted (x goes to $-x$, y goes to $-y$, and z goes to $-z$), and again the description of the familiar forces remains unchanged. Such interactions are said to be parity-conserving. For a long time physicists thought that all basic interactions must be parity-conserving. But if an interaction depends on the "screw"-like behavior of particles, its description will not be invariant under a parity transformation. Consider the screw in Figure 1; it has right-handed threads, so when it is rotated as shown, it advances in the $+z$ direction (up, in the figure). Its mirror image, however, has left-handed threads. Further, as the screw rotates, its mirror image rotates in the same direction but advances in the $-z$ direction.

Screws can be machined with either right-handed or left-handed threads, so their handedness (the relationship between the direction of rotation, or spin, and the direction of motion) is not an intrinsic property of nature. However, if an elementary particle with intrinsic spin has a fixed handedness (a fixed relationship between its spin direction and its direction of motion), the description of the particle will change under a parity transformation and it is said to violate parity conservation. Likewise if a basic interaction between particles involves only the left-handed or only the right-handed "screw"-like behavior of the particles, the interaction is said to violate parity conservation.

In 1957 it was demonstrated that the interaction responsible for the beta decay of a neutron into a proton, an electron, and a neutrino violates parity conservation. Specifically, when cobalt-60 nuclei, spinning in the same direction around the *z*-axis in the presence of a magnetic field in the +*z* direction, underwent beta decay, they emitted more electrons with a component of momentum in the –*z* direction than in the +*z* direction. This result is not invariant under a mirror reflection, indicating that the interaction responsible for the decay process, called the weak interaction, does not conserve parity. Further, the direction and amount of asymmetry indicated that the weak interaction is left-handed and violates parity in a maximal way. The reason that the weak interaction is left-handed is because the carriers of the weak force (the particles that are exchanged in weak processes), namely the W^+, W^-, and Z^0 bosons, interact with the left-handed component of particles and the right-handed component of antiparticles. (It is interesting to note that although the W^+, W^- , and Z^0 bosons were postulated much earlier to unify the description of the electromagnetic and weak interactions, they were first directly created and observed at the European Center for Nuclear Research (CERN) in 1982.)

Neutrinos and antineutrinos are particles that participate in, as far as we know, only weak interactions. Thus, it is not a surprise that these



$F_{\text{gravity}} \propto \frac{1}{d^2}$ and is therefore unaffected by mirror reflection.

$F_{\text{electrostatic}} \propto \frac{1}{d^2}$ and is therefore unaffected by mirror reflection.

Figure 1. Effects of a Mirror Reflection
A mirror reflection is one type of parity transformation. In the figure the reflection inverts the *z* axis. The distance between points A and B is unchanged in the mirror image, so that the descriptions of the gravitational and electrostatic forces between, say, two electrons located at A and B would would also be unchanged. In contrast, the mirror image of the right-handed screw is a left-handed screw. When turned in the direction indicated by the red arrows, the screw advances in the +*z* direction whereas its mirror image advances in the °*z* direction. Note that if you curl the fingers of your right hand along the red arrow, your thumb points up, in the direction of motion of the right-handed screw. Alternatively, if you curl the fingers of your left hand along the red arrow, your thumb points down, in the direction of motion of the screw's left-handed mirror image. Forces that depend on the relationship of spin rotation to direction of motion violate parity conservation.

massless particles with intrinsic spin show a fixed handedness. A neutrino always appears to spin clockwise when it is coming toward the observer and is therefore a left-handed particle, whereas the antineutrino appears to spin counterclockwise when it is coming toward the observer and is therefore a right-handed antiparticle.

The fact that the neutron decays via the weak interaction into a proton plus an electron and antineutrino means that the neutron and proton, which are known to interact through the strong force, must also interact through the weak force; otherwise they could not be involved in weak decay processes. One can therefore ask how the weak force affects the interaction between two nucleons (the neutron and the proton look the same to the strong force and are both called nucleons). The strong

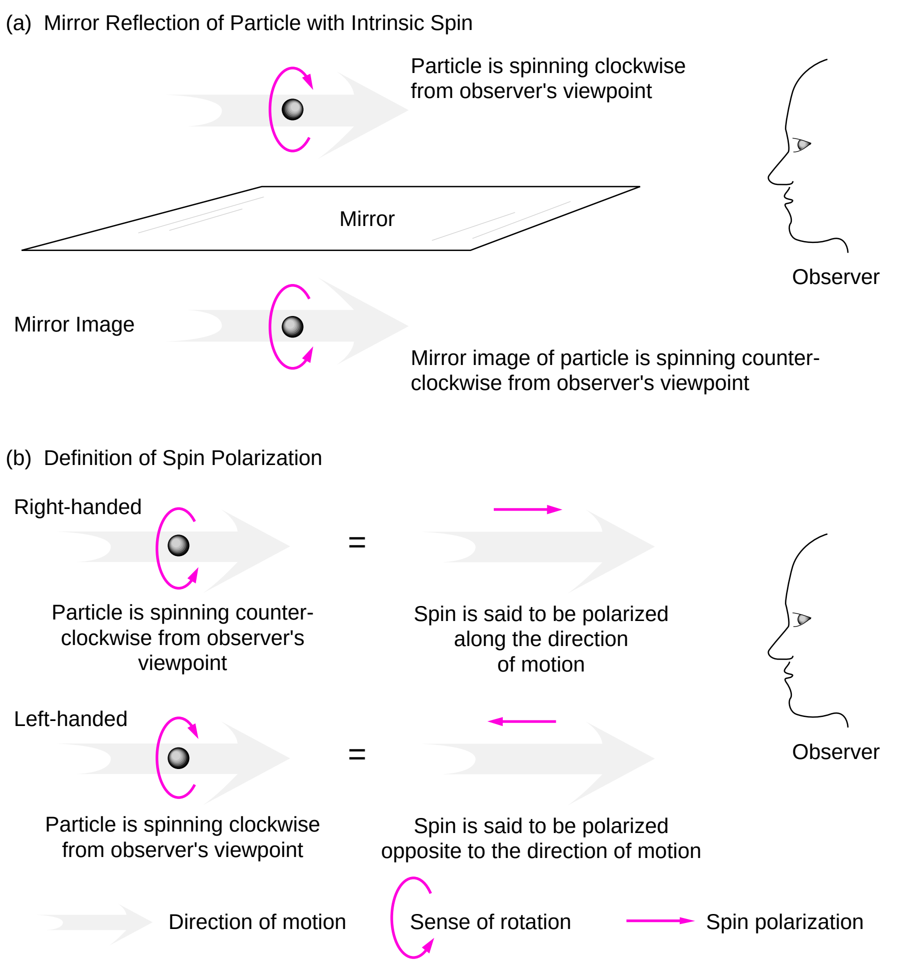


Figure 2. Mirror Reflection of a Particle with Intrinsic Spin
(a) A proton is moving toward the observer and is spinning around the direction of motion as indicated by the red arrow. The mirror image of the proton is also shown. To the observer, the proton spins clockwise, whereas the mirror image of the proton spins counterclockwise as both are moving toward the observer. Therefore if parity is conserved, the probability that a proton is scattered by a target should be independent of its spin direction, provided that the target nuclei are spinning in random directions. (b) The direction of spin is often represented by a vector that is along the axis of spin rotation. Here the axis of spin rotation of a proton is parallel to its direction of motion. According to convention, when the relation between the rotation and the motion is like that of a right-handed screw, the spin vector points in the same direction as the direction of motion and the proton is said to have its spin polarized along the direction of motion. When the relation between the rotation and the motion is like that of a left-handed screw, the spin vector points opposite to the direction of motion and the proton is said to have its spin polarized opposite to the direction of motion.

force dominates the interaction between two nucleons; specifically, the ratio of the strength of the weak interaction to that of the strong interaction is about 1 to 10 million. Typically the effects of such a small interaction would be next to impossible to detect, but the weak interaction has a unique signature in that it is the only interaction in the standard model that violates parity. Hence measurement of the amount of parity violation in a given process is a direct measure of the role played by the weak interaction in that process.

As was mentioned earlier, parity violation was discovered in 1957, but it was not until seven years later that the first clear parity-violating effect was measured in processes other than weak decays of nuclei. In 1964 a group headed by Yuri G. Abov in the then Soviet Union observed parity violation in the capture of polarized neutrons by the nucleus cadmium-113. The gamma rays emitted following the neutron capture were emitted preferentially in the direction of the neutron polarization, which indicated that parity was violated. Thus the weak interaction between the nucleons within the nucleus was producing measurable effects. Unfortunately the complexity of the relative motions of the nucleons in the cadmium nucleus made it impossible to determine the strength of the weak interaction between pairs of nucleons from that experiment.

In 1970 a Los Alamos group led by Hans Frauenfelder, Dick Mishke, and Darrah Nagle began investigating parity violation in the scattering of protons from protons. For their first experiments they used the polarized ion source installed by Joe McKibben in the tandem Van de Graaff accelerator. The polarized ion source and its subsequent versions were essential to the experiment because they produce a beam of protons all of which are spinning in the same direction.

The proton has intrinsic spin but no intrinsic handedness, so its spin direction can be changed relative to its direction of motion. Figure 2 illustrates that the ability to manipulate the proton's spin direction in a known and controlled way is valuable in investigating the degree of parity violation in scattering two protons from one another. Figure 2a depicts a fast-moving proton such as would be found in a proton beam from an accelerator, as well as its mirror image. The proton is moving to the right and appears to be spinning clockwise to the observer at right. (It is behaving like a left-handed screw.) In the mirror image the proton is again moving to the right but appears to be spinning counterclockwise to the observer. (It is behaving like a right-handed screw.) Thus, if the principle of parity conservation applies, protons rotating clockwise or counterclockwise relative to their direction of motion (or, as defined in Figure 2b, with their spins polarized either along or opposite the direction of motion) should be scattered identically from a target composed of protons that are spinning in random directions.

A container of hydrogen provides a suitable target because the average spin of the protons (hydrogen nuclei) in the target is zero. In the Van de Graaff experiment protons polarized along the direction of motion were scattered from the target, and then protons polarized in the opposite direction were scattered. The total scattering cross section, or probability of scattering, was measured in each case. The Van de Graaff experiment began in the early seventies and was not concluded until the end of the decade. It was the first scattering experiment anywhere in the world in which parity violation was observed. Protons with spins polarized along their direction of motion were scattered slightly more often than those with opposite polarization. The reason the experiment took so long to carry out was that the measured asymmetry between the two neutron polarizations was very small. The protons scatter mainly as a result of strong interactions, so the difference (resulting from the weak interactions) between the fraction of particles scattered for two different spin polarizations was only 2 parts in 100 million. Thus a variety of new techniques had to be invented to make the measurement possible. After the Van de Graaff experiment was completed, the research group carried out a further measurement of parity violation using a much-higher-energy polarized proton beam available at LAMPF. The observed effects

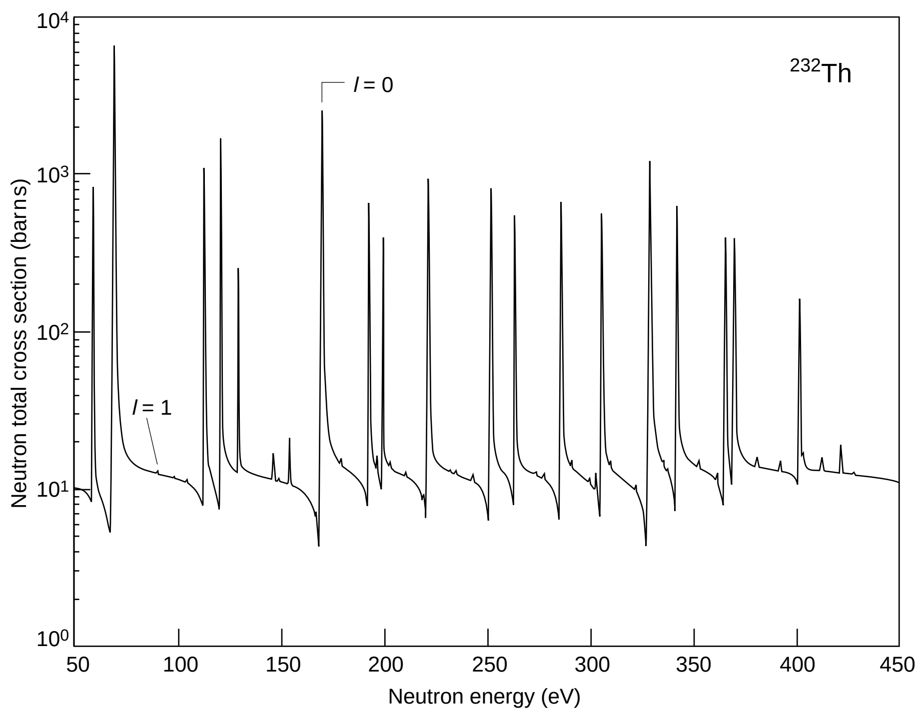
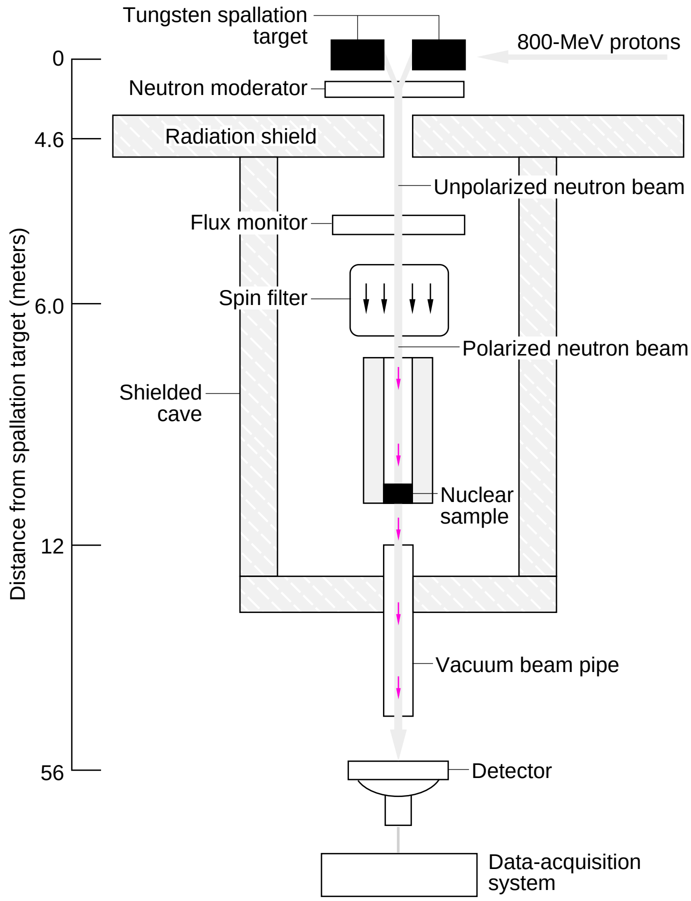


Figure 3. Total Cross Section for Scattering and Absorption of Neutrons by ^{232}Th

The total cross section (or probability) for the interaction of neutrons with ^{232}Th is plotted as a function of neutron energy. The many sharp peaks in the cross section are called resonances and occur when the neutron energy equals the energy of an excited state of the compound nucleus ^{233}Th and can therefore be absorbed by ^{232}Th. The tall peaks occur at energies of nuclear states with orbital angular momentum equal to zero ($l = 0$) and the small peaks (lower by two orders of magnitude) occur at energies of nuclear states with $l = 1$. Both types of resonances can be studied with great sensitivity for the parity-violation effects, which are expected for $l = 1$ but not for $l = 0$ resonances.

Figure 4. Setup for Parity-Violation Experiment at LANSCE
Neutrons are produced by interaction of 800-MeV protons with a split tungsten target shown at the top of the figure. The energies of the neutrons so produced range from almost zero to nearly 800 MeV. The neutrons pass first through a moderator that reduces the energy of the neutrons to the eV or keV range. Because the neutrons are produced in pulses and because the time required to produce and moderate the neutrons is small compared to their time of flight to a detector 56 meters away, the energy of each detected neutron can be measured from its measured time of flight. The beam of moderated neutrons passes through a spin filter—a material in which the proton spins have been aligned in the same direction as the direction of motion of the neutrons (large red arrows). Those neutrons whose spin directions are opposite that of the protons in the spin filter are absorbed or scattered out of the beam. The neutrons with the same spin direction as the protons interact more weakly with the protons and remain in the beam. The neutron beam emerging from the spin filter contains neutrons with only one spin direction (small red arrows) rather than both and is thus polarized. As the polarized neutron beam passes through the sample, its intensity is reduced as neutrons are absorbed by or scattered from the nuclei in the sample. A detector measures the number and the times of arrival of the neutrons that are transmitted through the sample. The polarization of the neutron beam can be reversed (by reversing the polarization of the protons in the spin filter) and the experiment repeated. If the measured fraction of neutrons transmitted through the sample at a given resonance energy is different for one neutron polarization than for the other, then that resonance exhibits parity violation.



were only somewhat larger, at the level of a few parts in 10 million, which is near the expected value at that energy. However, because the parity-violating effect was so small, experimental errors were about the same size as the effect and so the precise strength of the weak force between two nucleons could not be determined.

In the meantime research groups in the Soviet Union were reporting parity-violating effects one million times larger in the absorption of very-low-energy polarized neutrons by certain heavy nuclei. Neutrons carry the same amount of intrinsic spin as protons do and, like the proton, their spins can be polarized along or opposite to the direction of motion (see Figure 2). The cross sections differed depending on the polarization of the incident neutrons, indicating parity violation. It was again, however, impossible to deduce the strength of the weak interaction between two nucleons from the observed degree of parity violation in these experiments because even though the effect of the weak force was amplified many times by nuclear motions, the amount of amplification could not be quantified.

Figure 3 shows the measured probability of a low-energy neutron interacting with a thorium-232 nucleus as a function of energy of the incident neutron. The series of large bumps evident in the data appear at the energies of quantum states of the compound thorium-233 nucleus. When the energy of the incident neutron corresponds to the energy of one of these states, the neutron is said to be at a resonance and the incident neutron can readily share its energy with the neutrons and protons in the thorium target nuclei. In other words, at a resonance there is a large probability for the neutron to be absorbed into a thorium nucleus. At other energies the neutron is much less likely to interact, and when it does interact, it is simply deflected from the ^{232}Th nucleus without sharing its energy. If the probability of absorption at a resonance depends on the spin polarization of the incident neutron, then the resonance process exhibits parity violation.

The members of the Triple Collaboration* realized that if parity violation could be measured at several resonances of the same nucleus, one could determine an average value of the magnitude of the parity violation that would be independent of the statistical properties of the nuclear motions and therefore could be used to estimate the strength of the underlying nucleon-nucleon weak interaction. As indicated above, the effect of parity violation on nucleon-nucleon interactions is very small. However, the effect is amplified by a factor of about 1 million by the motion of many nucleons in the nucleus to yield the large parity violations observed for the scattering of a neutron by a nucleus at a resonance. The size of the amplification depends on statistical properties of the quantum states of the nucleus and therefore the amplification has a random distribution. By averaging the parity-violating effects observed at many resonances, the average value of the amplification can be determined. Our theoretical models of the nucleus are sufficiently detailed to deduce from the observed average amplification in a nucleus containing many nucleons a fairly good estimate of the strength of the weak interaction between two nucleons.

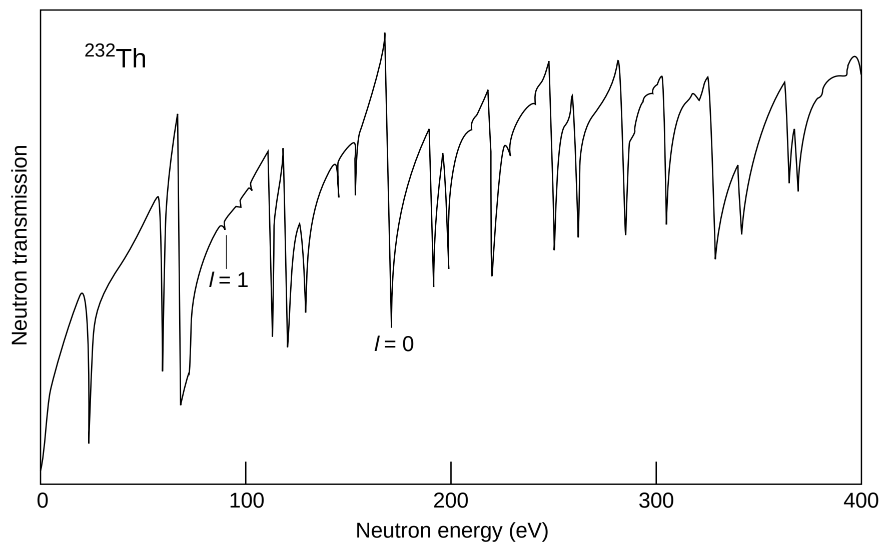


Figure 5. Neutron Transmission Spectrum of ^{232}Th
Shown here are the measured values for the fraction of neutrons transmitted through a ^{232}Th sample as a function of neutron energy. The fraction of neutrons transmitted decreases sharply when the neutron energy is at a resonance of ^{232}Th. That is, the neutrons can be absorbed by ^{232}Th to form an excited state of the compound nucleus ^{233}Th. The large dips are at the energies of the l = 0 states of the compound nucleus ^{233}Th; these resonances do not exhibit parity violation. Smaller dips, such as the three between 80 and 110 eV, are at the energies of the l = 1 states of the compound nucleus ^{233}Th, which can exhibit parity violation and are therefore of interest in these experiments.

The Triple Collaboration also realized that the Los Alamos Neutron Scattering Center (LANSCE) was an ideal facility at which to carry out the necessary measurements. One of the principal advantages of the LANSCE neutron source over reactor neutron sources is that the range of available neutron energies is much greater. A typical reactor neutron source has a very limited flux of neutrons with energies above 10 eV, whereas Figure 3 shows that measuring the parity violation at a number of resonances of ^{232}Th requires the availability of neutrons at energies ranging up to several hundred eV.

Since the features of the LANSCE facility are essential to the successful experimental program undertaken by the Triple Collaboration, we describe it briefly. The major elements of the facility are LAMPF, which provides 500- to 700-microsecond-long trains

*The Triple Collaboration is a collaboration between Los Alamos National Laboratory, North Carolina State University, Duke University, Triangle Universities Nuclear Laboratory, TRIUMF (Canada), University of Technology at Delft (The Netherlands), KEK (Japan), and the Joint Institute for Nuclear Research (Russia). The collaboration was formed to study fundamental symmetries using polarized neutrons at LANSCE.

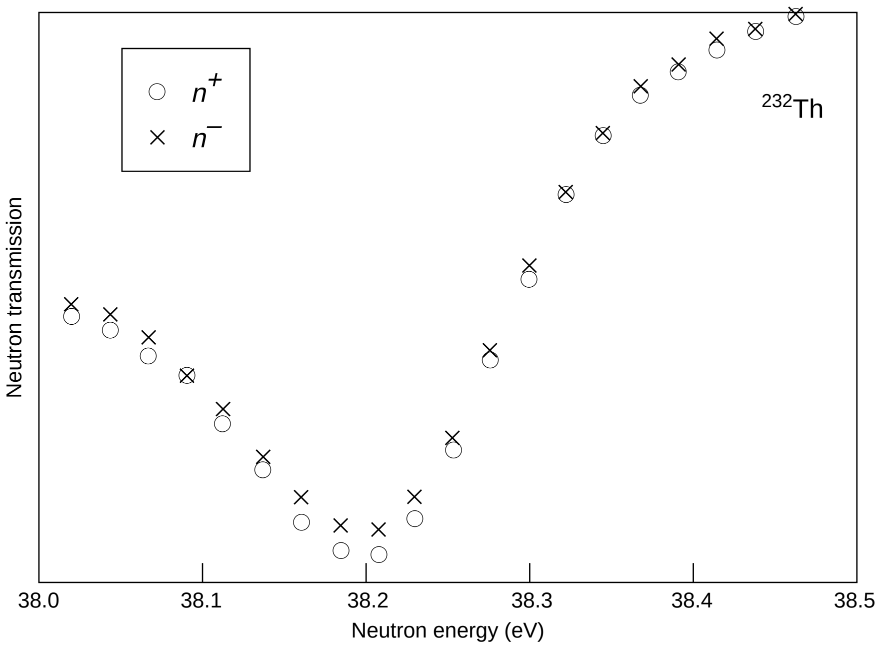


Figure 6. Parity Violation in a Neutron Resonance of ^{232}Th

Transmission data for two different neutron polarizations are shown near the $l = 1$, $J = \frac{1}{2}^{-}$ resonance of ^{232}Th at 38.2 eV. Results for neutrons polarized along the direction of motion (n^{+}) are designated by circles. Results for neutrons of the opposite polarization (n^{-}) are designated by crosses. The circles and crosses are close together over most of the energy range shown. At the resonance (the dip in the transmission spectrum) the circles clearly fall below the crosses; that is, more neutrons polarized along the direction of motion are absorbed by ^{232}Th than neutrons polarized opposite to the direction of motion. Thus this resonance exhibits parity violation.

of 250-nanosecond pulses of 800-MeV H^{-} ions; the Proton Storage Ring, which combines the many short H^{-} pulses in each train into a single, intense 250-nanosecond proton pulse; and finally the LANSCE spallation source, in which the incredibly intense pulse of protons is converted to neutrons through the process of spallation. That is, when the 800-MeV protons impact the tungsten target, each proton liberates about 20 neutrons from the neutron-rich tungsten nuclei. As a result each intense pulse from the storage ring creates about 10^{13} neutrons all within a quarter of a microsecond. This time is very short compared to the time it takes the neutrons to travel to a detector some 50 meters away from the area where they are produced. Therefore the measurement of a neutron's time of flight over the known distance from source to detector gives a direct measurement of the neutron's speed and hence its energy. Thus LANSCE not only generates large numbers of neutrons over the necessary range of energies but also produces neutrons whose energies can be accurately measured. Only one requirement for studying parity violation is missing: the neutrons coming from the spallation source do not have a definite polarization, or spin direction.

Figure 4 shows a schematic setup of the experiment and indicates how a beam of polarized neutrons is produced. The neutrons produced in the tungsten target are not polarized; that is, the spin of each has an equal probability of pointing along or opposite its direction of motion. They are passed through a spin filter, a special material in which the protons in water (or hydrocarbon) molecules are polarized along the direction of motion of the incoming neutron beam. The neutrons with spins polarized opposite to those of the protons in the spin filter interact most strongly with those protons and are therefore scattered out of the beam. Thus the neutrons that pass through the spin filter are those with spins polarized along the direction of motion. In this way a polarized neutron beam is produced. By changing the direction of proton polarization in the spin filter, neutrons can be polarized either along or opposite their direction of motion.

A beam of neutrons polarized in one direction is passed through a particular nuclear sample (for example, ^{232}Th), and the experimenters measure the fraction that are transmitted through the sample. The experiment is repeated with neutrons polarized in the opposite direction. If parity is conserved, the fraction transmitted through the target sample would be the same for both experiments. Figure 5 shows neutron transmission through a ^{232}Th sample as a function of energy. The neutron transmission is reduced when the neutrons have the same energy as a resonance because then the probability of neutron absorption by the nuclei is greatest. Figures 6 and 7 show examples of the data obtained for the transmission of neutrons with opposite polarizations through a ^{232}Th target and a ^{238}U target, respectively. Figure 6 shows the transmission for the two neutron polarizations for a $J = \frac{1}{2}^-$, l = 1 resonance of ^{232}Th at 38.2 eV, where J is the total angular momentum of the resonance, l is the orbital angular momentum, and the minus sign denotes the parity of the resonance. (The parity of a resonance can be + or –, depending on whether the wave function of the resonance remains unchanged or reverses sign under a parity inversion. Because the sign of the wave function is unmeasurable, negative-parity wave functions do not violate parity conservation.) The top graph in Figure 7 is a plot of the asymmetry parameter ε as a function of neutron energy for transmission through ^{238}U. The asymmetry ε is a measure of the difference in transmission for the two neutron polarizations and makes small differences easier to detect. The the figure shows a small asymmetry for the $J = \frac{1}{2}^-$, l = 1 resonance of ^{238}U at 64 eV and no asymmetry at the large $J = \frac{1}{2}^+$, l = 0 resonance at 66 eV. This dependence of neutron transmission on the total angular momentum of the resonance and its parity is a well-understood feature of the *strong* force, but beyond the scope of this article to explain.

The first experiment by the Triple Collaboration was performed on uranium-238. Seventeen resonances were examined and five showed measurable asymmetries indicating parity violation. The experiment on ^{238}U was the first in which a research team had seen parity violation in more than one resonance of a single nucleus. A later experiment on the isotope ^{232}Th (whose resonances are depicted in Figures 3, 5, and 6) studied twenty-three resonances of which seven had measurable asymmetries ranging between 1 and 10 percent. This data sample is sufficiently large that a value can be extracted for the average strength of the weak interaction between a single nucleon and all the nucleons in the ^{232}Th nucleus. This result is serving to refine our knowledge of the weak interaction between two nucleons and should prove far more useful as experimental techniques are improved and additional data are taken. The weak interaction between nucleons will eventually be understood in terms of

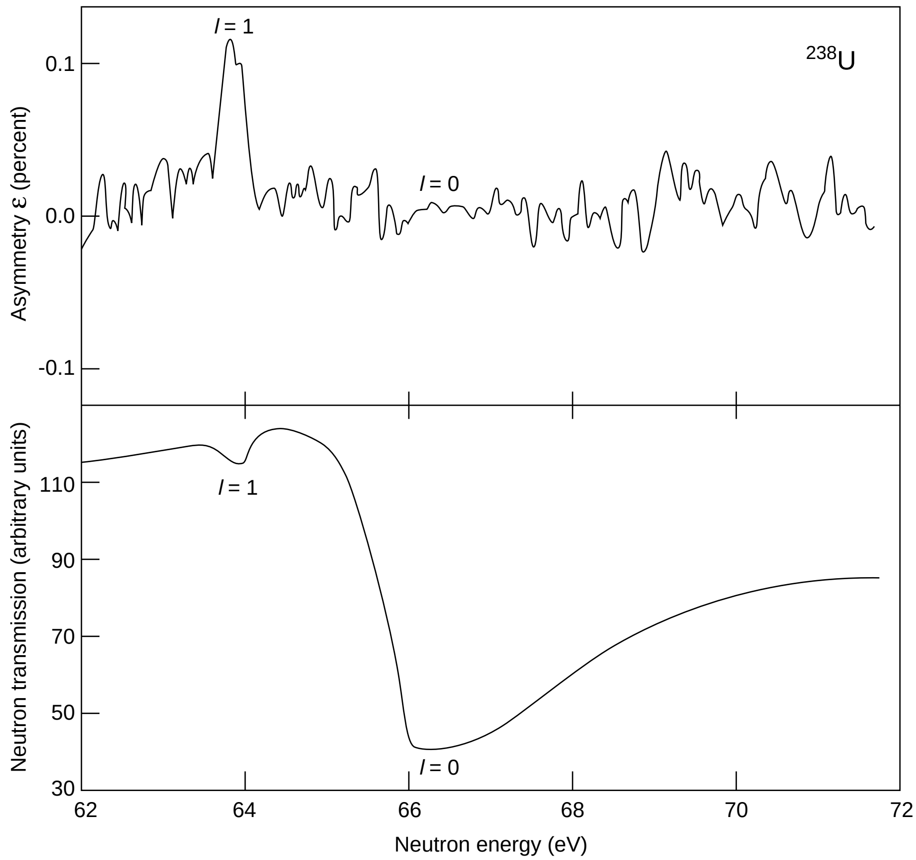


Figure 7. Parity Violation in a Neutron Resonance of ^{238}U

The parity-violating effect at the 64-eV, l = 1 resonance of ^{238}U is smaller than that shown in Figure 6 and is more easily detected by plotting the results for the two neutron polarizations in terms of the asymmetry parameter $\varepsilon \equiv (T^+ - T^-)/(T^+ + T^-)$, where T^+ is the transmission for neutrons polarized along the direction of motion and T^- is the transmission for neutrons polarized opposite to the direction of motion. The asymmetry parameter ε plotted in the top graph has statistical fluctuations over the energy range shown except at the 64-eV resonance where the asymmetry between the two neutron polarizations is 0.1 percent. The sum of the transmissions for both polarizations ($T^+ + T^-$) is shown in the bottom graph. The parity-violating resonance at l = 1 appears as a small dip in transmission at 64 eV, whereas an l = 0 resonance at 66 eV appears as a large dip in transmission. As expected the plot of the asymmetry parameter ε shows no asymmetry at the energy of the l = 0 resonance.

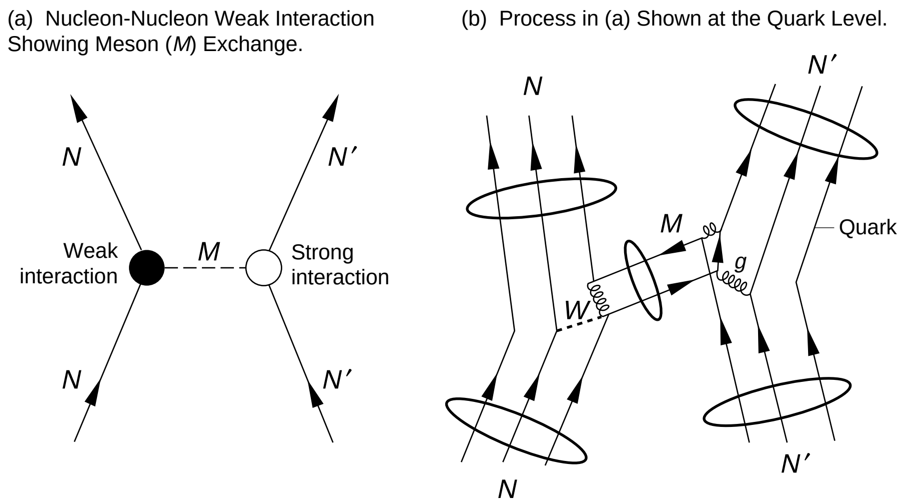


Figure 8. The Nucleon-Nucleon Weak Interaction at the Quark Level
(a) The weak interaction between nucleons N and N' is often described as an exchange of the meson M, in which the open circle is a strong interaction meson-nucleon vertex, and the solid circle is a weak interaction meson-nucleon vertex. (b) This cartoon of the weak interaction between nucleons N and N' shows the possible interactions that might take place involving the three quarks composing each nucleon and the quark-antiquark pair composing the meson. The dotted line is a W boson, which carries the weak force. The spiral lines are gluons, the carriers of the strong force.

interactions among the quarks composing the nucleons and the carriers of the weak force (Figure 8).

One feature of the observed results is difficult to understand. There seems to be a mysterious preference for a positive sign to the asymmetry. That is, neutrons with spins polarized along the direction of motion tend to be scattered more readily than neutrons with opposite polarization. In fact, in the case of ^{232}Th, all seven observed asymmetries showed this preference. Several papers dealing with this issue have been published, but as yet no satisfactory explanation has emerged.

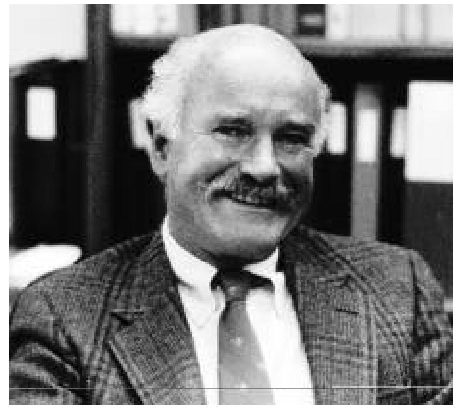
It is clear that the facilities and personnel at Los Alamos in conjunction with the world scientific community continue to contribute to the store of scientific knowledge that will represent one of the great legacies of the last half of this century. ■

Further Reading

C. M. Frankle, J. D. Bowman, J. E. Bush, P. P. J. Delheij, C. R. Gould, D. G. Haase, J. N. Knudson, G. E. Mitchell, S. Penttilä, H. Postma, N. R. Roberson, S. J. Seestrom, J. J. Szymanski, S. H. Yoo, V. W. Yuan, and X. Zhu. 1992. Parity nonconservation for neutron resonances in ^{232}Th. *Physical Review C* 46: 778.

J. D. Bowman, C. D. Bowman, J. E. Bush, P. P. J. Delheij, C. M. Frankle, C. R. Gould, D. G. Haase, J. Knudson, G. E. Mitchell, S. Penttilä, H. Postma, N. R. Roberson, S. J. Seestrom, J. J. Szymanski, V. W. Yuan, and X. Zhu. 1990. Parity nonconservation for neutron resonances in ^{238}U. *Physical Review Letters* 65: 1192.

V. W. Yuan, C. D. Bowman, J. D. Bowman, J. E. Bush, P. P. J. Delheij, C. M. Frankle, C. R. Gould, D. G. Haase, J. Knudson, G. E. Mitchell, S. Penttilä, H. Postma, N. R. Roberson, S. J. Seestrom, J. J. Szymanski, and X. Zhu. 1991. Parity nonconservation in polarized-neutron transmission through ^{139}La. *Physical Review C* 44: 2187.

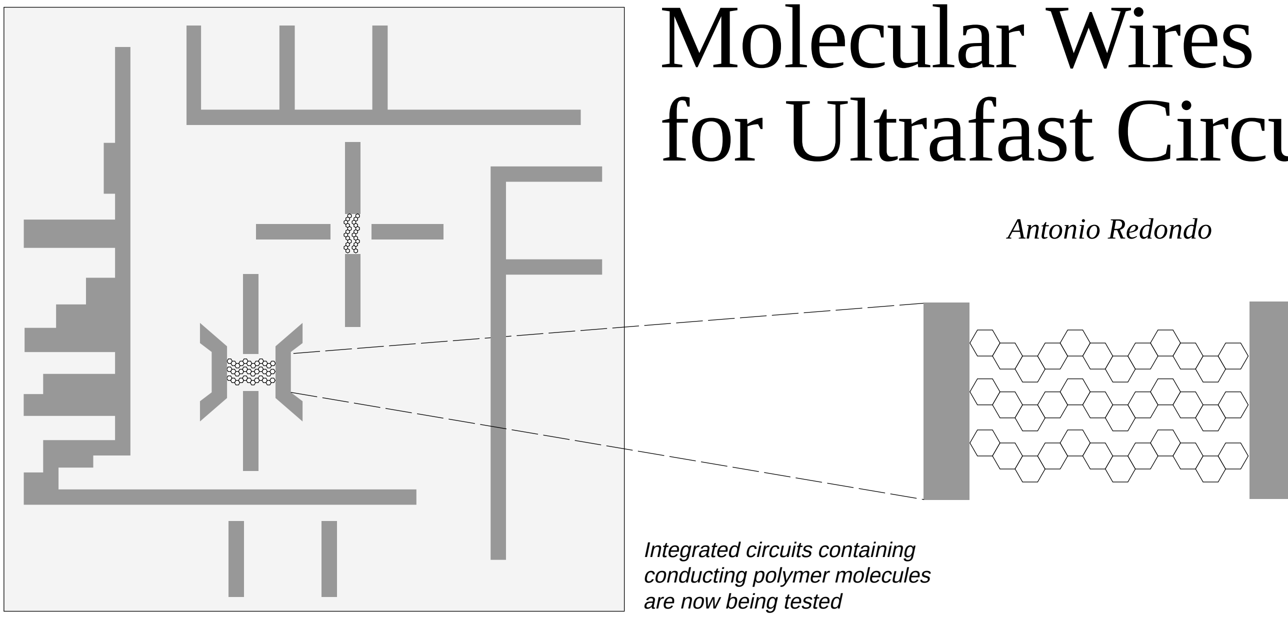


Gerald T. Garvey, currently a Laboratory Senior Fellow, was director of LAMPF from November 1985 to August 1990. Among his past positions are directorship of Argonne National Laboratory's Physics Division and professorships at the Universtiy of Chicago and Princeton University. He is an American Physical Society Fellow, a recipient of a Humboldt Award, and editor of *Comments on Nuclear and Particle Physics*. He received a B.S. from Fairfield University in 1956 and a Ph.D. from Yale Univeristy in 1962.

Susan J. Seestrom is a staff member in the Neutron and Nuclear Science Group of the Laboratory's Physics Division, where she has been studying fundamental and applied nuclear physics with neutrons since 1986. She received her B.S. and Ph.D. degrees from the University of Minnesota. She was a Director's Postdoctoral Fellow at Los Alamos from 1981 to 1983 and a Research Associate at the University of Minnesota from 1983 to 1986, working in medium-energy nuclear physics using pions and protons as probes of nuclear structure and reaction dynamics.

Molecular Wires for Ultrafast Circuits

Antonio Redondo



Integrated circuits containing conducting polymer molecules are now being tested

Electronic circuitry has undergone continuous miniaturization since transistors replaced vacuum tubes in the 1960s. A dominant force driving the miniaturization is the incessant demand for faster processing of ever-increasing amounts of information. Shown in Figure 1 are two measures of the rapid pace of miniaturization. The minimum dimension of devices that can be fabricated on a semiconducting chip (such as transistors, diodes, and switches) has decreased over the last twenty-five years from about 16 microns (1 micron = 10^{-6} meter) to about 0.5 micron, and the number of devices that can be included on a single chip has increased from 1 in 1958 to over 10 million. The remarkable increase in the scale of "integration" is due not only to development of techniques for producing smaller devices but also to improvements in circuit design.

The past rates of integration would seem to imply that gigascale integration (1 billion components per chip) may be reached by the year 2000. However, such an extrapolation ignores limitations on the level of integration achievable with current semiconductor technology. Many of the limitations arise from inescapable physical laws and properties of matter. For example, as a device is made smaller, the electric field produced by a fixed voltage applied across its ends becomes larger. But the electric field cannot be allowed to exceed the value E_{max} at which the material composing the device suffers electrical breakdown. The electric field can be kept below E_{max} by reducing the voltage applied to the device. But that approach faces another fundamental limit. The voltage cannot be reduced below a certain minimum imposed by the thermal energy of the electrons in the device, which is given by kT, where k is Boltzmann's constant and T is the temperature in kelvins. In particular, if an applied voltage is to produce a detectable response (an increase in current flow, say, or an electronic transition resulting in emission of light), the minimum applied voltage V_{min} must be several times larger than the so-called thermal voltage given by kT/e, where e is the electronic charge. At room temperature kT/e is about 0.025 volt.

The limits E_{max} and V_{min} are universal, applying to all materials. Another limit—one applicable to semiconductors and other solid materials—was expounded in the mid 1960s by Edward O. Johnson. He argued that there exists a maximum speed v_{max} at which an electron can travel a distance L in a solid and that therefore the travel time t must be greater than L/v_{max}. The existence of a maximum speed is due to the inevitable presence in any real solid of thermal excitations and lattice defects such as impurity atoms, vacancies, and randomness in the locations of the atoms that constitute the material. Defects in crystalline semiconductors are not only inevitable but also, in fact, deliberately introduced to achieve desirable properties. Collisions between such defects and an electron moving through a semiconductor cause the electron to lose kinetic energy and to undergo random changes in the direction of its motion. (Such collisions are

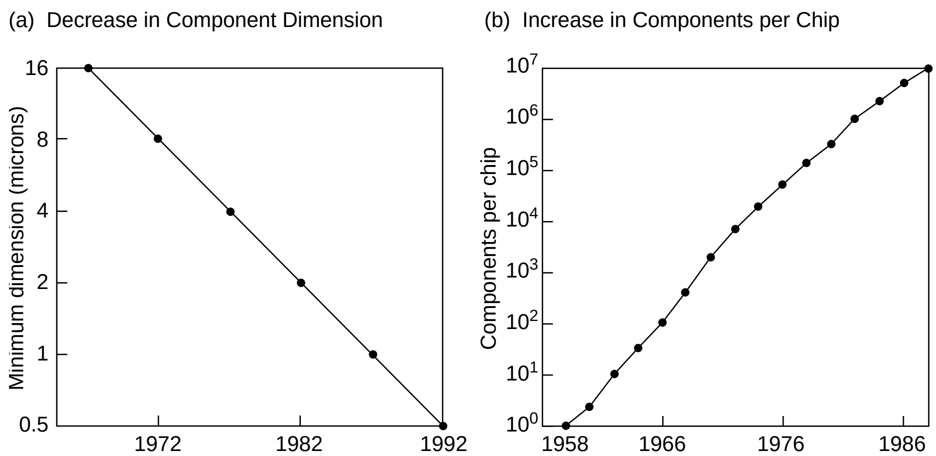


Figure 1. Measures of Electronic Miniaturization
(a) Continuing developments in photolithographic techniques have brought about an exponential decrease in the size of electronic components that can be fabricated on chips. (b) The decrease in component dimensions and improvements in circuit design have led to amazing increases in the number of components that can be included on a single chip. Although the rate of increase has decreased somewhat from an exponential rise during the 1960s, the number of components per chip has now reached 10 million.

the main cause of the electrical resistance of solid materials; the energy lost by the electrons appears in the form of heat.) If an electric field in a given direction is created in the semiconductor (by applying a voltage of fixed value across the semiconductor), then, between collisions, the electron is subject to a force in the direction of the electric field and thereby gains kinetic energy. Over time the electron moves as if it possessed a zero average velocity perpendicular to the field and a nonzero velocity component, a "drift velocity," in the direction of the field. Thus the electron executes a not-quite-random walk through the semiconductor. The drift velocity, or v_{max}, of electrons in a semiconductor increases with the applied voltage and also varies with the composition of the semiconductor, its temperature, and the number and nature of the defects it contains. When 1 volt is applied across 1 millimeter of a typical silicon crystal, electrons traverse the 1 millimeter with a drift velocity of about 100 meters per second and thus in a time of about 10 microseconds. (In contrast, if a crystal with no lattice defects were available and if 1 volt were applied across 1 millimeter of that "ideal" crystal, electrons would traverse the 1 millimeter at an average speed of about 6000 meters per second.)

Since the existence of a v_{max} for a semiconducting material is an inescapable fact, the relation $t = L/v$ implies that the only way to reduce t below the value given by L/v_{max} is to reduce L, the dimension of the device.

* It should be noted that another solution to the problem of increasing the speed of electron transfer along interconnects may lie in the existence of semiconducting devices in which, during sufficiently short times and over sufficiently short lengths, electron motion is not impeded by collisions and is instead "ballistic."

As mentioned above, however, the process of miniaturization can be carried only so far. In particular, since the electric field E created by applying a voltage V across a length L of semiconducting material is given by $E = V/L$ and since E cannot exceed the breakdown field E_{max} nor can V be less than V_{min}, there exists a minimum L given by V_{min}/E_{max}. Hence the travel time t must be greater than $V_{min}/E_{max}v_{max}$.

How then can the speed of electronic circuits be increased beyond that dictated by the fundamental limit given above? One approach,* and the subject of this article, focuses (initially at least) on the connections along which electrons are transferred from one device to another. Today the "interconnects" on chips are made of semiconducting or metallic materials along which the speed of electron motion is limited by the drift velocity. Perhaps, though, the nearly random walk of electrons in such conventional interconnects can be replaced by a different, more rapid type of motion, one based on quantum-mechanical effects. The approach we are pursuing hinges on the fact that delocalized states are accessible to electrons in certain polymeric molecules.

The molecules in question belong to a class of organic compounds that are said to be conjugated. One of the simplest conjugated organic molecules, benzene, is of course not a polymer and therefore not a candidate for interconnects. However, consideration of that familiar compound can help define the terms "conjugated organic molecule" and "delocalized electron." As shown in Figure 2, the conventional representation of benzene—its Kekulé structure—includes a pathway of alternating single and double carbon-carbon bonds. Such a Kekulé structure suggests that one pair of electrons is confined, or localized, between carbon atoms 1 and 2;

two pairs are localized between carbon atoms 2 and 3; one pair is localized between carbon atoms 3 and 4; and so on. However, the Kekulé structure of benzene (or of any other molecule) is a simplification and hence leads to an idealized picture of where the nine electron pairs that bind its carbon atoms are likely to be found. An alternate picture is that each electron pair has the same probability of being found at any point around the benzene ring as does any other pair. Such electrons are said to be delocalized (in this case over the entire ring), and electron delocalization is a characteristic feature of all conjugated organic molecules.

Also shown in Figure 2 are the Kekulé structures of two conjugated organic polymers, polyacetylene and polythiophene. Like the Kekulé structure of benzene, the Kekulé structure of any conjugated organic polymer includes a sequence of alternating single and double carbon-carbon bonds, a "backbone." Furthermore, like the pairs of electrons that bind the carbon atoms of benzene, the electron pairs that bind the backbone of a conjugated organic polymer are delocalized, in some instances over extensive regions of the molecule.

Polyacetylene is an example of a conjugated organic polymer with a straight-chain backbone. Although the conductivity of very pure polyacetylene approaches that of copper, polyacetylene is not a candidate for interconnects because, like other straight-chain conjugated polymers, it oxidizes easily in air. Polythiophene is an example of a cyclic conjugated polymer. A number of such polymers, although they exhibit lower conductivities than that of polyacetylene, are extremely stable in air and are therefore possible candidates for interconnects.

Cyclic conjugated organic polymers may offer several advantages over semiconducting materials as interconnects. First, the voltage that must be applied across such a "molecular wire" to bring about the transfer of electrons from one end to the other may be very low. (Like any material, the polymer cannot be subjected to an electric field greater than its breakdown field.) Second, the number of electrons that need to be transported along the polymer in order for it to function as an interconnect may be very low (possibly one or a few at a time). If so, the power consumption of circuits containing molecular wires can be reduced, together with the concomitant problems associated with heat dissipation. Finally, the transfer of electrons along a molecular wire may be very fast. An electron "moves" along a conjugated organic polymer primarily by undergoing transitions from one delocalized electronic state to another. Whereas times for the motion of electrons in the fastest semiconducting interconnects are on the order of picoseconds (1 picosecond = 10^{-12} second), times for electronic transitions are on the order of femtoseconds (1 femtosecond = 10^{-15} second). The optimism inspired by a simple comparison of those times must be tempered, however, by the fact that electronic transitions, though intrinsically fast,

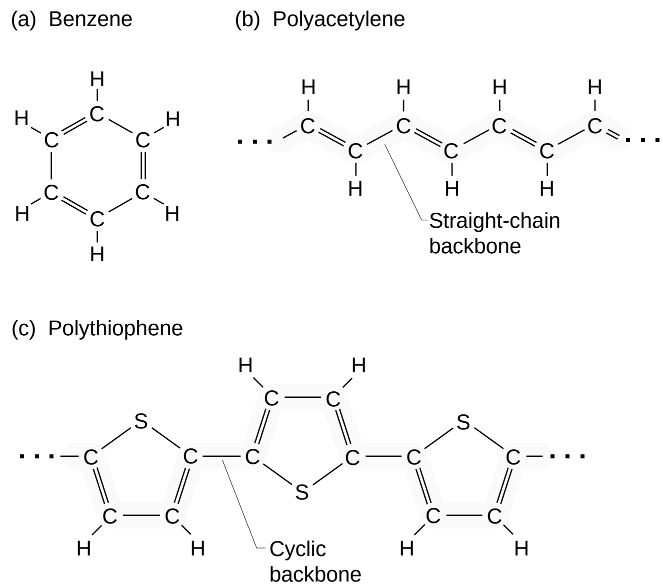


Figure 2. Kekulé Structures of Some Conjugated Organic Molecules
(a) The alternation of single and double carbon-carbon bonds in the Kekulé structure of benzene suggests that each of the nine electron pairs that bind its six carbon atoms into a cyclic configuration is localized between a pair of carbon atoms and that the number of localized electron pairs alternates around the ring between one and two. An alternate picture is that each of the nine electron pairs is delocalized around the entire ring. That is, each electron pair has the same probability of being found anywhere around the ring as does any other pair. (b) and (c) Single and double carbon-carbon bonds also alternate along the straight-chain and cyclic backbones of the polymers polyacetylene and polythiophene. The electron pairs that bind the polymer backbones, like the electron pairs that bind the benzene ring, are delocalized.

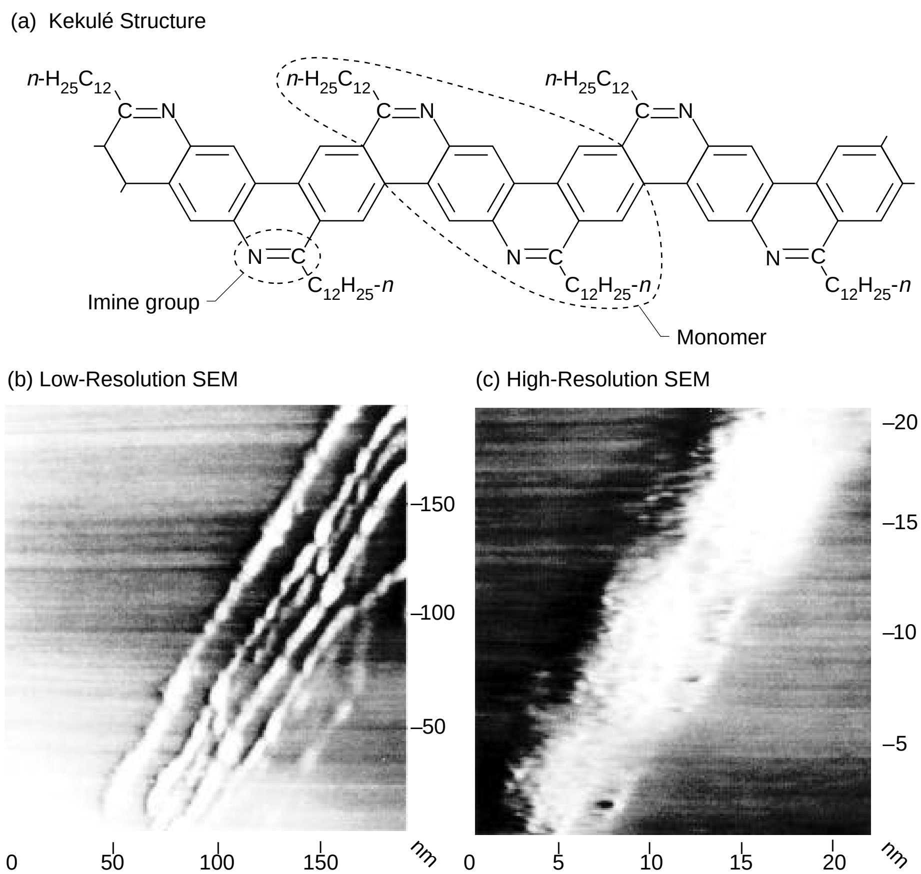


Figure 3. A Candidate Molecular Wire

(a) A conjugated organic polymer that may transfer electrons rapidly and efficiently from one electronic device to another has been synthesized by James M. Tour and Jay Lamba at the University of South Carolina. As shown by its Kekulé structure, the polymer consists of linked benzene residues (more precisely, linked para-phenylene residues) bridged by imine groups. (The carbon symbols have been omitted in the benzene residues.) Note the alternating single and double carbon-carbon bonds along the polyphenylene backbone. All the benzene rings along the backbone are coplanar. A monomer of the polymer is highlighted, and the symbol *n* indicates that the -$C_{12}H_{25}$ groups are unbranched. The advantages of using imine groups as bridges are discussed in the text. (b) This scanning tunneling electron micrograph of molecules of the polymer whose Kekulé structure is given in (a) was prepared by depositing a solution of the molecules on a substrate and allowing the solvent to evaporate. Note that the polymer molecules tend to form stacks whose orientation is driven by the rigid planarity of the polymer. That tendency will be ideal for establishing interconnects between electrodes. (c) This higher-resolution scanning electron micrograph shows that the polymer molecules are about 7.5 nanometers long (1 nanometer = 10^{-9} meter), a length that corresponds to the presence of eight monomers in each molecule. Both electron micrographs were taken at the University of South Carolina by Jay Lamba and John Cooper under the direction of M. Myrick.

may not occur very often, particularly at the interfaces of the polymer with the devices it connects.

During the past year researchers at Los Alamos National Laboratory, the University of South Carolina, and Yale University have started a collaborative experimental and theoretical study of the viability of conjugated organic polymers as components of ultrafast electronic circuits. The work is concentrating initially on the use of cyclic conjugated organic polymers as interconnects between the simplest of devices, namely metallic electrodes. The key to success will be the design of polymers possessing suitable electron-transfer properties and capable of being incorporated into real circuits. Figure 3 shows scanning tunneling electron micrographs of a candidate polymer designed and synthesized by James M. Tour and his graduate student Jay Lamba at the University of South Carolina. Each polymer molecule, which is only eight monomers long, consists of a polyphenylene backbone (a series of linked benzene rings) bridged with imine (>C=N–) groups. Although the polyphenylene backbone can, in principle, conduct electrons, it is not a particularly good intrinsic conductor because each benzene ring is free to rotate relative to its neighbors. Such rotation leads to a polymer with a nonplanar configuration that is less than optimally favorable for electron transfer along the backbone. But the imine bridges assure a configuration in which all the benzene rings are coplanar. The use of imine groups as bridges was motivated, in addition, by the fact that the imine groups, unlike other bridging groups that would also assure a strictly planar polymer, do not disturb the conjugation of the backbone.

Yale University's contribution to the collaborative effort will consist of experimental studies of the polymer shown in Figure 3 and of other candidate polymers synthesized by Tour's group. In particular, Mark A. Reed and his graduate student Rachel Lombardi will measure the conductivities of the polymers in specially designed chips.

Los Alamos National Laboratory's contribution to the project is theoretical analysis. Our aim is not only to understand the experimental results, but, just as important, to predict and provide feedback to the experimenters. For example, as mentioned above, the interfaces between a polymer and the electrodes (or other devices) it connects are crucial to efficient electron transfer. In fact, the polymer molecules must be designed so that they have terminal chemical fragments that efficiently transfer electrons from and to the electrodes. Unfortunately, experimental design and synthesis of a candidate polymer requires a minimum of six months. As a result, the number of candidate polymers that can be investigated experimentally is seriously curtailed by manpower and cost considerations. One role of theorists will be to make educated guesses about which polymers are the most promising candidates and should be synthesized for experimental study.

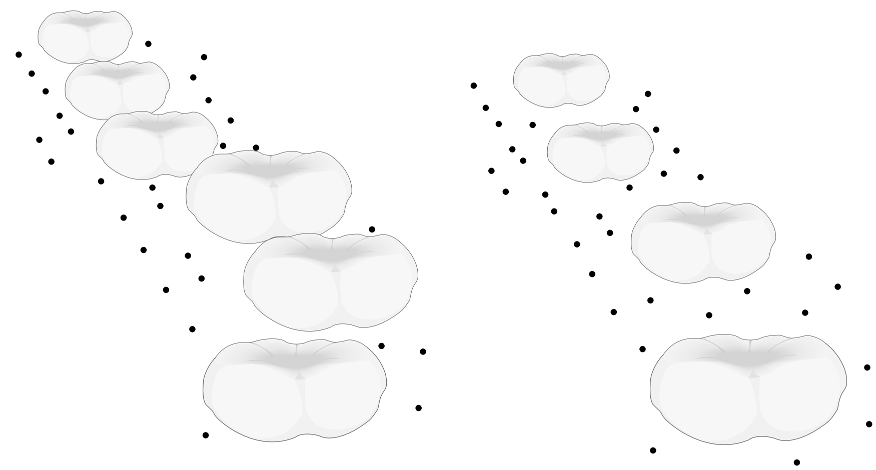


Figure 4. Orbitals for the Polymer of Figure 3
Shown here are an artist's renditions of two molecular orbitals. Each orbital is occupied by a certain pair of electrons along the polyphenylene backbone of the trimeric version of the polymer shown in Figure 3. The surfaces that make up each orbital define a volume within which the absolute value of the wave function is equal to or greater than a fixed low value (0.05 atomic units) and hence the volume within which the electron pair is most likely to be found. The shape of the orbital on the left indicates that the electron pair has equal probability of being found within a doughnut-like region around any of the six benzene residues in the trimer. The shape of the orbital on the right indicates that the other electron pair has an equal probability of being found within a doughnut-like region around only four of the six benzene residues in the trimer. The dots located near but outside each orbital denote atoms other than the carbon atoms of the backbone.

One of the properties of conjugated organic polymers we are trying to calculate is their capability to transfer electrons—their conductivities. The theoretical techniques available for studying the transfer of one or a few electrons at a time along a molecule are at the cutting edge of current research, and therefore their application to the question at hand is not cut and dried. Our calculation of the conductivities of conjugated organic polymers illustrates the problems that must be confronted.

Under many circumstances a macroscopic material obeys Ohm's law. That is, the current density J (the current per unit area) established in a material by an electric field E is directly proportional to the electric field. Symbolically, $J = \sigma E$; the constant of proportionality σ is the conductivity. When the flow of current induced in a material has the same direction as the applied electric field (as is true in many materials), the conductivity is a scalar, a single number. In some materials, however, the current does not flow in the same direction as the electric field. For example, some layered crystalline materials can exhibit a current parallel to the layers irrespective of the direction of the electric field. The conductivity of such a material is not a scalar but a tensor, that is, a 3-by-3 matrix. The conductivities of conjugated organic polymers are expected to be tensors.

One of the approaches we are using to calculate the conductivity of a polymer molecule is based on macroscopic analogues. We start by assuming that the polymer satisfies Ohm's law. Although we expect that assumption to be valid at sufficiently small electric fields, we are not certain about its validity at the fields that will be imposed on the polymer molecules. Our criterion for acceptance of the applicability of Ohm's law and any other assumption

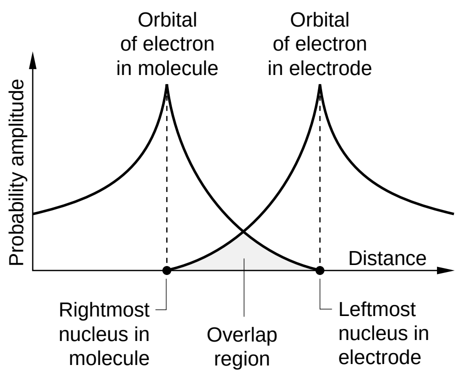


Figure 5. Overlapping Wave Functions and Transition Probabilities
The wave function on the left in this schematic illustration represents an orbital of an electron near the rightmost end of a molecule, which contacts the leftmost end of a metallic electrode. The local maximum in the wave function at the position of the nucleus of the rightmost atom in the molecule indicates that the electron has a high probability of being found close to the positively charged nucleus. The exponential decrease of the wave function toward the right indicates that the electron has an exponentially decreasing probability of being found to the right of the rightmost nucleus. The wave function on the right represents an orbital of an electron in the electrode near the interface between molecule and electrode. That wave function has a local maximum at the position of the nucleus of the leftmost atom in the electrode and decreases exponentially toward the left. The probability that an electron will undergo a transition from one orbital to the other increases with the extent of the overlap between the two orbitals.

will be a comparison of calculated and measured conductivities. If Ohm's law is applicable, we can calculate the conductivity tensor of a polymer molecule by using a theory developed by the Japanese physicist Ryogo Kubo during the 1960s. Kubo's theory is called linear response theory because it relates the response of a system (here the current density) to the external agent causing the response (the electric field) by a general equation in which both the response and the agent are raised to the first power. Kubo developed a set of formulas for the conductivity tensor of a macroscopic body; we have had to extend the theory to produce a different set of formulas for the conductivity of a single polymer molecule.

Our method starts with the solution of Schrödinger's equation for a polymer molecule. We solve Schrödinger's equation by using standard techniques widely available and thoroughly tested during the last twenty years. The solution of Schrödinger's equation, the wave function, provides us with a set of orbitals, one for each pair of electrons along the backbone, and a set of corresponding electron energies. Each orbital is a function that describes the probability of finding the electron pair at any point in three-dimensional space; the corresponding energy indicates how tightly the electrons are bound to the molecule. Typical orbitals for the trimeric version of the polymer molecule shown in Figure 3 are depicted in Figure 4. The complete set of orbitals, one for each pair of electrons, can be combined to give the total wave function for the molecule.

Our approach to obtaining the conductivity tensor requires calculating the wave function of the polymer molecule for each of a sequence of values of the applied electric field. From the resulting sequence of wave functions, we then calculate the dipole-moment matrix for the molecule as a function of the applied electric field. (The dipole-moment matrix is a measure of how much the wave function distorts when a voltage is applied across the ends of the molecule.) We then enter the dipole-moment matrix into the modified Kubo formulas we have developed to calculate the conductivity tensor of the molecule.

Although the procedure described above is conceptually straightforward, many as yet unanswered questions will have an impact on how our research will evolve. For example, our calculation of the dipole-moment matrix involves the assumption that the electric field is constant over the entire molecule. That assumption may or may not provide a calculated conductivity with the same level of accuracy as the measured conductivity but is most certainly invalid at high electric fields. Furthermore, the natural extensions of Ohm's law to high electric fields include nonlinear terms; that is, the current density is given by an equation such as $J = \sigma_1 E + \sigma_2 E^2 + \sigma_3 E^3 + \ldots$. As a consequence, the formulas based on Kubo's formalism would have to be modified accordingly.

Other questions of a more practical nature also need to be addressed. Already mentioned is the question concerning transfer of electrons from electrode to polymer molecule and vice versa. Electrons may be transferred along the length of the polymer molecule very efficiently but face bottlenecks at the electrode-polymer interfaces. We know, however, that the probability of an electron "hopping" from an orbital in the molecule to an orbital in the electrode increases with the overlap between the orbitals (Figure 5). The hopping probability also depends in a dramatic manner on the difference between the energies of the electron in the two orbitals: the smaller the difference the larger the probability. Therefore, a crucial part of our research is the study of the conductivity between a polymer molecule and a metallic electrode and the exact dependence of the conductivity on the overlap between the wave functions. Fortunately, the methods we have been developing are as applicable to polymer-electrode conductivity as they are to the conductivity of the polymer itself.

Another problem concerns depositing the polymer molecules at the desired locations on a chip, a task that cannot be accomplished by standard techniques. The most promising approach is "self-assembly," which involves designing the polymer to include terminal chemical groups that selectively bind to metallic electrodes. For example, the thiol group, –SH, is known to selectively bind to gold. The terminal groups must be chosen not only on the basis of their binding selectivity but also on their ability to provide maximum conductivity between the polymer and the metallic electrode. After the polymer molecules are equipped with appropriate terminal groups, a small quantity of a solution containing the molecules is deposited on the regions of a chip between electrodes. By natural chemical affinity the molecules then self-assemble at the correct positions.

After the feasibility of molecular wires has been demonstrated, a number of more advanced applications can be investigated. For example, introducing the appropriate chemical fragment (a porphyrin, say) in the middle of a conjugated organic polymer may create a device that acts as a switch. Or, if the fragment acts as a barrier to the passage of electrons, the result may be a tunnel device that allows passage of electrons only when the applied voltage is above a given threshold. A number of other relatively simple electronic devices that work at the molecular level can be envisioned. Indeed, at the present time, the possibilities seem limitless. ■

Further Reading

Forrest L. Carter, editor. 1982. *Molecular Electronic Devices*. New York: M. Dekker.

Charles H. Bennett and Rolf Landauer. The fundamental physical limits of computation. *Scientific American*, July 1985.

Mordehai Heilblum and Lester F. Eastman. Ballistic electrons in semiconductors. *Scientific American*, February 1987.

James D. Meindl. Chips for advanced computing. *Scientific American*, October 1987.

Hermann Grabert and Michel H. Devoret, editors. 1992. *Single Charge Tunneling: Coulomb Blockade Phenomena in Nanostructures*. NATO ASI Series B (Physics), volume 294. New York: Plenum Press.

Robert W. Keyes. The future of solid-state electronics. *Physics Today*, August 1992.

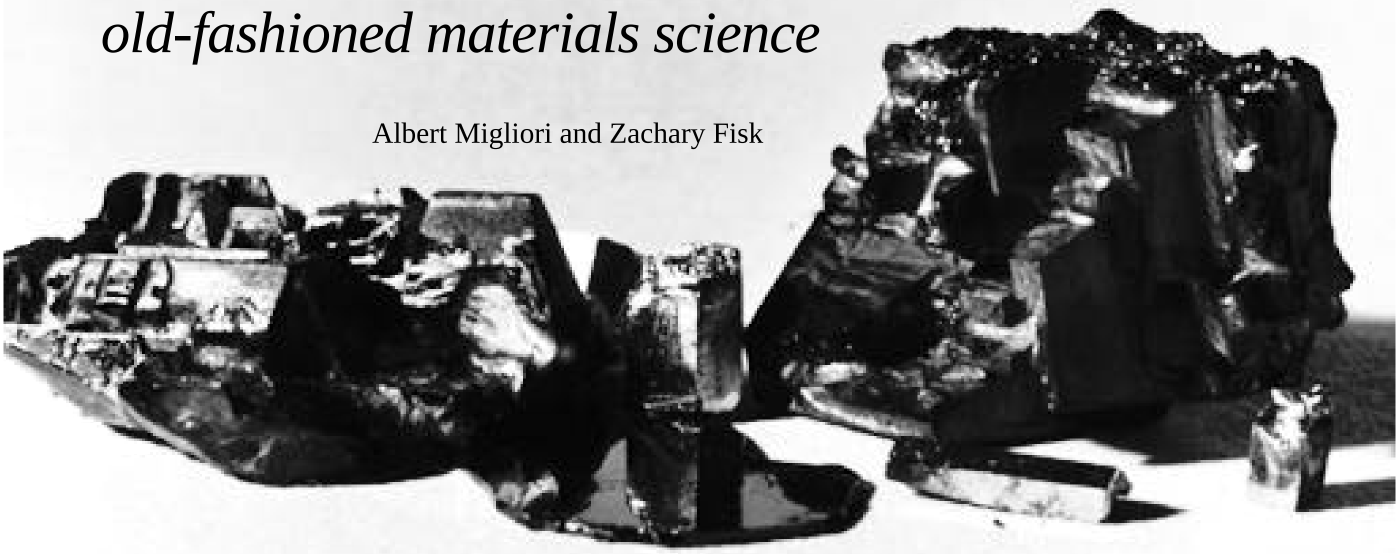
Konstantin K. Likharev and Tord Claeson. Single electronics. *Scientific American*, June 1992.

Antonio Redondo received a B.S. in physics from Utah State University in 1971 and a Ph.D. in applied physics from the California Institute of Technology in 1976. After teaching for three years at the University of the Andes in Venezuela, he returned to Caltech in 1980 as a visiting associate in the Chemistry Department. In 1983 he joined Los Alamos National Laboratory as a staff member in what is now the Electronic Materials and Device Research Group. His research involves collaborative efforts between experimental and theoretical scientists in molecular electronics, conducting polymers, and catalysis.

Crystals and Ultrasound

old-fashioned materials science

Albert Migliori and Zachary Fisk



Stone Age, Bronze Age, Iron Age, the age of exploration and the search for gold, the industrial revolution and the steel that made it possible, and, of course, the present age of electronics—totally dependent on tiny chunks of ultrapure, single-crystal silicon—it seems that entire chapters in human history are strongly connected to certain special materials that enabled massive bursts of what has come to be called progress.

But the pursuit of human progress is not the driving force behind physicists studying new materials today. We have much simpler motivations, more akin to the curiosity of the scientist who four thousand years ago first noticed the brilliant luster revealed by scratching the surface of the glob of bronze pulled from a state-of-the-art, charcoal-driven, high-temperature, controlled-atmosphere furnace. That scientist had, by serendipity, produced some sort of very imperfect alloy. The appearance and the properties of even that poor excuse for bronze were so remarkable that many, many others spent lifetimes advancing available technology to improve that material to the point where the fully developed and vastly improved copper alloys were to the first bronze nugget as that nugget was to the stone it replaced.

Just as important as the production of high-performance bronzes were the parallel developments necessary to understand completely the properties of such alloys. And so goes the story for many other materials. The entire business of materials science, from the first discovery of a new substance to the full development of its properties, is a tremendously attractive intellectual puzzle of sufficient complexity to interest almost anyone. Moreover, the incredible importance of new materials, the qualitative changes in human life that they produce, their intrinsic physical attractiveness, their appeal to the intellect, and our total inability to foresee how important their impact will be make the study of new materials exciting, risky, difficult, and what seems best described as fun.

Essential to the richness of the field is an understanding of the laws of physics and chemistry that enable detection and realization of new properties of matter. But despite having its roots in basic science, condensed-matter physics is viewed as a kind of applied science: After all, we know all of the fundamental physical laws that describe the individual atoms in solids. The interesting fact, however, is that the collective properties of many atoms assembled together are too complex to be predicted from first principles. For example, the sharp transition from nonmagnetic to magnetic behavior in iron at 1043 kelvins cannot be predicted from a calculation of the properties of an assemblage of a few or even a few hundred iron atoms. Such calculations indicate no abrupt transition to ferromagnetism but only a gradual change from weakly magnetic to more strongly magnetic. Nobel laureate P. W. Anderson has most clearly laid out this idea in his article "More is Different": He makes the point that phenomena arising as more and more particles interact are not simple extrapolations of the behavior of a few particles. Further, we cannot, nor do we want to, solve the most basic equations for the many-atom system because those equations are too complex and provide too much information. No one cares where a particular atom is at a particular time in the magnetic powder of a cassette tape. What is important is the overall magnetization of millions of atoms. Thus most of the useful theoretical descriptions of solids rely on statistical averages to predict measurable quantities. There is, then, a theoretical gap between our knowledge of the basic laws for each atom and our ability to predict what large numbers of interacting atoms do.

In the last few years some attempts have been made to control macroscopic material properties by constructing artificial materials through the deposition of sequential layers of carefully controlled composition and thickness. Attempts are being made also to produce nanostructures, materials assembled atom by atom through such processes as molecular-beam epitaxy and chemical or vapor deposition. Both approaches produce very small, but not quite microscopic, building blocks, or repeated units. Such designed materials are applicable to the production of practical devices as well as the study of quantum physics, and they have properties that are roughly predictable from theory. Still, the theoretical tools needed to predict their material properties necessarily include statistics and macroscopic approximations.

Recent work on "materials by design" contrasts strongly with the work of what one could call "old-fashioned" materials scientists, who work semi-empirically and produce structures that are largely only what nature allows; that is, the materials are microscopically homogeneous and theoretically intractable. The payoff from this old-fashioned work is the discovery of completely new phenomena, which often arise when the number of atoms is large (on the order of 10^{18}).

A particular innovation that arose from the "large-number effect" coupled with "old-fashioned" intuition was the inadvertent discovery by K. Mueller in 1987 that certain cuprates (namely, those copper-oxide compounds that are doped with transition-metal and other impurities) are high-temperature superconductors, that is, they become superconducting at temperatures well above absolute zero (above 30 kelvins). Mueller knew that those cuprates are quite unusual solids. He also knew that in many materials, if a smaller atom is deliberately substituted for a larger atom at the center of each unit cell in the crystal lattice, then when the solid cools the smaller atom and its cloud of electrons would not have a stable resting spot in the center of its symmetric cage. As a result it must move to some other position in the unit cell and spontaneously break the crystal symmetry. The effect (called a Jahn-Teller instability) is driven by large symmetric arrays of atoms and is only poorly understood in detail ("large numbers" + "intuition" = "mysterious phenomena"). Mueller's brilliant conjecture was that, in electrical conductors, the distortion of the crystal lattice resulting from the Jahn-Teller instability would produce new material properties provided the energy associated with the lattice distortion is comparable to one of the energy scales of the conduction electrons. As it turned out, the marvelous superconducting properties of the cuprates were not attributable to Mueller's conjecture, but nevertheless his very creative idea led to their discovery.

The unpredictable effects found when large numbers of various atoms are assembled thus provide both motivation and justification for the empirically based search for new physics and chemistry through the study of new materials. Fortunately, the production of any single new material is often accomplished by a very few scientists working with a small budget and a limited collection of inexpensive equipment. The demonstration of superconductivity in heavy-fermion compounds by F. Steglich at Universität zu Köln, and the discovery of the high-temperature superconductors by K. Müeller at IBM Zurich are typical examples of small, successful efforts. However, underlying all of the apparently small efforts is a powerful information and technology base easily accessible to those working within the umbrella of a large research laboratory or university. That access is crucial; without it small groups of scientists, driven by whatever odd notions motivate them, could not succeed.

The Remarkable Effects of Impurities

Also important to success is a clear focus on some particular set of material properties. Many groups today are focused on modifying the electronic and magnetic properties of solids by doping them with impurities and studying the effects of those impurities as the solids are cooled to low temperatures. Often the presence of impurities causes the solids to undergo a drastic change in some physical property as they cool, from conducting to superconducting, from nonmagnetic to magnetic, from paraelectric to ferroelectric, and so on. These drastic changes are called phase transitions.

It is a curious finding that some seemingly intrinsic properties of solids

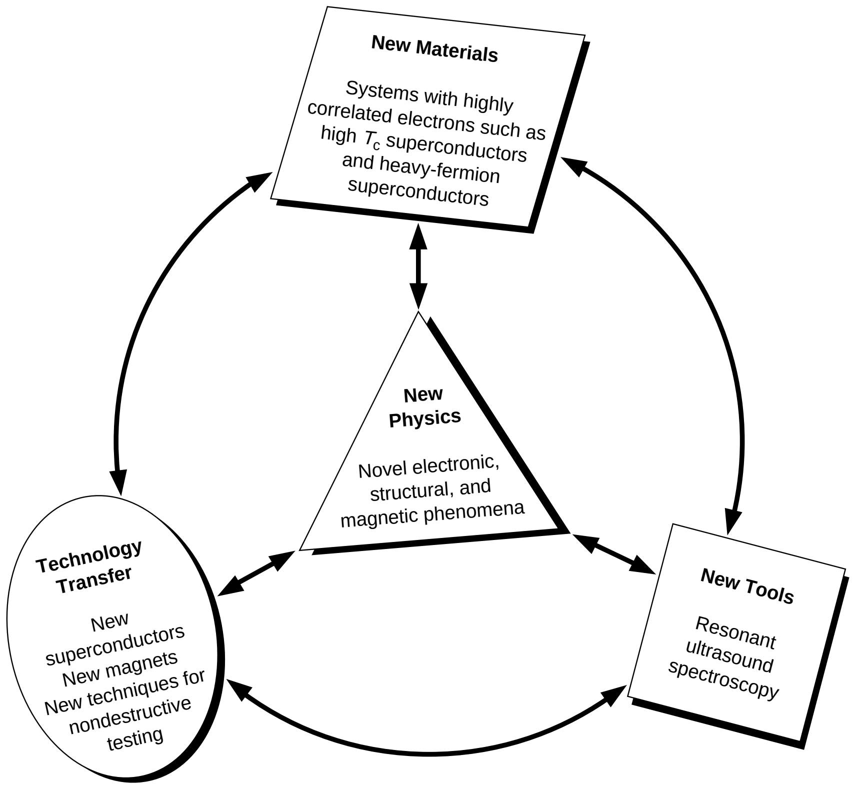


Figure 1. Materials Science at Work
The production of new materials with exotic properties leads to the development of new tools to study them, the use of those tools to advance the understanding of the new materials and the application of the new materials, the new tools, and the new science in the commercial sector. Here we elaborate on this paradigm with examples from our own work.

owe their existence to small amounts of impurities. For example, off-the-shelf tungsten metal is typically hard and brittle. When pains are taken to remove the various trace impurities dissolved in nominally pure tungsten, it becomes quite soft. The change in hardness is so pronounced that the purity of a tungsten sample can be reliably estimated with an ordinary Dremel grinding tool. The very pure metal is softer than annealed copper, and its x-ray pattern becomes blurred, no doubt because of the large deformations that easily occur in its soft condition. Perhaps a more surprising case is that of salt (NaCl). Usual NaCl contains a small amount of hydroxl, OH^-. When great care is taken to prepare OH^--free NaCl, the resulting solid can be spread like butter on bread.

The most important example ever in which the effects of impurities were understood and then exploited was in the development of transistors. In the early decades of this century, work by Thompson, Drude, Pauli, Fermi, Dirac, and Sommerfeld led from the discovery of the electron to an understanding of the special quantum physics governing conduction processes in pure and impure (doped) semiconductors.

This revolution in understanding the electronic properties of solids had, by the late 1930s and early 1940s, germinated into an idea in the minds of John Bardeen, William Shockley, and Walter Brattain at AT&T Bell Laboratories. They realized that the intrinsic properties of very pure (but still doped) semiconductors would allow the electrical conductivities of these materials to be controlled and varied by externally applied voltage, just as the flow of electrons is in a vacuum tube. Therefore such materials might serve as replacements for the reed relays in telephone switchboards. In order to obtain this effect, they knew they needed to produce single crystals of germanium and silicon with the impurity content kept to the unprecedented low level of 1 part per million—and they used every piece of technology available to reach their goal.

Along the way, when state-of-the-art was inadequate, their team developed new methods of growing and characterizing crystals, such as zone-refinement, a simple method for removing impurities while a crystal is growing. In the end they were able to make the first transistors. In the process of succeeding, they significantly advanced the theory and methods of solid-state physics. Not only did they predict the need for purity so high as to seem unnecessary, unachievable, and up until then unmeasurable, they also developed methods to produce crystals of the desired purity, developed characterization schemes, grew the crystals, and then produced practical devices whose function was totally reliant on the remaining and carefully controlled impurities. Their odyssey from the purest of basic research to the stupendous discovery of one of the most important practical devices ever was one of the greatest achievements of twentieth century physics.

Second-Order Phase Transitions and the Measurement of Elastic Properties

We have been interested in producing and understanding materials that undergo second-order phase transitions. In some metals the transition from the normal conducting phase to the superconducting phase is second order, as are many other structural, magnetic, and electric phase transitions. In this type of phase transition, the symmetry and material properties of the solid change gradually rather than abruptly (as in first-order phase transitions), but the changes nevertheless begin at a very distinct temperature called the critical temperature, T_c. The solid is in a more symmetric phase above T_c than below.

In a crudely parallel way the introduction of various amounts of impurities can also break the symmetry of a solid, at least locally. The combination of changes in temperature and control of impurity levels can thus be used to tune, in more or less continuous ways, important physical properties in metastable materials. Therefore, the study of such systems is widespread in materials science. It involves many tools, requires science, intuition, and luck, and it produces returns in both fundamental physics and very practical applications.

The tale we are about to tell concerns the development of a very old technique into a powerful and essentially new one for determining the changes in the elastic stiffness of crystals, particularly as they undergo second-order phase transitions. This now mature technique, called resonant ultrasound spectroscopy, has led to a simple procedure for making what once was a very difficult measurement. Our work illustrates the synergism among the production of new materials, the invention of new tools to study them, the subsequent use of those tools to advance the physics of other materials, and the application of the tools, the science, and the materials to industrial applications. In other words, our story illustrates a paradigm of how materials science actually works (Figure 1).

Measuring the elastic stiffness, or elasticity, of a solid as a function of temperature has been a traditional and very important technique for studying

Discontinuities at Second-Order Phase Transitions

Three thermodynamic quantities, the specific heat C_P, the thermal-expansion coefficient α, and the elastic-stiffness tensor c_{ij}, can be discontinuous at the critical temperature T_c of a second-order phase transition. Each is proportional to a second derivative of ΔG, the change in the Gibbs free energy per unit mass across the boundary separating two phases. Although ΔG is always zero in any phase transition, in a second-order phase transition the first derivatives of ΔG are always continuous and the first sign of any discontinuities occurs in the second derivative of ΔG.

The relationships between ΔG and C_P, α, and c_{ij} are particularly simple to write down for a liquid because the elastic-stiffness tensor has only one component, namely, the bulk modulus, B.

In terms of temperature T, pressure P, and volume V, the thermodynamic relations are

$$\frac{\partial^2 \Delta G}{\partial P^2} = \frac{\partial \Delta V}{\partial P} = -\frac{1}{B}; \qquad B = \text{bulk modulus}$$

$$\frac{\partial^2 \Delta G}{\partial T^2} = -\frac{\partial \Delta S}{\partial T} = -\frac{C_P}{T}; \qquad C_P = \text{specific heat}$$

$$\frac{\partial^2 \Delta G}{\partial P\,\partial T} = \frac{\partial \Delta V}{\partial T} = \alpha. \qquad \alpha = \text{thermal-expansion coefficient}$$

ΔV is the change in volume per unit mass across the phase boundary and ΔS is the change in entropy per unit mass across the phase boundary. Although ΔG, ΔV, and ΔS are zero and the phase transition is continuous, B, C_P, and α can exhibit discontinuities at the critical temperature T_c.

The relationships for a solid are more complicated and are usually written in terms of the stress tensor σ_{ij} and the strain tensor ϵ_{kl} instead of P and V. Further, the bulk modulus B becomes the elastic-stiffness tensor c_{ijkl}, which relates the stress to the strain in terms of the fundamental relation

$$\sigma_{ij} = \sum_{kl} c_{ijkl}\epsilon_{kl}.$$

second-order phase transitions in materials doped with impurities. Nature helps us with the measurement because the elasticity of a solid is discontinuous at T_c; that is, it jumps to a different value when that solid begins to undergo a second-order phase transition. The jump is perhaps surprising because, as we mentioned above, the solid exhibits no obvious microscopic changes at the critical temperature. The atoms do not suddenly change position, magnetism and ferroelectricity do not suddenly appear, and a conductor is not suddenly able to carry super-large currents. However, three thermodynamic quantities—the elastic stiffness, the specific heat, and the thermal-expansion coefficient—do exhibit discontinuities at T_c and those abrupt changes can be measured. The accompanying box defines these three quantities in terms of variations with respect to temperature and pressure of the change in the Gibbs free energy per unit mass across the boundary between the two phases.

Discontinuities are a boon to the experimentalist because they are often the most unambiguous of measured quantities. Of the three discontinuous quantities mentioned, elastic stiffness is particularly informative because in a solid this quantity is a fourth-rank tensor, c_{ijkl}, with 81 components. The elastic-stiffness tensor relates the stresses (forces) to the strains (displacements) in the solid through the fundamental relation

$$\sigma_{ij} = \sum_{kl} c_{ijkl}\varepsilon_{kl},$$

where σ_{ij} is the stress tensor and ε_{kl} is the strain tensor. This relation between stress and strain is simply the generalization of Hooke's law, $F = -kx$, where F is the force developed in a spring if it is stretched a distance x. Thus the elastic-stiffness tensor c_{ijkl}, in analogy with the spring constant k, describes the stiffness of the bonds holding the solid together. Although c_{ijkl} has 81 components (or moduli), in crystalline structures that do not produce external magnetic fields, the tensor has at most 21 independent components. Consequently the elastic stiffness tensor is commonly written as c_{ij}, where i and j run from 1 to 6, for pragmatic rather than mathematical reasons.

It is also important to realize that the stiffness of a metallic solid comes not only from the chemical bonds that hold the ions together but also from the degeneracy pressure of the Fermi sea of electrons. Thus a great deal of information about the crystal is contained in a complete description of its elasticity. Because of this wealth of information, if one knows the crystal structure and can measure the changes in the elastic-stiffness tensor as a function of temperature through a phase transition, then it is possible to infer the detailed changes in the crystal lattice or in the electronic structures that are driving the phase transition. Few other measurements can reveal as much of the physics of phase transitions. Consequently we would like to be able to make this measurement in high-T_c superconductors, heavy-fermion superconductors, and other exotic materials exhibiting second-order phase transitions.

The Need for a New Measurement Technique

Unfortunately, nature, while permitting us to produce single crystals of high-temperature superconductors, heavy-fermion superconductors, and other materials in variously doped versions, has often restricted the dimensions of the crystals that can be easily grown to the millimeter range. That size is entirely adequate for many measurements. But for the crucial determination of elastic stiffness, a thermodynamic measurement that provided the first verification of the BCS theory of ordinary low-temperature superconductivity, millimeter-sized crystals are a disaster.

What we describe below is a solution to this measurement problem. It involves measuring the resonances (or natural frequencies of vibration) of a crystal and often makes measurement of the elastic-stiffness tensor more trivial than measurement of the resistivity tensor.

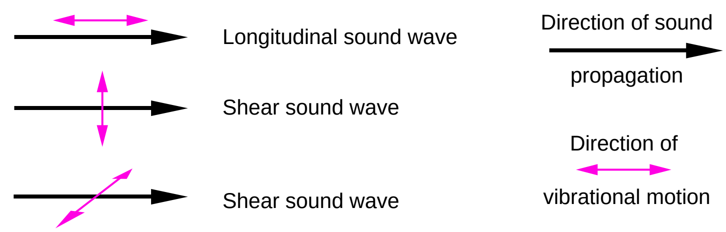
Prior to development of this the new resonance approach, elastic moduli were determined mostly from measurements of the speeds of longitudinal and shear sound waves along different directions in a sample. Figures 2 shows the simple relationships among these sound speeds and the elastic moduli for a cubic crystal. For a large chunk of isotropic material such as ordinary glass, there are only two independent sound speeds, one for longitudinal waves in which the atoms vibrate along the direction of the sound wave (these are the waves we hear), and shear waves, in which the atoms vibrate in a direction perpendicular to the direction of the sound waves. A simple measurement of the pulse-echo time of each type of sound is easy to do and yields sound speeds for glass accurately and completely. As shown in Figure 2, the pulse-echo time is the time for a short sound pulse to travel the distance from one face of the sample to the opposite face and back again.

In applying the pulse-echo technique to single crystals of more exotic materials, the first catch is that the speed of sound in nearly every solid that interests us is a blazing 5 millimeters per microsecond or so (Mach 15 in jet fighter units). The second catch is that we need to use very-high-frequency ultrasound (several hundred megahertz or

Figure 2. Sound Speeds, Elastic Moduli, and Pulse-Echo Measurements

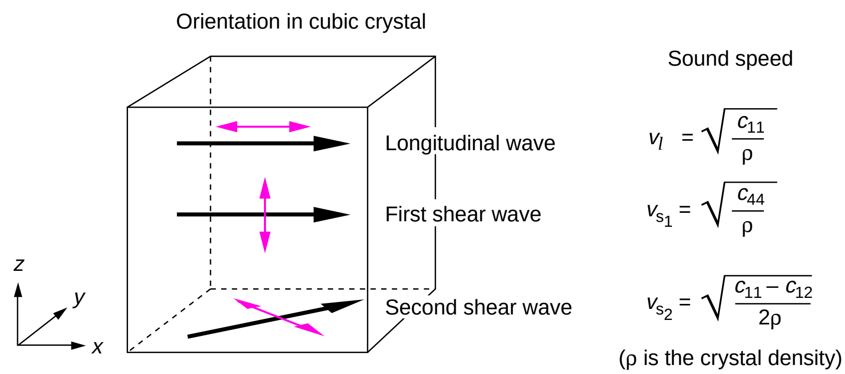
(a) Longitudinal and Shear Sound Waves

In longitudinal sound waves atoms vibrate along the direction of wave propagation. In shear sound waves atoms vibrate perpendicularly to the direction of wave propagation.



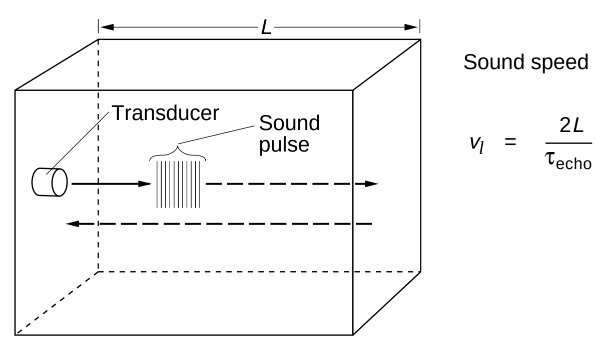
(b) Sound Speeds and Elastic Moduli

The elastic-stiffness tensor of a cubic crystal has three indepedent elastic moduli (c_{11}, c_{12}, and c_{44}) and therefore there are three sound speeds in a cubic crystal. The elastic modulus c_{11} is determined from the speed v_l of the longitudinal sound wave shown at right. The two shear moduli c_{12} and c_{44} are determined from two speeds v_{s_1} and v_{s_2} of the two shear waves shown at right.



(c) Pulse-Echo Measurements of Sound Speed

Conventional measurements of the elastic stiffness tensor c_{ij} are made by measuring the sound speeds of longitudinal and shear sound waves along various directions in a large crystal. In the example shown here a short pulse of longitudinal sound is generated by a transducer on one crystal face; the sound pulse travels as a narrow beam through the crystal and bounces back from the opposite face to produce an echo that is picked up by the transducer. The measured time between the inititation of the pulse and the echo (called the pulse-echo time, τ_{echo}) is equal to $2L/v_l$, where L is the distance between the two crystal faces and v_l is the longitudinal sound speed. Thus the sound speed can be calculated directly from τ_{echo}. Further, if the density of the crystal, ρ, is known, the measurement also provids a direct determination of c_{11}, one component of the elastic-stiffness tensor. By placing the transducer at various locations on the cyrstal and using both longitudinal and shear sound waves, all the components of the elastic-stiffness tensor can be determined.



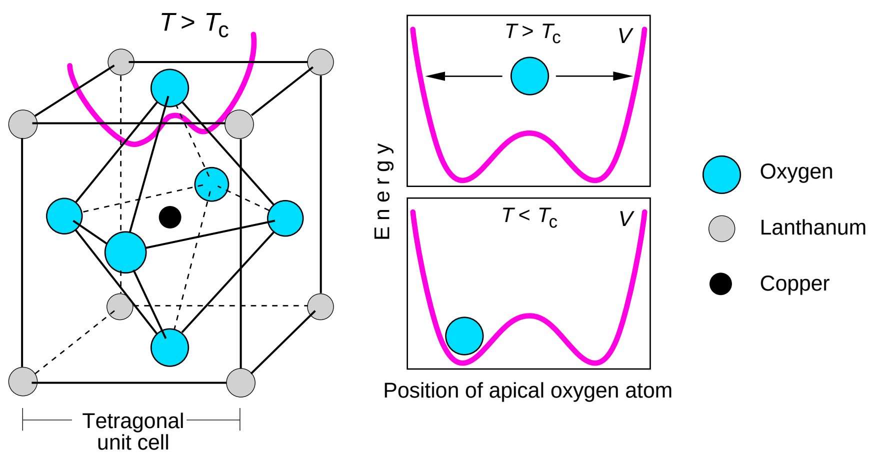
higher) to measure the sound speed in samples whose largest dimension is about a millimeter. But often ultrasonic attenuation increases to prohibitive levels at such frequencies.

Let's say that we want to measure the speed of sound in a 1-centimeter cube (big!) of La_2CuO_4. Above 530 kelvins this cuprate has a tetragonal structure (all three crystal axes of the unit cell are at right angles, two axes are equal in length, and one is longer or shorter). La_2CuO_4 is an interesting material because when doped with strontium it becomes a high-temperature superconductor and its doped versions undergo a structural phase transition as they cool, from a tetragonal to an orthorhombic structure (see Figure 3). Therefore, we're interested in the variation in sound speed (or equivalently, elastic-stiffness tensor) as a function of temperature through the structural phase transition. In our 1-centimeter sample, a single cycle (wavelength) of 50-megahertz sound is 0.1 millimeter long. To obtain a good signal-to-noise

Figure 3. Structural Phase Transition of La_2CuO_4

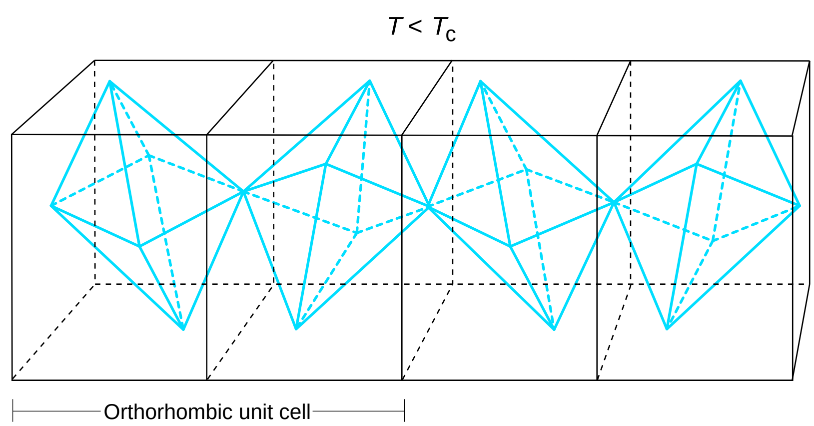
(a) Tetragonal Structure of La_2CuO_4 above T_c

The unit cell of La_2CuO_4 above the critical temperature, T_c = 525 kelvins, is tetragonal; that is, the axes are at right angles, the *x* and *y* axes are of equal length, and the *z* axis is longer. Note that the oxygen atoms form an octahedron. Each oxygen atom at the apex of the octahedron sits in a double-well potential, *V* (shown here in red). Thermal motions of the apical oxygen atoms are sufficiently large that the equilibrium position of each is at the center of its potential well.



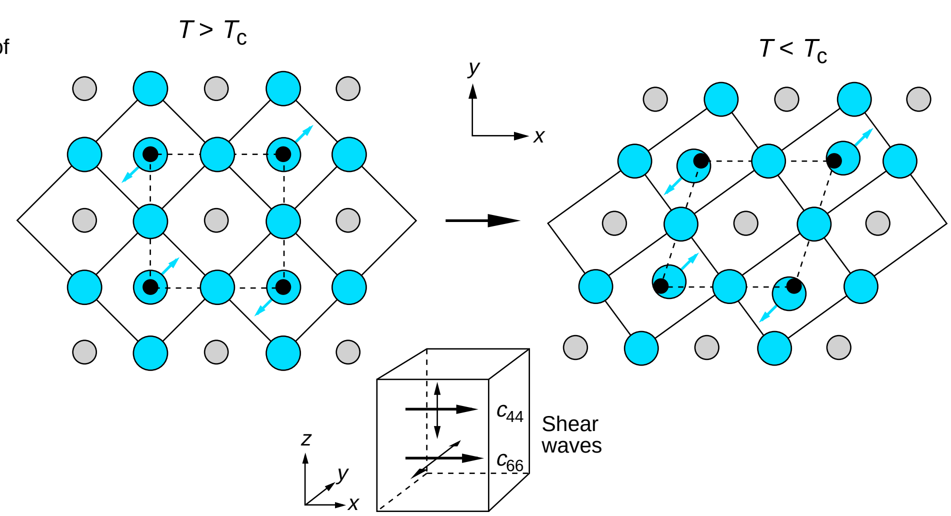
(b) Orthorhombic Structure of La_2CuO_4 below T_c

At temperatures below the critical temperature, each apical oxygen atom falls into one or the other minima of its double-well potential so that the octahedron of oxygen atoms in each unit cell has a static tilt. Since octahedra in adjacent cells tilt in opposite directions, two of the old unit cells form the unit cell of the new structure. Thus the crystal now has an orthorhombic structure (all three axes of the unit cell are unequal in length).



(c) Top View of Phase Transition from Tetragonal to Orthorhombic Structures of La_2CuO_4

As the temperature decreases below T_c and oxygen octahedra develop permanent tilts, shear forces develop and change the square array of copper atoms to a rhombus. Here the distortion is exaggerated for the purposes of illustration. Theory predicts that the shear modulus c_{66}, which characterizes shear forces in the *x*-*y* plane, will shift abruptly during the phase transition whereas the shear modulus c_{44}, which characterizes shear forces in the *z*-*y* and *z*-*x* planes, will remain unaffected by the phase transition.



ratio, we send in a pulse consisting of many cycles of sound and bounce the pulse from the inside walls of the crystal. Such a pulse might last, then, 0.2 microseconds and be 1 millimeter long, 10 percent of the width of the sample.

To determine the relative change in sound speed to 1 part per million, a not unreasonable goal, we would need to time the pulse echo to an accuracy of 2 picoseconds, corresponding to about 1/1000 of a wavelength of sound. All of this is not wildly difficult to do for the frequencies appropriate to big samples. But for a 1-millimeter crystal we must increase the frequency, and therefore the timing accuracy, by a factor of 10. At 500 megahertz and 0.2-picosecond timing accuracy, things get tough. Even worse, the orthorhombic phase of La_2CuO_4, whether pure or doped, has nine independent elastic moduli and therefore nine different sound speeds, each requiring a separate measurement. And each measurement requires that a small transducer be glued to the sample and that it not fall off as we cool the sample from room temperature down to a temperature well below the critical temperature. The pulse echo is an intermittent signal, and the signal can easily become so greatly attenuated that barely one echo can be detected.

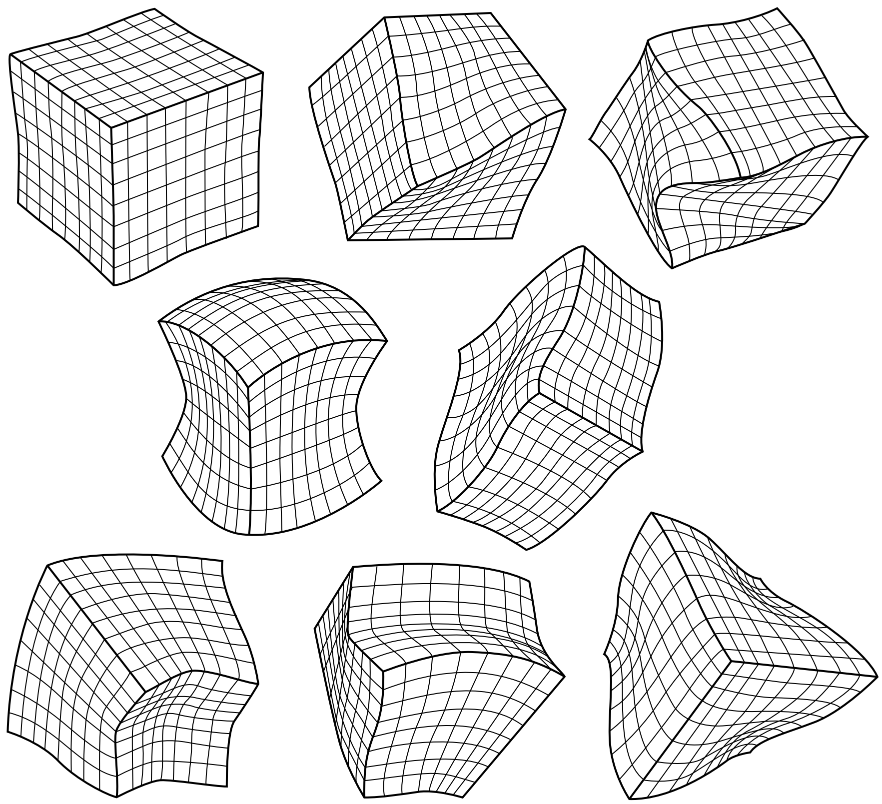


Figure 4. Distortions of a Cube-Shaped Sample on Resonance
Each figure above is an example of the distortions that a cubical object undergoes when it is driven at a natural (resonant) vibrational frequency. At each resonant frequency the vibrational motion is dependent on complicated linear combinations of all the elastic moduli as well as the exact shape of the object. Because of this complexity, Rayleigh, Love, and others were unable to compute the resonances of such short, fat objects from their elastic moduli. With the advent of big computers and some clever algorithms, such computations are now easily done.

The Development of Resonant Ultrasound Spectroscopy

Direct measurement of the pulse-echo time (or sound speed) to determine elastic stiffness would seem to be nearly hopeless for small crystals. But usually, when nature makes the measurement of time tough, the measurement of frequency is much easier. In fact, if we apply continuous sound to a solid sample, the solid will resonate, or ring just like a bell, provided the applied sound frequency is one of the solid's natural vibrational frequencies. (Application of this technique to long, thin rods is perhaps hundreds of years old!) Because the resonant frequencies are related to the pulse-echo times, we can measure resonant frequencies instead of pulse-echo times to determine elastic-stiffness tensors.

We measure the frequency of a resonance by driving the sample continuously with ultrasound and slowly changing the frequency of the sound until the sample suddenly starts to resonate (1-millimeter samples resonate at 1 megahertz or so, a wonderfully low frequency). The resonating sample acts like a natural amplifier, greatly increasing the amplitude of the vibrations (a factor of 10,000 is not uncommon). We now need only measure frequency. We can take as long as we like to do it, and the naturally amplified signal is present during the entire time we are doing the measurement. Therefore, a resonance measurement can easily have a signal-to-noise ratio that is a million times higher than a measurement of the echo time of short pulses of sound.

You might, at this point, wonder why anyone would have used echo-times rather than resonant frequencies to determine the components of the elastic-stiffness tensor. The catch is that the resonant frequencies are hard to interpret because they have a somewhat complicated relationship to the elastic moduli.

Over a hundred years ago John William Strutt, Baron Rayleigh, attempted to calculate the resonant frequencies of cubes, short cylinders, and other short fat objects from known elastic moduli. Unlike the corresponding calculation for thin rods and plates, this wonderfully tantalizing and seemingly simple problem stymied the brilliant Rayleigh, who finally concluded that "the problem . . . has for the most part, resisted attack". Figure 4 illustrates the origin of the complexity: When a cube resonates, it exhibits significant distortion typically involving all the elastic moduli in complicated linear combinations. A. E. H. Love, Willis Lamb, and others were also stymied by this problem. But in the 1960s Orson Anderson of Columbia University and his postdoc, Harold Demarest, hit on a fast, accurate numerical algorithm for obtaining the solution. The algorithm requires computations only at the surface of the object and achieves an accuracy much greater than the relatively crude "finite-element" techniques, which compute throughout the volume of a sample and are the only alternate method. With this fast, accurate algorithm Anderson and his coworkers were able to make the first fully interpreted resonant ultrasound measurements on large, high quality mineral crystals. That is, they were able to match measured resonance frequencies to predicted ones, but they had big crystals, big signals, and almost perfectly known answers for the elastic moduli before they started. In contrast, our small and completely uncharacterized samples could not be so easily attacked, so we had to develop new hardware and refined analysis procedures.

To measure the resonant frequencies of a 1-millimeter object, we had better (1) not disturb the sample's resonances with the measuring device and (2) not generate resonances in the measuring system that might confuse the issue. To not change what we wish to measure in the process of measuring it requires that our transducers make extremely weak contact with the sample. So if our sample is approximately cubical, we lightly contact its corners with transducers, using no glue or any other coupling medium, and apply just a gram or so of force (pardon the unit).

With that light contact we lose a factor of 1000 in our signal-to-noise ratio. We must also drive the sample lightly so as not to destroy it. Fortunately, the natural amplification at resonance recovers most of what is lost with a gentle drive. The resonances generated in the measurement apparatus are another matter. We are not ourselves small enough to make transducers much smaller than 1 millimeter or so; therefore, most transducers would ring in just the same frequency range as the sample and spoil the signal. The way around this problem is to make the transducers of a composite structure consisting mostly of single-crystal diamond. Diamond, with a sound velocity of 17 millimeters per microsecond (Mach 50!) has such a high sound speed that its resonance frequencies are much higher than those of the crystals of comparable size that we wish to measure.

When we hooked all the pieces together and connected them to electronics specially designed (with the help of Thomas Bell) to maximize the signal-to-noise ratio, John Sarrad, a student working with us on his thesis research, was able to measure all of the lower fifty or so resonant frequencies of a crystal with a largest dimension of 1 millimeter or so. The entire experiment fits on the top of a desk and costs less than a car. Figure 5 shows the measurement apparatus and a typical resonance spectrum.

Analysis of Resonant Ultrasound Data

To our surprise, the algorithm to compute resonant frequencies from elastic moduli and our simple technique to measure accurately the resonances of a useful sample were not quite enough to produce a robust technique for determining elastic moduli from measured resonances. (Note that this is the inverse of the problem solved by Orson Anderson.) In real life the measured resonant frequencies have some errors relative to the resonant frequencies of a perfect sample. The errors arise from many sources including anomalies in the geometry of the sample (such as chips and rounded edges) and transducer loading effects. Typically, the resonances of the mathematical model of the solid and those measured for the real object differ in frequency by 0.1 percent, more or less. Further, one or two resonances out of the thirty or forty that exist in a given frequency range may be missing from our data because some resonances, by accident, may have no component of vibrational motion along the driving direction of the transducer. The experimental data are fed into a computer program that tries to find a set of elastic moduli consistent with the measured resonances, the sample dimensions, and the symmetry of the sample's crystal lattice. If one or two resonances are missing from the measured resonance spectrum, the computer will invent an answer that is "not even wrong."

We have not yet developed a computer program for guessing at the absence of a resonance; however, those guesses can be made by looking at the data and relying on the intuition that comes from experience. After several iterations and applications of William M. Visscher's pioneering computer algorithms, we usually do guess which resonance is missing, and then the computer instantly determines the elastic moduli of the sample to an accuracy of 0.05 percent or better. The algorithms also determine the true dimensions of the sample to an accuracy as high as 0.01 micron. The results are the most accurate and complete measurements to date of the elastic-stiffness tensor, and the many elastic moduli are obtained simultaneously from a single measurement of the resonance spectrum.

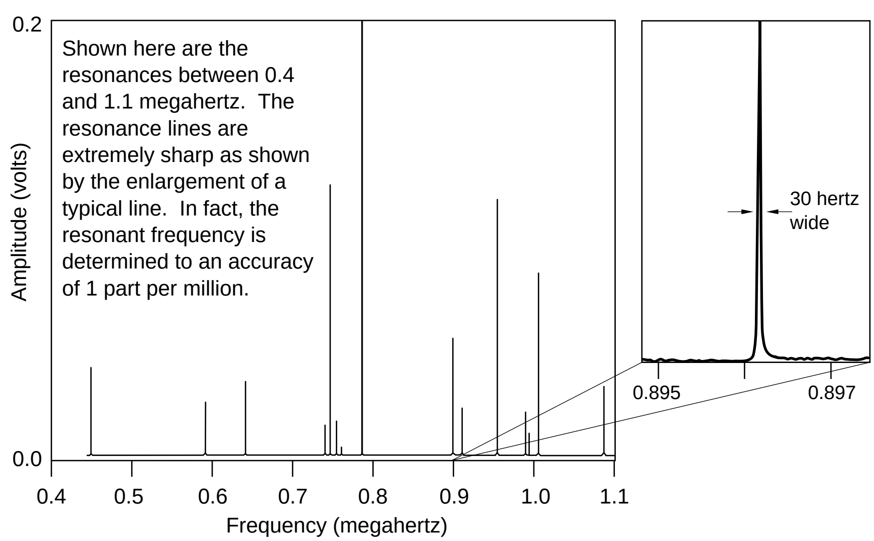
Our ability to infer the existence of missing resonances in our data was critically dependent on our being able to produce some samples that were perfect small single crystals and then using their known perfection to refine our experiment for less perfect crystals. This McLuhanesque bootstrapping of the development of a measurement technique by using known properties of near-perfect materials is not uncommon and is often the only safe route. Without the collaboration among crystal growers, instrument developers, and theorists, the resonant ultrasound spectrometer could not have been developed to its full potential. Without the instrument our perfect small single crystals would have remained strangers to the ultrasound and their elastic properties would still be a mystery. Although the project has involved many types of expertise, the size of the overall effort has remained small. The required components of the effort, however, were not predictable when we began developing the instrument. Thus the work really needed to be done at a national laboratory, which could serve as a technology supermarket with a large and available stock from which to select just the right things.

Figure 5. Ultrasound Measurements of $La_{1.86}Sr_{0.14}CuO_4$

(a) Measurement Technique

The photograph shows a sample of $La_{1.86}Sr_{0.14}CuO_4$ held between two transducers, one of which drives the sample over a continuous range of frequencies in the megahertz range. The apparatus is placed in a cryostat, which can cool the sample to the desired temperatures. While the temperature is held fixed, the frequency of the driver is gradually and automatically changed by an electronic signal generator until the sample resonates and thereby amplifies the applied signal. The amplified signal is picked up by the second transducer and recorded automatically by a specially designed electronic detector. The frequency of the driver then continues to change again until another resonant frequency is reached. All the resonant frequencies of the sample are thus recorded during a single experimental run. The entire resonance spectrum is measured in several seconds. To study a second-order phase transition in the sample, the experiment is repeated at various temperatures between room temperature and the critical temperature of the phase transition.

(b) Resonance Spectum of $La_{1.86}Sr_{0.14}CuO_4$ at T = 297.2 kelvins



Applications of the Resonant Ultrasound Spectrometer

With our resonant ultrasound spectrometer and a wonderful collection of crystals doped with various impurities, we, our postdoc, and our students have been able to make steady progress in expanding our knowledge of second-order phase transitions. Our resonant ultrasound measurements on pure and doped La_2CuO_4 single crystals near their structural phase transition from tetragonal to orthorhombic are particularly interesting.

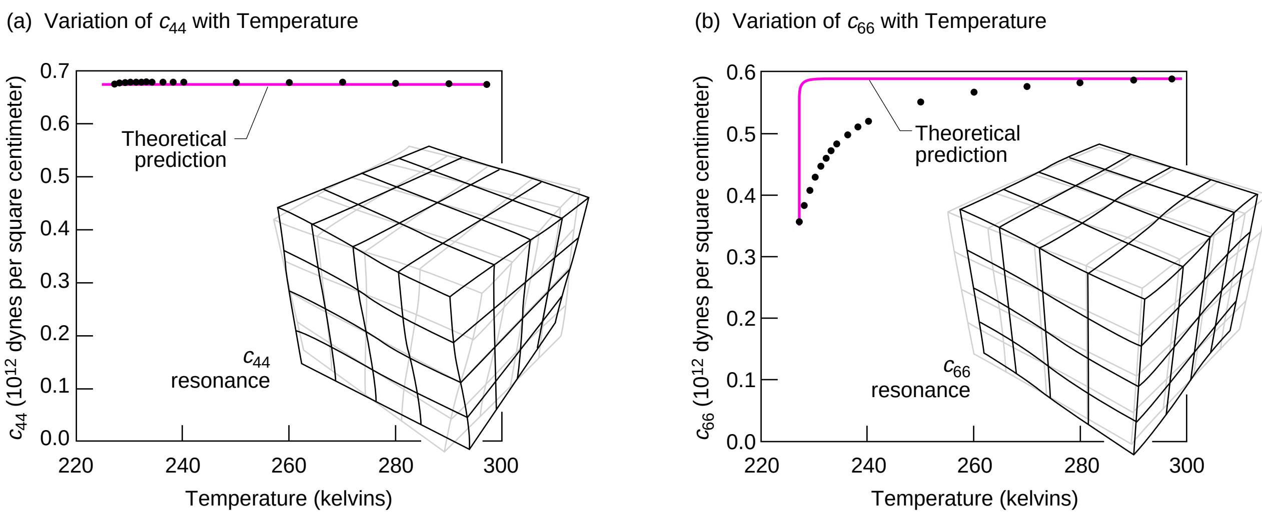


Figure 6. Surprising Results from Resonant Ultrasound Measurements
To study the transition of the superconducting cuprate $La_{1.86}Sr_{0.14}CuO_4$ from a tetragonal structure to an orthorhombic structure, we measured the elastic moduli of an approximately cubical, strontium-doped sample as a function of temperature using the resonant ultrasound technique. Resonance spectra were taken at various temperatures from room temperature down to the critical temperature of the sample, 223 kelvins. The values of the elastic moduli at the various temperatures were determined from the resonance spectra by using a computer program discussed in the main text. As explained in Figure 4 and the main text, theory predicts that the shear modulus c_{44} should remain constant through the transition, whereas the shear modulus c_{66} should remain constant but then undergo an abrupt change at (or slightly above) T_c. Above are shown the results from the resonant ultrasound measurements. (a) shows that the value of c_{44} remains constant, as expected, and (b) shows that the value of c_{66} begins decreasing at a temperature roughly 80 kelvins above T_c and continues to decrease as the temperature decreases to T_c. The departure from theoretical predictions (shown in red) indicates that this second-order phase transition is less well understood than was thought. The dotted and dashed lines of the cube inset in each graph indicate peak distortions at opposite ends of the vibrational cycle for each resonance.

When we started the measurement, the structural phase transition was thought to be well understood. As shown in Figure 3a, at high temperatures each apical oxygen atom in the unit cell sits in the center of a double-well potential (a potential with two minima). More precisely, fast thermal motions cause each oxygen atom to vibrate so that it appears to fill the available space symmetrically. Near the critical temperature of the structural phase transition, the energy of thermal vibrations is just about equal to the height of the bump (local maximum) in the center of the potential well. Thus the thermal motion is now too weak to keep the atom buzzing in the center (the atom can no longer make it over the bump of the potential well). As the solid cools further, the atom gradually drops into one of the two small wells at the bottom of the potential. Thus the arrangement of atoms is altered; that is, the solid experiences a structural phase transition.

Figure 3b shows the rearrangement. Each oxygen atom in the unit cell undergoes a displacement such that the octahedron formed by the oxygen atoms develops a permanent tilt. Because the octahedra in neighboring unit cells develop tilts in alternate directions, the two neighboring unit cells are no longer identical. In fact, the unit cell of the new structure is now composed of two of the old unit cells and has orthorhombic symmetry.

In Figure 3c, a projection of the crystal lattice onto the *x*-*y* plane shows that as the oxygen octahedra develop alternating tilts during the phase transition, they pull apart the corners of the unit cells and shear the square array of atoms into a rhombus (or diamond shape). Thus, according to theory, the shear modulus c_{66} should drop abruptly as the temperature drops below the critical temperature, whereas the shear modulus c_{44} should remain unaffected. In fact, very general symmetry arguments can be used to predict exactly how each of the six independent elastic moduli of the tetragonal structure should change as a function of temperature near the critical temperature.

Much to our surprise, the behavior of the elastic moduli as a function of temperature determined from our resonant ultrasound measurements differed markedly from theoretical predictions. Rather than changing abruptly at the critical temperature, the shear modulus c_{66} changes smoothly over a temperature drop of 80 kelvins or so (Figure 6). Because our simultaneous measurements of the six elastic moduli are extremely precise, there can be no doubt as to the validity of our results. What we are left with is a clear indication that this structural phase transition is not as well understood theoretically as was thought. Moreover, measurements on pure La_2CuO_4 as well as samples doped with oxygen, barium, and strontium showed similar results. Whether these results are related to the mysteries of high-temperature cuprate superconductors remains to be seen.

Our results for the La_2Cu0_4 system are typical of the surprises we find with resonant ultrasound measurements. The surprises provide motivation for improving our understanding of the basic physics, which in turn enables us to see the way to make new materials that have desirable structural, magnetic, or other properties. Thus, the physics we uncover has the promise that it will eventually apply directly to the real engineering aspects of materials.

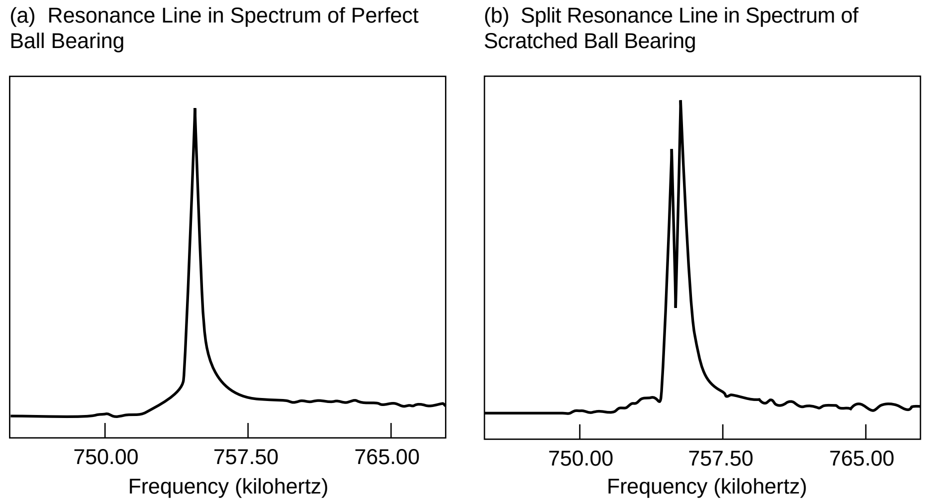


Figure 7. Fast, Accurate Detection of Flaws in Ball Bearings
The resonant ultrasound technique is accurate enough to detect tiny flaws in high-precision objects. The figure compares two resonance lines, one from a perfect silicon nitride ball bearing and the other from a similar ball bearing with a tiny surface scratch. The flaw causes the single line to split into two lines separated in frequency by 824 parts per million. The resonant ultrasound measurement can detect errors in sphericity as small as 0.005 microns in seconds, whereas commonly used optical inspections typically require one hour. The resonant ultrasound spectrometer is now being readied to test these newly developed ceramic ball bearings, which can withstand very high temperatures and acidic environments and are compatible with dry lubricants. A resonant ultrasound spectrometer specifically designed to test ball bearings is now available from Quatro Corporation as a result of a collaborative technology-transfer effort between that company and Los Alamos National Laboratory (see Figure 8).

A Breakthrough in Nondestructive Testing

Because of our justifiable confidence in the accuracy and precision of resonant ultrasound measurements, when the measured resonance spectrum of a sample of known structure cannot be made to fit the mathematical model to within 1 percent or so, as happened with a few brittle samples we attempted to study, we were able to conclude that the samples were not the perfect cube-like chunks that the microscope revealed but were, in fact, cracked. The cracks altered the resonance spectrum, and therefore, the data could not be fit to any object shaped like the sample. From this simple effect, discovered accidentally during our research on single-crystal samples, we developed new nondestructive testing approaches that subsequently received a 1991 RD100 award and a 1993 Federal Laboratory Consortium Award for Excellence in Technology Transfer. In particular, together with Raymond D. Dixon and others we have shown how certain anomalies in the resonance spectrum can be used to detect cracks and other flaws in small, precision objects including aluminum plates, ball bearings, high-strength permanent magnets for lightweight motors, and more.

One test for cracks involves identifying the presence of second harmonics in the resonance spectrum under dry conditions and the absence of those second harmonics under wet conditions. We, together with George W. Rhodes, discovered that a crack produces second harmonics when dry but not when filled with fluid because the fluid prevents the crack from banging

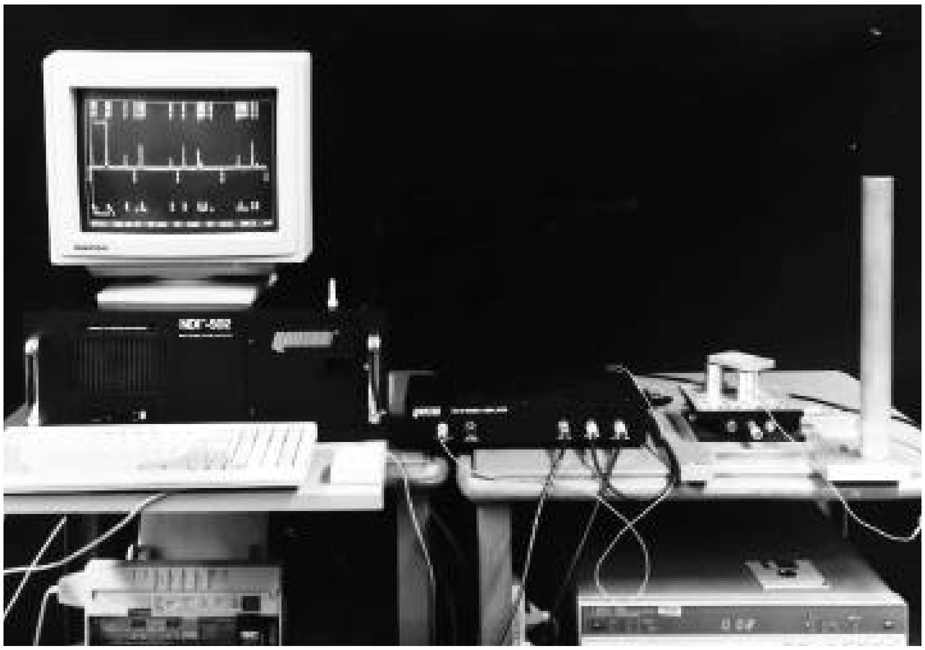


Figure 8. Resonant Ultrasound Spectrometers on the Market
The photograph shows a commercial resonant ultrasound spectrometer complete with computer, receiver, transducers, and sample. This system, produced by Quatro Corporation, is designed to perform nondestructive testing of ball bearings and other small objects. After having devised applications of resonant ultrasound technology to the detection of flaws in precision objects, we sought through a public advertisement a commercial company that could develop the technology for a wide variety of applications. Quatro came to us in 1991 and in 1992 obtained a license to develop, manufacture, and market systems for nondestructive testing based on our resonant ultrasound measurement techniques and computer software. The systems under development by Quatro will perform nondestructive inspections of metal, ceramic, composite, and rigid plastic parts in a high-volume manufacturing environment. Quatro is presently working with numerous clients to design custom-engineered systems that meet the particular needs of each application. The Quatro system for testing ball bearings is being adapted to handle thousands of ball bearings per hour and is sensitive to geometry errors as small as 2 parts per million. Another system tests the integrity of oxygen sensors. The commercial production of research electronics, hardware, and software by Quatro has, in turn, assisted us in expanding research efforts and government applications of resonant ultrasound spectroscopy..

shut when mechanically excited on resonance. Ceramic (Si_3N_4) ball bearings, which can withstand very high temperatures, are being developed for use in naval and military aircraft operations. Accurate resonance measurements of these precision objects can reveal tiny internal and surface flaws. As shown in Figure 7, these flaws shift and break the symmetry of the resonances of the ball bearings. As a result, 0.005-micron errors in sphericity of a ball bearing can be detected in seconds, instead of the hour needed with present optical studies. These nondestructive testing techniques, fathered by basic research in materials science at the Laboratory, have been transferred to private industry (Figure 8). Thus new jobs are now being created in Albuquerque as a direct result of the technology we developed here. ■

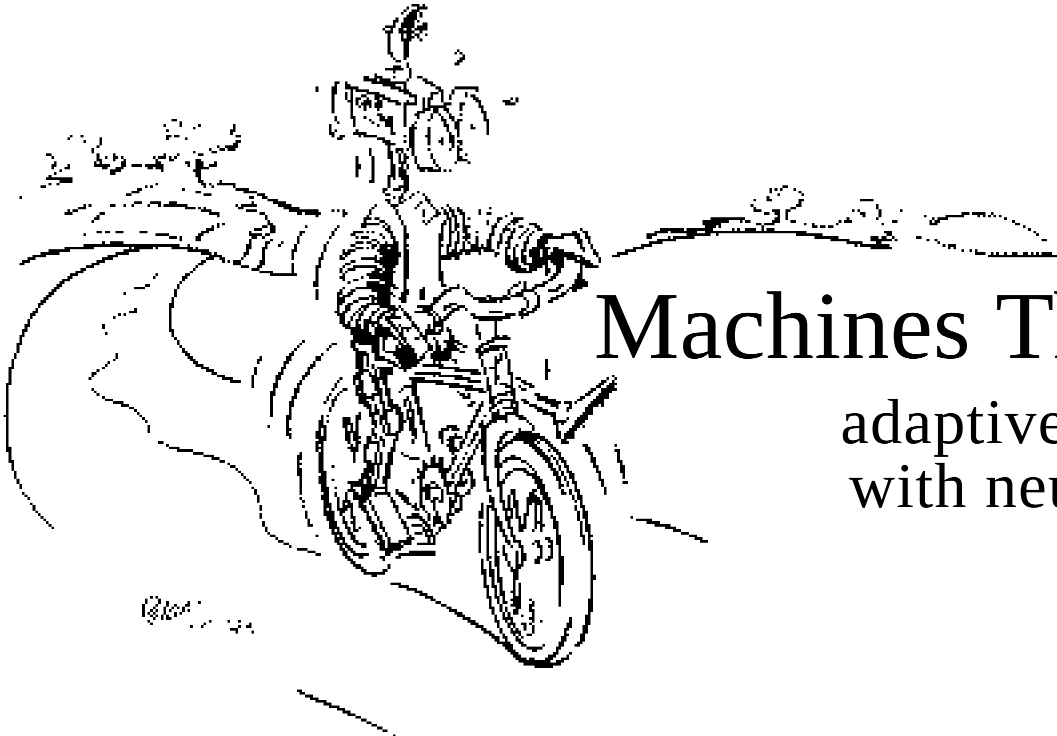
Further Reading

William M. Visscher, Albert Migliori, Thomas M. Bell, and Robert A. Reinert. 1991. On the normal modes of free vibration of inhomogeneous and anisotropic elastic objects. *Journal of the Acoustical Society of America* 90: 2154.

A. Migliori, J. L. Sarrao, William M. Visscher, T. M. Bell, Ming Lei, Z. Fisk, and R. G. Leisure. 1993. Resonant ultrasound spectroscopic techniques for measurement of the elastic moduli of solids. *Physica B* 183: 1–24.

Albert Migliori (left) received his B.S. in physics from Carnegie-Mellow University and his M.S. and Ph.D. in physics from the University of Illinois. He came to Los Alamos as a Laboratory Postdoctoral Fellow, remained as a National Science Foundation Postdoctoral Fellow, and then accepted his current position as a Staff Member in the Condensed Matter and Thermal Physics Group. He was honored by the Laboratory in 1989 with a Distinguished Performance Award. He has authored numerous publications and been awarded forteen patents.

Zachery Fisk was educated at Harvard University and the University of California, Dan Diego. He received his Ph.D. in physics at the latter in 1969. A postdoctoral year at Imperial College, London, was followed by a year as Assistant Professor of Physics at the University of Chicago. He returned to UCSD, becoming Research Physicist and Adjunct Professor of Physics before joining the Laboratory in 1981 as a staff member in the Physical Metallurgy Group of what is now the Materials Science and Technology Division. His research interests include the low-temperature electrical and magnetic properties of metals and the growth of single crystals of these materials. Fisk is a Laboratory Fellow.



Machines That Learn

adaptive computation with neural networks

Roger D. Jones

Early in 1989 I found myself on a team that was responsible for automating the control and tuning of the source for a negative-ion accelerator. The purpose of the source was to create a rapidly pulsed ion beam to be injected into an accelerator. The pulses had to have high current and low noise and had to be reliably similar to each other.

The task presented a number of challenges. The machine was complex. The chemistry and plasma physics in the ion source were sufficiently complicated that modeling was very difficult. Moreover, at given control settings the machine's output drifted significantly in times as short as an hour. Therefore the controller had to adapt constantly to changing conditions inside the machine. Finally, since the source was ultimately intended to operate in space, humans could not be involved in adjusting it.

Earlier attempts at automating the source had failed. On the other hand, despite the complexity of the problem and the drift of the source's response to the control settings, experienced human technicians routinely tuned the machine. Their success suggested the path to the automation of the problem. We had been thinking at the time about computational algorithms, called artificial neural networks, that learn and adapt in crude imitation of the ways humans learn. The ion-source control problem seemed like an ideal test bed for some of our ideas on machine learning. In fact, Bill Mead, P. S. Bowling, and Stan Brown ultimately designed and implemented a controller that used machine-learning techniques to successfully tune and control the beam. The performance of the controller was comparable to that of a human control engineer who had a few months' training on the device.

The control of the negative-ion source is a representative of a class of technical problems that is growing at a tremendous rate. These problems are characterized by the requirement that action be taken as information flows to the controller. The amount of data available is often enormous. The controller is often required to identify the relevant facts from the fire hose of available information and to shape those facts so that meaningful control actions can be taken. Human beings accomplish this task regularly. We constantly learn, adapt, and revise our internal models of the world in order to control our environments more effectively. The challenge is to automate the capability.

In this article we will present some of the basic ideas in machine-learning research and discuss how the techniques can be used to solve practical problems such as the control of the negative-ion source. We will also discuss the direction research in this field seems to be heading.

What Is Machine Learning?

In setting up the ion-source problem for a computer, we represented the control settings by a vector, **x**, whose components corresponded to hydrogen flow rate, arc voltage, cathode temperature, and anode temperature. We defined a fitness function, f, to measure the quality of the beam. Our fitness function was a linear combination of the beam's current, its noise level, and a statistical measure of the variability of the pulses. A large value of the fitness function indicated a desirable beam. The beam quality depended on the control settings. For example, at one instant the flow rate might be set to x_1, the arc voltage set to x_2, and the cathode and anode temperatures set to x_3 and x_4 respectively. For these settings, the beam quality would be $f(x_1,x_2,x_3,x_4)$. The problem was to find a value of the control settings that maximized $f(\mathbf{x})$.

Our method was to reconstruct $f(\mathbf{x})$ from its values at various control settings (about seventy), or more precisely to find a function $\Phi(\mathbf{x})$ that approximates $f(\mathbf{x})$ by interpolating between and extrapolating from those values. The values of $\mathbf{x}$ and $f(\mathbf{x})$ that determine $\Phi(\mathbf{x})$ are called training examples because each pair of values of $\mathbf{x}$ and $f(\mathbf{x})$ is incorporated one at a time into the formula for $\Phi(\mathbf{x})$ and the quality of the approximation should improve with each additional training example. When a good approximation $\Phi(\mathbf{x})$ has been found, standard search techniques can be used to find the value of the control settings that maximizes $\Phi(\mathbf{x})$. Those control settings are a good approximation to the desired control settings that maximize $f(\mathbf{x})$. When a computer goes through this process, we say it "learns" to control the ion source.

Simple methods for function approximation from training examples are well known. For instance, a smooth function can be approximated by a series of sinusoidal func-

$$\Phi(\mathbf{x}; a, \delta) = \sum_{n=0}^{\infty} a_n \cos(n\omega \cdot \mathbf{x} + \delta_n).$$

tions known as a Fourier series:

The cosines are known as the *basis functions* of the Fourier series. The a_js and b_js are adjustable parameters. The frequency, ω, is determined by the spacing of the examples and by the boundary conditions. Suppose we have a number, N, of training examples, $[\mathbf{x}^p, f(\mathbf{x}^p)]$, for p=1 to N, at evenly spaced values of $\mathbf{x}$. Then the best approximation f of the form Φ is given by choosing the adjustable parameters, a and δ, according to

$$a_n = \sqrt{\alpha_n^2 + \beta_n^2}$$

$$\delta_n = \arctan(\beta_n/\alpha_n);$$

and

$$\alpha_n = \frac{1}{\pi}\sum_{p=1}^{N} f(x_p)\cos(n\omega x_p)$$

$$\beta_n = \frac{1}{\pi}\sum_{p=1}^{N} f(x_p)\sin(n\omega x_p).$$

where
and

If the function to be reconstructed varies slowly on scales of the order of the example spacing, then Φ is a good reconstruction of f from N examples. On the other hand, if f varies more rapidly, Φ can be a very poor approximation.

The Fourier series is convenient to compute. Because we can employ a separate computational hardware element to calculate the parameters associated with each sinusoid, we can perform the $2N$ computations for the evaluation of Φ simultaneously in parallel. That method is much faster than waiting for the parameters for each sinusoid to be evaluated in turn. Since each hardware element performs a similar calculation, the construction of each element can be the same. Input information is broadcast to many elements and output information is collected from many elements. Figure 1 shows the architecture of a computer built to perform those calculations; such a computer is known as a network. Network architecture is a powerful concept. It permits the construction of a large class of mappings, $\Phi(\mathbf{x})$, from a large class of examples, $\mathbf{x}$ and $f(\mathbf{x})$, using fixed arrays of similar or identical hardware elements.

Many biological systems take advantage of the network concept. In particular, neurons are simple computational elements which occur in great numbers. Neurons in the brain and spinal cord typically receive information from many other neurons, including input sensors (sensory nerves in the eyes, skin, and so forth), and pass information to many other neurons and to output transducers (muscles). Very simple systems of neurons can carry out sophisticated control actions. More complex arrangements can perform very sophisticated processing such as that involved in reading and understanding this paragraph.

The Fourier series can approximate smooth integrable functions well. It is very useful for handling many simple computational tasks, partly because it can be calculated by a network. However, Fourier approximation, like all approximation techniques based on orthogonal functions, also has three unpleasant properties that prevent its use in the control of the negative-ion source: (1) In the ion-source problem as well as in many other problems of interest, the data are not sampled evenly in the input space. Fourier approximation requires even sampling. (2) We needed to make decisions on how to adjust the controls as the input data came in, whereas Fourier methods require all the data to be in before an approximation is made. (3) We expected to have to

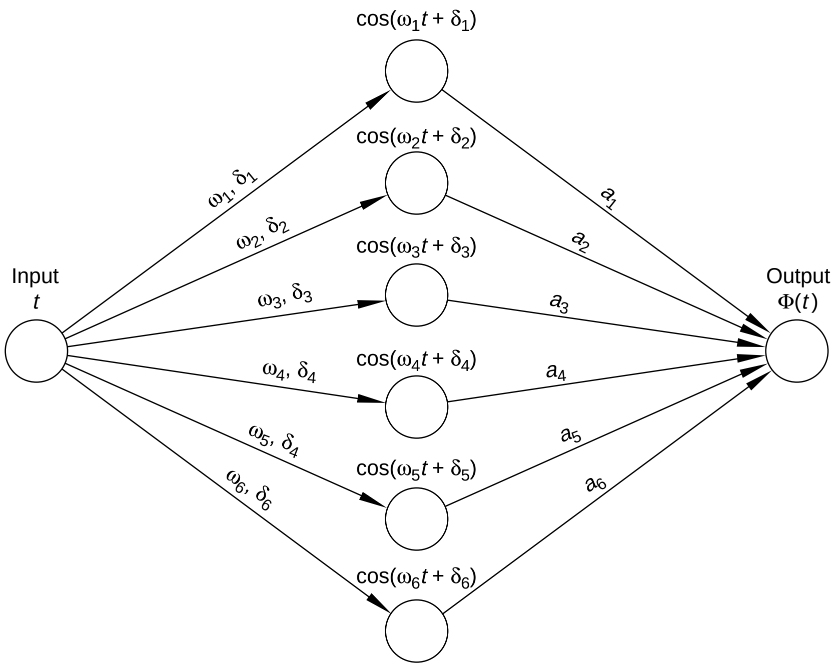


Figure 1. The Network Structure of the Fourier Series

The input and output are each associated with a single node. The sinusoids are identified with the layer of hidden nodes. The frequencies, ω_j, must be calculated before the sinusoids are known, so they are associated with the links between the input node and the hidden nodes. (The frequencies are given by $\omega_n = n\omega$, where ω is calculated from the spacing of the examples and from the boundary conditions.) The adjustable weights of the sinusoids, a_n and δ_n, are associated with the links between the hidden nodes and the output node because the weights for each sinusoid are calculated from the sinusoid and the training examples.

deal with many controls and consequently many input dimensions. The number of basis functions in the Fourier method increases exponentially with the number of dimensions, making Fourier approximation of high-dimensional functions impractical. A further disadvantage of the Fourier series is that it cannot approximate functions that are not smooth, but we would like to approximate such functions in many problems.

The CNLS Network

How do we design a computer-learning method that avoids the disadvantages of the Fourier series and still retain the advantages of multiple similar computational elements? Fourier series have the valuable properties that the basis functions are orthogonal and span the input space. The orthogonality property guarantees the existence of an algorithm to compute the Fourier coefficients in one step. The fact that the space is spanned guarantees that any smooth function can be accurately approximated. It turns out that in many real problems those guarantees are more than we require. If we choose a basis set that "mostly" spans the space and is designed to approximate functions "of the type in which we are currently interested," then we can reduce the number of basis functions in high-dimensional approximations by orders of magnitude. The penalty for using "almost" orthogonal functions is that the adaptive parameters often must be calculated by an iterative process rather than in a single step. However, in typical problems solved by computer learning, the data arrive one example at a time, so the calculation must be iterative anyway.

The use of networks that can be trained was originally suggested by the ways humans learn. The networks were inspired by the networks of nerves in the brain and the process of training from examples was inspired by human learning from examples. Such computational systems are called artificial neural networks because of their analogies to animal nervous systems; biologists insist on the word "artificial" because of the many differences.

To find an appropriate set of nonorthogonal basis functions for an artificial neural network, we start with the identity,

$$f(\mathbf{x}) = \frac{\sum_{j=1}^{M} f(\mathbf{x})\rho_j(\mathbf{x})}{\sum_{j=1}^{M} \rho_j(\mathbf{x})};$$

where M is the number of basis functions. We choose each basis function $\rho_j(\mathbf{x})$ to be a localized function (such as a Gaussian) whose center is at $\mathbf{c}_j$; that is, each basis function is large in the neighborhood in input space around its center and is small outside the neighborhood. If $f(\mathbf{x})$ is a smooth function, it

$$\varphi(\mathbf{x}) = \sum_{j=1}^{M} [a_j + (\mathbf{x} - \mathbf{c}_j)^T \mathbf{d}_j] \frac{\rho_j(\mathbf{x})}{\sum_{k=1}^{M} \rho_k(\mathbf{x})}:$$

can be approximated by its Taylor expansion about $\mathbf{c}_j$.

Approximating $f(\mathbf{x})$ by discarding the quadratic and higher-order terms of the Taylor expansion gives
The superscript T indicates a transpose in input space; in other words, the second term in brackets is a dot product or inner product. Here $\mathbf{a}_j$ corresponds to the zero-order term in the Taylor expansion and $\mathbf{d}_j$ to the gradient term. In principle these terms could be approximated directly from the data. We will, however, regard them as adjustable or adaptive parameters; that is, they will be changed in the process of training. In some cases we will also regard

$$u_j(\mathbf{x}) = \frac{\Omega_j(\mathbf{x})}{\sum_{k=1}^{M} \Omega_k(\mathbf{x})};$$

the basis-function centers, $\mathbf{c}_j$, and the widths of the basis functions as adaptive parameters; in other cases we will simply specify them. In this approximation the basis functions are

The adjustable parameters are changed each time a training example is shown to the network. A *learning rule* computes new parameters from the previous parameters so that the new $\Phi(\mathbf{x})$ approximates $f(\mathbf{x})$ better than the previous $\Phi(\mathbf{x})$. The examples need not be evenly spaced in input space.

What is the appropriate form for Ω? All we have required so far is that the function be localized. We would like to choose Ω in a manner that maximizes the accuracy of the mapping and the ability of the network to generalize, that is, to interpolate between and extrapolate from the data points in a way that uses as much of the information they contain as possible. Information-theory arguments indicate that Gaussians optimize accuracy and generalization. Therefore our network uses Gaussians for the Ω_j functions. A diagram of the network appears in Figure 2. We call it the CNLS network, for Connectionist Normalized Local Spline (alluding to the Laboratory's Center for Nonlinear Studies, where some of the development of the network was done).

Unlike the sinusoids in the Fourier transform, the set of Gaussian basis functions does not span the input space. Therefore not every smooth function can be approximated by this network. However, simple functions (those without too many wiggles) are not a problem. Furthermore the number of basis functions required for approximation of simple functions does not explode when the number of input dimensions is high, as it does when the Fourier series is used.

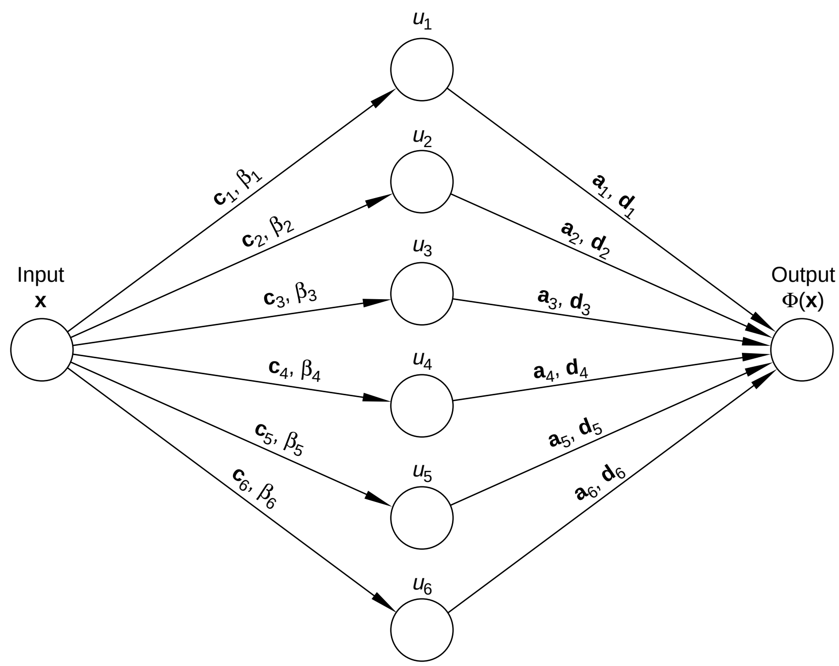


Figure 2. The Network Structure of the CNLS Net
The basis functions, u_j, are normalized Gaussians. Because the basis functions depend nonlinearly on their inverse widths, Ø$_j$, and centers, $\mathbf{c}_j$, those parameters are associated with the links between the input node and the hidden nodes. The linear weights, $\mathbf{a}_j$ and $\mathbf{d}_j$, are calculated after the basis functions are known, so they are associated with the links between the hidden nodes and the output node. In contrast to networks that calculate Fourier approximations, the CNLS net calculates adjustable parameters by an iterative process. Each step of the process alters the parameter values resulting from the previous step to produce new parameter values that better approximate the training examples. The process is complete when additional steps do not appreciably improve the approximation.

Applying the CNLS net to control of the negative-ion source. Given the network architecture and the basis functions, how do we train the adjustable parameters? The best way is to define a cost function that measures the error of the approximation and then minimize that function by adjusting the parameters. Per-

$$E = \frac{1}{2}\sum_{p=1}^{N}[f(x_p) - \Phi(x_p;\alpha)]^2 :$$

haps the simplest and most common cost function is the sum of the squares of the differences between the measured output, f (of the negative-ion source in the present case), and the network output, Φ:
The summation is over the set of training examples and Æ is shorthand for the set of adjustable parameters. This least-squares cost function has a minimum value of 0 when Φ exactly matches f at the training points.

Figure 3 shows the architecture of the controller for the negative-ion source. It consists of an ion-source adjuster, a neural network, a neural-network trainer, and a comparator. Each time a new training example, $[\mathbf{x}_i, f(\mathbf{x}_i)]$, is generated by running the ion source at a particular control setting, $\mathbf{x}_i$, the training example is presented to the network. The network uses standard iterative numerical methods to find the parameter values that minimize the cost function. Because the parameters $\mathbf{a}_j$ and $\mathbf{d}_j$ appear linearly in the equation for ¡, they can be optimized by a simple parallel calculation. The centers and widths of the basis function are less easy to optimize, so in the control of the negative-ion source and many other problems, we choose those parameters from initial data (acquired in an exploratory run or previous runs) and leave them fixed. We choose the starting values for the adjustable parameters from a previous run as well. Once the network has found new $\mathbf{a}_j$'s and $\mathbf{d}_j$'s that minimize the cost function, the adjuster finds a new value of the control settings, $\mathbf{x}_{i+1}$, that maximizes Φ. The calculation then proceeds to the next iteration, in which the pair $[\mathbf{x}_{i+1}, f(\mathbf{x}_{i+1})]$ is the new training example. Thus the data are taken near the true optimum settings, where knowledge of the fitness function is most useful. The controller iterates the process of measuring the fitness function, training the network, and

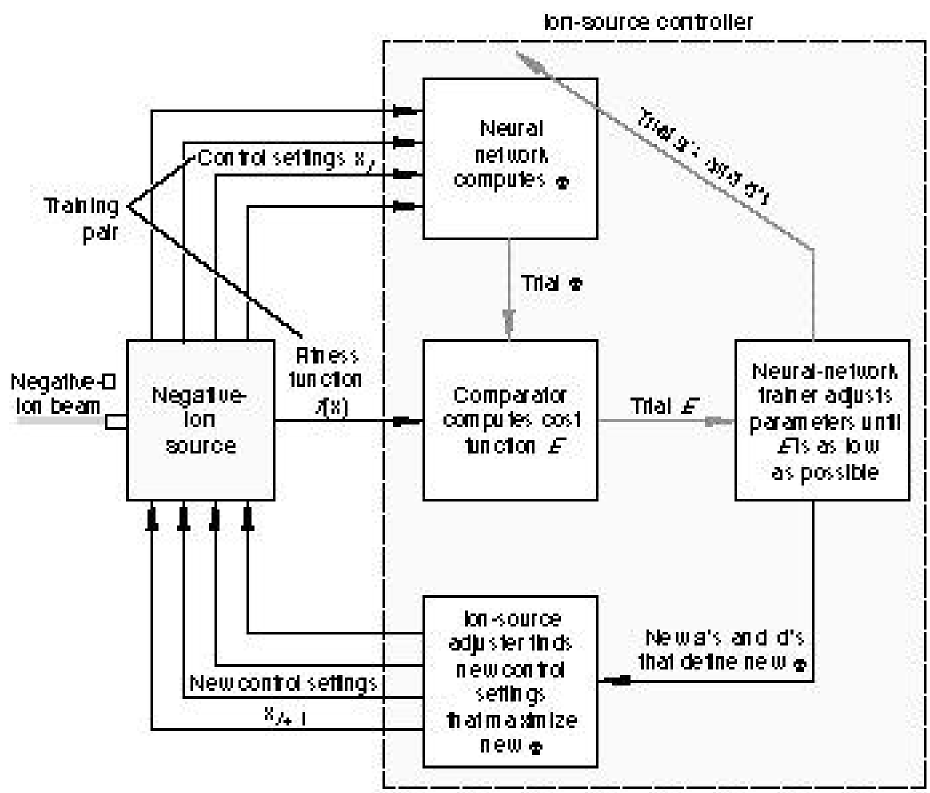


Figure 3. How the CNLS Net Controls the Negative-Ion Source
The figure shows the overall architecture of the negative-ion-source controller. When the negative-ion source runs at control settings $\mathbf{x}_i$, the fitness function (or beam quality) $f(\mathbf{x}_i)$, at those settings is measured and the pair $[\mathbf{x}_i, f(\mathbf{x}_i)]$ is sent to the controller. The neural-network parameters, **a** and **d** (the subscript is omitted for simplicity), are then adjusted by an inner loop (red arrows) as follows. First the neural network calculates $\Phi(\mathbf{x};\mathbf{a},\mathbf{d})$ at all values of **x** at which $f(\mathbf{x})$ is known blah blah blah blah blah. The comparator uses those values of ¡ and all the known values of $f(\mathbf{x})$ to calculate the cost function, E. The neural-network trainer, given that value of E, generates new trial values of **a** and **d** intended to give a smaller E. (In many other applications previous training examples are not stored and the parameters are trained by changing them slightly so that ¡ reproduces the latest example.) The process is iterated until the parameters minimize E; those parameters, a and d´, are sent to the ion-source adjuster. The ion-source adjuster finds the control settings, $\mathbf{x}_{i+1}$, that maximize $\Phi(\mathbf{x};\mathbf{a},\mathbf{d})$ by the straightforward method of computing ¡ at all 7000 possible control settings. The adjuster then begins the next iteration of the outer loop by operating the source at the settings $\mathbf{x}_{i+1}$.

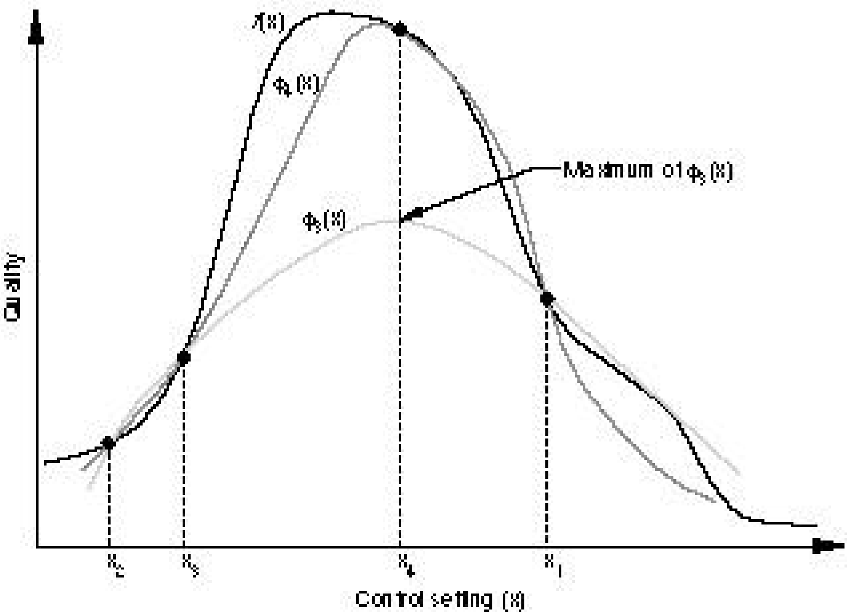


Figure 4. Convergence on the Optimum Control Settings
The graph is a mock example in one dimension of how successive approximations ¡ approach the fitness function $f(\mathbf{x})$ (black) and how the control settings approach their optimal values. Black dots represent training examples, $[\mathbf{x}, f(\mathbf{x})]$. The network approximation, Φ_3, produced by the third iteration is shown in gray. The next training example consists of $\mathbf{x}_4$—the control settings that maximize Φ_3—and $f(\mathbf{x}_4)$. A new approximation, Φ_4 (red), is calculated by the neural network to reproduce as well as possible all the training examples, including the latest one. The control settings for the next iteration, $\mathbf{x}_5$, will be the settings that maximize Φ_4. Thus in each step the network is trained and the true optimum value of $\mathbf{x}$ is approached.

locating the best settings at which to measure the fitness function again, approaching the true optimum settings with each step. Figure 4 illustrates the approach to the optimum. Very near the optimum the controller switches to conventional optimization techniques to fine-tune the beam.

During the experimental runs the dependence of *f* on the inputs continuously changed, but the network successfully tracked the optimum settings as they drifted. In a few instances the machine crashed through no fault of the controller. The controller was able to reoptimize the beam in a short amount of time using the parameter values saved from before the crash, which were generally fairly close to the optimum after the source was restarted.

The CNLS network is the workhorse for many of our applications. There are also other networks and learning algorithms in use. The most popular is known as the multilayer perceptron with backpropagation learning. It is derived from the perceptron, one of the first artificial neural networks, and is more closely based on biological networks than the CNLS network. The multilayer perceptron contains several hidden layers of neurons. When the network is trained, the adjustable parameters of the last hidden layer are adjusted according to the difference between the output and the desired function, *f*. The parameters of each other hidden layer are adjusted according to the adjustments to the following layer. Thus information about the difference between the network's output and the desired output is propagated from the output layer to the first hidden layer, a process called backpropagation.

We have not yet addressed the issue of function smoothness. Fortunately, this issue did not arise in the negative-ion source problem, although it has arisen in a number of other applications. We will discuss this issue in more detail in the section on future directions.

What Else Can They Do?

The potential value of neural networks is illustrated by the variety of problems that Laboratory researchers are already solving with them. For instance, the success of the negative-ion-source controller suggested that similar networks could be applied to process control in general. In fact, we have projects under way with DuPont to develop for chemical processing a set of controllers based on the ion-source controller. The economic impact of improving the efficiency and reducing the waste products from industrial processing could be tremendous. For example, a group of outside consultants performed a benchmarking effort for DuPont. Their results indicate that improved process control could result in an annual savings of $500 million for DuPont alone. These techniques can also be used within the DOE laboratories. Artificial neural networks might improve the efficiency and materials accountability of uranium processing. The Laboratory's Nuclear Chemistry and Analysis Group has developed an adaptive scheduler based on the multilayer perceptron that can optimize the work throughput of automated radionuclide-assay equipment.

Artificial neural networks are by no means limited to process control. For instance, the External Information Technology Group has successfully used the multilayer perceptron to classify ordnance that might be found in the field. In that situation often only partial information (such as a piece of a tailfin) is available, but classification must be very reliable.

Many potential applications of adaptive computing involve predicting a point in a time series based on the previous points. In the late 1980s Alan Lapedes and Rob Farber of the Complex Systems Group demonstrated that the multilayer perceptron could predict the behavior of chaotic time series (generated by the logistic map and the Mackey-Glass equation) with an accuracy comparable to that obtained from more conventional methods. The method required significantly more computational resources, however.

In the summer of 1992, researchers in the Inertial Fusion and Plasma Physics Group and the Center for Nonlinear Studies turned to the problem of predicting highly correlated time series in which the data arrive rapidly. They devised a new learning algorithm that was sensitive to the small differences between consecutive data points. The technique was immediately combined with the CNLS net to attack a control problem of concern to everyone. Researchers in the Applied Theoretical Physics Division were already working with the automotive industry on adaptive controllers. The new algorithm allowed them to start developing a controller that can learn and adjust to road conditions in a fraction of a second. Such a controller could significantly improve automobile safety.

Another interesting application, studied in the Computer Research and Applications Group, is the prediction of tongue and mouth motions of speakers given the sounds of speech. This work could be valuable in speech therapy and possibly in speech understanding by computers. The researchers who performed that work received an R&D 100 Award in 1992. Time-series prediction has also been used to predict successfully when floods will occur in Venice Lagoon.

Database Mining. With the increase in the availability of computation has come an explosion in the amount of information that flows around the world and in near-earth space. This flow is overwhelming the humans who are tasked with extracting information from it. It is also severely taxing the communications networks that support the flow. Some projects at the Laboratory are concerned with the construction of information filters which can automatically and adaptively make some rudimentary sense of very large, possibly noisy data streams.

One industry where information filtering is especially important is the banking industry. In the electronic age money flow is reduced to information flow. There are people at the Laboratory who never touch a checkbook. Their paychecks are deposited automatically and their monthly bills are paid automatically. Goods and services are paid for by bits flowing into and out of accounts. U.S. banks could analyze the manner in which money/information flows through their nodes in order to improve products and services and consequently build up the bit counts in their own accounts instead of seeing the bit counts rise in the accounts of their foreign competitors. On a larger scale, policy makers would like to extract information from the flow in order to have better control over the economy. The Laboratory has a project with the banking industry to build adaptive computers for purposes such as forecasting and individual profitability prediction. This technology will be applied to information extraction from arms-control and nonproliferation databases to identify problems of nuclear proliferation.

Overwhelming amounts of data will also be produced by the Superconducting Super Collider. Information will be flowing from the experiments at a rate several times that at which the human visual system receives information. Researchers in the Complex Systems and Fluid Dynamics groups are developing adaptive techniques to quickly sort through this data explosion.

Another source of overwhelming information flow is visual images. Much more imagery is being generated by satellites and other sources than humans can analyze or even store in our memories. Quite often one would like information to be extracted before the data are transmitted. Researchers in the Analysis and Assessment Division and the Inertial Fusion and Plasma Physics Group are developing an adaptive tracking system which can automatically identify and track interesting items in noisy and cluttered images.

Where Are We heading?

The CNLS net approximates smooth functions reasonably well. In many problems, however, the function to be approximated contains discontinuities and singularities. A common way to calculate discontinuous functions is by iteration with feedback. Figure 5 depicts an iterated network with feedback. Such a network has two inputs, one for the input vector **x** and the other for the output of the previous iteration. After a number of iterations (usually five to twenty), the output settles down to a fixed point; that is, when the output is fed back into the network in combination with the input vector **x**, the network gives the same output that was fed in. The fixed point is the result of the entire calculation. One can think of finding a fixed point as stepping downhill on a potential function; a stable fixed point corresponds to a minimum of the function. Therefore a very small difference in the input can produce a large change in the output by affecting which side of a peak the calculation starts on and consequently which valley it ends in. Figure 6 shows that the CNLS network with feedback can closely match a discontinuous function (a fitness function for a free-electron laser beam) that the unmodified CNLS net approximates rather poorly.

In most of the applications discussed in the previous section, an artificial neural network learns a single-valued function of the inputs, **x**. Quite often, however, real systems can have several states under a single set of conditions, depending on their previous states. Networks with feedback can approximate such multivalued functions. The network is trained so that at a given value of **x**, each of the possible output values is a fixed point. In the first iteration, the two inputs are **x** and a guess supplied by the user. The calculation chooses the fixed point that is reached by going downhill from the initial guess.

To attack still more complex problems, we can generalize the idea of feedback by designing arrays of networks that communicate with each other dynamically. This method permits very sophisticated computation. For instance, a simple array of networks, developed by researchers in the Center for Nonlinear Studies, learned how to control in simulation the balancing of two inverted broomsticks, one on the end of the other. The sticks were confined to move in a vertical plane. The input to each network consisted of the angles and motions of the broomsticks and the outputs of adjacent networks in the array; the output of each network also controlled the motion of a cart to which the bottom end of one broomstick was fixed. The networks were trained only by "punishing" them

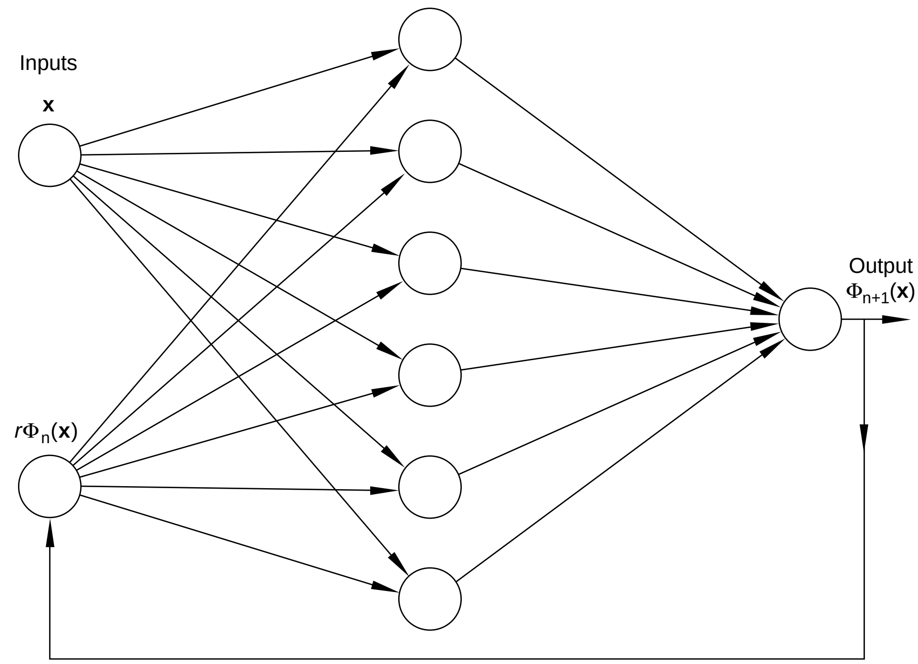
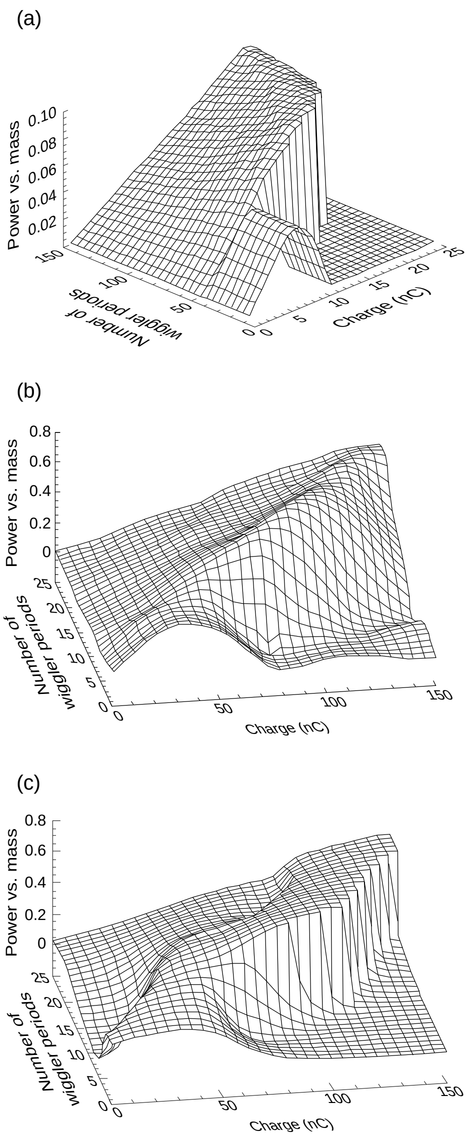


Figure 5. A Network with Feedback
To make an artificial neural network with feedback, we add an input node to the network for each output node. We run the network a number of times, each time using the input vector **x** as input to the original node and the output of the previous run, $\Phi_n(\mathbf{x})$, as input to the new node (or nodes). Before the previous output becomes the input, it is multiplied by a constant, r, that is between 1 and 0. We let Φ_0, the input to the first iteration, be the $\Phi(\mathbf{x})$ produced by a version of the network that does not have feedback. We train a network with feedback by giving it, for each training example [**x**, $f(\mathbf{x})$], the inputs **x** and $rf(\mathbf{x})$. The output is trained to be $f(\mathbf{x})$ using the same methods as for networks without feedback. Thus we make each training example a fixed point of the network.



every time the brooms fell. This problem is too difficult for the networks previously discussed, but the networks in the array arrived at sets of parameters that allowed them to balance the broomsticks. Similar arrays of networks will be required for more difficult learning problems such as language processing, image interpretation, large-scale simulation interpretation, and design of experiments with many parameters.

Figure 6. The Effect of Feedback on Network Approximation of a Discontinuous Function
(a) A two-dimensional cross section of a simulated fitness function for the beam of a free-electron laser. The goal is to find the control settings that maximize the fitness function. Since the maximum of the fitness function is very close to the edge of the "cliff," modeling the cliff accurately is important. (b) The output of a CNLS net without feedback trained with the function shown in (a). The control settings that maximize the approximation differ from the true optimal control settings by 20 percent. (c) The output of a CNLS net with feedback after two iterations, also trained with the function shown in (a). The cliff is much better reproduced. The control settings that maximize the approximation are very close to the true optimal settings.

Concluding Remarks

I had an eerie feeling when I watched the network control the negative-ion source. Pumps and heaters turned off and on and voltages were adjusted while four human observers watched. Watching the beam traces appear on the screen was like watching a rat find its way through a maze. When Bill gave the controller the command to shut the machine down, I expected to see a message on the screen like those from the insubordinate computer HAL in Arthur C. Clarke and Stanley Kubrick's *2001: A Space Odyssey*: "Sorry, Bill, I cannot do that. I am in charge now." I have had similar experiences while watching the networks learn in other applications. It is hard not to think about networks anthropomorphically. The field is still in an early stage of development and it is not clear yet what the capabilities and limitations of machine learning will finally be.

Further Reading

G. M. Bruce, K. Lee, L. A. Lee, W. C. Mead, M. K. O'Rourke, L. E. Thode, Y. C. Lee, G. W. Flake, and I. J. Poli. 1990. Nonlinear adaptive networks: A little theory, a few applications. Los Alamos National Laboratory unclassified release LA-UR-90-273.

R. Hecht-Nielsen. 1989. *Neurocomputing.* Reading, Mass.: Addison-Wesley.

J. Hertz, A. Krogh, and R. G. Palmer. 1991. *Introduction to the Theory of Neural Computing*. Reading: Addison-Wesley.

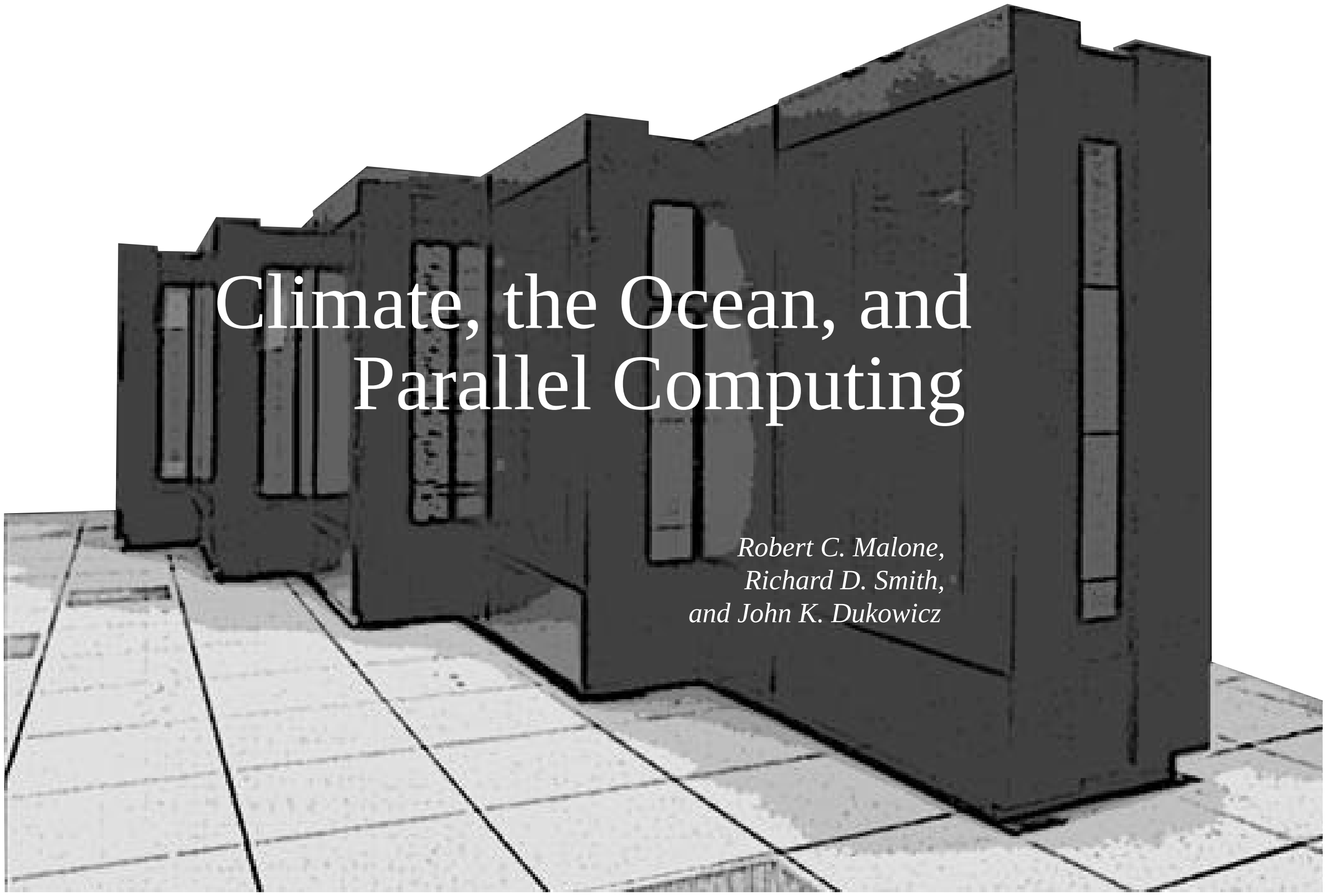
R. D. Jones, Y. C. Lee, C. W. Barnes, G. W. Flake, K. Lee, P. S. Lewis, and S. Qian. 1990. Function approximation and time series prediction with neural networks. In *Proceedings of the International Joint Conference on Neural Networks, San Diego, California, June 17–21, 1990. Volume 1.* New York: IEEE.

P. D. Wasserman. 1989. *Neural Computing: Theory and Practice*. New York: Van Nostrand Reinhold.

Roger D. Jones earned a Ph. D. in physics from Dartmouth College in 1979. He has worked at the Laboratory as a staff member since then. He worked for ten years on inertial-confinement fusion as a designer and physicist. He has made various theoretical contributions to adaptive computation, including participating in the design of the CNLS network as well as studies in image processing and making inferences from databases. He is a member of the Executive Committee of the Center for Nonlinear Studies at the Laboratory. Roger lives in Española and plays bass with local jazz bands.

Climate, the Ocean, and Parallel Computing

Robert C. Malone, Richard D. Smith, and John K. Dukowicz



The Laboratory's most powerful Connection Machine, a CM-5

The climate of the earth is controlled by an interplay among many competing physical processes operating in the atmo-sphere and the ocean and on land. Of the many questions facing today's climatologists, two seem particularly urgent. Is the balance among those processes being affected by human activities? And if so, how large are the resulting climatic changes relative to natural climatic variations? The high priority accorded internationally to answering those questions is impelled in part by the growing list of troubling environmental problems such as greenhouse warming, ozone depletion, pollution of the atmosphere and ocean, and tropical deforestation.

Past observations of the atmo-sphere and the ocean have contributed greatly to our knowledge of the climate system but still constitute only a short and incomplete baseline of data with which future changes can be compared. More extensive observations are in progress or being planned: NASA's Earth Observing System satellite will produce the most comprehensive picture to date of the present state of the earth's climate; the Tropical-Ocean Global-Atmosphere Program is investigating the impact of El Niño and the Southern Oscillation on weather patterns in mid-latitude regions; and the decade-long international World Ocean Circulation Experiment will probe to great depths the circulation in all the major ocean basins.

Such projects are an essential aspect of climate studies, but a theoretical framework is also needed as a basis for interpreting the accumulated data. The computer models known as general-circulation models (GCMs) provide such a basis by simulating the temporal evolution of the atmosphere or the ocean in three dimensions (latitude, longitude, and altitude or depth). In addition, GCMs are indispensable tools for investigating parts of the climate system, such as the deep ocean, that are very difficult to observe and for estimating the effects of natural and human-induced environmental changes on climate. Three-dimensional GCMs were first developed in the 1960s. Their fidelity has since been greatly increased but is still limited by a shortage of data for validating the models and by the capabilities of the computers on which they are implemented.

Atmospheric and oceanic GCMs each contain mathematical representations of the dominant relevant physical processes. Included in an atmospheric GCM are transport of heat and moisture by the winds; exchange of momentum, moisture, and heat between the atmosphere and oceanic and terrestrial surfaces; condensation of moisture into clouds and precipitation; and absorption and scattering of incident sunlight and emission and absorption of infrared radiation by clouds, atmospheric gases, and oceanic and terrestrial surfaces. Also included are factors that affect those processes. For example, sea ice, snow, and vegetation affect energy exchange through their influence on the fraction of incident sunlight absorbed by the ocean and land masses, and the earth's rotation strongly influences the circulation patterns of the winds. Included in an oceanic GCM are interaction of the ocean surface with the winds and with solar and infrared radiation; exchange of fresh water and heat with the atmosphere through evaporation and precipitation; convection driven by temperature and salinity variations; interaction with the edges of continents and islands and with the ocean bottom; and the effect of the earth's rotation on the ocean's circulation. Clearly a reasonably complete description of the climate system—one that couples the dominant physical processes operating in both the atmosphere and the ocean—is enormously complex, and therefore climate simulation taxes the capabilities of even the most powerful of today's supercomputers.

Development of detailed and realistic atmospheric GCMs has been spurred by their use in weather prediction and by the extensive array of satellite- and ground-based equipment put in place over the last few decades to observe atmospheric conditions. Improvements in atmospheric GCMs for weather prediction are directly applicable to GCMs for climate prediction (and vice versa) because the physical processes involved in both are the same. However, prediction of climate requires simulations extending over much longer time intervals (decades to centuries) than does prediction of weather (days to weeks). Therefore the ocean, which varies much more slowly than the atmosphere, can be held fixed in a weather model but must be treated as a dynamical component in a climate model.

Climate can be thought of as the statistical aspects of weather averaged over a period of many years. For example, a weather forecast might tell us whether rain is likely in Los Alamos during the next several days, whereas a climate forecast might tell us whether the springtime precipitation in the midwestern United States, averaged over a decade, will increase or decrease and by how much compared with the present-day average. Although an atmospheric GCM can address both of those questions, the model must be used differently in each case. An atmospheric GCM calculates the temporal evolution of various atmospheric variables (such as temperature, wind velocity, and humidity) at a number of regularly spaced grid points. When an atmospheric GCM is used for weather prediction, the number of grid points must be large (that is, the grid points must be closely spaced, typically less than a hundred kilometers apart horizontally and a kilometer or less apart vertically) to achieve the most accurate prediction possible for a region the size of an average American state. Achieving such fine spatial resolution is very costly in terms of computer time, but a weather prediction need extend only a short time, say a week, into the future. When an atmospheric GCM is used for climate studies, the computing time is kept within reasonable bounds by sacrificing spatial resolution in favor of simulating time intervals of decades or longer.

The atmospheric component of a climate model is essentially a weather-prediction model with a coarse horizontal spatial resolution, typically several hundred kilometers. The atmospheric model generates a time sequence of simulated atmospheric states that can be analyzed statistically to obtain the time averages, variances, and covariances of the atmospheric variables used to describe climate. However, as noted above, the atmosphere is strongly influenced by and interacts with other components of the climate system that evolve more slowly (primarily the ocean but also soil moisture, vegetation, sea ice, and glaciers). One of the challenges of climate modeling is to develop and validate models that adequately represent the more slowly varying components, which can then be coupled to the atmospheric model to form a complete model of the interactive climate system. Another challenge is to determine what spatial scales must be resolved to realistically model the long-term dynamics of climate.

Oceanic GCMs calculate the temporal evolution of oceanic variables on a three-dimensional array of grid points spanning the global ocean domain. Validation of ocean models has been hampered by the scarcity of oceanic data; that situation will be greatly improved by the data on temperature, salinity, and currents to be gathered by the World Ocean Circulation Experiment. (Primarily because they are relatively easy to observe, much is known about the ocean's surface currents. But currents at lower depths are less well explored. Figure 1 shows the major surface currents of the ocean.) Another difficulty of ocean modeling is that the dynamics of the ocean involves longer time scales and smaller spatial scales than does the dynamics of the atmosphere. The deep ocean takes far longer (decades to centuries) to react to external changes than does the atmosphere (months), and oceanic eddies are much smaller horizontally (less than 100 kilometers) than are atmospheric eddies (1000 kilometers or larger). Atmospheric eddies are known to play an important role in the transport of heat and momentum; oceanic eddies are believed to play a similarly important role, although the evidence for that is less conclusive. Therefore ocean simulations not only must extend over long time intervals but also should probably be finely resolved spatially. A question that remains to be answered is whether eddy-resolving

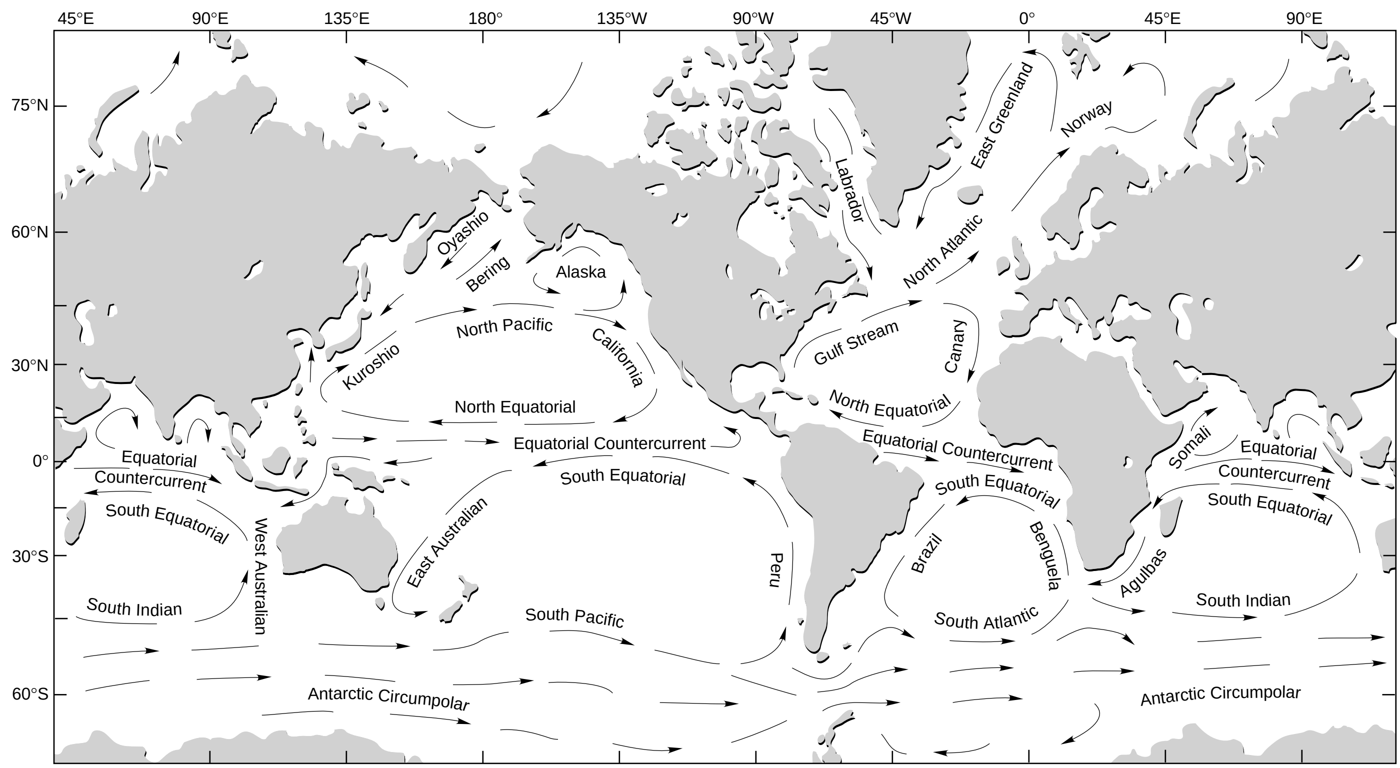


Figure 1. Major Oceanic Surface Currents

Prominent features in this map of the major oceanic surface currents include the subtropical gyres centered on 30° latitude in each of the major ocean basins. The earth's rotation and the change in wind direction with latitude (from the east in the tropics and from the west at mid latitudes) cause the circulation of the gyres to be clockwise in the Northern Hemisphere and counterclockwise in the Southern Hemisphere. The well-known Gulf Stream in the Atlantic and its counterpart in the Pacific, the Kuroshio Current, are strong currents that carry heat northward from the tropics. Both currents are evident in the model simulations displayed in Figures 3 and 4. The Antarctic Circumpolar Current flows around Antarctic in a band of ocean centered on latitude 60°S that is uninterrupted by continents. The Antarctic Circumpolar Current can be clearly seen in Figure 4. (Adapted with permission from a figure in *Principles of Ocean Physics* by John R. Apel, Academic Press, 1987.)

ocean models are necessary to understand climate.

That the ocean is a major component of the climate system is well documented. For example, as shown in Figure 2, the amount of heat transported from the tropics to the polar regions by the ocean is comparable to the amount transported by the atmosphere. In addition, because the ocean, particularly the deep ocean, has such a tremendous heat capacity, it acts as a "thermal flywheel" for the climate system by moderating changes occurring in the atmosphere. For example, by sequestering heat trapped by greenhouse gases such as carbon dioxide and methane, the ocean may be delaying the onset of global warming due to production of those gases by human activities. The ocean also acts as a reservoir for carbon dioxide. It is estimated that the ocean presently holds fifty times as much carbon dioxide as the atmosphere and takes up half of the carbon dioxide released into the atmosphere each year. The ocean's buffering of carbon dioxide plays a central role in the carbon cycle of the earth.

Atmospheric and oceanic GCMs must be coupled to model interactions between the atmosphere and the ocean. Although ocean conditions can be held fixed for short-range weather prediction, for seasonal forecasting (three to six months in advance) and longer-term variations such as El Niño and the Southern Oscillation, atmosphere-ocean interactions must be modeled by treating the ocean as a dynamical entity. El Niño is the name given to the dramatic warming of surface waters of the eastern tropical Pacific Ocean that occurs every three to seven years. Such warmings are now understood to be closely linked physically to episodic shifts, called the Southern Oscillation, in atmo-spheric circulation linking the Indian and eastern Pacific Ocean regions. The combined El Niño–Southern Oscillation phenomenon is an unforced free oscillation of the atmosphere-ocean system. The interactions between ocean and atmosphere that are responsible for an El Niño event are not fully understood, nor are the precursor conditions that initiate an event. El Niño events are of great interest because they not only affect weather in the tropical Pacific region but also seem to influence atmospheric conditions beyond the tropics in significant and potentially predictable ways.

Paleoclimatic data and computer simulations both suggest that shifts in ocean-circulation patterns are associated with changes in climate. At present a global-scale ocean-circulation pattern known as the conveyor belt warms the climate of northern Europe by carrying warm water into the North Atlantic Ocean via the Gulf Stream. That circulation pattern is driven by thermohaline effects (effects related to oceanic temperature and salinity gradients). The warmth of the water being carried into the North Atlantic Ocean enhances its evaporation, which increases the water's saltiness and hence its density. It cools and sinks, forming North Atlantic deep water (NADW). A current of NADW flows southward in the Atlantic Ocean, around Africa, and into the Indian and Pacific oceans, where it slowly mixes and rises toward the surface. Water near the surface flows back through the Indian Ocean, around Africa and northward in the Atlantic Ocean, thus completing the global conveyor belt. Evidence from ice cores and deep-sea sediments suggest that the North Atlantic branch of the thermohaline circulation became active at the end of the last ice age and that fluctuations in its geographical extent and intensity are correlated with changes in atmospheric conditions. Thus the paleoclimatic data, as well as computer simulations with coupled atmospheric and oceanic GCMs, suggest that more than one mode of ocean circulation may exist. In the present-day mode NADW is formed in the North Atlantic and transported to other oceans by the conveyor belt. In another, glacial mode, production of NADW is either partially or completely shut off. An important question is how the northward extent of the conveyor belt might be affected by the warming induced by greenhouse gases. Answering that question demands the best computer models that can be developed.

Using coupled atmospheric and oceanic GCMs to simulate climate is expensive because the simulations must extend over the long times required for the deep ocean to adjust. Many simpler models were developed in the past when computers were much less powerful, and research on simpler models continues today. Unfortunately, simpler models typically include ad hoc assumptions that are often difficult to justify or validate. Furthermore, they may fail to include the effects of feedback within or between components of the climate system that arise from the complexity or nonlinearity of the processes. Thus, although simpler models can be

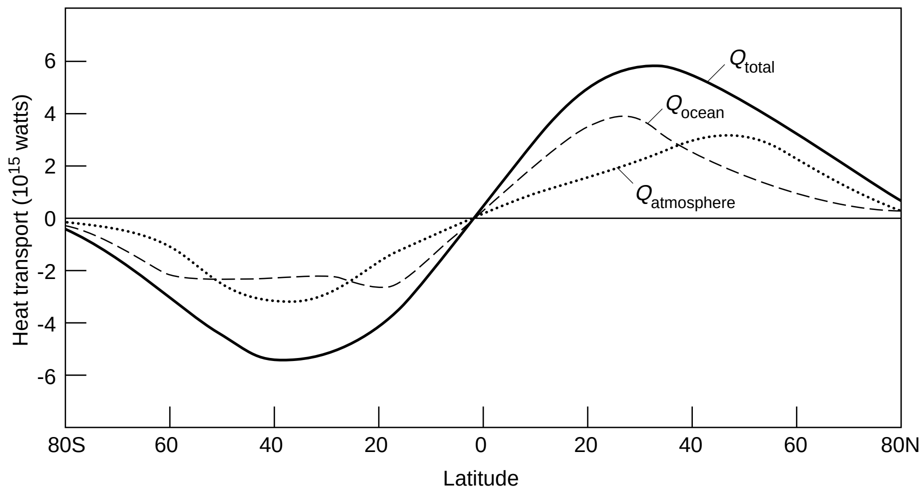


Figure 2. Heat Transport by the Atmosphere and the Ocean
Shown here are the annual mean values of the heat transported by the atmosphere ($Q_{atmosphere}$), by the ocean (Q_{ocean}), and by the atmosphere and the ocean together (Q_{total}), each as a function of latitude. The $Q_{atmosphere}$ values are measured values; Q_{ocean} cannot be measured and must be determined indirectly. To do so, Q_{total} is deduced by computing the total heat transport required to offset the imbalance at various latitudes between incoming solar radiation (dominant in the tropics) and outgoing infrared radiation (dominant in polar latitudes). Q_{ocean} is then obtained by subtracting $Q_{atmosphere}$ from Q_{total}. Positive and negative heat-transport values indicate northward transport and southward transport, respectively. (Adapted with permission from a figure in *Physics of Climate* by José P. Peixoto and Abraham H. Oort, American Institute of Physics, 1992.)

very useful in preliminary investigations, their predictions must ultimately be compared with those of the most realistic models available. Therefore GCMs are still the tool of choice for simulating climate.

The advent of massively parallel computers such as the Connection Machine has provided a new and potentially much more powerful approach to computing. The Department of Energy established the CHAMMP (Computer Hardware, Advanced Mathematics, and Model Physics) Program to pursue development of a new generation of global climate models to be implemented on such computers. It is anticipated that the power and capacity of massively parallel computers will enable future models to include more realistic representations of a greater number of climate-system components. However, fully utilizing the potential of massively parallel computers may require that the mathematical representations of the processes included in climate models be extensively reformulated or that entirely new representations be developed.

A brief explanation of what a massively parallel computer is and how it differs from traditional computers may be helpful here. The words "massively parallel" refer to the fact that such a computer contains hundreds or thousands of processors, all performing their allocated share of the computational work more or less simultaneously. A local memory unit attached to each processor holds the data on which the processor is operating, and a high-speed network connects each processor to the others so that data can be exchanged among the processors whenever required. In contrast to a massively parallel computer, a traditional supercomputer contains only a small number of powerful processors (four to sixteen), all of which share direct and equal access to a global memory bank through a very-high-speed network. Unfortunately, as the number of processors increases, the cost of such a "shared-memory" network becomes prohibitive. Therefore designers of massively-parallel computers are forced to distribute memory so that each processor has fast access to its local memory unit but slower access to all other memory units. Thus, the large number of cooperating processors and the distribution of memory and data across all of the processors are the key features that distinguish a massively parallel computer from traditional supercomputers.

Designing codes to run efficiently on massively parallel computers is more difficult than designing codes to run efficiently on traditional computers, but the possibility of using a very large number of processors is a strong incentive for the extra effort. Codes for massively parallel computers must use mathematical algorithms that divide the work as equally as possible among all the processors; the data must be organized so that most of the data needed by each processor is stored in its local memory; and when data must be exchanged between processors, the least data necessary must be transmitted as efficiently as possible. Developing computer codes with those characteristics is a challenging task.

The availability several years ago of a Connection Machine (a CM-2) in the Laboratory's Advanced Computing Laboratory motivated a long-term project to develop the first global ocean model for massively parallel computers. Albert Semtner of the Naval Postgraduate School in Monterey, California, and Robert Chervin of the National Center for Atmospheric Research in Boulder, Colorado, generously gave us a copy of their ocean model, which had been designed to run on traditional Cray supercomputers. Semtner and Chervin have used their model to perform what are, to date, the highest-resolution simulations of global ocean circulation. The simulations were performed on a grid whose points are 0.5 degree apart in latitude and longitude and located at twenty vertical levels. Such a grid, hereafter referred to as the 0.5-degree grid, is sufficiently fine to begin resolving the oceanic eddies. The Semtner-Chervin model is a variant of the highly regarded and widely used Bryan-Cox-Semtner model, which was originally developed in 1969 by Kirk Bryan of the Geophysical Fluid Dynamics Laboratory/NOAA in Princeton, New Jersey.

Our approach was to first develop a version of the Semtner-Chervin model for use on the CM-2 without changing any of the basic numerical algorithms. Such an approach would allow us to verify that the resulting model functioned properly after being moved to a computer with a different architecture and would provide us with a performance baseline with which future improvements could be compared. However, moving the code from a computer with a few processors and shared memory to one with thousands of processors and distributed memory made certain changes obligatory. The data structures were completely reorganized to improve usage of the CM-2's processors, and the code was entirely rewritten in data-parallel FORTRAN to improve its organization and structure. (Data-parallel FORTRAN extends the old standard, FORTRAN 77, to include new features such as array syntax, which is a simpler and more compact way of expressing operations on data arrays. These features are part of the new standard, FORTRAN 90.) After making the necessary modifications to the code, we found that the performance of the model on the

2048 floating-point processors of the CM-2 was about the same as its performance on a 4-processor Cray X-MP. However, the portion of the code that calculates the vertically averaged ("barotropic") velocity field did not function efficiently on the CM-2, so we were led to reformulate the equations for that portion. We also implemented more efficient algorithms for solving the reformulated equations on parallel computers.

Details of the changes to the model's "barotropic solver" are presented in the sidebar "New Numerical Methods for Ocean Modeling on Parallel Computers." Only their benefits to the physical realism and computational efficiency of the model will be discussed here. First, all eighty of the islands that can be resolved on the 0.5-degree grid can be included in the revised model at the same computational cost required to include the three "islands" used in the original model (Antarctica, Australia, and New Zealand). Second, unlike the original model, the revised model can be executed without smoothing the topography of the ocean bottom to remove steep depth gradients. And third, the revised model does not impose an artificial condition on the ocean surface (the "rigid-lid" condition) that was needed in the original model to eliminate surface waves.

As indicated above, our revisions to the model's barotropic solver have also increased its efficiency. The revised barotropic solver is many times faster than the original (when each is executed on the 0.5-degree grid) even though it treats eighty islands and the original treats only three. (The difference in running times would be even greater, of course, if the original model treated more than three islands.) In its entirety the revised code runs about two and a half times faster on the CM-2 than did the original implementation. The 0.5-degree simulation now runs on 512 floating-point processors (one-fourth) of the CM-200 (an upgraded Connection Machine that is more than 40 percent faster than the CM-2) at about the same speed as the original code runs on a 4-processor Cray X-MP.

We have begun using the new model in global ocean simulations. The simulations are initiated by setting the temperature and salinity to approximate climatological values and the velocity to zero everywhere. The real ocean is driven at the surface by the atmospheric winds and by exchange of heat and fresh water with the atmosphere. We would like to drive the model with accurately measured values of both wind velocities and fluxes of heat and fresh water. The velocity data are available but the flux data are not, at least not on a global basis. However, reasonably complete measurements of temperature and salinity have been made across the surface of the global ocean and, with less accuracy, even in the deep ocean. Therefore measured wind velocities are applied at the model's upper surface, and the temperature and salinity values in the model's topmost layer are continually "nudged" toward climatological values to compensate for the lack of data about heat and fresh-water fluxes. The nudging, which forces the solution toward observed values over a time scale of a month, causes the model's predictions for slowly changing aspects of surface temperature and salinity to correspond closely to the climatological data, but more rapidly changing aspects are hardly affected by the nudging. Because the deep ocean evolves so slowly, a common practice is to nudge the model's temperature and salinity fields there also toward observed values but to do so much more slowly than in the surface layer. That technique was used to produce the simulations presented here because it greatly reduces the computer time required to obtain fairly realistic conditions in the deep ocean.

Figures 3 and 4 show examples of output from the revised ocean model, as executed on the 0.5-degree grid and after 12 simulated years. Figure 3 shows the simulated temperature of the ocean surface. As expected, the large-scale temperature distribution closely resembles the climatological distribution imposed by nudging. But the velocity field and smaller, rapidly evolving features of the temperature distribution can be considered to be predicted by the model. Examples of such features that are evident in Figure 3 are two narrow, meandering western-boundary currents: the Gulf Stream off the eastern coast of the United States and the Kuroshio Current east of Japan. Figure 4 displays the magnitude of the vertically integrated velocity. Predictions of the model that correspond to known ocean features are pointed out in the captions to Figures 3 and 4.

Much more work must be done to thoroughly characterize the circulation patterns predicted by the model and to compare them quantitatively with observations. We hope to use the model to gain insight into aspects of the real ocean that are difficult to observe and, in the future, to couple it to an atmospheric model to study some of the atmosphere-ocean interaction phenomena described earlier.

Although our revisions to the ocean model were motivated primarily by the desire to improve its performance on massively parallel computers, most of the revisions are advantageous even when an ocean model is executed on a traditional supercomputer. Our free-surface representation, for example, is currently being implemented in versions of the Bryan-Cox-Semtner model used at the Naval Postgraduate School and the Geophysical Fluid Dynamics

New Numerical Methods for Ocean Modeling on Parallel Computers

The Bryan-Cox-Semtner ocean model is a three-dimensional model in Eulerian coordinates (latitude, longitude, and depth). The incompressible Navier-Stokes equations and equations for the transport of temperature and salinity, along with a turbulent eddy viscosity, are solved subject to the hydrostatic and Boussinesq approximations. The model includes a rigid-lid approximation (zero vertical velocity at the ocean surface) to eliminate fast surface waves; the presence of such waves would require use of a very short time step in numerical simulations and hence greatly increase the computational cost. The equations of motion are split into two parts: a set of two-dimensional “barotropic” equations describing the vertically averaged flow, and a set of three-dimensional “baroclinic” equations describing temperature, salinity, and deviation of the horizontal velocity components from the vertically averaged flow. (The vertical velocity component is determined from the constraint of mass conservation.) The barotropic equations contain the fast surface waves and separate them from the rest of the model.

The baroclinic equations are solved explicitly; that is, their solution involves a simple forward time-stepping scheme, which is well suited to parallel computing and presents no difficulty on the Connection Machine. On the other hand, the barotropic equations (two-dimensional sparse-matrix equations linking nearest-neighbor grid points) must be solved implicitly; that is, they must be solved at each time step by iteration. For historical reasons the barotropic equations in the Bryan-Cox-Semtner model are formulated in terms of a stream function. Such a formulation requires solving an additional equation for each island, an equation that links all points around the island. The extra equations create vectorization difficulties when the model is implemented on a Cray and serious communication difficulties when it is implemented on a Connection Machine because a summation around each island is required for every iteration of the implicit solver. Therefore all but the three largest islands had been deleted from the original model, even though eighty islands are resolvable at the horizontal resolution employed (0.5 degrees latitude and longitude). Even so the barotropic part of the code consumes about one-third of the total computing time when the model is executed on a Cray and about two-thirds of the total computing time when the model is executed on a Connection Machine.

The above considerations led us to focus our efforts on speeding up the barotropic part of the code. We developed and implemented two new numerical formulation of the barotropic equations, both of which involve a surface-pressure field rather than a stream function. The surface-pressure formulations have several advantages over the stream-function formulation and are more efficient on both parallel and vector computers.

The first new formulation recasts the barotropic equations in terms of a surface-pressure field but retains the rigid-lid approximation. The surface pressure then represents the pressure that would have to be applied to the surface of the ocean to keep it flat (as if capped by a rigid lid). The barotropic equations must still be solved implicitly, but the boundary conditions are simpler and much easier to implement. In addition, islands then require no additional equations, and therefore any number of islands can be included in the grid at no extra computational cost. Furthermore, and perhaps more important, the surface-pressure, rigid-lid formulation, unlike the stream-function, rigid-lid formulation, exhibits no convergence problems due to steep gradients in the bottom topography. The matrix operator in the surface-pressure formulation is proportional to the depth field H, whereas the matrix operator in the stream-function formulation is proportional to $1/H$. As a result, the latter matrix operator is much more sensitive than the former to rapid variations in the depth of waters over the edges of continental shelves or submerged mountain ranges, where the depth may change from several thousand meters to a few tens of meters within a few grid points. Because such a rapidly varying operator may prevent convergence to a solution, steep gradients were removed from the stream-function formulation by smoothing the depth field. The surface-pressure formulation, on the other hand, converges even in the presence of steep depth gradients. Smoothing of the depth field could significantly affect the accuracy of a numerical simulation of the interaction of a strong current with bottom topography. For example, the detailed course and dynamics of the Antarctic Circumpolar Current (the strongest ocean current in terms of total volume transport) is greatly influenced by its interaction with bottom topography.

As we worked with the surface-pressure, rigid-lid model, we noticed a problem in shallow isolated bays such as the Sea of Japan. In principle, we should have been able to infer the elevation of the ocean surface (relative to the mean elevation) from the predicted surface pressure. We found, however, that the surface heights so inferred were quite different from those expected due to inflow or outflow from the bays. Removing the rigid lid solved that problem, but of course it also brought back the undesirable and unneeded surface waves. We were able to overcome that new difficulty by treating the terms responsible for the surface waves implicitly, which artificially slows down the waves, whereas the rigid-lid approximation artificially speeds up the waves to infinite velocity. (Either departure from reality is acceptable: Climate modeling does not require an accurate representation of the waves because they have little effect on the ocean circulation.)

Those considerations led us next to abandon the rigid-lid approximation in favor of a free-surface formulation. The surface pressure is then proportional to the mass of water above a reference level near the surface. The benefits of the surface-pressure, free-surface model are greater physical realism and faster convergence of the barotropic solver. In particular, the revised barotropic part of the code, including eighty islands, is many times faster than the original, including only three islands (when both are implemented on the 0.5-degree grid). In addition, the surface pressure is now a prognostic variable that may be compared to global satellite observations of surface elevation to validate the model, and satellite data may now be assimilated into the model to improve short-term prediction of near-surface ocean conditions.

None of our revisions, of course, changed the fact that the large matrix equation in the barotropic solver must be solved implicitly. We chose to use conjugate-gradient methods for that purpose because they are both effective and easily adapted to parallel computing. Conjugate-gradient methods are most effective when the matrix is symmetric. Unfortunately, the presence of Coriolis terms (terms associated with the rotation of the earth) in the barotropic equations makes the matrix nonsymmetric. By using an approximate factorization method to split off the Coriolis terms, we retained the accuracy of the time-discretization of the Coriolis terms and produced a symmetric matrix to which a standard conjugate-gradient method may be applied. We also developed a new preconditioning method for use on massively parallel computers that is very effective at accelerating the convergence of the conjugate-gradient solution. The method exploits the idea of a local approximate inverse to find a symmetric preconditioning matrix. Calculating the preconditioner is relatively expensive but need be done only once for a given computational grid.

Further Reading

Kirk Bryan. 1969. A numerical method for the study of the circulation of the world ocean. *Journal of Computational Physics* 4: 347–376.

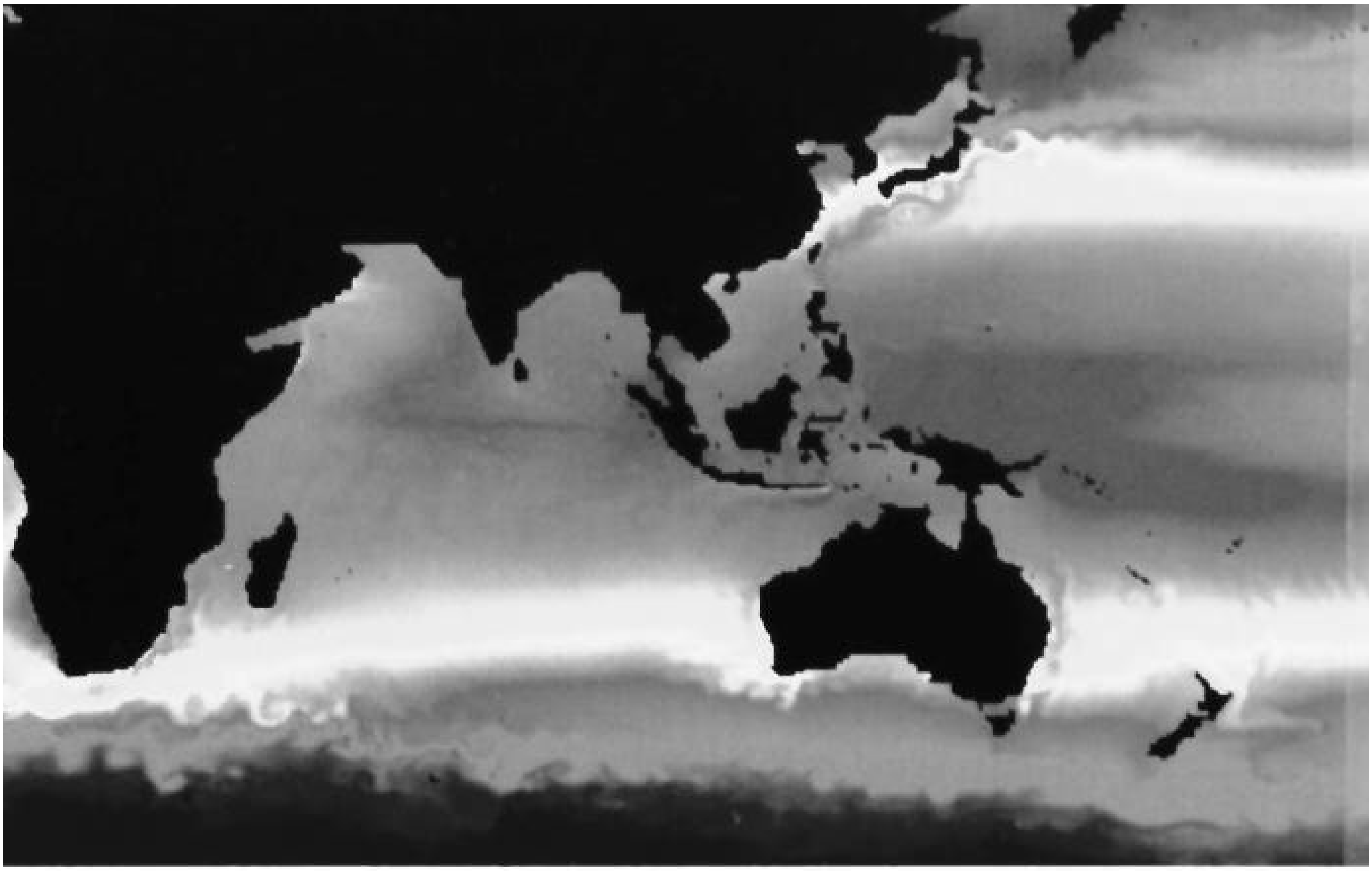
John K. Dukowicz and Arkady S. Dvinsky. 1992. Approximate factorization as a high order splitting for the implicit incompressible flow equation. *Journal of Computational Physics* 102: 336–347.

J. K. Dukowicz, R. D. Smith, and R. C. Malone. 1993. A reformulation and implementation of the Bryan-Cox-Semtner ocean model on the Connection Machine. *Journal of Atmospheric and Oceanic Technology* 10: 195–208.

Albert J. Semtner, Jr. 1976. Finite-difference formulation of a world ocean model. In *Advanced Physical Oceanographic Numerical Modeling*, edited by J. J. O'Brien, 187. Dordrecht, The Netherlands: Reidel.

Albert J. Semtner, Jr., and Robert M. Chervin. 1988. A simulation of the global ocean circulation with resolved eddies. *Journal of Geophysical Research* 93: 15502–15522.

R. D. Smith, J. K. Dukowicz, and R. C. Malone. 1992. Parallel ocean general circulation modeling. *Physica D* 60: 38–61.



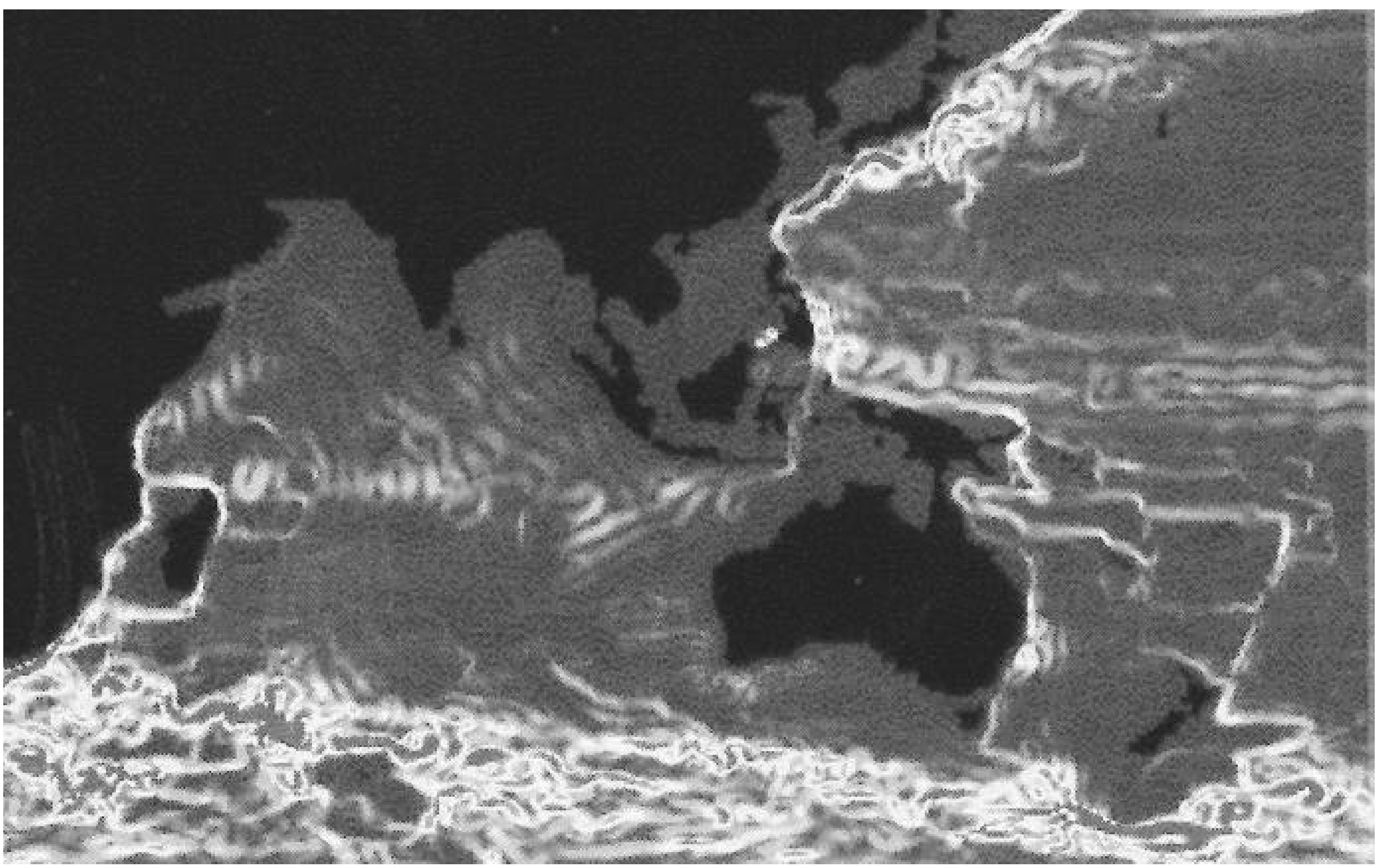


Figure 3. Simulated Oceanic Surface Temperature
The sea-surface temperatures shown here were simulated with our revised ocean model on the CM-200. The temperatures are color-coded from dark red for the hottest to dark blue for the coldest. Continents and islands are black. Meanders and eddies are evident in the warm water being transported poleward in the Gulf Stream along the east coast of North American and in the Kuroshio Current east of Japan. Another interesting feature is the progression of waves in the tropical Pacific Ocean; in movies we have made from the model output, those waves are seen to propagate westward. Similar westward-propagating waves have been observed in satellite measurements of sea-surface temperature. The spatial resolution of the computer model is 0.5° in latitude and longitude with 20 vertical levels; realistic ocean bottom topography is used.

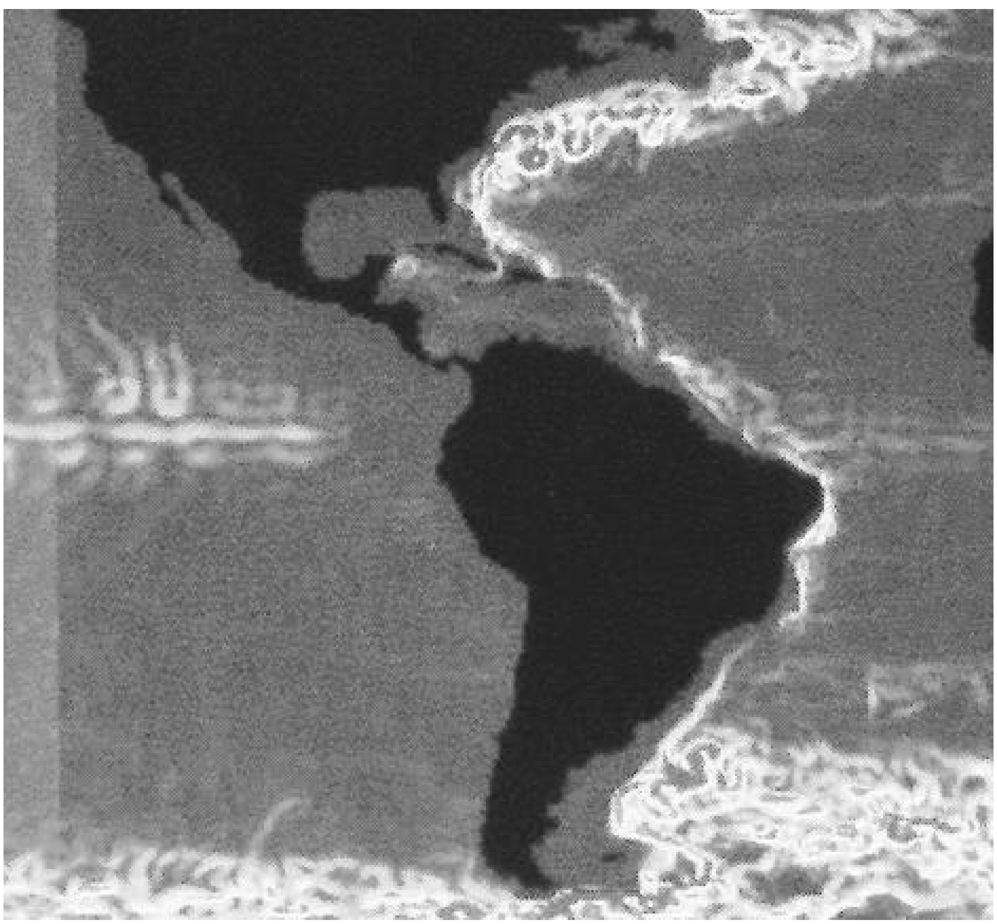


Figure 4. Simulated Vertically Integrated Ocean Currents
Shown here is the magnitude of the vertically integrated horizontal velocity field simulated with our revised ocean model on the CM-200. The speeds are color-coded from dark red for the highest to dark blue for the lowest. Continents and islands are black. Intense flows are evident in the Antarctic Circumpolar Current, the Gulf Stream, and the Kuroshio Current. The influence of submerged topography on the flow is particularly evident east of Australia, south of Alaska, and at several locations along the path of the Antarctic Circumpolar Current.

Laboratory. And, in collaboration with scientists at the latter institution, we are developing a more comprehensive data-parallel version of the model that includes more options for physical parameterizations and is designed to run on a variety of computer architectures.

Further Reading

Wallace S. Broecker. 1987. Unpleasant surprises in the greenhouse? *Nature* 328: 123–126.

Wallace S. Broecker, Dorothy M. Peteet, and David Rind. 1986. Does the ocean-atmosphere system have more than one stable mode of operation? *Nature* 315: 21–26.

Christopher C. Charles and Richard G. Fairbanks. 1992. Evidence from Southern Ocean sediments for the effect of North Atlantic deep-water flux on climate. *Nature* 355: 416–419.

Arnold L. Gordon. 1986. Interocean exchange of thermocline water. *Journal of Geophysical Research* 91: 5037–5046.

Scott J. Lehman and Lloyd D. Keigwin. 1992. Sudden changes in North Atlantic circulation during the last deglaciation. *Nature* 356: 757–762.

S. Manabe and R. J. Stouffer. 1988. Two stable equilibria of a coupled ocean-atmosphere model. *Journal of Climate* 1: 841–866.

S. G. H. Philander. 1983. El Niño Southern Oscillation phenomena. *Nature* 302: 295–301.

Warren M. Washington and Claire L. Parkinson. 1986. *An Introduction to Three-Dimensional Climate Modeling*. Mill Valley, California: University Science Books.

Robert C. Malone (center) received his B.A. in physics from Washington University and his Ph.D. in theoretical physics from Cornell University. He has been a Staff Member at the Laboratory since 1973, carrying out research on laser-fusion and magnetic-confinement-fusion theory, atmosphere and climate modeling, and computer science. He has also fulfilled various managerial roles. He has been honored by the Laboratory with a Distinguished Performance Award and currently directs model development for the DOE's Computer Hardware, Advanced Mathematics, and Model Physics Program.

Richard D. Smith (right) received his B.A. in liberal arts from St. John's College and his Ph.D. in theoretical physics from the University of Maryland. He was a Postdoctoral Fellow at Lawrence Livermore National Laboratory before coming to Los Alamos in 1987. He became a Staff Member in 1990 and has since been honored by the Laboratory with a Distinguished Performance Award. His research interests include computational fluid dynamics, parallel computing, and climate modeling.

John K. Dukowicz (left) received his B.A.Sc., M.A.Sc., and Ph.D in aerophysics from the University of Toronto. After advancing to the position of Principal Aerophysicist at Cornell Aeronautical Laboratory, he transferred first to the Mitre Corporation as a Technical Staff Member and then to General Motors Research Laboratory as a Senior Research Engineer. He joined the Laboratory as a Staff Member in 1977 and has since been honored by the Laboratory with a Distinguished Performance Award and an appointment as a Laboratory Fellow.

New Numerical Methods for Ocean Modeling on Parallel Computers

The Bryan-Cox-Semtner ocean model is a three-dimensional model in Eulerian coordinates (latitude, longitude, and depth). The incompressible Navier-Stokes equations and equations for the transport of temperature and salinity, along with a turbulent eddy viscosity, are solved subject to the hydrostatic and Boussinesq approximations. The model includes a rigid-lid approximation (zero vertical velocity at the ocean surface) to eliminate fast surface waves; the presence of such waves would require use of a very short time step in numerical simulations and hence greatly increase the computational cost. The equations of motion are split into two parts: a set of two-dimensional "barotropic" equations describing the vertically averaged flow, and a set of three-dimensional "baroclinic" equations describing temperature, salinity, and deviation of the horizontal velocity components from the vertically averaged flow. (The vertical velocity component is determined from the constraint of mass conservation.) The barotropic equations contain the fast surface waves and separate them from the rest of the model.

The baroclinic equations are solved explicitly; that is, their solution involves a simple forward time-stepping scheme, which is well suited to parallel computing and presents no difficulty on the Connection Machine. On the other hand, the barotropic equations (two-dimensional sparse-matrix equations linking nearest-neighbor grid points) must be solved implicitly; that is, they must be solved at each time step by iteration. For historical reasons the barotropic equations in the Bryan-Cox-Semtner model are formulated in terms of a stream function. Such a formulation requires solving an additional equation for each island, an equation that links all points around the island. The extra equations create vectorization difficulties when the model is implemented on a Cray and serious communication difficulties when it is implemented on a Connection Machine because a summation around each island is required for every iteration of the implicit solver. Therefore all but the three largest islands had been deleted from the original model, even though eighty islands are resolvable at the horizontal resolution employed (0.5 degrees latitude and longitude). Even so the barotropic part of the code consumes about one-third of the total computing time when the model is executed on a Cray and about two-thirds of the total computing time when the model is executed on a Connection Machine.

The above considerations led us to focus our efforts on speeding up the barotropic part of the code. We developed and implemented two new numerical formulation of the barotropic equations, both of which involve a surface-pressure field rather than a stream function. The surface-pressure formulations have several advantages over the stream-function formulation and are more efficient on both parallel and vector computers.

The first new formulation recasts the barotropic equations in terms of a surface-pressure field but retains the rigid-lid approximation. The surface pressure then represents the pressure that would have to be applied to the surface of the ocean to keep it flat (as if capped by a rigid lid). The barotropic equations must still be solved implicitly, but the boundary conditions are simpler and much easier to implement. In addition, islands then require no additional equations, and therefore any number of islands can be included in the grid at no extra computational cost. Furthermore, and perhaps more important, the surface-pressure, rigid-lid formulation, unlike the stream-function, rigid-lid formulation, exhibits no convergence problems due to steep gradients in the bottom topography. The matrix operator in the surface-pressure formulation is proportional to the depth field H, whereas the matrix operator in the stream-function formulation is proportional to $1/H$. As a result, the latter matrix operator is much more sensitive than the former to rapid variations in the depth of waters over the edges of continental shelves or submerged mountain ranges, where the depth may change from several thousand meters to a few tens of meters within a few grid points. Because such a rapidly varying operator may prevent convergence to a solution, steep gradients were removed from the stream-function formulation by smoothing the depth field. The surface-pressure formulation, on the other hand, converges even in the presence of steep depth gradients. Smoothing of the depth field could significantly affect the accuracy of a numerical simulation of the interaction of a strong current with bottom topography. For example, the detailed course and dynamics of the Antarctic Circumpolar Current (the strongest ocean current in terms of total volume transport) is greatly influenced by its interaction with bottom topography.

As we worked with the surface-pressure, rigid-lid model, we noticed a problem in shallow isolated bays such as the Sea of Japan. In principle, we should have been able to infer the elevation of the ocean surface (relative to the mean elevation) from the predicted surface pressure. We found, however, that the surface heights so inferred were quite different from those expected due to inflow or outflow from the bays. Removing the rigid lid solved that problem, but of course it also brought back the undesirable and unneeded surface waves. We were able to overcome that new difficulty by treating the terms responsible for the surface waves implicitly, which artificially slows down the waves, whereas the rigid-lid approximation artificially speeds up the waves to infinite velocity. (Either departure from reality is acceptable: Climate modeling does not require an accurate representation of the waves because they have little effect on the ocean circulation.)

Those considerations led us next to abandon the rigid-lid approximation in favor of a free-surface formulation. The surface pressure is then proportional to the mass of water above a reference level near the surface. The benefits of the surface-pressure, free-surface model are greater physical realism and faster convergence of the barotropic solver. In particular, the revised barotropic part of the code, including eighty islands, is many times faster than the original, including only three islands (when both are implemented on the 0.5-degree grid). In addition, the surface pressure is now a prognostic variable that may be compared to global satellite observations of surface elevation to validate the model, and satellite data may now be assimilated into the model to improve short-term prediction of near-surface ocean conditions.

None of our revisions, of course, changed the fact that the large matrix equation in the barotropic solver must be solved implicitly. We chose to use conjugate-gradient methods for that purpose because they are both effective and easily adapted to parallel computing. Conjugate-gradient methods are most effective when the matrix is symmetric. Unfortunately, the presence of Coriolis terms (terms associated with the rotation of the earth) in the barotropic equations makes the matrix nonsymmetric. By using an approximate factorization method to split off the Coriolis terms, we retained the accuracy of the time-discretization of the Coriolis terms and produced a symmetric matrix to which a standard conjugate-gradient method may be applied. We also developed a new preconditioning method for use on massively parallel computers that is very effective at accelerating the convergence of the conjugate-gradient solution. The method exploits the idea of a local approximate inverse to find a symmetric preconditioning matrix. Calculating the preconditioner is relatively expensive but need be done only once for a given computational grid.

Further Reading

Kirk Bryan. 1969. A numerical method for the study of the circulation of the world ocean. *Journal of Computational Physics* 4: 347–376.

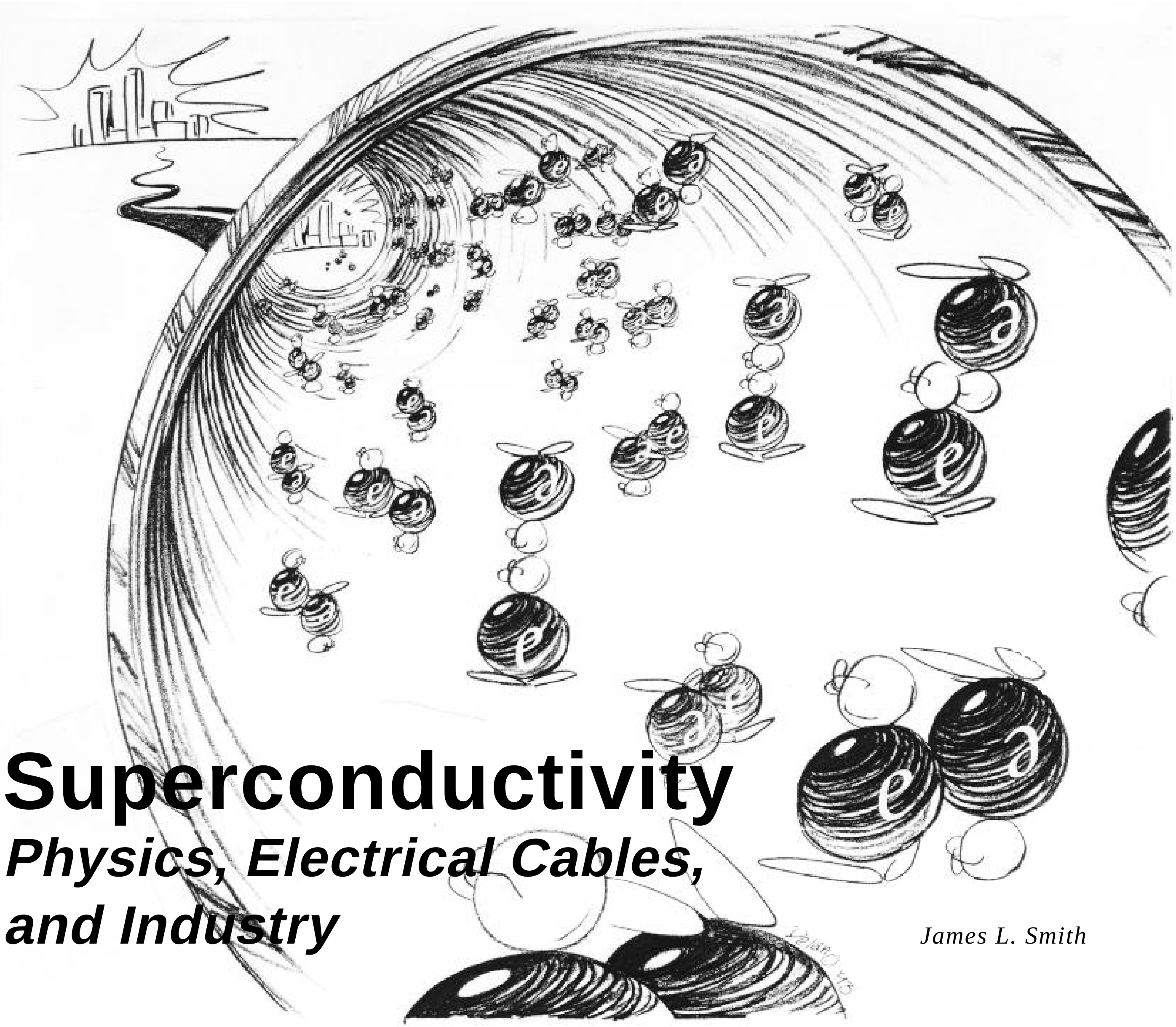
John K. Dukowicz and Arkady S. Dvinsky. 1992. Approximate factorization as a high order splitting for the implicit incompressible flow equation. *Journal of Computational Physics* 102: 336–347.

J. K. Dukowicz, R. D. Smith, and R. C. Malone. 1993. A reformulation and implementation of the Bryan-Cox-Semtner ocean model on the Connection Machine. *Journal of Atmospheric and Oceanic Technology* 10: 195–208.

Albert J. Semtner, Jr. 1976. Finite-difference formulation of a world ocean model. In *Advanced Physical Oceanographic Numerical Modeling*, edited by J. J. O'Brien, 187. Dordrecht, The Netherlands: Reidel.

Albert J. Semtner, Jr., and Robert M. Chervin. 1988. A simulation of the global ocean circulation with resolved eddies. *Journal of Geophysical Research* 93: 15502–15522.

R. D. Smith, J. K. Dukowicz, and R. C. Malone. 1992. Parallel ocean general circulation modeling. *Physica D* 60: 38–61.



Superconductivity

Physics, Electrical Cables, and Industry

James L. Smith

It does not surprise us that we can see through glass. After all, we say, glass is transparent to the light that travels to our eyes from objects on its far side. But what does it mean to say that glass is "transparent" to photons of visible light? It means that the photons interact very little with the electrons in glass. And why do they interact so little? Because the spectrum of possible energy levels of the electrons has a gap extending over the range of energies possessed by photons of visible light. Similarly, the electrons of a superconducting material that move through the material under the influence of an applied voltage do not undergo collisions within the material because the spectrum of possible energy levels of the mobile electrons has a gap at an energy such that the collisions cannot occur. A superconductor is "transparent" to the flow of electricity.

Since collisions between the mobile electrons and the fixed electrons within a conductor are the source of the conductor's electrical resistance, superconductivity is certainly a property to be desired of wires used to carry electricity. But superconductivity has its drawbacks: The phenomenon occurs only at low temperatures, in low magnetic fields, and when currents are low. However, during the eighty-two years since mercury was found to become superconducting at about 4 kelvins, many superconducting materials have come to light for which those limitations are reduced. For example, in 1987 a material was discovered that becomes superconducting at a temperature above the boiling point of liquid nitrogen (77 kelvins, or –320°F). The discovery of "high-temperature" superconductivity was a momentous event because liquid nitrogen is very much cheaper than the helium required to cool previously known superconductors to their transition temperatures.

By July of 1988 Congress was infused with high-temperature-superconductivity fever—a feeling that the technology of high-temperature superconductors, unlike that of semiconductors, must not be lost to foreigners, that the United States had another chance to do it right, that here was an opportunity for government to help industry. The efforts of New Mexico Senator Pete Domenici and Laboratory Director Sig Hecker led to legislation that established Superconductivity Pilot Centers at Argonne, Los Alamos, and Oak Ridge national laboratories. Special agreements between the pilot centers and industry were allowed that gave more rights to industry, more protection to ideas. This was to be a test case of marriage between government and industry.

But making wire capable of carrying practical amounts of current from the high-temperature superconductors proved daunting. Nature had charged a price for superconductivity at higher temperatures. Among the possible wire materials are the copper-oxide compounds $YBa_2Cu_3O_7$, $Bi_2Sr_2Ca_2Cu_3O_{10}$, and $Tl_2Ba_2Ca_2Cu_3O_{10}$ and many doped versions of those compounds. They all exhibit "weak links" or "flux creep." Weak links—the imperfect bits of material between the perfect crystalline grains that make up all wires—are so called because their superconductivity is more subject to degradation by magnetic fields and currents than that of the perfect grains. The solution is to find a processing protocol that will at least partially align the imperfect grains and reduce the amount of intergranular material. Some progress has been made. Flux creep refers to the motion of magnetic-flux quanta—tiny whirlpools of current that exist throughout a current-carrying superconducting material. The motion occurs when the material is carrying a practical current. Flux creep has the same effect as electrical resistance and thus defeats the purpose of using a superconductor. Progress has been made in combating flux creep with line-like (one-dimensional) imperfections that "pin" the flux quanta against movement. Acting on ideas from IBM, workers at Oak Ridge and Brookhaven national laboratories have made tiny holes in superconductors with energetic, highly ionized atoms and found that such holes do pin flux effectively. The fact that each known high-temperature superconductor exhibits either weak links or flux creep but rarely both is tantalizing, as is the lack of any scientific reason to rule out the possibility of a high-temperature superconductor that exhibits neither problem. It is important to note that over the last five years incremental success at licking those difficulties has occurred around the world. There have been no touchdowns, but progress has been steady.

The disappearance in high-temperature superconductors of both weak links and flux creep at temperatures well below their transition temperatures demonstrates that those phenomena are indeed the price that must be paid for high-temperature superconductivity. At temperatures in the vicinity of the transition temperatures of the older superconductors (say below 20 kelvins), the high-temperature superconductors do retain their lack of electrical resistance in record high magnetic fields. That property may prove to be very useful when the cost of cooling to the lower temperatures must be paid for other reasons, and we are working to optimize wire for such applications. For example, the ability of high-temperature superconductors to maintain their superconductivity in very high magnetic fields at temperatures below their transitions temperatures might prove advantageous for the Superconducting Super Collider, which must be cooled to very low temperatures to maintain its high vacuum. At the Laboratory last winter, we used very high magnetic fields (produced by compressing copper carrying a large current with explosives) to measure the field required to drive $YBa_2Cu_3O_7$ out of its superconducting state at the lower temperatures. We found that the field required is very high indeed—140 teslas, or 3 million times the earth's magnetic field. Only the Russians at Arzamas-16 (one of their nuclear-weapons laboratories) have carried out a similar experiment; they reported a slightly higher magnetic field. Apparently the competition between the weapons laboratories in the United States and Russia has not stopped with the end of the Cold War.

The excitement created by the discovery of high-temperature superconductivity has now faded, and Edison's 99-percent perspiration is somewhere near half spent. American Superconductor and Intermagnetics General, industrial partners of the three national laboratories, are producing strips of flexible, silver-clad tape more than a hundred meters long that can carry over 10 amperes even after being wound on a drum and then unwound. It is clear that the cost of feeding electricity to a major new building (one requiring, say, 20 megawatts of power) in a crowded city (Boston or New York, say) with an underground cable of such superconducting wire would be similar to the cost of feeding with an oil-cooled cable, the type of cable now used despite the risk of oil leaks. In a city where electricity generation is physically distant from new construction (such as Baltimore, where electricity generation is to the south and growth is to the north), superconducting cables with twice the current-carrying capacity could replace existing underground cables and thus eliminate the need for overhead lines around the edge of the city or for miles of new excavation through the city. Furthermore, construction of new high-voltage overhead lines may be halted by concerns of the American public about their aesthetics and safety, and underground superconducting cables now offer a solution at the right time.

In the future, when weak links and flux creep are licked and materials are available that maintain their high-temperature superconductivity in higher magnetic fields, electric motors can be lighter and more efficient and electric energy can be accumulated at night as a magnetic field in large coils and used in the day when demand is greater. Here much work and the risk of failure remain. Even the effects of mechanical strain and the details of propagation of thermal transients through coils made of high-temperature superconductors are subjects that remain largely untouched by researchers. The modeling capabilities of the national laboratories and their abilities to field large-scale tests can help industry, and we have been talking to companies about those areas of concern.

There remain "cultural" differences between national laboratories and corporate America. We tend to want to understand down to the last detail; they need understand only enough to get a product on the market. And researchers at the national laboratories are not accustomed to paying attention to proprietary information. Yet company representatives all seem to want access to the enthusiasm, ideas, and capabilities of our scientists and seem willing to put up with managerial and legal obstacles to gain that access. They do understand that high-risk development can be the charge of the Laboratory, that the design engineering of a competitive product comes later, and that the Laboratory can therefore make most of its findings available to all.

The experiment of government helping industry is continuing, and apparently government has the most to learn, just as we expected. Steadily albeit slowly, the problems of high-temperature superconductivity are being solved. ■

James L. Smith received his B.S. in physics from Wayne State University and his Ph.D. in physics from Brown University. He is currently Chief Scientist of the Laboratory's Superconductivity Technology Center. During his twenty years at the Laboratory, he has received numerous awards, including two Distinguished Performance Awards from the Laboratory, the International Prize for New Materials from the American Physical Society, and an E. O. Lawrence Award from the DOE. He is a Laboratory Fellow, a Fellow of the American Physical Society, and the North American Editor of the venerable *Philosophical Magazine*. His continuing research interest is the study of superconducting materials from high to extremely low temperatures and in very high magnetic fields.

Science Policy and the Role of the National Laboratories

Robert W. Seidel

Federal science policy and funding have supported Los Alamos for fifty years. Now that the Cold War has ended, the Laboratory faces a dramatically different world. Government leaders either will set forth new missions for the Department of Energy (DOE) laboratories or will suffer their obsolescence. What is the future science policy of the United States? How will the national laboratories adapt to changing roles? An examination of the fortunes of science and technology in the United States in general, and in the Laboratory in particular, may help us to answer those questions.

Science and the Federal Government before World War II

Scientific activity is granted certain federal provisions in the United States. The Constitution provides for federal regulation of patents and trademarks, establishment of standard weights and measures, and promotion of science and useful arts. Before World War II, however, those powers were used sparingly. The progress of science was incidental to the pursuit of other national ends.

One such end was exploration. Lewis and Clark's overland expedition to the Pacific, for example, was supported with the aim of providing strategic, geographical, and scientific knowledge about the new nation. Charles Wilkes' nautical expedition to Antarctica and Pacific Islands from 1838 to 1842 helped to develop the nation's maritime economy. Those early expeditions incidentally enriched the nation's storehouse of scientific information. The American Philosophical Society of Philadelphia and the Smithsonian Institution were among the beneficiaries of that increase in scientific knowledge.

Another end that entailed scientific progress was national security. Scientific and technological education found a home in the military academies. West Point, modeled on the French military school *Ecole Polytechnique*, supplied the bulk of engineering talent for the country's canal and railroad systems in early 19th-century America. The Federal arsenals were also a seedbed for technological innovation. Eli Whitney's idea of using interchangeable parts in arms manufacture quickly spread to the manufacture of machine tools, pocket watches, agricultural implements, sewing machines, bicycles, and automobiles. The "American System of Manufactures," was an early, important, but by no means unique, example of technology transfer from the military to the civilian industrial sector.

The Civil War marked a turning point in federal science policy. The Morrill Land-Grant College Act initiated federal support for higher education and the Department of Agriculture, created in 1862, explicitly sponsored scientific research. The department developed programs related to the cure of plant diseases, introduced new crops, and established a system of agricultural experiment stations to promote scientific farming techniques. By World War I the Department of Agriculture had almost two thousand employees engaged in scientific research, far more than any other federal agency. On the eve of World War II, the department was the largest patron of scientific R&D in the federal government. Research flourished in conjunction with the primary economic activity of the nation.

In contrast to its contributions to agriculture, science as a whole had little to offer military and industrial technology. The National Academy of Sciences (NAS) was chartered in 1863 to advise the government on possible applications of science to war, but made no substantial contributions to the Civil War effort. The growing industrial enterprises of the nation supported their own research; laissez faire economics did not endorse government intervention in industrial development. Thomas Edison's small laboratory at Menlo Park, Standard Oil of New Jersey's petroleum research, and Alexander Graham Bell's work on the telephone were all forerunners of industrial research laboratories and instruments of monopolies that resisted any attempts at government interference.

The federal agencies created in the 19th century supported science unevenly. No overall science policy governed the work of those agencies. A proposal made in 1884 to establish a federal department of science won no support from Congress. The nation's patronage of science was shaped by the exigencies of settling a continent, preparing for war, supporting agriculture, and extending commerce.

That fragmented national science policy began to change at the turn of the century. By that time, the growth of American manufactures had created a pressing need for standardized weights and measures. Consequently, in emulation of other industrialized powers, Congress created the National Bureau of Standards (NBS) in 1901. The NBS brought science to bear on some new technologies of the industrial revolution.

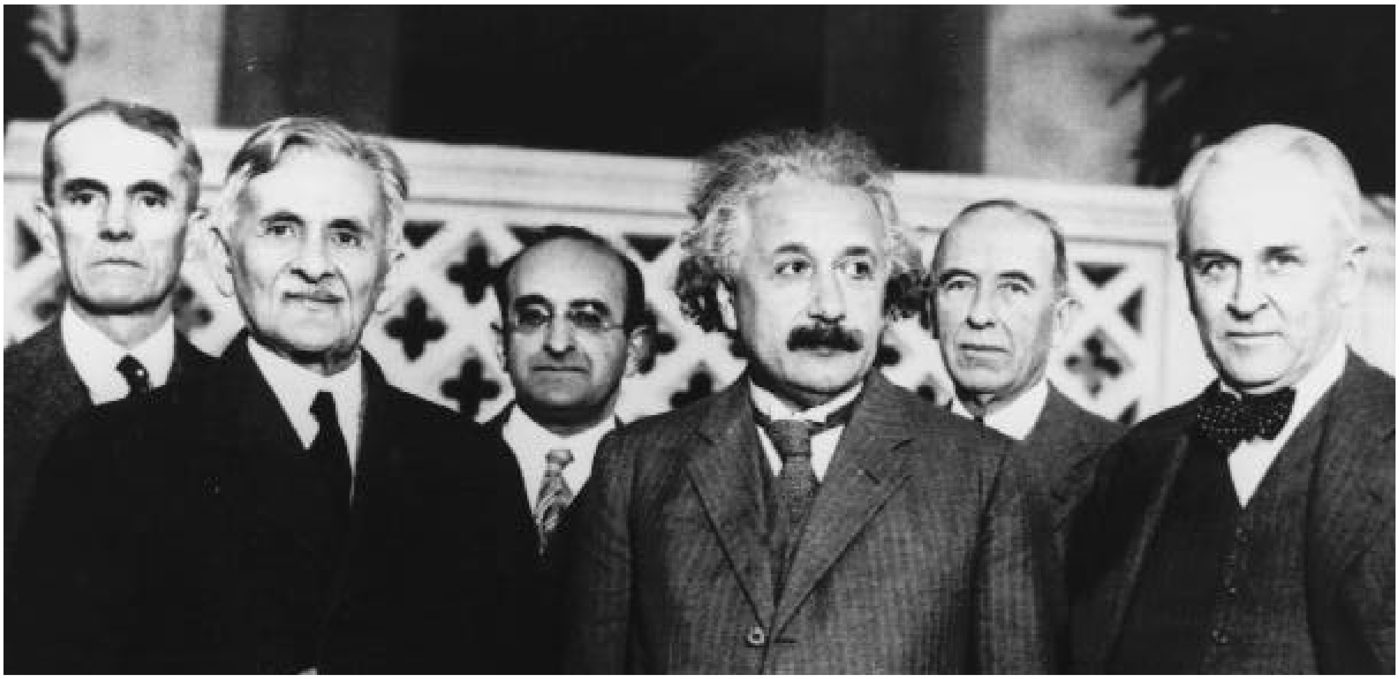
Federal scientific research expanded with American entry into World War I. The NBS undertook research in aeronautical engineering, optics, radio, and armaments design. In addition, the National Academy of Sciences mobilized university research through its National Research Council (NRC) which investigated problems related to sonar, artillery range-finding, and chemical warfare. The Army's Chemical Warfare Service gave academic chemists particular prominence. A Naval Consulting Board led by Thomas Edison recommended the formation of the Naval Research Laboratory (NRL) which transformed a service traditionally reluctant to develop new technologies into an active promoter of research. Although isolationism and the Great Depression reduced both interest in and funding for military R&D, the War and Navy departments, on the eve of World War II, accounted for approximately one-fifth of federal research expenditure.

Eli Whitney first proposed the principle of interchangeable parts in gun manufacture to President Thomas Jefferson early in the 19th century. That technique, developed in Federal armories, was transferred to many other industries.

Science at War

Science was mobilized during World War II in an unprecedented way. The war spawned a large defense research establishment to develop radar, the atomic bomb, and other defense technologies. That infrastructure persisted after the war. Thousands of government and federal contract laboratories, including the national laboratories of the Department of Energy (DOE), were legacies of the wartime mobilization of science. New alliances between the military and academia, and new ways of making national science policies were additional legacies.

The mobilization began in 1939. President Franklin Delano Roosevelt created a "Uranium Project" after



America's first generation of scientific statesmen are pictured here at the California Institute of Technology in about 1930. Albert A. Michelson (left) measured the speed of light, Albert Einstein (center) conceived of the theory of relativity, and Robert A. Millikan (right) led Caltech to the front rank of technical schools. All three men were Noble Laureates in Physics and greatly assisted in the rise of the United States to a world leadership position in science.

The leaders of the wartime mobilization of science are shown here at the Berkeley Radiation Laboratory in 1940. Shown from left to right is Ernest O. Lawrence, leader of the electromagnetic separation effort, Arthur H. Compton, leader of the reactor development effort, Vannevar Bush, leader of the early phases of the Uranium Project, and James Bryant Conant, chief scientific advisor to (?). Also present in the photo are Karl T. Compton, President of MIT, and Alfred T. Loomis, who oversaw the radar research effort.

being alerted by Albert Einstein to the threat posed by Nazi acquisition of heavy water and uranium. The Project was led by NBS Director Lyman Briggs and was designed to investigate nuclear fission.

American defense research proceeded slowly until Vannevar Bush founded the National Defense Research Committee (NDRC) and the Office of Scientific Research and Development (OSRD) which consolidated the research being conducted in universities and in government bureaus. Under contract to those agencies, civilian scientists worked at academic laboratories to perfect radar, sonar, and other defense technologies.

Vannevar Bush recognized that the development of nuclear weapons required a much larger effort. In late 1942, President Roosevelt, spurred on by Bush, decided to pursue nuclear weapons on a high-priority basis. The Army's Manhattan Engineer District (MED) constructed plants at Oak Ridge, Tennessee, and Hanford, Washington, to produce enriched uranium and plutonium to fuel the weapons. Several universities operated the research laboratories supporting that effort, including the new weapons laboratory at Los Alamos. This new "partnership" between the military and university laboratories, in which the universities themselves were hardly knowledgeable, was to become an enduring fixture of the scientific landscape. Secrecy precluded ordinary democratic decision-making as well as academic participation in research policy making. Billions of dollars were spent, and research policies formulated without any substantial knowledge in the Congress as to the nature of the project. Decisions were made by a Top Policy Committee which included Secretary of War Henry Stimson, MED Commander General Leslie Groves, and OSRD scientific leaders. Appropriations were cloaked in War Department budgets.

Although the secret policy carried the risk of failure and condemnation, Stimson and Groves successfully kept Congress in the dark until the atomic bomb was dropped on Hiroshima. The decision to end the war with nuclear weapons was also made in secret, despite agitation by Niels Bohr, Leo Szilard, James Franck, and other Manhattan Project scientists for a non-military demonstration of the weapon's power. Although the British were informed of American plans, French and Soviet allies were not told. Both nations accelerated their own atomic energy programs once they became aware of the American success, leading to the Cold War nuclear arms race.

An early postwar colloquium at the Laboratory saw a rare assemblage of wartime leaders. Left to right are Norris Bradbury, J. Robert Oppenheimer, John Manley, Richard Feynman, Enrico Fermi, and J. M. B. Kellogg.

Science in the Cold War

At the end of the war, Henry Stimson recommended that a new federal agency be set up to control nuclear research and Vannevar Bush, in his influential tract *Science, the Endless Frontier*, advised that a national foundation be established to support basic research. Extended Congressional debates, however, over the exact natures of those two agencies left federal patronage of science in the hands of the military immediately following the war.

MED Commander General Groves took advantage of that opportunity and founded “national laboratories” at Brookhaven, Argonne, and Oak Ridge, funded university laboratories, and built the first permanent facilities at Los Alamos. Basic research in nuclear physics and related fields was undertaken at those laboratories because the prewar stock of knowledge was thought to be exhausted and in need of replenishment, and because such a program of basic research was required to recruit and retain trained scientists.

The Atomic Energy Commission (AEC) took over the MED laboratory system in 1947. The AEC was a new experiment in American science policy—a single government bureau with virtually unlimited control over all nuclear energy R&D. The AEC's supervisory committee, the General Advisory Committee (GAC), included many representatives of AEC laboratories and provided guidance on research policies. Congressional oversight was provided by the Joint Committee on Atomic Energy (JCAE) which expedited appropriations for AEC programs. Despite civilian control of the nuclear research program, the military retained an interest in that research. After the war, in an attempt to gain a foothold in the field, the Navy founded the Office of Naval Research (ONR) to promote nuclear physics

In 1946 Norris Bradbury, General Leslie R. Groves, and Eric Jette (seated, left to right), Colonel L. E. Seeman, and E. E. Wilhoyt (standing, left to right) plan a new technical area for the Laboratory.

research in universities. In the absence of a national science foundation, the ONR played a major role in funding basic scientific research and established a pattern of extramural research support later followed by the Air Force and the Army.

Nuclear weapons were the most important of the AEC's interests. The Baruch Plan of 1946 sought to vest in the United Nations control of all nuclear weapons, but the Soviet Union resisted the Plan's provisions for inspections and resisted limits to the U.N. Security Council's veto in atomic energy matters. Meanwhile the Soviets pursued their own nuclear weapons program and, in 1949, exploded an atomic bomb. Three scientists who had been involved in the Manhattan Project––Edward Teller of Los Alamos, and Ernest Lawrence and Luis Alvarez of the University of California Radiation Laboratory—then lobbied the JCAE and the DOD to pursue the Super, or hydrogen bomb, against the advice of the GAC. In January, 1950, President Harry S. Truman decided to proceed with development of the Super.

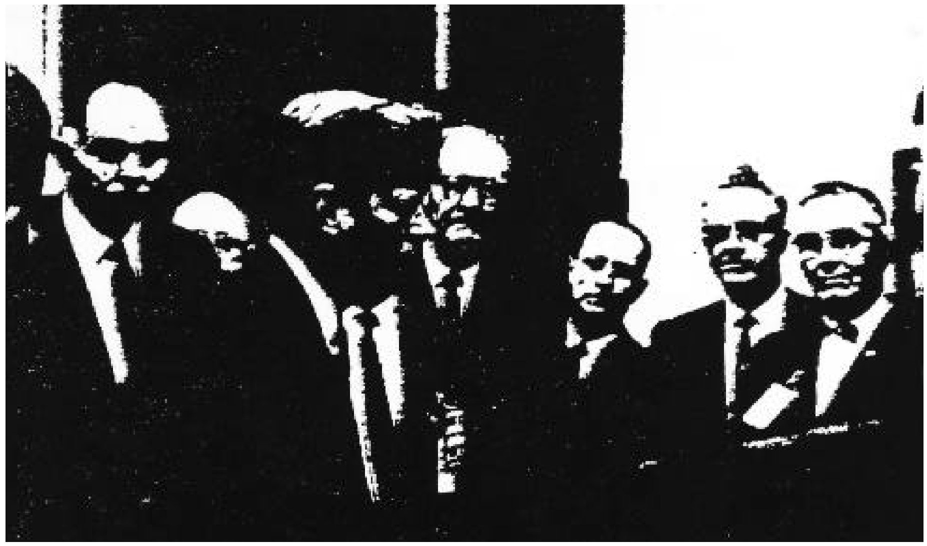
When the Laboratory failed to make what they considered to be adequate progress in developing the hydrogen bomb, Lawrence, Alvarez, and Teller created a second weapons laboratory, again in the face of GAC opposition. The Livermore branch of the Radiation Laboratory at the University of California quickly established itself as a competitor with Los Alamos in the development of innovative weapons concepts, although Los Alamos succeeded in producing the first thermonuclear weapons. Both of the AEC's weapons design laboratories were operated by the University of California. Sandia National Laboratory, which engineered nuclear weapons, was spun off by Los Alamos in 1949 and operated by AT&T thereafter. Originally constituted for the sole purpose of creating nuclear weapons, the laboratories were given new missions by the AEC that included nuclear propulsion studies, biomedical studies of radiation, peaceful applications of nuclear explosives, and development of reactors.

The AEC had few rivals as a supporter of basic scientific research. The AEC's support even dwarfed that of the agency supposedly responsible for supporting basic research, the National Science Foundation (NSF). Vannevar Bush's proposal for such a foundation had languished in Congress during the first half-decade after World War II. When the NSF was finally established in 1950, the AEC and other federal bureaus had already assumed control of much of the basic scientific research that Bush had hoped the NSF might support. The NSF's first appropriation, $225,000, was only a miniscule fraction of the AEC's research budget. Rather than dominating government support for basic research, the NSF tried to fill in gaps left by programmatic agencies.

The remobilization of science following the detonation of the Soviet atomic bomb in 1949, the fall of China to Mao Tse-Tung's Communist forces in 19 **(?)**, and the start of the Korean War in 19 **(?)** led American companies to develop innovative technology to support DoD procurement contracts and the AEC's ambitious expansion program. The record of technology transfers to the civilian sector was not nearly as successful. The National Bureau of

Standards launched Project Tinkertoy, an advanced manufacturing process using modular assembly of ceramic wafers carrying printed circuits under the direction of a punched card computer. Attempts to transfer this technology to the civilian electronics industry failed, although it did inspire Jack Kilby of Texas Instruments to invent the integrated circuit. But theNBS, which lost four major divisions to the DOD in 1953 could not, like the DOD or the AEC, use the leverage of large defense contracts to push major industrial initiatives thereafter.

In the postwar period the AEC and the DOD dominated the patronage of science and technology, except for the life sciences, where the National Institutes of Health (NIH) held sway. Agencies conceived at the end of the war to support basic research received only a fraction of total funding. The mission-oriented agencies of the federal government supplied the bulk of federal money for research, and much of that research was done in government-owned, contractor-operated laboratories—including Los Alamos, Livermore, and Sandia laboratories. As "in-house" laboratories, they were in a good position to develop ties with industrial contractors supplying military needs, but were less effective in influencing the growth of civilian technologies.



On December 7, 1962, President John F. Kennedy visited Los Alamos after announcing the end of nuclear testing in the atmosphere. He is flanked by AEC Chairman Glenn T. Seaborg (left) and New Mexico Senator Clinton P. Anderson (right). A member of the Joint Committee on Atomic Energy (JCAE), Anderson championed Laboratory interests for many years.

Glenn T. Seaborg served as chairman of the AEC from 1961 to 1972. A strong advocate of nuclear power, he helped to redirect the national laboratories towards civilian technologies. Seaborg won the Nobel Prize in 1951 for the discovery of eight transuranic elements, including plutonium.

Science in the Space Age

The Russian launch of the Sputnik satellite in 1957 gave new impetus to scientific activity in the United States. The National Aeronautics and Space Administration (NASA) and the Defense Advanced Research Projects Agency (DARPA) were created in response to a perceived national deficit in science. The passage of the National Defense Education Act (NDEA) also helped to strengthen science education at many levels. Government programs to attract young people into science proliferated.

DARPA helped to put the Department of Defense at the forefront of high technology. The agency transcended the interservice rivalries that had persisted after World War II by funding research relevant to all the military services. Focusing on computer technology, materials science, nuclear test detection, semiconductor research, and other fields of advanced technology, the small DARPA staff stimulated industrial innovation through judicious funding of contract research. In addition, NASA spawned a series of civilian space projects that included Apollo—the program that put a manned rocket on the moon. That project mobilized American science and technology much as the MED had done during World War II.

Federal support for university research—including the granting of fellowships and the construction of new facilities—increased greatly during the space race, amounting to three-quarters of university research budgets and one-quarter of one percent of the U.S. Gross National Product by 1968. The National Defense Education Act provided $1 billion of support for graduate education; the budgets of the NIH, NSF, and NASA shot upwards and helped to support academic science. The expansion of those agencies between 1960 and 1970 sharply reduced the share of university research supported by DOD funding. The Mansfield Amendment, which passed during the Vietnam War, further limited the DOD's support to projects of demonstrable military interest.

The space race also had a great impact on the AEC laboratories. In 1961, when President John F. Kennedy announced the Apollo program, he also told of plans to develop a nuclear rocket which might take men to Mars or beyond. Project Rover—a program to design a nuclear-powered rocket—was begun as a cooperative venture between Los Alamos and the Air Force, and was later taken over by NASA. Los Alamos led the new research effort, while Aerojet General and Westinghouse Corporations cooperated in the engineering development of the project.

Although Los Alamos successfully developed suitable prototype reactors for nuclear pro-pulsion, the objectives of the Rover project were progressively scaled down by the President's Science Advisory Committee (PSAC), which Eisenhower had established in 1958. Concerned that there was no clearly defined mission for a nuclear-powered spacecraft, PSAC successfully lobbied to reduce Rover from a flight test to a research and technology program. The diminished Rover program continued until 1973. Other space programs at the Laboratory found more uses: heat pipes, radioisotopic generators, and Vela satellites (used to detect nuclear tests in the atmosphere and in space) were successful applications of space technologies.

The space program followed the pattern established in World War II and in the postwar arms race. Mission-oriented agencies were created to respond to an external threat by providing advanced scientific research and technological development in contractor-operated national laboratories of the DOD and the AEC. By making use of those laboratories, DARPA and NASA capitalized on the investment those agencies had made in the preceding generation. However, the rise of a new national challenge raised the question of whether the old ones had been satisfactorily met, and the institutions devised to meet them had been rendered obsolete.

Environment, Energy, and Antiscience

The publication of Rachel Carson's *Silent Spring* in 196 **(?)** and Ralph Nader's *Unsafe at any Speed* in 196 **(?)** and the general rise of an anti-science and technology movement heralded in Theodore Roszak's *The Making of a Counter Culture* in 196 **(?)** undermined the cultural and political consensus upon which the AEC and Department of Defense depended for support. The AEC had been in the ambivalent position of both generating and regulating nuclear power since its creation in 1946; radioactive fallout caused by nuclear-weapons tests in the 1950's made the agency a target of environmentalists. The Limited Test Ban Treaty of 196 **(?)**, which restricted nuclear tests to underground, relieved the pressure for a time, but the promotion of nuclear power by Glenn Seaborg and Dixie Lee Ray, successive chairmen of the AEC, created an industry which presented environmental hazards of its own. The energy crisis in the 1970s ended a ninety-year spiral of growth in American energy consumption that coincided with growing concerns about nuclear energy.

The Debate over the Role of the AEC Laboratories

Changes in the AEC's mission and the advent of the Limited Test Ban Treaty raised new questions as to the purposes and potential of the national laboratories. After the Atomic Energy Act of 1954 replaced the government nuclear power monopoly with a licensing regime for private industry to develop reactors, Oak Ridge and Argonne National Laboratories diversified into new fields such as high-energy physics. There was, however, a tendency for those laboratories to become job shops as they sought new work for industry and other governmental agencies to replace the reactor mission.

President Eisenhower's Atoms for Peace program, designed to explore peaceful international uses of atomic energy, and the Test Moratorium from 1958 to 1961 raised further questions about the future of Los Alamos and Livermore laboratories. A study by the Congress's Joint Committee of Atomic Energy in 1960 sought to redefine the role and missions of the national laboratories: Congress eventually expanded

Vice-President Hubert H. Humphrey visited the Laboratory on September 9, 1966, and discussed LAMPF with Norris Bradbury (center) and Louis Rosen (right). New Mexico Senator Clinton P. Anderson succeeded in winning appropriations for the facility despite opposition from the Johnson administration.

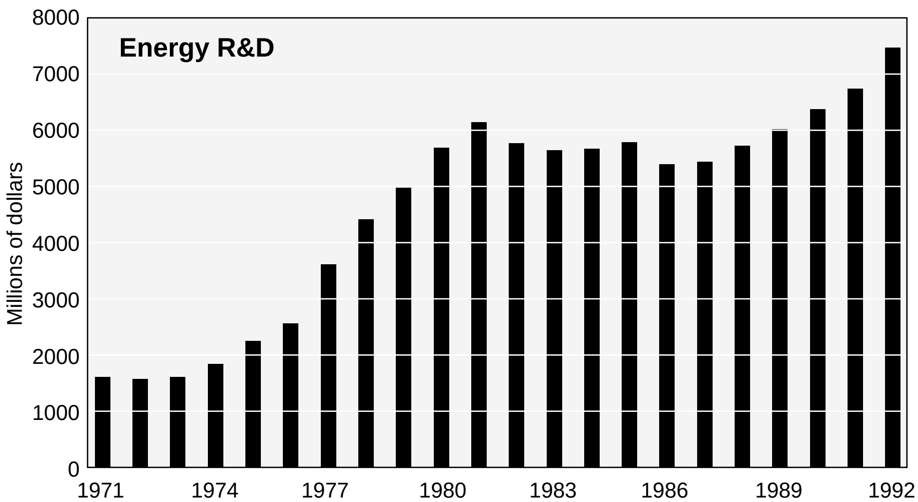
the laboratories' roles to include environmental and safety research, and non-nuclear energy R&D. Throughout the 1960's and 1970's, Congressional committees on science policy sought to redefine and extend the missions of the federal laboratories. However, the new programs that were generated accounted for only a small percentage of the laboratories' total budgets.

Problems in the nation's energy policy became apparent during the 1970s. By 1973 nuclear energy supplied only 1 percent of national needs. Demand for transportation fuels-–which had once been cheap, abundant, and readily supplied by foreign nations-–became acute. The Nixon Administration formulated a more comprehensive energy policy, which led to the absorption of the AEC into ERDA and ERDA into the DOE.

The national laboratories turned their attention to non-nuclear energy sources. At Los Alamos, for example, a broadening of energy-related research in the1970s spawned projects like Hot Dry Rock, a program to tap geothermal energy. Other projects involved research in solar energy and in superconducting electrical transmission and storage. The Laboratory also provided tools like crosswell seismic surveys and computer simulations of oil flow for the petroleum industry. The nation's search for energy independence, however, did not long survive the energy crisis.

Ronald Reagan campaigned on a platform of abolishing the DOE and, after his election, redirected the agency's efforts to deregulation, support for nuclear energy, and nuclear arms development. Although the Reagan administration did not destroy the DOE, it did undermine the department's morale. Morale in the laboratories rose, however, when President Reagan announced his Strategic Defense Initiative (SDI). Based in part upon work on neutral particle beam and rail-gun technologies undertaken at Los Alamos, SDI revived interest in anti-ballistic missile technologies which had been for the most part shelved after the Anti Ballistic Missile Treaty of 1972. In response to Reagan initiatives, the Laboratory also increased its conventional defense research, investigated armor and armor- penetrators for the DOD, and took steps to modernize the nuclear arsenal.

In 1986 Reagan and Soviet President Mikhail Gorbachev signed an agreement to limit strategic warheads to 6,000 for each nation. In June, 1992, their successors, George Bush and Boris Yeltsin, agreed to reduce their respective nuclear arsenals to between 3,000 and 3,500 warheads within a decade. Those agreements, and the dissolution of the Soviet Union, marked the end of an arms race that lasted nearly a half-century.



This chart displays national energy R&D expenditures over a twenty-year period. The steepest rise is between 1971 and 1981 when expenditures grew from $1,064 million to $6,125 million—an increase of only $1,468 million in real dollars. Those funds diverted the DOE laboratories from their traditional nuclear energy missions, but did not replace weapons research as the primary focus of Los Alamos, Livermore, or Sandia Laboratories.

Resuming the Debate

The Department of Energy has not fully defined the role that the national laboratories will play in the future. New government initiatives in environmental and health safety have drastically affected the laboratories' operations, and the national laboratories have found themselves operating in a fluid policy environment where response is difficult and directions unstable. Consequently, the laboratories have looked beyond the DOE for guidance to Congress, which has formulated new policies in the area of technology development and which may afford a new mission for the laboratories.

Recent history suggests there is reason for optimism. In the1970s and 1980s Congress passed technology transfer legislation, created the Human Genome Project, and supported other federal interagency initiatives in high-performance computing, materials science, biotechnology, and mathematics and science education that were formulated by the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET). Those initiatives enriched DOE research budgets, which rose 29% under the Bush administration. Similar increases have been recorded in the NIH, NSF, NASA, and DOD budgets from which the Laboratory received funding. The end of the Cold War suggests that these kinds of activity will become more important in DOE and defense laboratories which, in the past, have embraced nuclear propulsion, reactor development, non-nuclear energy research, and other initiatives. Those old initiatives are now being succeeded by new ones, to which the laboratories must be able to respond. Their primary mission, the stewardship of nuclear weapons, has not been supplanted by any single one of these activities but probably will be in the future.

In the next four years, the Laboratory may be called upon to respond to new policy initiatives such as a national research program that would be the civilian equivalent of the Defense Advanced Research Projects Agency, work on non-proliferation, health care and the environment, a fiber optic network connecting major supercomputer centers together, and other Clinton-Gore initiatives. These opportunities are appropriate to the capabilities of the Laboratory, if they can be mobilized. The history of the Laboratory, in responding to past departures in science policy, suggests that they will. A larger mission, comparable to the Manhattan Project, Rover, or SDI, may afford the kind of challenge to which the Laboratory has responded best—a large and complex problem that requires scientific creativity and engineering know-how. ■

Robert W. Seidel is a Senior Staff member of the Center for National Security Studies at the Laboratory. He holds a bachelor's degree in chemistry and an M.A. and a Ph.D. in the history of science from the University of California, Berkeley, and is the author of numerous articles on the history of the national laboratories run by the Department of Energy and on the history of military research and development. He wrote *Lawrence and His Laboratory*, Volume I of *A History of the Lawrence Berkeley Laboratory*, published by the University of California Press in 1990.

Los Alamos
beginning the second fifty years

Siegfried S. Hecker

The following is based on an address presented at Los Alamos on April 16, 1993, at the close of a week-long seminar series held in celebration of the Laboratory's fiftieth anniversary.

Where do we go from here? This has been a fascinating week, very much in keeping with the motto of our fiftieth anniversary, "A proud past and an exciting future." We can and should be proud of our past. And many things presented this week should make us excited about our future. I'm going to talk about future directions. In doing so, I will try to act as if Washington gave me the keys to the car, remembering, of course, that they control the money for gas. And so we'll see how far we can get.

During the past fifty years we have been fortunate to receive generous government support to build and maintain one of the finest scientific institutions in the world. To fight the Cold War, we developed an incredibly broad science and technology base. We were allowed and encouraged to contribute to other important problems such as nuclear power, nuclear propulsion, alternative energy technologies, and biomedical applications. Our sponsors recognized the importance of basic research in maintaining our science and technology base. Consequently, we have become a world-class laboratory working on complex scientific and technological problems of national importance. Basic research and evolving missions helped to rejuvenate our institution for fifty years.

Today, several fundamental changes strain not only the foundations of our laboratory, but also the nation's entire science and technology enterprise. The collapse of the Soviet Union has made us the world's only military superpower. At the same time the rise of Japan, Germany, and other industrial countries to economic parity means we are no longer the world's only economic superpower. These profound changes in military and economic status require the United States to adopt a new approach to technology policy. In their recent book, *Beyond Spinoff,* Alic et al. make a strong case for changing the nation's policies, institutions, and habits of mind.

The new priorities are primarily domestic and economic. Defense spending, the mainstay of the Laboratory's existence for the past fifty years and formerly the primary engine driving the nation's R&D effort, will decrease. Such fundamental changes require us to re-examine how we do science and technology in this country and how laboratories like ours should contribute to the new national order.

The central question today is not: Can we still do good science? We *must* do good science. And our seminar series highlighted plenty of examples. Doing good science, however, is necessary but not sufficient. Since we're a billion-dollar R&D institution, the real question is: How can we help the nation solve the pressing problems of tomorrow?

In retrospect, defense was easy. Government control of defense makes sense because the government is the customer. Our principal role—that of designing nuclear weapons—has been well matched to our capabilities, since from the beginning that has been our reason for being. We had the entire responsibility for the full life cycle of nuclear weapons, from generating the ideas for new designs all the way through production, deployment, and, finally, to retirement. And we did not let the nation down in that

Post-Cold-War Strategic Vision of the Laboratory

The end of the Cold War allows the Laboratory to address numerous other needs of the nation in addition to its continuing military responsibilities. Among the nonmilitary needs are those of our society and of our industry. Tapping the exceptional technical expertise and capabilities of the Laboratory will not only help tackle the new priorities but also keep the requisite defense expertise strong.

Military needs
- Stewardship of nuclear weapons
- Maintenance of nuclear know-how and judgment
- Limiting and countering nuclear proliferation
- Responding to nuclear accidents or acts of terrorism
- Reduction of the nuclear-weapons stockpile
- Clean-up of nuclear research and production sites
- Storage and disposal of nuclear waste
- Technology for intelligence gathering and non-nuclear weaponry

Societal needs
- Energy
- Environmental protection and restoration
- Infrastructure rehabilitation
- Affordable health care
- Basic scientific research
- Education
- Space exploration and exploitation

Industrial needs
- Cost-shared, market-driven research and development partnerships
- User facilities
- Technology assistance
- Start-up of new technical businesses

process. In return, the customer—the government—had staying power and patient capital. It has invested in this institution for fifty years.

By contrast, in the domestic and economic arenas the government is at best a stakeholder or one of many different customers. It can serve as catalyst or partner. It also serves as requirement setter, regulator, auditor, or often, merely as the barrier to getting things done. In fact, in issues like energy, environment, and economic development, all of which have a strong local component, it makes much less sense for the government to play the central role.

Nevertheless, there is no question that Los Alamos can develop technologies for the civilian sector, and I think we can do it very well. But making a major impact with that technology will be very difficult and will require forging new partnerships and relationships, not just with the DOE but also with agencies throughout the government, with industry, and with universities. We will have to be quick, responsive, and flexible, and we will have to bring together the best talent. This need is no less critical today than at the beginning of the Manhattan Project, when the government created Los Alamos as a contractor-operated, government-owned (GOCO) institution and asked the University of California to run it. The GOCO concept proved to be incredibly successful. Unfortunately, over the past fifty years our institutional agility has been severely compromised by the growth of bureaucracy both in Washington and in our own institution.

We do not face this problem alone. In fact, many have recognized that the government must change its relationships to become more effective. Osborne has argued for no less than reinventing government. Secretary of Energy Hazel O'Leary has included reinventing the Department of Energy as part of her high-priority agenda. We at Los Alamos are taking steps to reduce our internal bureaucracy and to make our operations more responsive to the challenges ahead. And we have a president who firmly believes that the federal government must play a strong role in domestic and economic matters and who also recognizes that technology holds the key to our nation's welfare.

Having presented this preamble, I wish to address the three principal missions I see for Los Alamos in the post-Cold-War era: defense technologies, civilian technologies, and commercial technologies.

Defense Technologies

Our defense needs have changed dramatically. This country, with help from its allies, has tamed the big bear—and I have proof of that. Last night, our colleagues from the Russian weapons institute gave us a beautifully carved bear as a symbol of friendship and, even more symbolic, a plaque containing a piece from one of their dismantled nuclear warheads. The inscription reads, “From Russia, with love.” So we’ve tamed the big bear, but still many dangers are lurking in the woods. People around the world face threats resulting from fractious nationalism, regional conflicts, proliferation of high-technology weapons, the breakdown of national law, increased terrorism, and increased influence of organized crime. Clearly our nation will continue to need new technology for defense. Here the federal government has the unquestioned central role, and Los Alamos must continue to play its special part in nuclear defense as well as some other defense missions.

In the nuclear-weapons business, we have four major priorities and challenges: (1) Stewardship of a much-reduced stockpile of nuclear weapons and, just as pressing, stewardship of nuclear competency—of the people and the capabilities needed to answer the questions that will inevitably come up as long as there are nuclear weapons on this planet. (2) Countering proliferation of weapons of mass destruction, an incredibly important challenge in today’s world. Only the nuclear-weapons laboratories—Los Alamos, Livermore, Sandia—have the full sweep of technological capabilities and the vertical integration required to evaluate such threats and develop

Advanced Conventional-Defense and Intelligence Technologies

Though the Cold War has ended, “The world is still a dangerous place,” as President Clinton has said. Political changes and advancement of technology everywhere may lead to advanced-technology threats and weapons, threats from countries that do not now threaten us, and terrorism on a scale we have not seen before. Maintaining a strong defense R&D effort that emphasizes effectiveness and quality over quantity of military hardware will help to dissuade potential future conflicts or, if necessary, to conclude future conflicts decisively and with low human casualties on both sides.

Areas of present and future contributions

- “Smart” weapons that can hit maneuvering and concealed targets
- Effective defenses against ballistic and theater missiles and other weapons
- Computing and communications hardware and software
  Command, control, communications, computation, and intelligence. The “forward edge” in future military actions will be information and command networks.
  Simulations of battles for training and tactical planning and simulations of the performance of new hardware for design.
- Intelligence, including global surveillance, and weapons expertise to avoid being surprised by advanced-technology foreign weapons.

Technologies for advanced systems (nearly all have civilian uses)

- Integrated modeling, simulation, and gaming techniques
- System-level integration (especially for command, control, communications, computation, and intelligence)
- Machines with useful levels of intelligence
- Powerful software-generation systems
- Other aspects of computing
- Communications
- Electro-optics and optical electronics
- Space technologies
- Sensors
- New materials for strucural use, electronics, and storage and conversion of energy, including biologically generated or copied materials
- Nanotechnology based on electronic, mechanical, and biological techniques
- Flexible manufacturing
- Unmanned automated vehicles for air, sea, and land
- Nonlethal weapons
- Stealth and counterstealth
- High-performance missiles

the means to counter them. (3) Facilitating the nuclear drawdown, which starts with warheads, the other nuclear materials out there in the system, and nuclear waste, but also includes the drawdown of the production complex itself, which has grown to enormous size over the years. Today the complex must be cut back, and we must make it more efficient and environmentally benign. (4) Cleaning up the legacy of fifty years of production of nuclear materials and nuclear weapons.

Just a few years ago, the superpowers had 75,000 weapons out there ensuring the peace. I contend that it was a rather uneasy peace. Now we will be dismantling tens of thousands of weapons, so all of us can sleep better. But nuclear weapons are still with us, and we at Los Alamos can't rest until each and every weapon that we have designed is actually retired and dismantled. We need to revamp the nuclear-weapons program to meet a dramatically changed situation, which, I should emphasize, requires much more than an oil change and a lube job! It requires science. Although the country is not terribly interested in nuclear weapons, our responsibilities have not gone away. It will be up to us to provide the intellectual challenges needed to retain the best people and to maintain their experience and judgement. The latter goal is our highest priority.

Two weeks ago in our classified seminar, Edward Teller and Dick Garwin were here participating on the same panel. Edward suggested that we develop very small, very-low-yield, tactical nuclear weapons. Dick Garwin believes our focus should be on figuring out a way to disable the plutonium pits that are going be stored by the tens of thousands over the next few years. I don't know who's right. But we must be prepared to respond quickly and appropriately to the changing needs of the nation. We must develop the capabilities to disable the pits, or to create small nukes, or to do both. We can afford to have fewer nuclear weapons, but we can't afford to be less smart. And we have to configure our jobs so that we are prepared to meet all eventualities. It's not up to this institution to decide how many weapons the nation ought to have—or how or whether they should be used. But it is up to us to support whatever comes along in national policy—and to provide technological options for whatever the country needs.

It's also our job to contribute to conventional-defense and intelligence technologies. This country will continue to rely on technology to gain through quality what we won't spend on quantity. Los Alamos is in an excellent position to help translate the latest scientific accomplishments into benefits to the nation's defense posture. We are active in almost all the areas needed for the development of the smart precision military technology of the future, and many of those technologies have the distinct advantage of being dual-use—that is, they are also applicable in the civilian sector.

Civilian Technologies

Now I will turn to the domestic and economic front. Societal needs that will benefit from better technology include clean, affordable, abundant energy; a cleaner environment; a refurbished public infrastructure; affordable health care available to all; continued innovation through research; a better educated work force; and finally the spirit and desire to explore the unknown. Our president believes that government must play a role in these areas, that technology is a key component of progress, and that progress is essential for improving the economic performance of the nation.

For Los Alamos to contribute in these areas, we need to forge new relationships, what Lewis Branscomb referred to earlier as "linkages." We must also stick to our strengths—those areas in which, as a result of missions in defense and basic research, we have developed special skills, facilities, and approaches to getting things done, approaches that are not found in universities or in industry. The talents of our scientists and our ability to form interdisciplinary teams have worked powerfully in the past and must now be applied to pressing domestic problems.

Our ability to translate science into applications was certainly demonstrated by the Manhattan Project, but there are many more recent examples. Just a couple of days ago Doyle Evans talked about our work on space sensors. We've designed over 300 of them that have worked. We have operated an accelerator in space, to the dismay of some, but to the delight of many of us. During the Gulf War we put together a LIDAR system for tracking chemical and biological agents with lasers. We accomplished that task in seventeen days, much of it during the Christmas holidays. It was a stunning example of the dedication and technical strength of the staff. Fortunately, we didn't have to use the system in the Persian Gulf, but we turned around and used LIDAR in Mexico City and in Barcelona to monitor air pollution and in the central Pacific to make some crucial measurements for global climate modeling.

Our strength in nuclear science and technology is unparalleled—and although things nuclear are a rather small part of the overall economic agenda, behind that strength lies a whole spectrum of expertise in mathematics, physics, modeling of complex physical systems, and large-scale computer simulation. We devote between $250 and $300 million a year to the broad subject of high-performance computing, which includes hardware, software, applications, and basic research. Having made that kind of investment, the country can expect us to be the best at applying computers to solve real problems, and we are. In the nuclear-weapons business, we've always had to outcompute everybody else in the world—particularly Livermore. That competition is still there, and it has made us world leaders.

We invest in excess of $100 million a year in materials science and technology. Of course, some of that work is very specific to uranium, plutonium, and other materials that people hope they never have to work with, but we're also working on materials problems for novel electronic and photonic devices, for nanotechnology, for application of high-temperature superconductors, for efficient fuel cells, higher-temperature structural materials, and much more.

It may be surprising to some, but environmental research is also a major activity at Los Alamos. Because of the legacy of fifty years of operations here, we are now supported to the tune of $200 million a year for environmental restoration, waste management, and environmental R&D. Our basic-research program means that we have the chemists, biologists, materials scientists, and computer scientists to contribute innovative solutions to environmental problems, and so we are very good at it.

Other areas of expertise include accelerators, lasers, sensors, and dynamic experiments. And, as I mentioned earlier, we are very strong in theory and the modeling of complex systems.

I am underscoring our strength in basic research because that strength underlies our ability to contribute to civilian technologies. Many in this country believe that we need less basic research and more applications. But as Charles Herzfeld pointed out yesterday, if one takes a rational look, even in the defense arena, it's clear that this country needs to cast the net as broadly as

Translating Basic Research into Practical Technologies

The laboratory's success in translating basic research into practical technology is demonstrated by the following examples.

- The KIVA computer program for modeling internal-combustion engines. A spinoff from groundbreaking developments in numerical hydrodynamics, KIVA is now used by all the Big Three automobile makers to design engines.
- The side-coupled cavity used in high-power radio-frequency commercial accelerators. Originally designed to create intense particle beams at the LAMPF accelerator, it is now used in high-energy x-ray machines for cancer therapy and industrial radiography. Such machines have a worldwide market of $400 million per year.
- High-temperature superconducting cables and electronics. Building on considerable expertise in the physics of high-temperature superconductors, the Laboratory is helping a total of seventeen U.S. companies develop the technology base required for competetiveness in thin-film electronics and in electric-power transmission.
- Flow cytometers for biology and medicine. The flow cytometer was developed at the Laboratory in the 1960s to study the effects of radiation on cells. This instrument, which uses lasers to separate cells or chromosomes according to their characteristics, is widely used in hospitals to diagnose diseases such as AIDS and leukemia, to monitor transplants and the effects of cancer therapy, and to assist biological research in the Human Genome Project and related work. The total market for three U.S. companies is $800 million per year.
- Resonant ultrasound spectrometry for nondestructive testing. Invented at Los Alamos to study basic structural properties of high-temperature superconductors and other new materials, the resonant ultrasound spectrometer is now available from Quatro Corporation as a tool for nondestructive testing of high-precision objects such as ball bearings.

High-Performance Computing at the Laboratory

Large-scale computing was a critical element in the the development at the Laboratory of the first fission weapons and the first thermonuclear weapon. Some of the best mathematical minds of that time, among them John von Neumann, Richard Feynman, Hans Bethe, Stan Ulam, Edward Teller, and Nick Metropolis, helped establish the foundations of what is today the most sophisticated computing center in the world.

Our extraordinary computing capabilities can help improve the economic position of the nation in many areas, including:

- exploration for and extraction of oil,
- efficiency of internal combustion engines,
- design of new materials,
- environmental restoration,
- modeling climate,
- medical data banks and diagnostic tools,
- manufacturing processes and materials.

The Laboratory's computing capabilities are currently being applied to "Grand Challenges" in science and technology such as:

- simulating properties of the global ocean over a hundred-year period at forty depth levels and a horizontal resolution of 0.25° latitude and longitude;
- simulating the three-dimensional flow of two immiscible liquids (such as oil and water) through porous media (such as oil-bearing rocks);
- simulating the behavior of materials on an atomic level at a spatial resolution of 1 cubic micrometer and a temporal resolution of 1 nanosecond;
- assessing the safety of nuclear weapons in various hypothetical accident scenarios.

These problems are grand challenges in the sense that their solutions require enormous numbers of computer operations and access to enormous amounts of data. To put the numbers below in perspective, consider that forty years of computing time on a Cray YMP 8/8 are required to carry out 10^{18} operations and that the Library of Congress contains the equivalent of 25 terabytes of data (1 terabyte = 10^{12} bytes). Solving the grand challenges will help to drive the leading edge of computational science just as designing nuclear weapons did in the past.

	Operations Required	Terabytes of Data Required
Global ocean features	10^{17}	20
Flow through porous media	10^{18}	4
Atomistic behavior of materials	10^{18}	6
Nuclear-weapon safety	10^{18}	20

possible—investments in research must remain broad and basic.

At Los Alamos we have some of the best people, and we have some very special and unique capabilities. Consequently, investment in basic research at Los Alamos results in multiple payoffs. We do not simply publish ideas that are picked up and applied by other countries. We have the people who can translate those ideas into technology, and we've done just that over and over. Basic research is also crucial to the health of the Lab itself. It is the fountain of ideas for technology, and it is our link to the academic community. For example, aside from the good physics done at LAMPF, the Los Alamos Meson Physics Facility, that facility has provided a most important connection to the academic community. Over a thousand scientists from around the world have come to do experiments at LAMPF. Many visiting scientists come from universities to work in other parts of the Laboratory. We also have 1300 students and 200 postdocs who are our link to the future.

Los Alamos always has advanced and must continue to advance the frontiers of science. Nuclear physics was *the* exciting frontier when this laboratory began and remains a challenging and productive field. But today we're seeing an explosion of knowledge in the biosciences, materials sciences, and computing/information sciences. For that reason we hope to refocus LAMPF from a nuclear-physics facility to one aimed at two of these areas. This change will depend on upgrading the accelerator at LAMPF and creating an advanced spallation neutron source, a facility that will make the U.S. once again a world leader in the fast-growing area of neutron scattering.

Environmental research, an expanding area at Los Alamos, is one of the best examples of a field in which civilian and defense needs are served simultaneously. It also has great potential for industrial spinoffs. The government will be its own customer—and the Department of Energy, in particular, a big customer. At the same time the technological developments will surely spur competitive industry in the environmental arena.

As an example of a specific, large civilian technology project suitable for Los Alamos, I would mention first the clean, or green, car. You might ask why. There is a real possibility of designing nonpolluting automobile engines that are twice as fuel-efficient as present combustion engines. Such a development would clearly benefit the entire nation, not just the automotive industry. Another major project is the application of accelerators to the efficient destruction of plutonium, the elimination of high-level nuclear waste, and the production of energy. Los Alamos is already exploring specific systems, and they look quite promising.

The building of electronic highways across the nation that integrate computing and communication hardware and software is a realistic goal and one that will be of great benefit to business, industry, education, and healthcare. Los Alamos and its sister national-security labs have the in-depth expertise to develop and test prototypes of such systems, and I expect we will be called upon to do so. Finally, I'd like to mention our growing contributions to the biosciences. Our work on the Human Genome Project, structural biology, and medical diagnostics has already demonstrated the value of the Laboratory's combined strength in biology, chemistry, physics, mathematics, and engineering.

Many of these civilian technology projects will contribute to the economic performance of U.S. industry. But as the government increases support of civilian (in contrast to defense) mission-oriented research, it should not rely simply on spin-offs. Such research must be deliberately planned so that it will contribute to the industrial technology base. A good start would be to fund all such research jointly with industry. In the past industry has been uninterested in or even hostile to the government's attempt to help. Today we are seeing a marked change in those attitudes.

Commercial Technologies

Los Alamos and its sister laboratories (both defense and energy multiprogram labs) can make substantial contributions to long-term industrial competitiveness by devoting up to 20 percent of their budgets to working directly with industry. But to make this work, it must first be acknowledged that the labs cannot "save" industry. We should also replace the concept of technology transfer with that of industrial partnerships—deliberate collaborations with U.S. companies driven by the needs of industry.

The Laboratory's relationship with industry is part of a much larger issue, the proper role of the government in economic competitiveness. Progress has been hindered by the spectre of a national industrial policy that places the government in the position of picking winners and losers. The federal government has always influenced the private sector through macroeconomic factors such

A New Framework for Government-Assisted R&D

Technical output doubles every ten to fifteen years. To keep U.S. industry competitive, the government must judiciously support R&D and facilitate access to knowledge and innovative capacity. In *Beyond Spinoff* Alic et al. suggest public investment in a range of technologies in the vast gray area between basic research and commercial-product process development. Technology diffusion must also be specifically encouraged and supported.

Appropriate areas for government investment

I. Pathbreaking technology—arising from new science and potentially leading to new industries or transformation of existing ones.
- Government funding: pivotal at early stages because of high technical risk, uncertain practical payoffs.
- Examples: biotechnology, satellite communications, nuclear power, nuclear medicine, massively parallel computing.
- Payoffs: typically after ten years.

II. Strategic technology—of great importance to sectors of existing U.S. industry, typically requiring industry-centered consortia.
- Government funding: necessary because of high business risk and difficulty of sustaining the technological edge
- Examples: semiconductor process tools (the SEMATECH consortium is a model for the support of strategic technology), high-performance computing, machine tools.
- Payoffs: typically after five years.

III. Infrastructural technology—for improving the productivity of a broad spectrum of firms by making the design and development of products and processes more efficient.
- Government funding: necessary because of high cost of infrastructure, provides institutional and technical support, but must be arranged so benefits cannot be captured predominantly by any one firm.
- Examples: national computer networks, national facilities (such as wind tunnels, light sources, and supercomputer centers), research emphasis in areas such as solid-state chemistry and manufacturing sciences and engineering.
- Payoffs: typically begin immediately.

Mechanisms of technology diffusion

- Better access to foreign technical knowledge and to technical information in the United States
- Technical services and industrial extensions
- Collaborative technical activities among firms
- Investment in human resources.

as tax policy, trade policy, and antitrust laws. It has also had a direct impact through mission-oriented research in agriculture, health, space, and defense, although those initiatives were not articulated as explicit technology policies.

Today, the major shift in world military- and economic-power balance requires a new paradigm—one in which government is more aggressive. The American public has asked the government to add economic performance to the well-accepted government responsibilities in defense, energy, environment, infrastructure, and healthcare. Furthermore, President Clinton has made economic development the nation's number-one priority. The authors of *Beyond Spinoff*[1] provide a useful framework for how the government can help industry. In the federal laboratory system, the DOE laboratories offer the most potent vehicle for helping industry because together they represent approximately one-third, or $6.5 billion, of the government's $21 billion investment in that system.

I envision four principal mechanisms for the DOE laboratories to work with industry on commercially useful R&D: collaborative R&D partnerships, user facilities, technology assistance, and small business start-ups.

Currently the principal mechanism for collaborative R&D partnerships is the Cooperative Research and Development Agreement (CRADA), a vehicle created by legislation. The collaboration is cost-shared, so no money need change hands. The work is driven by the needs of industry, and the industry's right to intellectual property is protected. The CRADA is much maligned these days, and that's unfortunate, because I think it's a superb vehicle. Industry not only gets a match for their R&D dollars but also buys into the Laboratory's capabilities. Moreover, our experience demonstrates that the agreements can address each of the three high-priority areas for government assistance outlined in *Beyond Spinoff*; that is, pathbreaking, strategic, and infrastructural technologies.

Xerox, for instance, is interested in working with us on pathbreaking technology. They want to develop new paradigms in computing, along the lines of the lattice Boltzmann technique that we developed for solving nonlinear differential equations. Mobil and Schlumberger are not primarily interested in developing the next technique, but they'd like to apply lattice Boltzmann to very complex problems involving three-dimensional, two-phase fluid flow, an example of a strategic technology. They have massively parallel computers, so they don't need our hardware. They do, however, want our help in modeling complex flows. As an example of infrastructural technology, Mark Murphy, who spoke earlier this week, wants us to help his independent oil company with specific experimental capabilities in microseismic detection and the calculations needed to interpret the data. Los Alamos has already established thirty-eight CRADAs, and the main problem has been overcoming the red tape that bogs down their implementation. If the Laboratory had more decision-making authority, we could simplify and speed up the industry interface.

User facilities provide the second, very effective avenue for helping industry. Early on, national user facilities, such as accelerators and reactors, were primarily tools for nuclear and high-energy physics. More recently, nuclear reactors have been used for neutron scattering studies in materials science and bioscience. Accelerators have been constructed as sources of neutrons and x rays for studies of materials. Transmission electron microscopes, combustion-research facilities, materials-processing facilities, and DNA-sequence databases have all been made available to users. These facilities are very attractive to private companies and will enhance our ability to work with them. The Computational Test Bed at Los Alamos is an example of a new type of user facility designed specifically for industry. It

Cooperative Research and Development Agreements

The Laboratory has executed thirty-eight CRADAs in a wide variety of fields. The following is a sample:

- High-speed sequencing of DNA by detecting individual base molecules with laser-induced fluorescence (Life Technologies).
- Removing uranium and plutonium from soil using high-gradient magnetic fields (Lockheed).
- Developing the next generation of models of electromagnetic processes in semiconductor chips, of large molecules (containing several thousand rather than several hundred atoms), and of global climate change (Cray Research).
- Constructing random-access computer memories using new techniques so that the memories do not lose information when power is interrupted (Radiant Technologies).
- Developing ways to produce very stiff materials for the aircraft and automotive industries by coating materials with diamond or diamond-like carbon (DuPont and Sandia National Laboratories).
- Optimizing complex metal-forming operations by expanding models of tool-workpiece friction and material behavior, with applications to forming aluminum sheet (Alcan) and corrosion-resistant pipe (Exxon).
- Developing high-temperature-superconducting materials, fabrication technology, and device applications: electronics (DuPont), microwave communication devices (Neocera).
- Adapting advanced pattern-recognizing neural networks for use in simplified and inexpensive quality-control systems (Ethicon).
- Isolating fetal cells from the mother's blood in order to test for congenital diseases without invading the womb (MediGene).

The Los Alamos Computational Test Bed for Industry

The DOE-sponsored High-Performance Research Center at Los Alamos's Advanced Computing Laboratory keeps the Laboratory at the leading edge of computing. We have established the Computational Test Bed for Industry to provide U.S. firms with access to the latest scientific computing environment. In particular, the Test Bed

- allows U.S. companies to use high-performance computer hardware, software, networks, storage devices, and visualization tools on a cost-shared basis;
- promotes interaction of U.S. companies with computational scientists and engineers at Los Alamos who have extensive experience with the most recent hardware and software;
- allows U.S. companies to test and improve their own programs;
- hosts industrial internships as well as conferences and workshops for industrial scientists and engineers;
- facilitates communication and collaboration between the Laboratory and industry on state-of-the-art hardware, software, and applications.

Educational workshops

- Nuclear Criticality Safety Workshop (completed fall, 1992)
  Applications of Monte Carlo transport techniques to the storage, environmental impact, transportation, and siting of fissile material
- Technology Commercialization for the Petroleum Industry Workshop (summer, 1993)
  Modeling of oil exploration, environmental impact, seismic imaging, refining, and remediation
- Environmental Modeling Workshop (under discussion)

Examples of industry-sponsored internships

- DuPont is participating in design and construction of a Gigabit Testbed.
- Electronic Data Systems is studying workstation clusters as well as the gasoline combustion/emission problem and is participating in benchmarking the Connection Machine 5.
- Schlumberger-Doll plans to use the CM-5 to study bio-remediation.
- The oil industry plans to implement seismic-imaging codes on the CM-200 and CM-5.
- Rocket Research plans to study design of arcjet-rocket thrusters.
- Xerox plans to model xerography.
- Texas Instruments plans to model semiconductors.
- BIOSYM Technologies plans to develop commercial materials-modeling software.

provides education and training as well as direct use, and it should serve as a model for other such facilities.

Technology assistance is the third avenue of collaboration. As you all know, much of the innovation and the job creation in this country is done by small companies. They incorporate new technology more quickly than large companies because their survival depends on it. Yet they often do not have their own research capabilities. For instance, the machine-tool industry in this country is only approximately 25 percent computerized and they need help to take better advantage of computers. We can provide that help. We could get on board with the National Institute for Standards and Technology (NIST) to work within the manufacturing extension centers. Through the nuclear weapons program and its legacy, many of our people have worked as liaisons in the DOE production plants. They developed the CAD/CAM systems for those plants and could do that for industry today.

Assisting small business start-ups, or creating them, is the fourth avenue and the most direct way of introducing new technology into the marketplace. It's even better than helping existing small companies. Although we have spun off thirty-eight companies in the past ten years, such start-ups have not been a large focus of ours in the past. They have, however, been successfully encouraged at universities in the areas of computer software and biotechnology. We have extensive software and biotechnology capability, and now we have a substantial environmental capability. So we view small business start-ups as a promising avenue for getting technology into the marketplace.

Clearly, there is no single magic bullet for improving the nation's economic performance. In the next few years the nation must experiment with several approaches. Progress will have to be monitored, and future funding of collaborations will have to be based on success with industry as well as the ability to demonstrate public good. But this is a very new area for the Laboratory and the government. Our emphasis should be on the creation of long-term relationships rather than on the production of short-term payoffs. We should continue to expose our industrial partners to the best of science, but we must also listen to them and allow their problems to drive our collaborative work. In the meantime, the Laboratory and the government must sharpen the national missions that will account for the bulk of our work. We should also involve industry in these missions so that as we accomplish the government's goals we also strengthen the nation's industrial science and technology base. Through fulfilling those missions, the Laboratory will also be able to sustain its core technical competencies, which, in turn, will provide the source of innovative contributions to commercial technologies.

Scientifically and technically we can rise to the challenge before us. Los Alamos has superb facilities and is still able to attract the best talent in the world. We are very strong in the vast area between basic research and product design, which many claim needs shoring up in the nation's science and technology base. Most of our bench-level skills are interchangeable among defense, civilian, and commercial missions, and we are capable of applying them to all three.

Our superb scientific capabilities notwithstanding, the management practices that have evolved over decades of defense dominance demand a major overhaul. As we maintain our emphasis on science and innovation, we must streamline the bureaucracy. I view this challenge as similar to that faced by most of U.S. industry over the past decade. We've spent a lot of time in the last few years studying institutions and corporations who've gone through their institutional renewals, companies such as Motorola, which started that process in 1979 and today is one of the most competitive companies in the world. We are beginning to simplify our organizational structure, change the way we manage, increase our focus on productivity, and become more responsive to a wider agenda. We are also working with the government to develop a partnership that will restore the institutional flexibility and agility necessary in a fiercely competitive environment with multiple missions and multiple customers.

Our world has changed, our missions are changing, and we must work together to change our institution. I'll close with the message Charles Herzfeld gave us yesterday: "There is plenty to do—let's go do it." ■

Further Reading

John A. Alic, Lewis M. Branscomb, Harvey Brooks, Ashton B. Carter, and Gerald L. Epstein. 1992. *Beyond Spinoff: Military and Commercial Technologies in a Changing World*. Boston: Harvard Business School Press.

Wayne Sandholtz, Michael Borrus, John Zysman, Ken Conca, Jay Stowsky, Steven Vogel, and Steve Weber. 1992. *The Highest Stakes: The Economic Foundations of the Next Security System*. New York: Oxford University Press.

Carnegie Commission on Science, Technology, and Government. 1993. *Science, Technology, and Government for a Changing World.* New York: Carnegie Commission.

Sam Nunn and Pete Domenici. 1992. *The CSIS Strengthening of America Commission.* Washington, D.C.: The Center for Strategic and International Studies.

David Osborne. 1992. Reinventing government: Creating an entrepreneurial federal establishment. In *Mandate for Change*, edited by Will Marshall and Martin Schram, page 263. New York: Berkley Books.

K. Finneran. 1986. The federal role in research and development: Report of a workshop. Washington, D.C.: National Academy of Sciences.