FALL 1982 VOLUME 3, NUMBER 3

Los Alamos Science

LOS ALAMOS NATIONAL LABORATORY

Inside This Issue

EDITOR'S NOTE

Hans Bethe has been a paternal figure to Los Alamos for its entire history. A German refugee in 1935, he first came to Los Alamos in 1943 to head the theoretical work of Project Y and has continued since the end of World War H as an active and much-prized consultant to the Laboratory. In 1954, about six months after the Oppenheimer hearings, he wrote an article refuting the notion, held by many at that time, that the development of thermonuclear weapons was delayed by the influence of Oppenheimer. This candid article, which we are honored to publish for the first time, documents the technical problems that in actuality dictated the pace of H-bomb development. In it Bethe expresses his own very strong reluctance to make this deadly weapon a reality. This story of the inner workings of a top-secret project necessarily reveals a somewhat unfamiliar picture of one of Bethe's oldest friends and associates, Edward Teller. We applaud Bethe's courage, integrity, and sense of responsibility in setting straight—personal considerations notwithstanding—the record of this important period in the history of American science and politics.

In a lighter vein we have an interview with Stan Ulam and Mark Kac, two outstanding mathematicians who, like Bethe, came to this country from Europe before the outbreak of World War II. The afternoon we taped this interview was one of the most delightful I have ever spent. These men spoke of their life and work with old-world wisdom, refreshing insight, and a sense of humor that engages the heart and the mind.

Mitchell Feigenbaum, whose idea it was to record these conversations. is himself a profoundly thoughtful man. His seminal work on chaos in deterministic systems reported in the first issue of *Los Alamos Science* has stimulated a surge of new activity in this challenging field, New results were reported at a conference entitled "Order in Chaos" sponsored by the Center for Nonlinear Studies at Los Alamos. This issue's report of the conference, unlike most such reports, introduces to the nonexpert the main concepts in this field and explains the significance of recent contributions. It is truly educational.

Quantitative theoretical immunology, a field that was born at Los Alamos in 1970 when George Bell applied a mathematical description to an animal's immune system, represents one of the few areas in biology in which mathematical descriptions are directly applied to biological experiments. In this issue we present theoretical work on one of the less fortunate aspects of the immune system. the allergic response. What turns this response on and off? Collaboration between theory and experiment has helped find mechanisms for desensitizing cells to the guilty allergens.

This issue starts out in the "wonderland" created by phase-conjugating mirrors. Acting like a time-reversal machine, these devices send a laser beam back along its original direction with all phase relationships preserved. Thus an incident laser beam, after suffering distortion as it passes through an amplifying system, can be returned through the system by the phase conjugator and re-emerge with its original beam quality. The developers of infrared and ultraviolet phase conjugators describe how these remarkable reflectors work and how they can revolutionize the rich field of laser optics.

Happy reading!

Necia Grant Cooper

Editor	Necia Grant Cooper
Art, Design, and Production	Vicki Hartford
Science Writer	Roger Eckhardt
Associate Editor	Nancy Shera
Assistant Editor	Judith M. Lathrop
Editorial and Production Assistant	Elizabeth P. White
Production Assistant	Judy Gibes
Illustrators	Anita Flores, Monica Fink (type)
"Alice" Illustration	Kathi Geoffrion Parker
Photography	John A. Flower, Henry F. Ortega, LeRoy N. Sanchez
Black and White Photo Laboratory Work	Dan Morse, Ernie Burciaga, Gary Desharnais, Chris Lindberg, Ken Lujan
Color Photo Laboratory Work	Cordie Roybal
Phototypesetting	Chris West, Alice Creek, Kris Mathieson, Kathy Valdez
Pasteup	Judy Gibes, Meredith S. Coonley
Printing	Robert C. Crook, Jim E. Lovato

Erratum: Los Alamos Science *apologizes for omission of credit to Ken Lujan for black and white photo laboratory workk in Volume 3, Number 2.*

Los Alamos Science

FALL 1982 · VOLUME 3, NUMBER 3

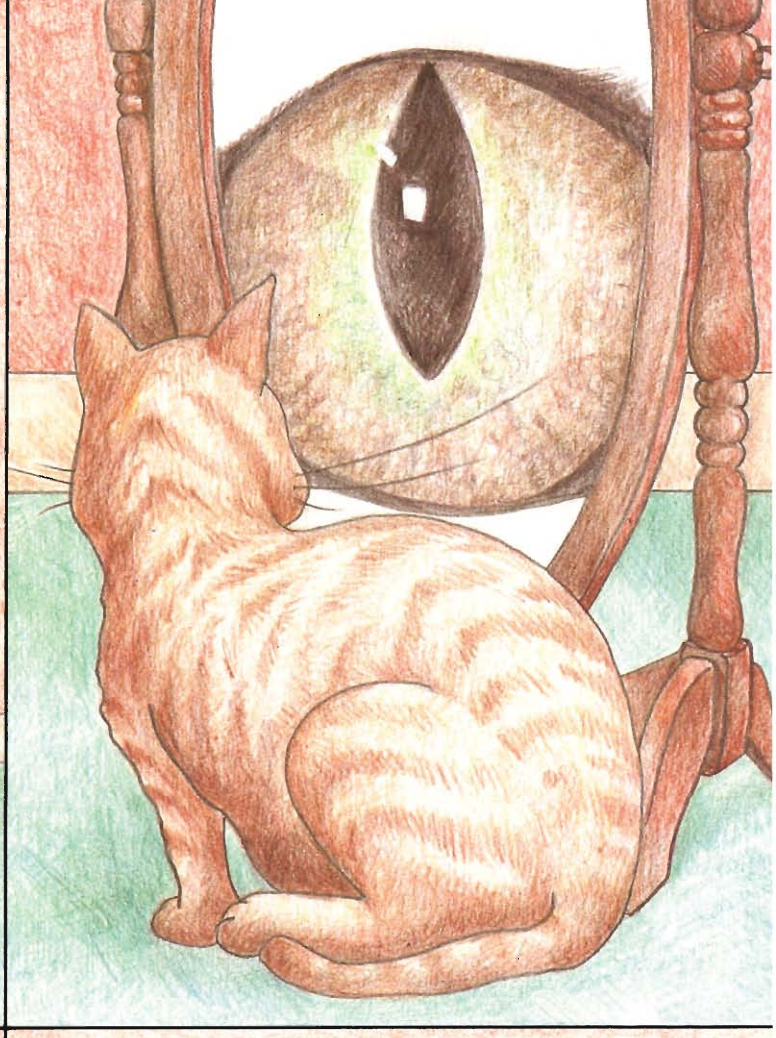
CONTENTS

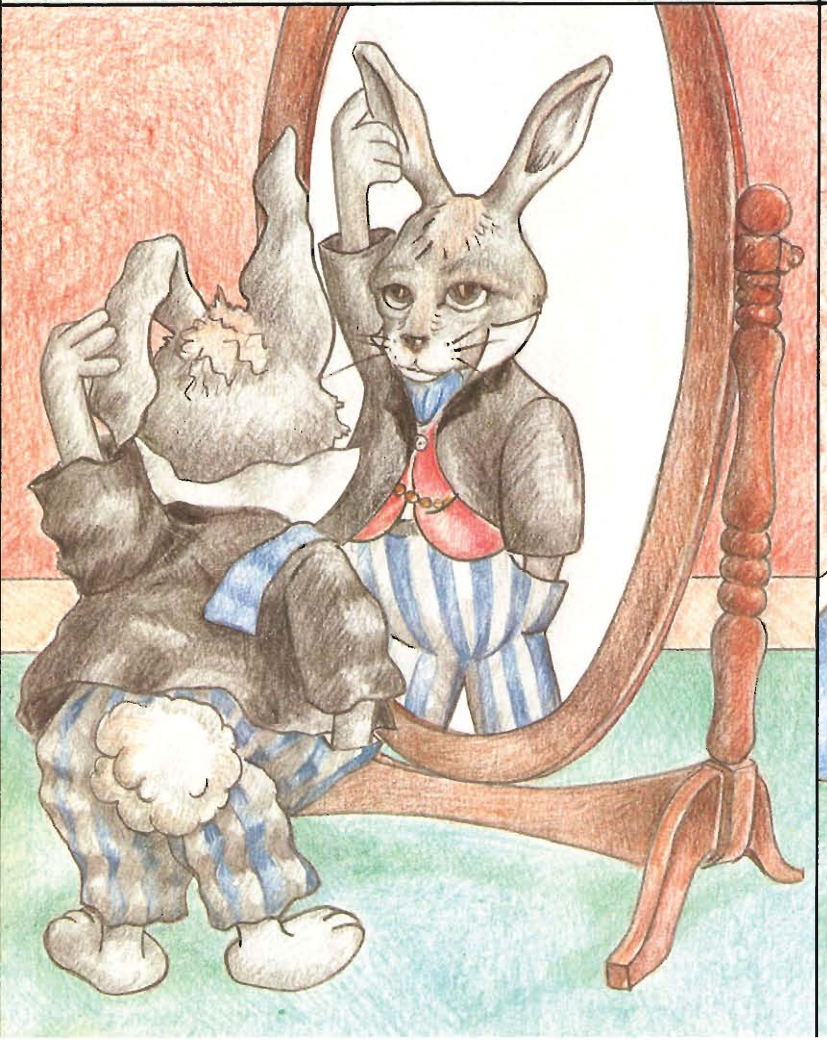
On the cover.

A unique "mirror' 'for lasers fluoresces as it reflects the phase-conjugate version of an incident laser beam. Ultraviolet light from a xenon fluoride laser enters a cell of liquid hexane from the left. There stimulated Brillouin scattering generates the reflected beam, which exactly retraces, in reverse, the path of the incident beam. (Photo by Henry Ortega)

Address mail to
LOS ALAMOS SCIENCE
LOS ALAMOS NATIONAL LABORATORY
MAIL STOP M708
LOS ALAMOS, NEW MEXICO, 87545

Los Alamos Science is published by Los Alamos National Laboratory, an Equal Opportunity Employer operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.






Kathi Geoffrion Parker

Through the Looking Glass

with phase conjugation

"I don't understand . . .," said Alice. "It's dreadfully confusing!"

"That's the effect of living backwards," the Queen said kindly: "it always makes one a little giddy at first—"

"Living backwards!" Alice repeated in great astonishment. "I never heard of such a thing!"

—Lewis Carroll

by Barry J. Feldman, Irving J. Bigio, Robert A. Fisher, Claude R. Phipps, Jr., David E. Watkins, and Scott J. Thomas

Imagine a mirror that reflects more light than was incident, that reflects a beam into the same direction regardless of the mirror's tilt, that eliminates image distortions by causing light rays to retrace their paths as if running backward in time, and that when looked at allows the observer to see absolutely nothing. Science fiction, you say? Well, such mirrors have been the subject of intense investigation both here at Los Alamos and at other research laboratories around the world. Not only do they exist, but their practical applications may be far-reaching.

At left. A whimsical look at four aspects of phase-conjugate reflection. These are (clockwise from upper left) backward-traveling wavefronts, light returning to its point of origin, time reversal, and restoration of beam quality.

The mirrors we refer to are called phase conjugators, and they reflect light in a manner radically different from conventional mirrors. Consider a beam of light incident on a conventional mirror (Fig. 1a). The incoming rays can be characterized by a wave vector **k** pointing along the direction of propagation. When a ray is reflected by a conventional mirror, only $k_\perp$, the component of the wave vector normal to the mirror surface, is inverted. Thus a light beam can be arbitrarily redirected by adjusting the orientation of the conventional mirror. In contrast, a phase conjugator (Fig. 1b) inverts all components of **k** and thus causes the wave vector to change sign, that is, to be reversed in direction. In this case, regardless of the orientation of the conjugator, the reflected beam exactly retraces the path of the incident beam. Surprising, perhaps, but there is more.

In addition to propagation direction, a complete description of a light beam requires information concerning its intensity and phase. The spatial and temporal dependences

of a beam's electric field E are separable, and typically the spatial component (at an instant in time) is described mathematically as the sum of many plane waves, each with a complex amplitude $\mathcal{E}_n$ and with an oscillatory factor $e^{i(\mathbf{k}_n \cdot \mathbf{r})}$ containing the phase information as a function of the spatial coordinate r. The electric field of an incoming beam, E_{in}, can be written as

$$E_{\text{in}} = \sum_n \mathcal{E}_n e^{i(\mathbf{k}_n \cdot \mathbf{r})} .$$

The intensity of the incoming beam, I_{in}, is then given by

$$I_{\text{in}} = |E_{\text{in}}|^2 = \sum_n \mathcal{E}_n \mathcal{E}_n^* ,$$

where $\mathcal{E}_n^*$ is the complex conjugate of $\mathcal{E}_n$. After reflection by a phase conjugator of amplitude reflectivity $\mathcal{R}$, the electric field of the outgoing beam, E_{out}, becomes

$$E_{\text{out}} = \mathcal{R} \sum_n \mathcal{E}_n^* e^{-i(\mathbf{k}_n \cdot \mathbf{r})} .$$

The components of the outgoing beam correspond to the components of the incoming beam, only with the amplitudes replaced by their complex conjugates and with the signs of the wave vectors reversed. This simple relationship between the incident and reflected beams should make it clear why the process is called phase-conjugate reflection.

So far we have ignored the temporal dependence of the electric field. To be complete, other oscillatory factors $e^{i\omega_n t}$ that depend on the frequencies ω_n of the component waves must be included in the equations for the incident and reflected beams. Taking these oscillatory factors into account, we have

$$E_{\text{in}} = \sum_n \mathcal{E}_n e^{i(\omega_n t + \mathbf{k}_n \cdot \mathbf{r})}$$

and

$$E_{\text{out}} = \mathcal{R} \sum_n \mathcal{E}_n^* e^{i(\omega_n t - \mathbf{k}_n \cdot \mathbf{r})} .$$

The fact that the sign reverses for the $\mathbf{k}_n \cdot \mathbf{r}$ term but does not reverse for the $\omega_n t$ term indicates that E_{out} is propagating opposite to the direction of E_{in}. Moreover, the complex conjugation of the amplitudes reverses the constant-phase wavefronts with respect to the propagation direction (for example, a lag in phase for E_{in} becomes an advance in phase for E_{out}, and so forth). Regardless of the value of the reflectivity, E_{out} can be thought of as having wavefronts that are everywhere in space coincident with those of E_{in} but that are traveling backward. It is as if time had been reversed: the reflected wave replicates—in reverse—the phase behavior of the incident wave.

Now we can understand one of the most important implications of this kind of reflection. Consider the situation in which a beam passes through an aberrator, or phase-dis-

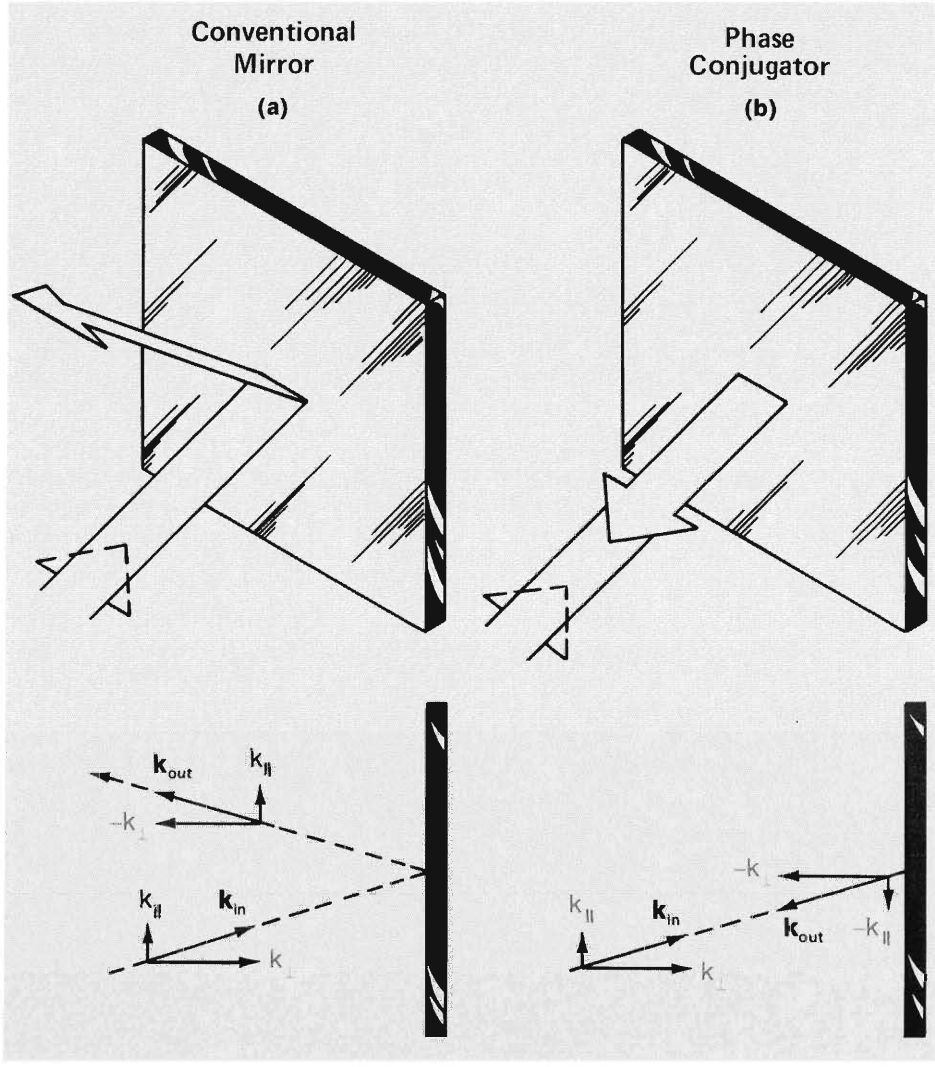


Fig. 1. (a) A conventional mirror reflects light by inverting only the normal component $\mathbf{k}_{\perp}$ *of the beam's propagation vector* k. *This process leads to the law that the angle of incidence equals the angle of reflection and allows the direction of the reflected beam to be altered by changing the tilt of the mirror. (b) A phase conjugator reflects light by inverting all components so that the propagation vector changes sign* ($\mathbf{k}_{\text{out}} = -\mathbf{k}_{\text{in}}$). *In this case, regardless of the tilt of the mirror, the reflected light exactly retraces the path of the incoming beam.*

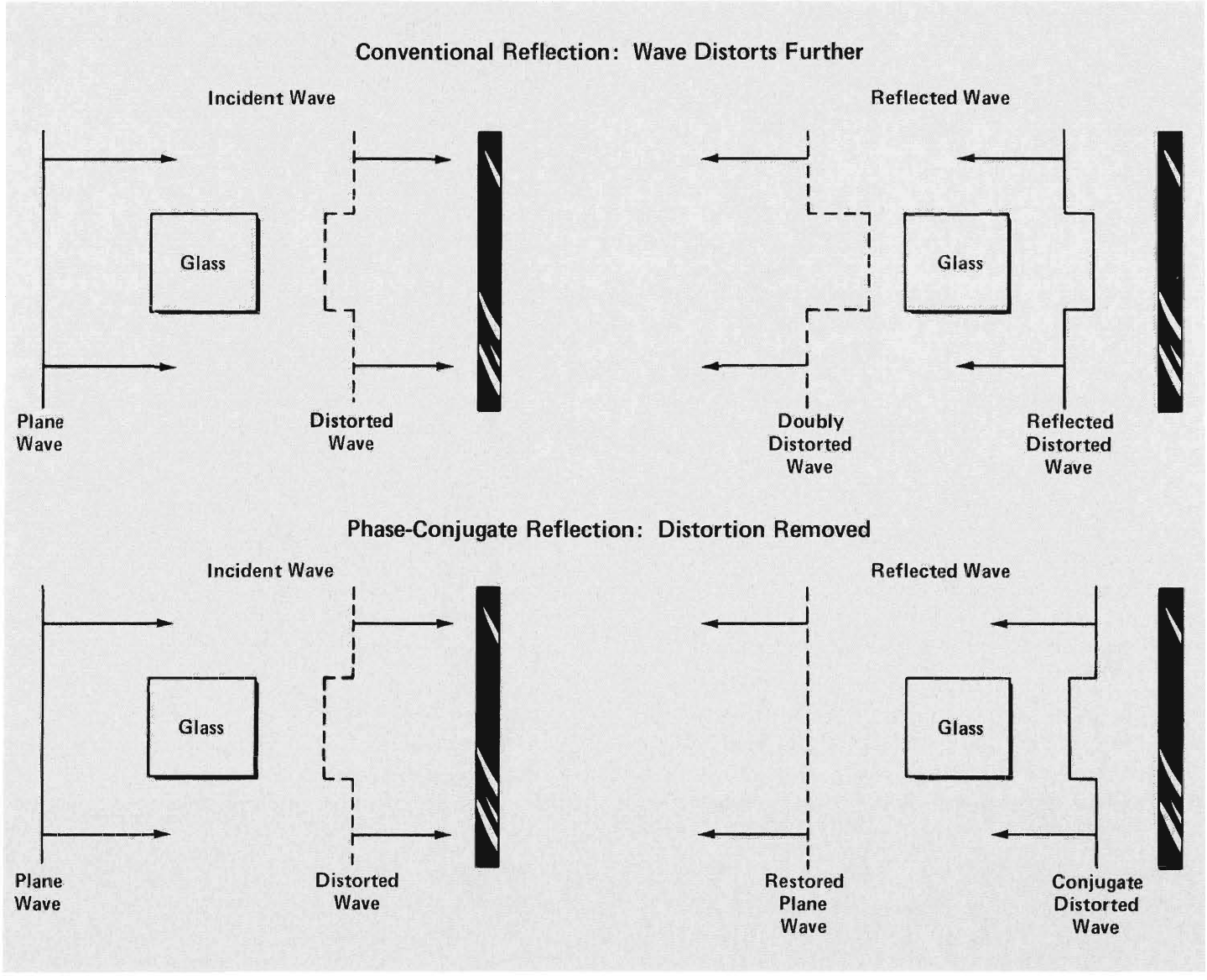


Fig. 2. Phase distortion with conventional and phase-conjugate reflection. In both cases the incoming plane wave (left side) encounters a block of glass, and a distorted wave is formed because the glass, with a different refractive index, retards the phase of the wave's central region. Conventional reflection (top right) retains this lag in phase so that the return trip through the glass doubles the distortion. On the other hand, phase-conjugate reflection (bottom right) changes the lag in phase to an advance in phase so that the return trip removes the distortion and a plane wave emerges, as if the wave had traveled backward in time.

torting medium, and then reflects from a phase conjugator (Fig. 2). The aberrator changes the beam into a distribution given by E_{in} that contains information about all of the phase distortions introduced by the medium. The phase conjugator then converts E_{in} into a new distribution E_{out} (by complex-conjugating all amplitudes and by reversing all wave vectors).This reflected beam is *exactly* programmed so that after passing backward through the aberrator it becomes a backward propagating replica of the original beam. The emerging beam does not contain any evidence that the aberrator existed! Thus, a high-quality optical beam can be double passed through a poor-quality optical system with no overall loss in beam quality. This double-passing technique can be applied to many problems in which a distorting medium, such as the turbulent atmosphere or a multi-mode optical fiber, would be

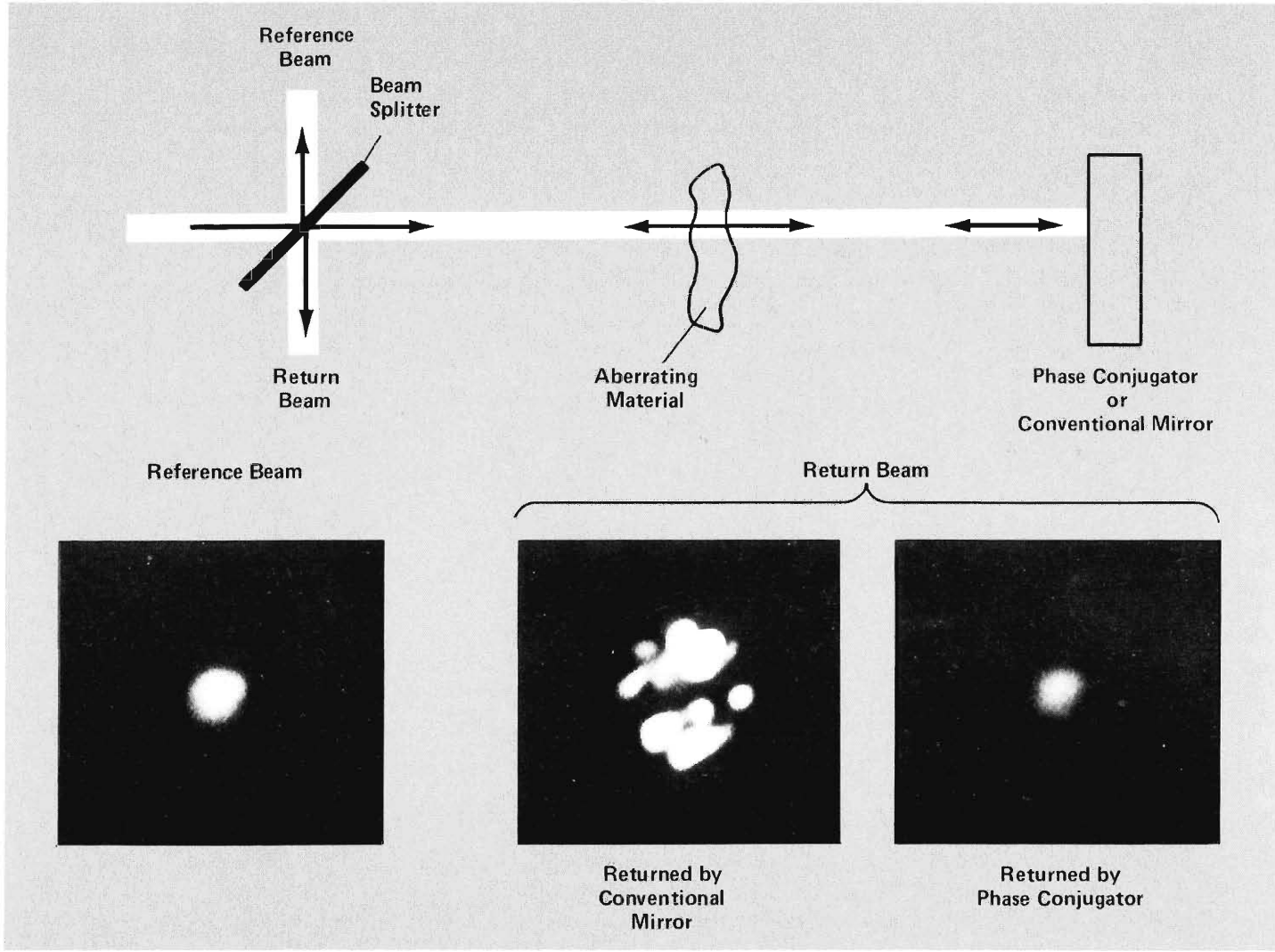


Fig. 3. Experimental demonstration of phase-conjugate reflection. An undistorted laser beam (left photograph) is double passed through an aberrating material. Conventional reflection for the return trip results in a highly distorted beam (middle photograph), whereas phase-conjugate reflection removes the distortions and only a uniform intensity change is obvious (right photograph).

detrimental to effective beam transport.

Figure 3 shows an experimental demonstration, with the aberrator pictured in Fig. 4, of this amazing feature of phase-conjugate reflection. Only two conditions are required to insure repair of the distorted beam. First, the phase-distorting aberrator must not undergo any changes during the time it takes for the beam to strike the phase conjugator and return; second, the light itself must not affect the physical properties of the aberrator.

It should now be clear why, when one looks at an ideal phase conjugator, one sees "nothing." All the light impinging on an ideal phase conjugator returns exactly on the path from where it came. Light glancing off one's nose, for example, is reflected directly back to one's nose, not into one's eyes. The only light an observer has a chance of seeing is that reflected off one's eyeball to the mirror and back. This is perhaps not quite nothing, but not much either. For those who believe that the eye is the "window to the soul," the phase conjugator allows the possibility of soul searching (patent pending), at least in the technical sense.

How Does One Make Such a Mirror?

In principle, if the phase distortions in a beam of light were known in advance, then one could design a mirror with a compensating surface to perform as a phase conjugator. Indeed, this is the principle behind the field of adaptive optics, in which a mirror surface is controlled and modified in such a manner as to reverse the phase front of an incoming beam (Fig. 5). Typically, the elements used for the shaping of this "rubber" mirror are piezoelectric crystals whose lengths change precisely when the voltages across their faces are changed. Such mirrors have been built, and research on improving their properties is proceeding in a number of laboratories. However, these mirrors suffer from slow response time (about 1 millisecond), imperfect correction due to the finite spatial resolution of each piezoelectric element, and expense in the construction and computer control of the large number of piezoelectric elements generally involved. In contrast, the phase conjugators discussed in this article (which invoke nonlinear optical techniques) need not suffer from such limitations.

NONLINEAR OPTICS. The methods to be discussed henceforth invoke processes en-

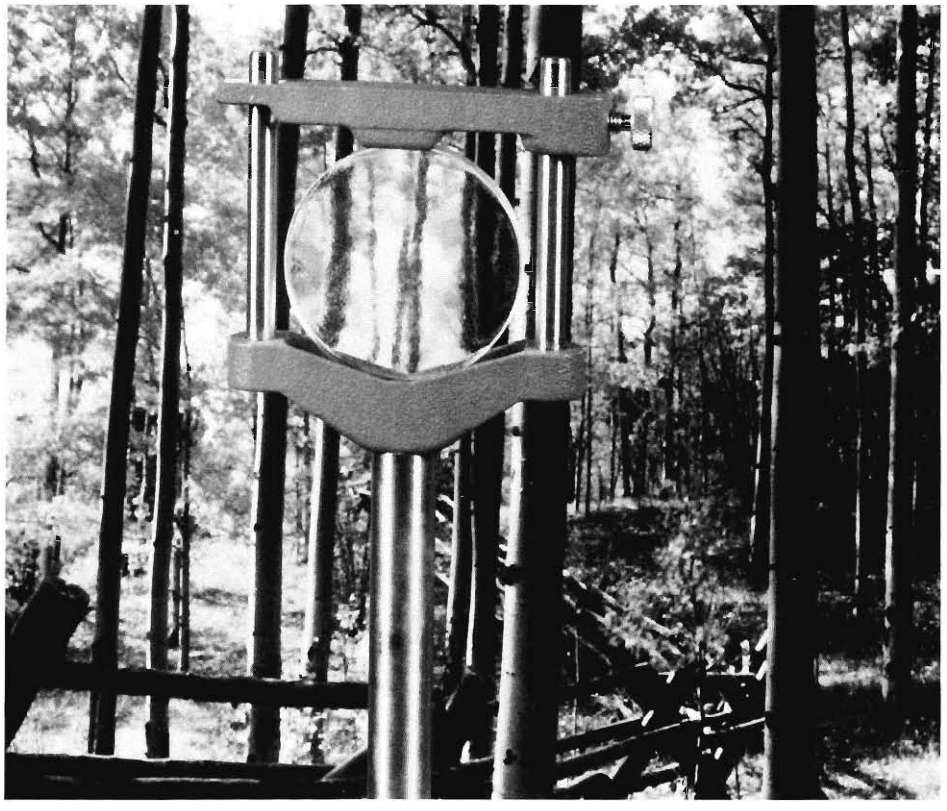


Fig. 4. Is this optical element useful? The distorted sodium chloride window in this picture was used as an aberrator in the experiment of Fig. 3 to illustrate the healing properties of phase-conjugate reflection. This technique becomes an attractive option when the quality of key optical components is limited by expense or technical considerations.

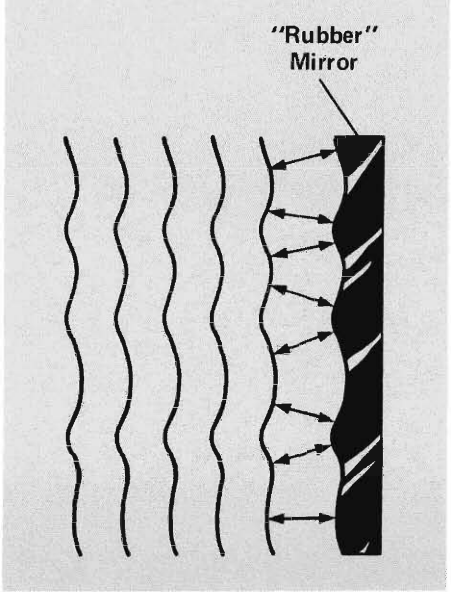


Fig. 5. Adaptive optics. If the phase distortions of the wavefront of an optical beam are known, a mirror surface can be shaped such that its surface is normal to the wave's propagation vector at every point. The reflected beam would then be the phase conjugate of the incoming beam because conventional reflection normal to the surface reverses the sign of the propagation vector.

tirely different from those of the flexible mirror described above, although the desired end result, formation of the conjugate wave, is the same. The research carried out at Los Alamos addresses the field that has become known as *nonlinear* phase conjugation. In this approach the processes that generate a phase-conjugate reflection depend upon the nonlinear response of matter to an optical field. (Generally, the nonlinearity of the response attains a useful magnitude only at the field intensities available from a laser beam.) There exists a plethora of these effects. In general, if a nonlinear response causes the refractive index of a medium to change with optical intensity, then the interference pattern formed by two or more laser beams can produce a volumetric index-of-refraction grating in the medium. Such gratings are the key to the magic of phase conjugators. But what is a refractive-index grating and why is it important?

First, it should be remembered that the refractive index is a relative measure of the speed of light through a material. As a result, the refractive index appears as a factor in the propagation vector ($|\mathbf{k}| = 2\pi n/\lambda$, where n is the refractive index and λ is the wavelength of the light in vacuum). The refractive index thus directly influences the oscillatory factor containing the phase information. Any physical process that alters the refractive index in a region of a material will, in turn, alter the phase of any light passing through that region. The trick, of course, is to alter the refractive index in just the right way so that the material scatters the light wave into its conjugate.

To further understand refractive-index gratings, we turn momentarily to holography. In fact, the true father of phase conjugation may well be the person who developed the notion of the hologram, Dennis Gabor (with help from W. L. Bragg). We say this because there are important similarities between holography and optical phase conjugation. One of the most important optical phase-conjugation techniques, which will be discussed later, is called degenerate four-wave mixing and is essentially real-time optical holography.

Consider the making of a holographic image (Fig. 6a). Typically, the light from a laser is split into two plane-wave beams. One, the reference beam, remains undistorted. The second is reflected diffusely off the object, causing the optical phase front to be distorted. The reference beam and the distorted beam are then directed from different angles onto a photographic film where they meet to form an interference pattern. All the phase information implicit in the interference is recorded as a fine pattern of silver grains in the developed film emulsion; the interference pattern has been "written" permanently into the film. Later, the pattern is "read" by directing at the film from the rear an undistorted plane wave (Fig. 6b). In this case, the grains of silver act as a grating and scatter the light to generate a distorted beam with the same phase relationships of the original distorted beam (when viewed from the same angle). This scattered beam is seen by the eye as a virtual image of the object.

The key to holography, of course, is the

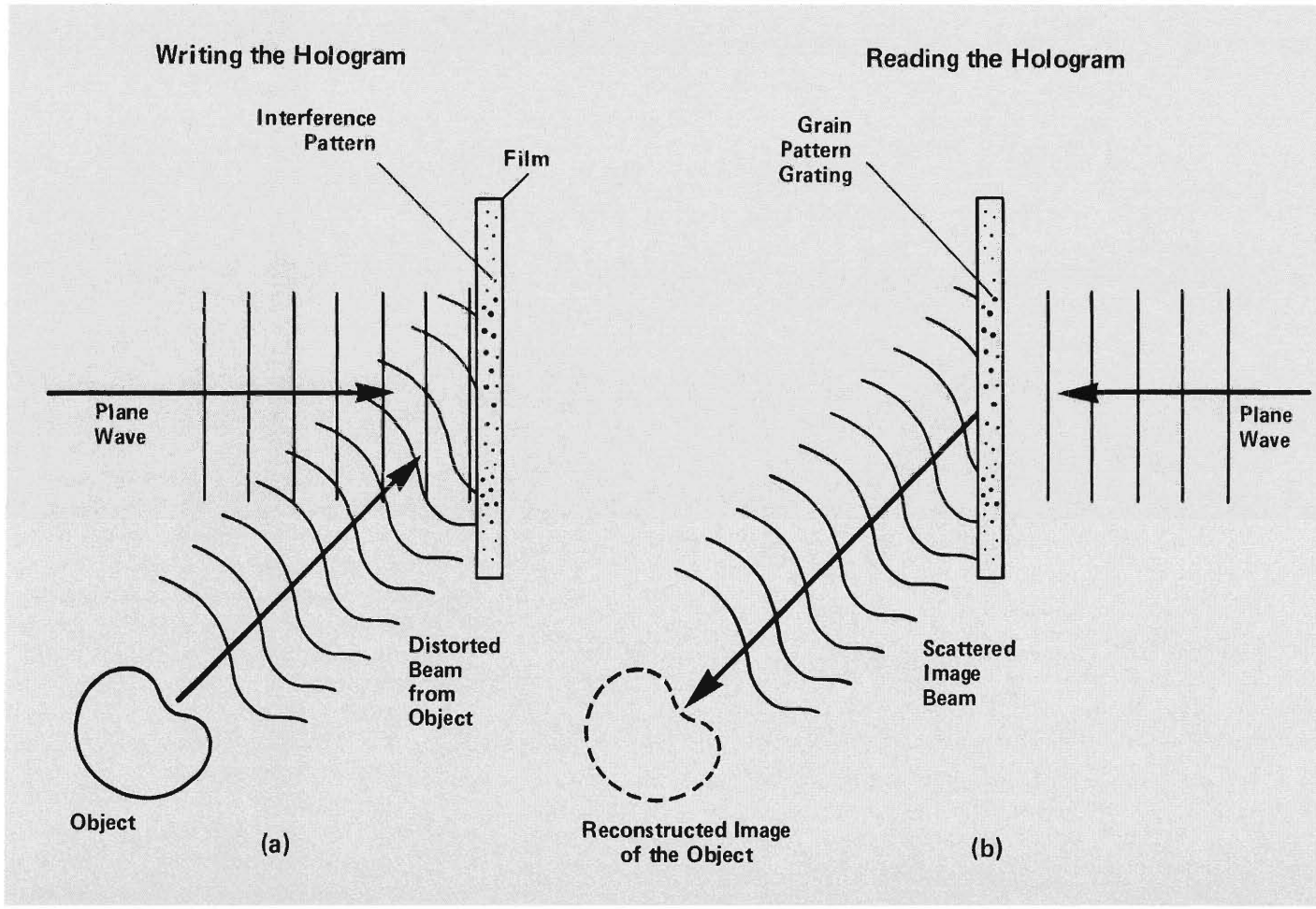


Fig. 6. Conventional holography consists of two distinct "write" and "read" steps. (a) First, the film is exposed to the interference pattern formed by an undistorted reference beam with a distorted beam reflected off the object. The result, after development of the film, is the hologram, a grain pattern in the emulsion. (b) A second undistorted reference beam, here traveling in the exact opposite direction, reads the hologram by scattering off the pattern of grains. Because the various scattered waves interfere with each other, these grains act as a heterogeneous grating. When viewed at the original angle, the phase relationships of the distorted beam will have been reconstructed, creating an image of the object.

pattern formed in the film emulsion. But this is a permanent grating. What is needed for phase-conjugate reflection is some medium in which a grating is written and read simultaneously; that is, the incident distorted beam generates a grating pattern that immediately scatters the reflected beam in the opposite direction with the conjugate phase relationships of the original. To set up such a grating we invoke nonlinear optics.

The nature and effectiveness of a refractive-index grating depend strongly on the nonlinear mechanism coupling the light and the material. Many such mechanisms are available. For example, if the optical wavelength corresponds to an absorption wavelength in the material, then the absorbed energy will give rise to heating of the material and a corresponding modification of the refractive index at that wavelength. If the absorption is bleachable (that is, if the material becomes more transparent as more energy is absorbed), then the index of refraction will change with intensity. However, if the material is nominally transparent, then other effects typical of nonlinear optics (such as those called stimulated Brillouin scattering, the optical Kerr effect, stimulated Raman scattering, and multiple-photon absorption) can be used to produce a refractive-index grating. The material itself can be a solid, liquid, gas, or plasma or more exotic systems such as liquid crystals, dielectric particles within a liquid, gaseous bubbles, or bulk plasma within a solid.

In this article we will discuss two types of nonlinear mechanisms for phase conjugators: those involving elastic photon scattering, in which the conjugating medium is left essentially unchanged by the process, and those involving inelastic photon scattering, in which the incident photons deposit some of their energy in the medium. We will treat an important example of each.

DEGENERATE FOUR-WAVE MIXING. An example of an elastic photon-scattering process in nonlinear optics is degenerate four-wave mixing, the phase-conjugation technique that corresponds to real-time holography. In this case the light and the material couple through a nonlinearity in the material's polarizability. When a light beam travels through a transparent material, its oscillating electric field generates a corresponding polarization wave by altering a number of properties (for example, the average position of the material's electrons). At low intensities the polarization can be taken to be directly proportional to the electric field ($P = \alpha E$). As a result, the induced polarization wave oscillates at the same frequency as the radiation but radiates its energy with a

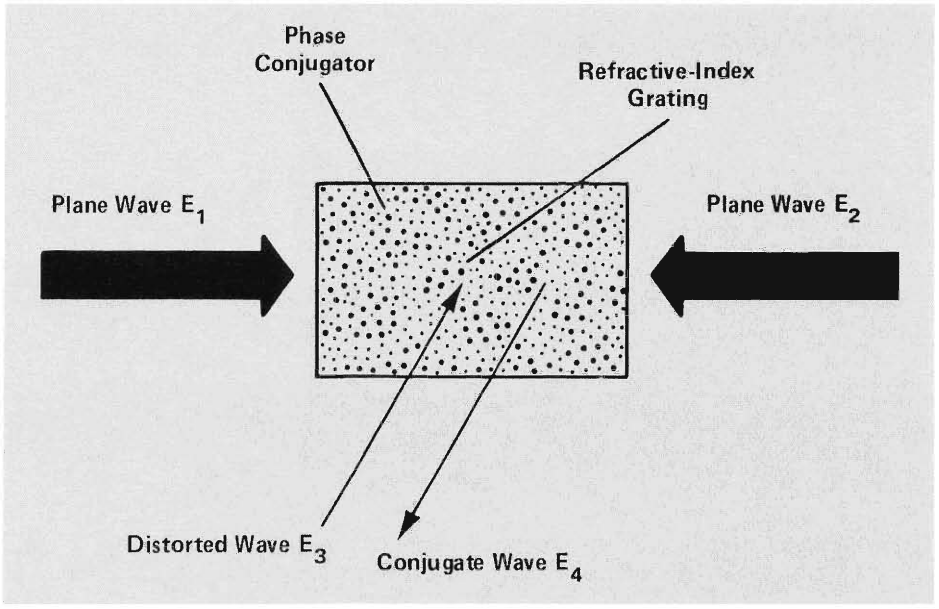


Fig. 7. In degenerate four-wave mixing the write and read steps of holography take place simultaneously. Interference of the intense plane wave E_1 ***and the distorted wave*** E_3 ***generates a refractive-index grating in the nonlinear optical material of the phase conjugator. The intense counterpropagating plane wave*** E_2 ***immediately scatters off this grating to form the reflected wave*** E_4 ***that is the phase conjugate of*** E_3***. Because the roles of*** E_1 ***and*** E_2 ***can be exchanged, a second wave, indistinguishable from*** E_4***, is also produced. The intensity of the reflected wave is a function of the three incident fields (*** E_1***,*** E_2***, and*** E_3***) and of the material properties of the phase conjugator (such as the magnitude of its nonlinear optical effect).***

time lag that retards the phase of the light beam (giving rise to the material's normal refractive index). At high enough intensities, however, the polarization becomes nonlinear and can be expressed as

$$P = \alpha_0 + \alpha_1 E + \alpha_2 E^2 + \alpha_3 E^3 + \ldots \ .$$

If the electric field is oscillatory ($E = Ae^{i\omega t}$), the higher-order terms in this equation cause the polarization wave to have a variety of frequency components that can radiate in new directions and at new frequencies and that alter the material's refractive index. The third-order term consists of a number of components, one of which is responsible for the polarizability changes used to generate the refractive-index grating in degenerate four-wave mixing.

In this process three fields of the same frequency impinge on a transparent or semitransparent material with a large third-order polarizability (Fig. 7). Two of the fields (E_1 and E_2) are counterpropagating, high-intensity plane waves (a reference field to help write the grating, another to read it), and the third (E_3) is the field one wishes to "reflect," or phase conjugate. In this environment the interference of the reference field E_1 with the field of interest E_3 generates a refractive-index grating. The other reference field E_2 experiences this bulk grating within the material and is partially scattered back along the direction of E_3. We refer to this scattered wave as E_4. However, the roles of E_1 and E_2 can be interchanged. Thus E_2 and E_3 establish a different refractive-index grating that partially scatters E_1 back along the direction of E_4. In general, the fields scattered from the two gratings are indistinguishable and both contribute to the phase-conjugate field E_4. Here we see that the sequential steps of normal holography—the formation of a grating and the subsequent scattering from it—are, indeed, accomplished simultaneously. It should be evident from this discussion, however, that in degenerate four-wave mixing E_4 is not really a reflection of E_3 but rather a scattering of E_1 and E_2.

Is the scattered field the conjugate of the incident field? The phase of a scattered beam is determined by the phase variations within the refractive-index gratings. Because of the unique phase relationships between the reference beam and the grating, the scattered field E_4 should be proportional to the complex conjugate of E_3. In fact, with the nonlinear polarization appropriate to degenerate four-wave mixing, Maxwell's equations give E_4 as everywhere *strictly* proportional to the phase conjugate of E_3. In degenerate four-wave mixing experiments it is crucial that E_1 and E_2 approximate plane waves within the interaction volume and that they be precisely counterpropagating; otherwise, the scattered radiation will not be exactly the conjugate of E_3.

Although degenerate four-wave mixing is a nonlinear optical effect generated by the interaction of three fields, the effect is nevertheless linear with respect to the field E_3 that is being phase conjugated. This means that a superposition of E_3 fields will generate a corresponding superposition of E_4 fields. Thus, accurate reconstruction of the original field (only propagating in the opposite direction) is possible.

If E_1 and E_2 are sufficiently intense and E_3 is weak, it is conceivable that E_4 will be more intense than E_3. Hence the phase-conjugate scattering can actually lead to a "reflectivity" greater than 100 per cent ($\mathcal{R} > 1$). This is accomplished, of course, not by generating light out of thin air but by scattering light from the intense fields E_1 and E_2 back into the direction of E_3, giving the appearance of amplification. Alas, energy is always conserved.

The origins of the concept of nonlinear optical phase conjugation are somewhat obscure owing to confused terminology and various incomplete demonstrations. Generally, B. I. Stepanov, E. V. Ivakin, and A. S. Rubanov of the Soviet Union are credited with the first demonstration in 1970 of distortion correction by degenerate four-wave mixing (similar work by J. P. Woerdman was nearly concurrent), and there is little doubt that the early pioneering in the field was by Soviet researchers. In particular, B. Ya. Zel'dovich, V. I. Popovichev, V. V. Ragul'skii, and F. S. Faizullov stand out as the first to recognize that nonlinear optical phase conjugation would also occur via stimulated processes such as our next example.

STIMULATED BRILLOUIN SCATTERING. This technique is an example of an inelastic photon scattering process. An intense laser beam is focused into a nearly transparent optical material where it Brillouin scatters off acoustic phonons (Fig. 8). As a result, the beam loses energy to the acoustic wave in the medium and is slightly reduced in frequency as it scatters back in the opposite direction. The high intensity of the focused laser beam literally drives the process to high efficiency by stimulating the scattering. Zel'dovich and coworkers were the first to demonstrate that this scattered beam was the phase conjugate of the incident beam. Here's how it works.

Intense optical radiation can interact with transparent media to produce material-density gradients by an effect called electrostriction. Electrostriction refers to the phenomenon in which a dielectric in an electric-field gradient experiences a force in the direction of increasing electric field. An analysis of this effect shows that the mechanical pressures in a liquid at the focal volume of commonly available lasers can exceed 100 atmospheres.

Now consider a strong incoming beam E_{in}, of frequency ω_{in}, moving through a material that exhibits electrostriction. Assume that the beam scatters off a sound wave (some acoustic noise always exists; for example, the laser beam itself can create such noise) to travel in the backward direction as E_{out} with lower frequency ω_{out}. The frequency shift of the light, $\omega_{in} - \omega_{out}$, is equal to the sound-wave frequency ω_s. Alternatively, assume that the incoming beam interferes with *optical* noise. If some of the optical noise happens to have the frequency ω_{out} and is propagating opposite to E_{in}, the two will interfere and produce a moving intensity grating. Because of electrostriction the intensity grating generates a sound wave, or density grating, of spacing $\lambda_s = v_s/\omega_s$, where v_s is the sound velocity.

Thus, there are two concurrent processes being described here. In one E_{in} interacts with a sound wave to produce E_{out}. In the other E_{in} interacts with E_{out} to produce a sound wave. For exactly the right set of frequencies and wave vectors, these two processes will reinforce each other by positive feedback and E_{out} will grow exponentially (until E_{in} is significantly depleted). Exponential growth will be fastest when E_{out} is precisely the phase conjugate of E_{in}, and thus non-phase-conjugate scattering is suppressed.

The acoustic wave generated in this process travels in the same direction and, most important, with the right phase fronts to conjugate the incident wave E_{in}. In essence, a "rubber grating" has been created in the conjugating medium whose scattering planes are always correctly aligned to reflect the conjugate wave.

Of course, the effectiveness of stimulated Brillouin scattering as a phase conjugating process is also dependent on the phase coherence of the incident beam, the extent of its phase disturbances, and the depth of the established grating. The details of these dependences are only now beginning to be appreciated.

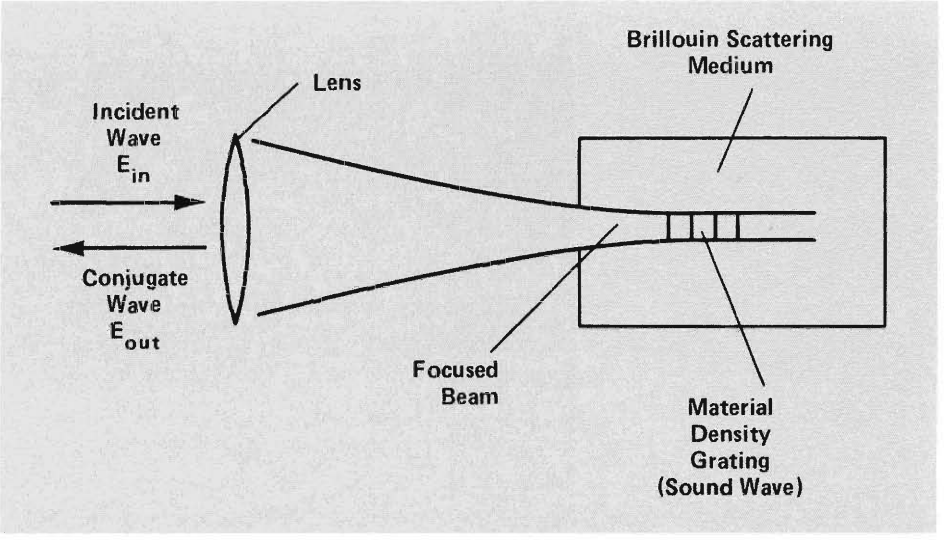


Fig. 8. Backward stimulated Brillouin scattering. The incoming field E_{in} *interacts with an acoustic wave to produce a backward scattered wave* E_{out} *and also interacts with* E_{out} *(through the medium) to produce an acoustic wave. The two processes positively reinforce each other only when* E_{out} *is the phase conjugate of* E_{in}.

Infrared Phase Conjugators

Work on nonlinear optical phase conjugation received a late start in this country, and it wasn't until the work on this phenomenon in 1976 by R. W. Hellwarth and in 1977 by A. Yariv and D. M. Pepper that phase-conjugation studies began in earnest in the United States. Not long thereafter, work began at Los Alamos in the laser fusion effort when it became apparent that nonlinear optical phase conjugation held promise not only for improvement of the beam quality of large-aperture lasers but also for improved target sighting and tracking of the tiny fusion pellets. (More will be discussed about applications later.)

Because the Los Alamos candidate in the laser fusion derby was the carbon dioxide (CO_2) gas laser operating in the infrared at a wavelength of 10 micrometers, the challenge was to find efficient nonlinear optical phase-

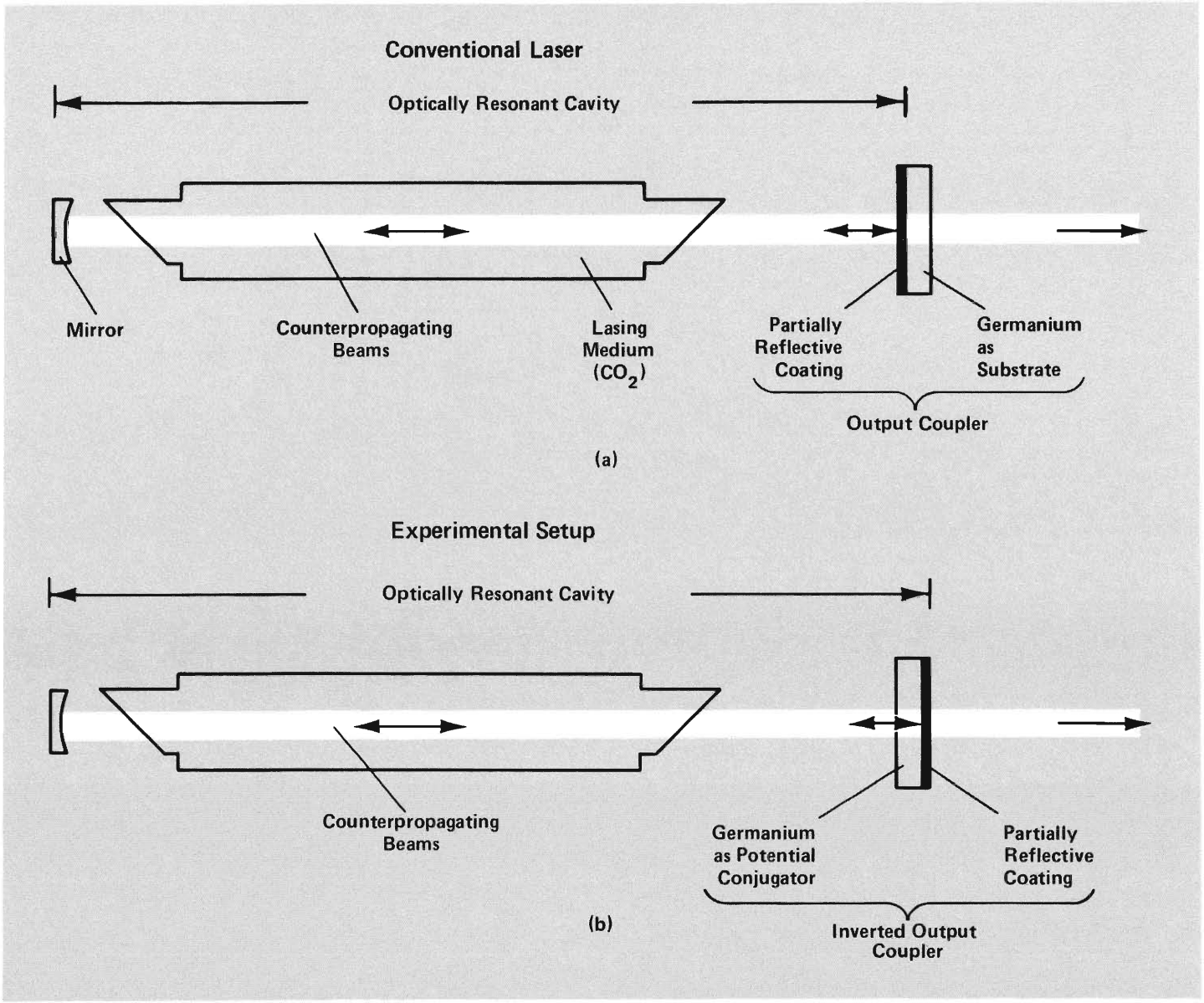


Fig. 9. Reconfiguring the carbon dioxide (CO_2) laser cavity for the degenerate four-wave mixing experiment. (a) The conventional CO_2 laser uses germanium as the substrate for the partially reflective coating on the output coupler. (b) Flipping the output coupler places the germanium, a highly nonlinear optical material, on the inside of the optically resonant cavity and thus within the intense, almost perfectly counterpropagating fields that constitute the laser cavity's standing wave.

conjugating materials at this wavelength. In March 1978, Ernest Bergmann, Irving Bigio, Barry Feldman, and Robert Fisher successfully produced the first demonstration of infrared optical phase conjugation with a CO_2 laser utilizing germanium as the nonlinear material.

Germanium had played an important role in CO_2 laser technology for many years. As an easy to grow, easy to polish, optically transparent material in the infrared, it had long been used as the substrate material that is coated with a partially reflecting, partially transmitting film to make it into a CO_2 laser mirror. Figure 9a shows the material in use as the substrate for a laser "output coupler" with the reflective coating toward the inside of the laser cavity. This device transmits part of the beam out of the laser and reflects the rest back into the optically resonant cavity where the counterpropagating beams form a standing wave. Note that the germanium material itself is outside the laser cavity.

With one of those welcome flashes of recognition, it was realized that a simple reversal of the output coupler (Fig. 9b) would immediately satisfy many of the requirements for degenerate four-wave mixing. This trivial operation placed the germanium substrate, which has a rather large nonlinear optical coefficient, inside the cavity, where it was exposed to the high-intensity intracavity electromagnetic field. Moreover, the two beams making up the standing wave inside an optical resonator are almost perfectly counterpropagating plane waves by design; the problems of misaligned and converging or diverging beams were thus readily avoided. All that was needed to complete the experiment was to redirect at an oblique angle the output of the laser back into the illuminated portion of the germanium substrate (Fig. 10). Lo and behold, phase conjugation occurred in the germanium. The reflectivity measured in that first experiment was only 2 per cent, but the work represented a breakthrough in CO_2 laser development and demonstrated that optical phase

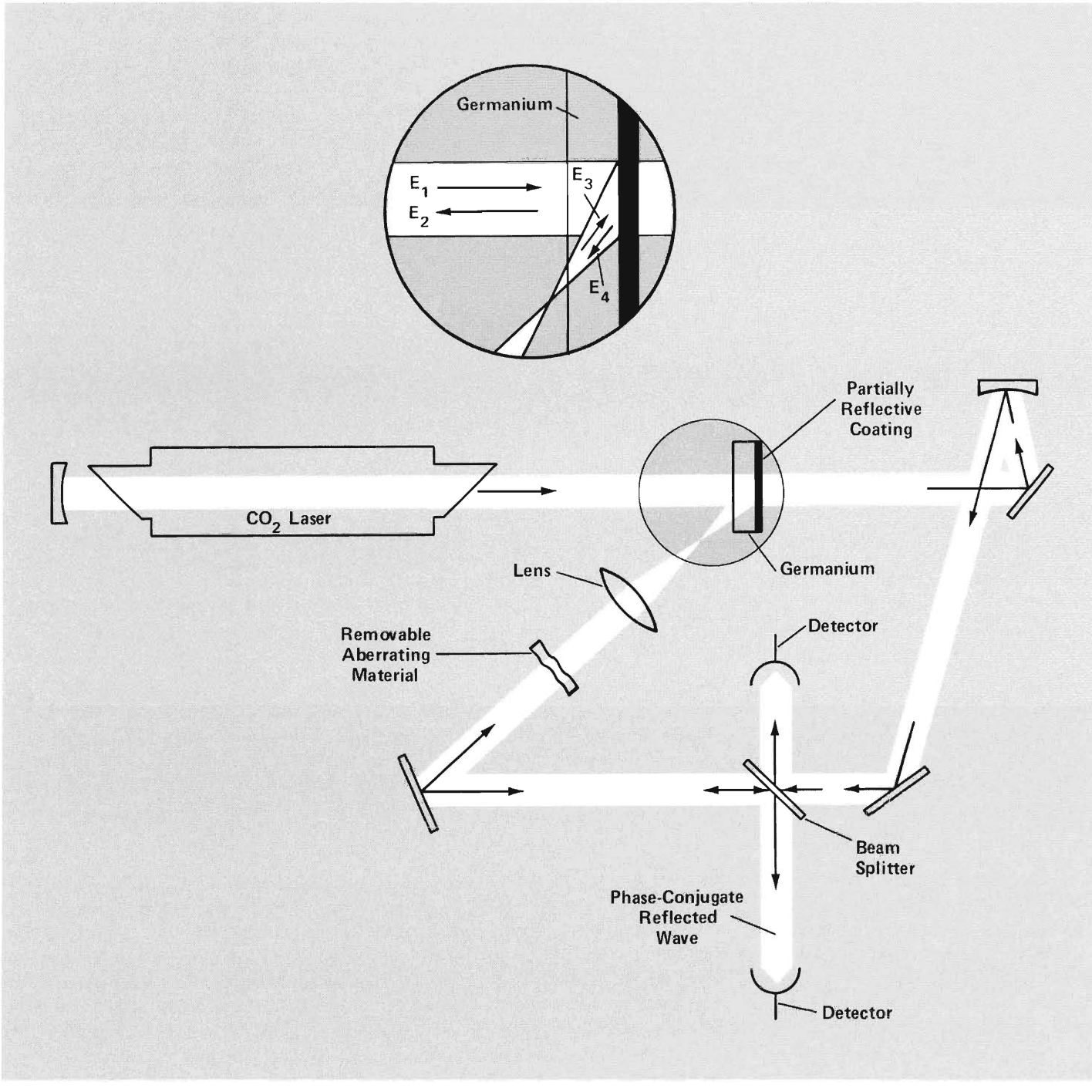


Fig. 10. Degenerate four-wave mixing in the infrared. Here the CO_2 laser shown in Fig. 9b with counterpropagating fields E_1 and E_2 has part of its output directed at an angle back into the germanium to form the field E_3. Since this arrangement provides the proper conditions for degenerate four-wave mixing (Fig. 7), the phase-conjugate wave E_4 is generated. The beam splitter allows both the original laser output beam and the reflected conjugate beam to be monitored by infrared-sensitive detectors. An aberrator can be placed in the beam to check the phase-conjugate properties of the reflected wave. The CO_2 laser has been simplified here for the sake of clarity.

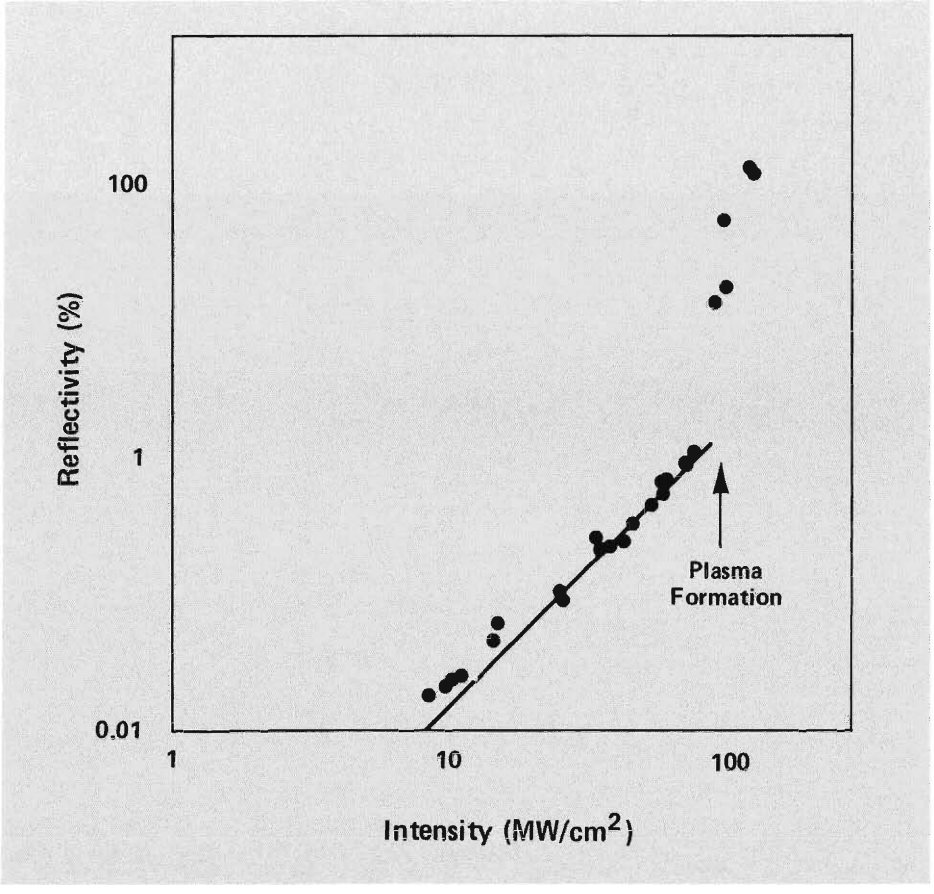


Fig. 11. Phase-conjugate reflectivity of germanium as a function of intensity. High field intensities in germanium give rise to a high-density electron plasma within that material. This creates large optical nonlinearities and phase-conjugate reflectivities of 200 per cent or greater.

conjugation was both possible and simple to achieve with materials already on hand in most laboratories involved in research on CO_2 lasers.

After these initial experiments, continued work on germanium by Claude Phipps and David Watkins revealed more surprises from this innocent looking material. In a carefully controlled experiment with germanium outside the laser cavity, they demonstrated, for field intensities of 100 megawatts per square centimeter and greater, that the phase-conjugate reflectivity increased dramatically. Apparently, at these high intensities free electrons were generated by multiple-photon absorption across the 0.6-electron-volt indirect band gap of germanium. This rapidly gave rise to a high-density electron plasma (2×10^{15} electrons per cubic centimeter) within the bulk germanium. Such a highly nonlinear process produced a dramatic increase in the phase-conjugate reflectivity of the material. Reflectivities greater than 200 per cent were demonstrated for germanium samples (Fig. 11).

Concurrently, Fisher and Feldman used the CO_2 gain medium itself as an optical phase conjugator by using the saturation properties of the excited CO_2 gas mixture to establish a field-dependent population grating. Because of larger interaction volumes and favorable gain conditions, effective phase-conjugate reflectivities greater than 400 per cent were obtained. At this same time, Fisher, Feldman, and Bergen Suydam carried out theoretical work on the pulse characteristics of optical phase conjugation.

Further CO_2 laser research was done by Watkins on a saturable absorber consisting of potassium chloride doped with rhenium tetroxide. This work confirmed many of the theoretical predictions about phase conjugation by ideal saturable absorbers.

Ultraviolet Phase Conjugators

Throughout 1979 substantial developments in the field continued worldwide for both the infrared and visible portions of the spectrum; there were, however, no observations of phase conjugation in the ultraviolet. Because of the increasing importance of ultraviolet lasers in photochemical and fusion research, Los Alamos researchers focused their attention on this part of the spectrum. Using pulses of 20-picosecond duration from a Nd:YAG laser whose emission had been quadrupled in frequency to yield light at a wavelength of 266 nanometers, Feldman, Fisher, and Stanley Shapiro set up the degenerate four-wave mixing experiment shown in Fig. 12. The increased complexity (when compared with the previously described experiment of Figs. 9 and 10) was required because great care had to be taken to insure temporal overlap of the very short pulses within the phase-conjugating medium by making the optical path lengths of each of the three interacting beams equal to within about 1 millimeter.

Liquid carbon disulfide (CS_2) was one of the most attractive conjugator candidates because of its large nonlinear optical coefficient. Although CS_2 is strongly absorbing in the ultraviolet, dilution with hexane produced a "window" between the two strong absorption peaks centered at 230 and 330

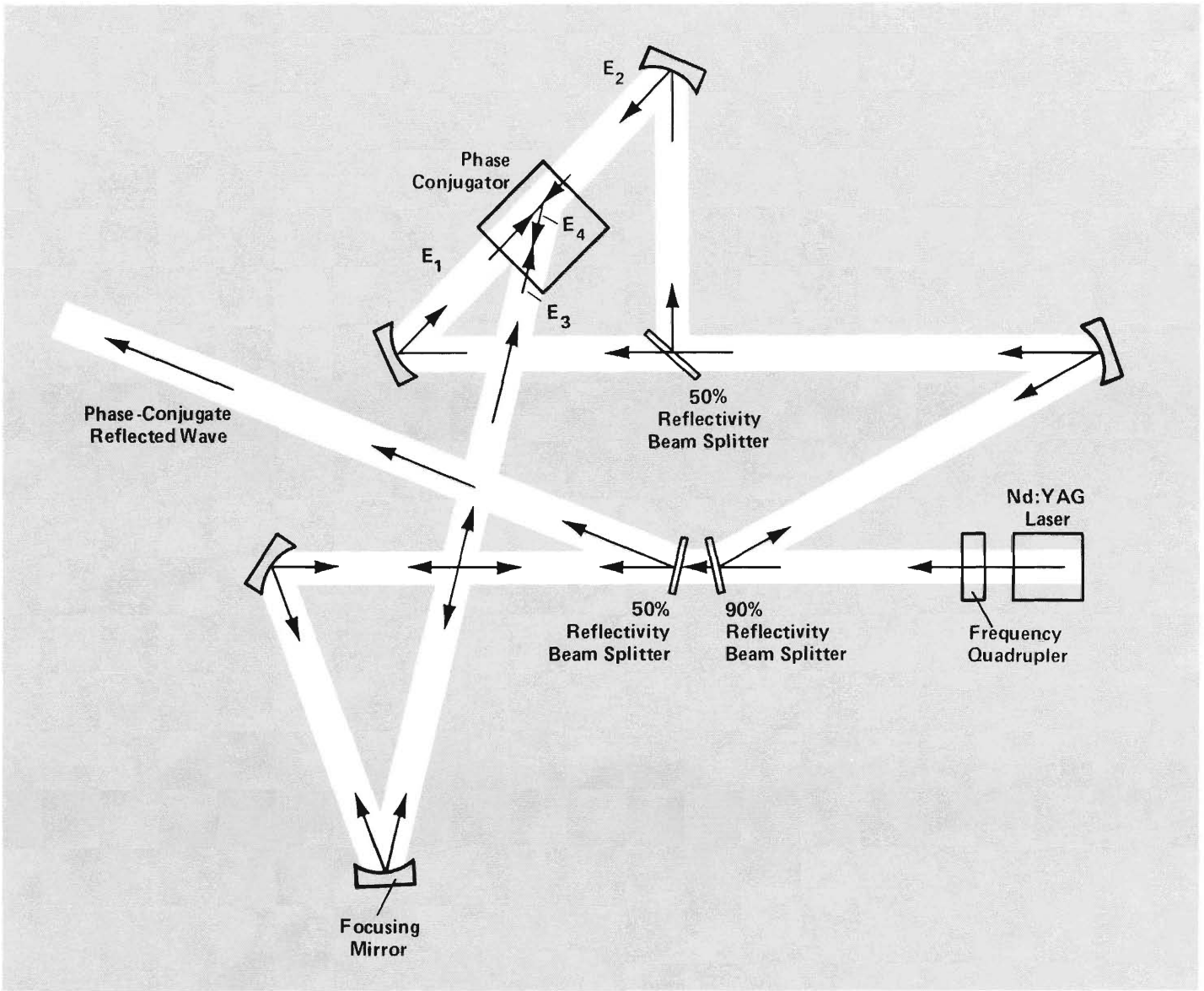


Fig. 12. Degenerate four-wave mixing in the ultraviolet. The frequency-quadrupled output (266-nanometer wavelength) of a Nd:YAG laser is split so that 90 per cent of the beam is directed to the phase conjugator as two counterpropagating beams (E_1 and E_2). Part of the remaining 10 per cent arrives as E_3 at the conjugator from a different angle and is phase-conjugate reflected (E_4).

nanometers. The transmission window had the remarkable property of being tunable as a function of CS_2 concentration in hexane (Fig. 13). A 40-per cent (by volume) mixture of CS_2 in hexane was chosen to optimize the nonlinear interaction at 266 nanometers. Although conjugate reflectivities of only 0.1 per cent and less were observed from the CS_2 -hexane mixture and from several other materials, these observations represented the first demonstration of nonlinear optical phase conjugation in the ultraviolet and gave impetus for further development.

Work in the ultraviolet continued at Los Alamos with several other notable achievements. This work was motivated by the development in the late '70s of a new class of lasers, the rare-gas halide excimers. The excimer lasers offered for the first time the possibility of high-power, high-efficiency emission at various wavelengths in the ultra-

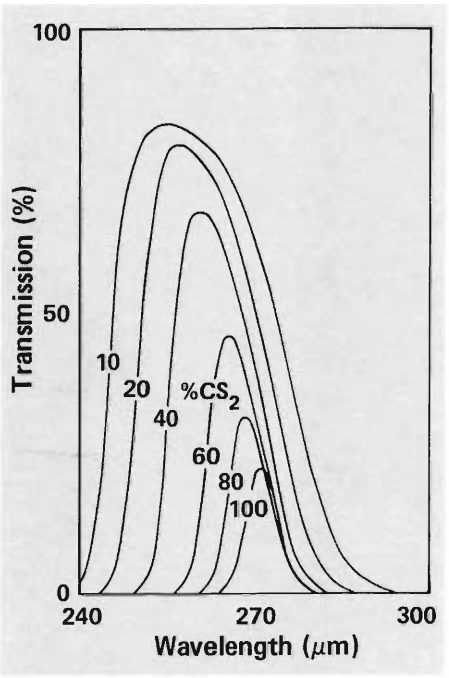


Fig. 13. The transmission spectra of various CS_2-hexane solutions. The transmission increases, broadens, and shifts as the percentage of CS_2 in the mixture decreases. These curves are for 1-millimeter path lengths through the sample.

violet. Using a high-power spectrally narrowed krypton fluoride laser at a wavelength of 248.6 nanometers, Bigio, Michael Slatkine, Feldman, and Fisher successfully demonstrated optical phase conjugation, again based on degenerate four-wave mixing in various liquid solutions. Similar successes were achieved with a xenon fluoride laser at 351 nanometers using backward stimulated Brillouin scattering in various organic liquids (Fig.14). In the latter case phase-conjugate reflectivities of over 70 per cent were clearly demonstrated. In another experiment nearly phase-conjugate reflectivities of about 30 per cent were observed using backward stimulated Raman scattering in liquid nitrogen. This process is, in essence, the same as stimulated Brillouin scattering except that rather than coupling with sound waves, energy from the incident beam is deposited into the vibrational energy levels of the nitrogen molecules. One of the remarkable features of this experiment, and of stimulated Raman scattering in general, is the large wavelength shift of the scattered beam with respect to the incoming beam. In this case the phase-conjugate beam at 382 nanometers was visible whereas the incoming beam at 351 nanometers was not. This wavelength shift precisely equals the difference between energy levels of the vibrational mode of the nitrogen molecule, a relatively large energy change.

In all cases involving these excimer lasers, whose emission is normally broad in frequency, phase conjugation could be observed only when the laser was constrained to operate within a narrow frequency bandwidth. Put simply, a broad range of frequencies results in a "smeared" interference pattern and a nondistinct refractive-index grating that fails to scatter the beam efficiently. The necessary bandwidth reduction was achieved by a process called injection locking in which a much weaker laser at the same frequency but with a narrow bandwidth controls the laser of interest. This technique was perfected at Los Alamos by Bigio and Slatkine. For example, the xenon fluoride laser was successfully injection locked using a weak, narrow-bandwidth argon-ion laser operating at a wavelength coincident with one of those of the xenon fluoride laser. As little as one watt from the argon-ion laser was sufficient to control the

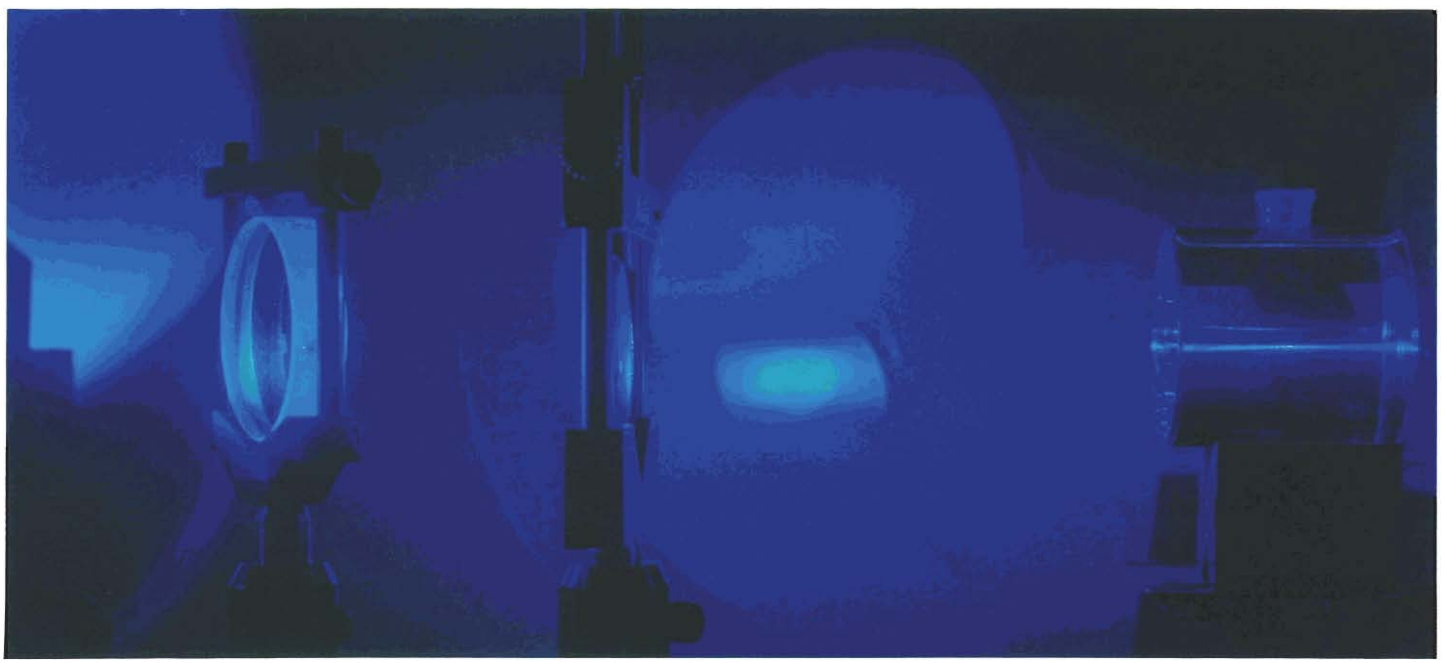


Fig. 14. In this photograph an ultraviolet light beam from a xenon fluoride laser passes through the optics from left to right and is phase-conjugate reflected by liquid hexane in the cell on the right via stimulated Brillouin scattering. The visible light beam in the cell is due to fluorescence. Part of the phase-conjugated return beam is diverted by the beam splitter on the left and appears as the spot in the background.

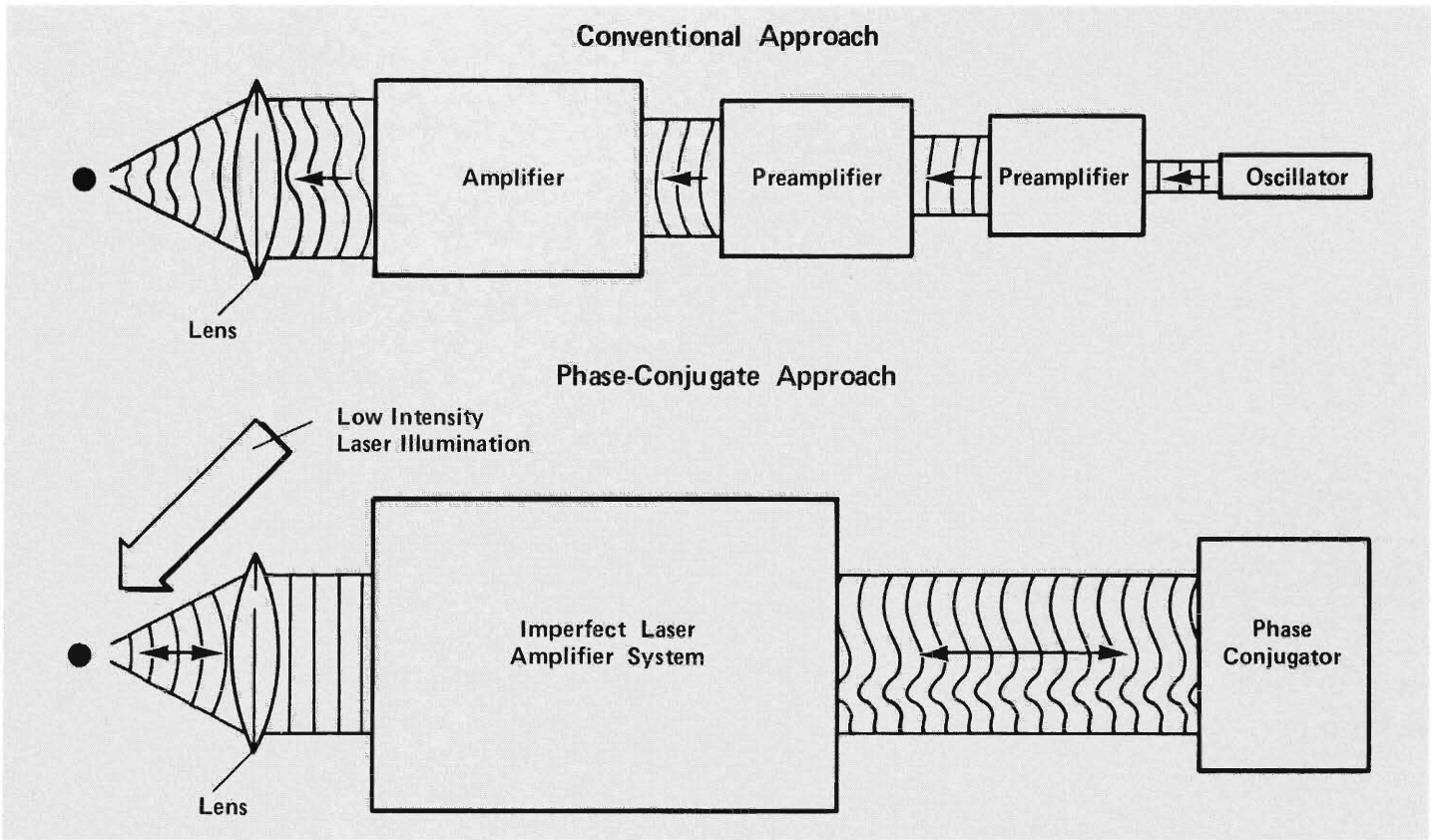


Fig. 15. Laser fusion systems. The conventional system (top) uses a long chain of laser amplifiers that may gradually introduce distortions in the beam arriving at the fusion target. In the phase-conjugate laser fusion system (bottom), a spatially broad, low-intensity laser illuminates the target. A small fraction of this illumination is reflected off the fusion target into the solid angle of the focusing optics and is amplified, phase-conjugate reflected, and further amplified on its return. Because the phase-conjugate beam exactly retraces its path, the amplified beam automatically hits the tiny fusion target. In addition, any phase distortions imparted to the beam by the complex amplification system will be removed on the return pass.

output bandwidth of the ten-million-watt xenon fluoride laser.

Applications of Optical Phase Conjugation

Although still in its infancy, the emerging field of nonlinear phase conjugation shows promise of revolutionizing the design of optical systems. As we have already discussed, the phase-conjugate beam has the remarkable property of emerging undistorted on its return pass through a distorting optical system. The advantages of this property for optical systems such as those involved in laser fusion, optical-fiber communication, and atmospheric propagation are enormous. Already the application of phase-conjugation techniques to the large fusion research lasers has resulted in their increased brightness on target. Moreover, the use of this technique (demonstrated in the Soviet Union) results in the automatic alignment of the beam on the fusion pellets. A schematic of such a phase-conjugating laser fusion system is shown in Fig. 15. Light from a low-intensity illumination laser is scattered off a fusion target. This illumination beam can be spatially broad and need not be critically aligned. Some of the scattered radiation is gathered in by a focusing system and undergoes amplification as it travels through the laser amplifiers. At the far end of the amplifier chain the radiation is returned by a phase conjugator through the laser chain for further amplification to exceedingly high intensities. Regardless of the optical distortions encountered on the first pass, the phase conjugator automatically redirects the beam back to its source, the fusion target. The amplified beam *cannot* miss! This technique allows the use of lower quality optics and eliminates much of the expense of the alignment systems usually required.

We now reconsider the scheme in Fig. 15, but this time with the laser and the target separated from each other by more than several hundred miles. Just as in the laser-fusion application of optical phase conjugation, similar aiming procedures could be used to direct laser light nearly instantaneously and accurately over long distances through the Earth's distorting atmosphere. These procedures could be extremely useful for communications systems.

Other potential applications of phase conjugation abound. The use of a phase conjugator as one of the cavity mirrors of a laser allows automatic cavity alignment and could lead the way to improved beam quality and stability. In fact, if a tunable laser is used to establish the counterpropagating beams for degenerate four-wave mixing, then external frequency control of the laser output is possible.

A phase conjugator has also been used as a fine optical frequency filter. In one of the injection-locking experiments described above, a xenon fluoride laser emitting radiation in roughly equal amounts at 351 and 353 nanometers was Brillouin scattered from

a variety of liquids. Because of injection locking by an argon-ion laser, the bandwidth of the radiation at 351 nanometers was much narrower than that of the 353-nanometer radiation. As a result, only the 351-nanometer light could form a distinct grating and only this radiation was efficiently backscattered. Thus all radiation but the narrow-bandwidth phase-conjugate component at 351 nanometers was filtered out by the scattering process.

Applications of phase conjugation have also been proposed in the use of photolithography. Potentially, the use of short-wavelength ultraviolet radiation should yield greater resolution and accuracy in the manufacture of microelectric circuits. However, distortions in the ultraviolet imaging systems have impeded the success of this application. Even with imperfect optics the unique imaging properties of the phase-conjugation process could result in far greater resolution and accuracy than heretofore has been possible.

Finally, a theoretical analysis of the quantum optical properties of a phase-conjugated beam arising from degenerate four-wave mixing indicates that a particular state (the so-called two-photon coherent state) of this radiation field possesses unique properties. These properties may allow substantial signal-to-noise improvements in certain light-detection schemes, improvements that would be especially pertinent to such applications as the detection of gravity waves.

In conclusion, optical phase conjugation is a rapidly expanding field that is radically altering the design of optical systems and their capabilities. Although not all of the proposed applications may prove to be more effective than other more conventional approaches, there is little doubt that some—and indeed many not yet even foreseen—will have a major impact on optical systems of the future. Much remains to be explored in this intriguing wonderland. ■

Further Reading

John Auyeung, D. Fekete, David M. Pepper, and Amnon Yariv, "A Theoretical and Experimental Investigation of the Modes of Optical Resonators with Phase-Conjugate Mirrors," IEEE Journal of Quantum Electronics **QE-15**, 1180-1188 (1979).

E. E. Bergmann, I. J. Bigio, B. J. Feldman, and R. A. Fisher, "High-Efficiency Pulsed 10.6 μm Phase-Conjugate Reflection via Degenerate Four-Wave Mixing," Optics Letters **3**, 82 (1978).

I. J. Bigio, B. J. Feldman, R. A. Fisher, and E. E. Bergmann, "High-Efficiency Wavefront Reversal in Germanium and in Inverted CO_2 (Review)," Soviet Journal of Quantum Electronics **9**, 1365-1369 (1979).

Irving J. Bigio and Michael Slatkine, "Transform-limited-bandwidth Injection Locking of an XeF Laser with an Ar-ion Laser at 3511 Å," Optics Letters **7**, 19-21 (1982).

B. J. Feldman, Robert A. Fisher, and S. L. Shapiro, "Ultraviolet Phase Conjugation," Optics Letters **6**, 84-86 (1981).

D. Gabor, "A New Microscopic Principle," Nature **161**, 777-778 (1948).

Concetto R. Giuliano, "Applications of Optical Phase Conjugation," Physics Today **34**, 27-35 (April 1981).

R. W. Hellwarth, "Generation of Time-Reversed Wave Fronts by Nonlinear Refraction," Journal of the Optical Society of America **67**, 1 (1977).

A. A. Ilyukhin, G. V. Peregudov, M. E. Plotkin, E. N. Ragozin, and V. A. Chirkov, "Focusing of a Laser Beam on a Target Using the Effect of Wave-Front Inversion (WFI) Produced as a Result of Stimulated Mandel 'Shtam-Brillouin Scattering (SMBS)," JETP Letters **29**, 328-332 (1979).

M. D. Levenson, K. M. Johnson, V. C. Hanchett, and K. Chiang, "Projection Photolithography by Wave-Front Conjugation," Journal of the Optical Society of America **71**, 737 (1981).

David M. Pepper, "Nonlinear Optical Phase Conjugation," Optical Engineering **21**, 156-183 (1982).

Michael Slatkine, Irving J. Bigio, B. J. Feldman, and Robert A. Fisher, "Efficient Phase Conjugation of an Ultraviolet XeF Laser Beam by Stimulated Brillouin Scattering," Optics Letters **7**, 108-110 (1982).

B. I. Stepanov, E. V. Ivakin, and A. S. Rubanov, "Recording Two-Dimensional and Three-Dimensional Dynamic Holograms in Transparent Substances," Soviet Physics Doklady **16**, 46-48 (1971).

D. E. Watkins, J. F. Figueira, and S. J. Thomas, "Observation of Resonantly Enhanced Degenerate Four-Wave Mixing in Doped Alkali Halides," Optics Letters **5**, 169-171 (1980).

D. E. Watkins, C. R. Phipps, Jr., and S. J. Thomas, "Observation of Amplified Reflection Through Degenerate Four-Wave Mixing at CO_2 Laser Wavelengths in Germanium," Optics Letters **6**, 76-78 (1981).

J. P. Woerdman, "Formation of a Transient Free Carrier Hologram in Si," Optics Communications **2**, 212-214 (1970).

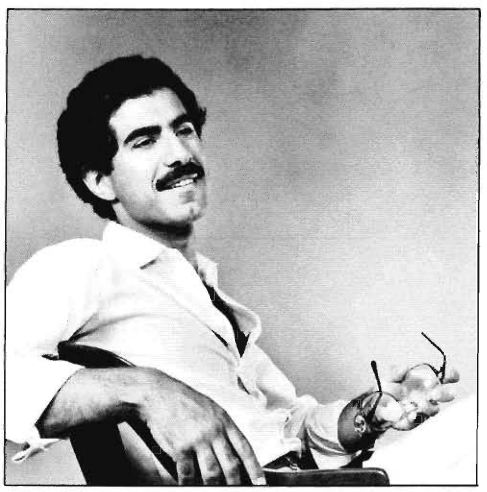
Amnon Yariv and David M. Pepper, "Amplified Reflection, Phase Conjugation, and Oscillation in Degenerate Four-Wave Mixing," Optics Letters **1**, 16-18 (1977).

Horace P. Yuen and Jeffrey H. Shapiro, "Generation and Detection of Two-Photon Coherent States in Degenerate Four-Wave Mixing," Optics Letters **4**, 334-336 (1979).

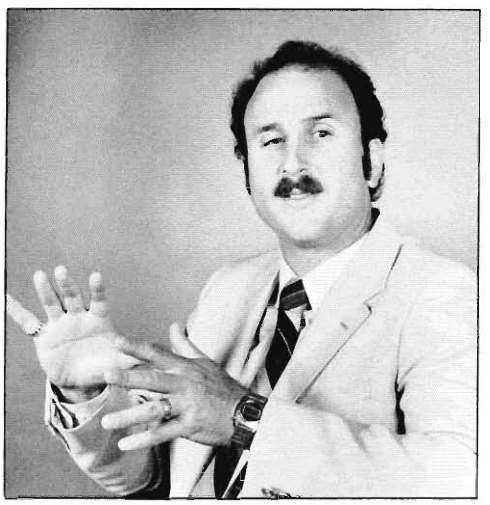
B. Ya. Zel'dovich, V. I. Popovichev, V. V. Ragul'skii, and F. S. Faizullov, "Connection Between the Wave Fronts of the Reflected and Exciting Light in Stimulated Mandel 'Shtam-Brillouin Scattering," JETP Letters **15**, 109-115 (1972).

AUTHORS

Barry J. Feldman received his Bachelor of Science from Brown University in 1965 and his Ph.D. in physics from Massachusetts Institute of Technology in 1971. It was at M.I.T., under the tutelage of Drs. Ali Javan and Michael Feld, that he first began his love affair with lasers. Upon graduation he came directly to Los Alamos where for several years he was involved in theoretical efforts related to laser fusion and laser isotope separation programs. His work included theoretical studies of laser coherence phenomena, laser pulse propagation, and Raman scattering. In 1976 he joined the CO_2 Laser Research and Applications Group as Associate Group Leader and was involved in the group's experimental efforts at ultrashort pulse generation, new laser development, optical phase conjugation, and nonlinear optics in the ultraviolet. Currently he has turned his attention to the study of nonlinear optical phenomena in organic and biological systems.

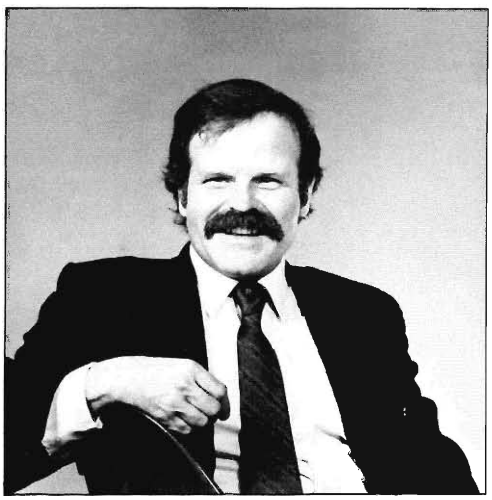


Irving J. Bigio received his B.S., M.S., and Ph.D. degrees in physics from the University of Michigan in 1969, 1970, and 1974, respectively. His doctoral work under John Ward and Peter Franken dealt with nonlinear optics, and he has maintained a broad interest in the field of quantum electronics ever since. He came directly to Los Alamos in April 1974 as a staff member in the laser isotope separation program and has also worked in the laser fusion program. In 1976 he received a Fulbright Senior Scholar Award and spent the 1976-77 academic year as a visiting professor at the Weizmann Institute of Science, Rehovot, Israel. During his tenure at the Weizmann Institute, he taught graduate courses in laser physics and nonlinear optics and helped direct graduate student research. Since returning to Los Alamos he has resumed his research and has taught courses at the University of New Mexico Graduate Center. Currently, he is working on a variety of topics in quantum electronics and has recently taken an interest in the application of laser techniques and nonlinear optics to the solution of biophysics problems.



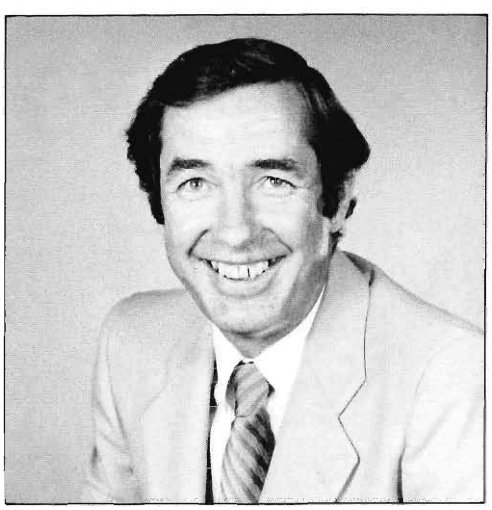
Robert A. Fisher received all of his schooling at the University of California, Berkeley, obtaining a B.A. in 1965, an M.A. in 1967, and a Ph.D. in 1971. He then joined the laser fusion effort at Lawrence Livermore Laboratory and concurrently taught at the University of California, Davis, before discovering New Mexico in 1974. While at Los Alamos he has worked in the Laser Fusion and Applied Photochemistry divisions. He was vice-chairman of the 1981 Gordon Conference on Lasers and Nonlinear Optics, and he served on the program committees for both the 1982 International Quantum Electronics Conference and the 1981 Annual Meeting of the Optical Society of America. He is the guest editor of a special issue on optical phase conjugation of the *Journal of the Optical Society of America* and is the editor of the soon-to-be-published Academic Press book entitled *Optical Phase Conjugation*. His professional interests include nonlinear optics, laser-related phenomena, optical phase conjugation, and molecular physics.

AUTHORS



Claude R. Phipps, Jr., ha[redacted]os Alamos since 1974. He received his B.S. and M.S[redacted]ng from the Massachusetts Institute of Technology in 196[redacted]3 and his Ph.D. in electrical engineering (plasma physics) from Stanford University in 1972. His research interests have ranged from superconductivity through Thomson scattering in plasmas to nonlinear optics at infrared wavelengths, particularly phase conjugation. He has also played a significant role in the measurement of infrared properties of optical materials. His wife, Lynn, is a commercial artist, and his son, David, is a physics major at Boston University. He is a member of the Society of Photo-Optical Instrumentation Engineers.

David E. Watkins earned his Bachelor of Science in 1975 from New Mexico Institute of Technology and his Master of Science in 1978 and his Ph.D. in 1981 from the University of Washington. He performed the research for his Ph.D. thesis, which involved phase conjugation by degenerate four-wave mixing, at Los Alamos as a graduate research associate. David has worked on high repetition rate CF_4 lasers and Raman conversion for the uranium enrichment program and maintains a strong interest in nonlinear optical phenomena.



Scott J. Thomas was born in Spruce Pine, North Carolina, on November 18, 1934. He joined the U.S. Air Force in 1955 and worked as an aircraft technologist in the Strategic Air Command. From 1961 to 1974 he was employed by Lawrence Livermore Laboratory in the Laser Fusion Division. He came to Los Alamos in 1974 and worked on laser research and development for the laser fusion program until 1981. Since then he has worked in the Applied Photochemistry and Chemistry divisions. He has published work on laser-produced plasmas, laser photochemistry, chemical lasers, dye lasers, gas lasers, nonlinear optical studies, and laser damage to optical surfaces. His present position as a staff member entails work on laser research and development.

the On & Off of Human Allergies

by Byron Goldstein and Micah Dembo

A study of the interaction between allergen molecules and a cellular surface bristling with antibodies has revealed details about the on-off switches for release of chemicals that cause allergic symptoms.

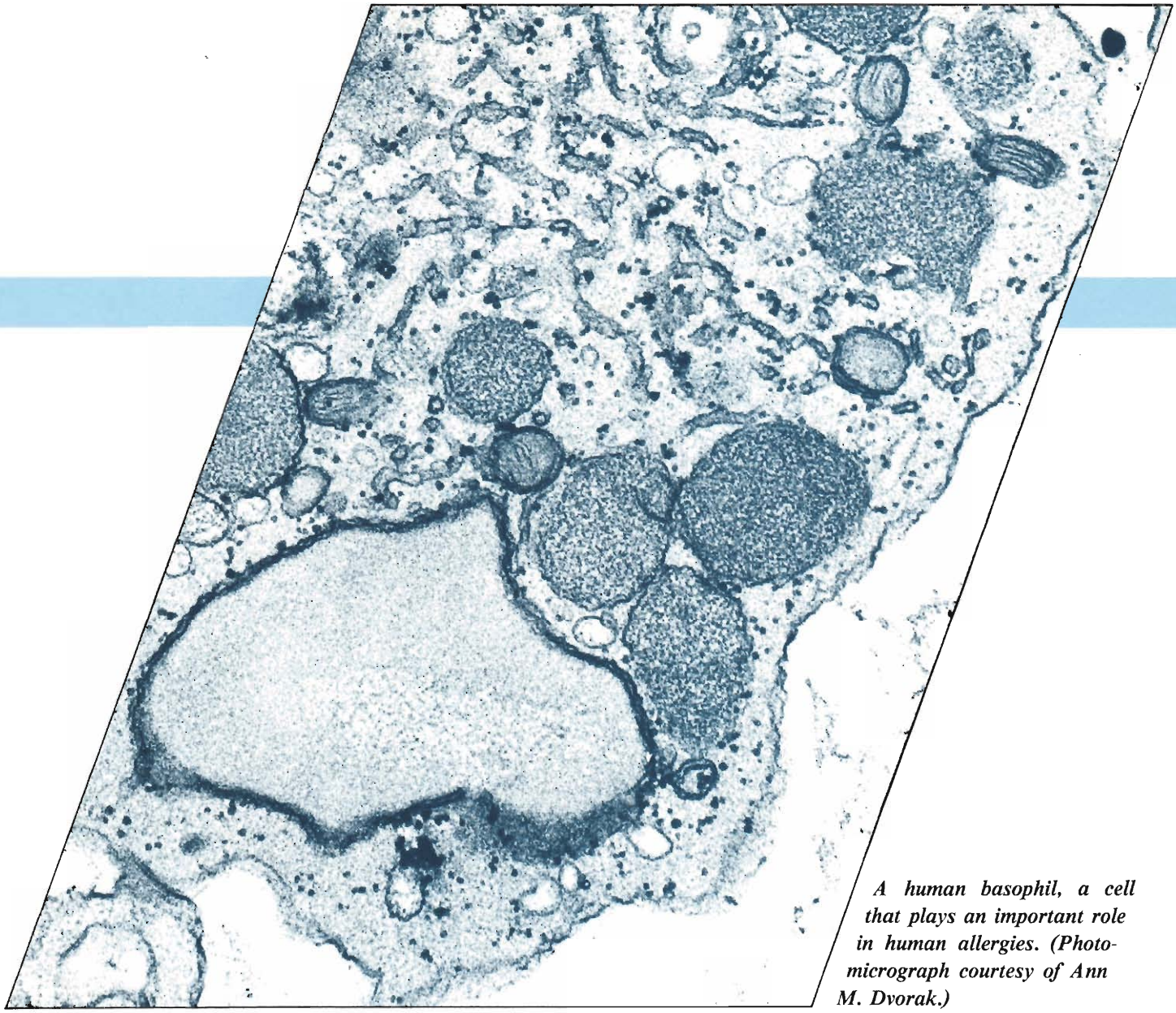
When spring arrives in New Mexico, invisible, diffuse clouds of juniper pollen drift through the air leaving behind a wake of red eyes, runny noses, and sneezing, suffering people. How is it that such a tiny, seemingly harmless substance can wreak such havoc? Why are some people victims, others not? How might this and other allergies be brought under control? Regrettably, allergies are the result of complex biochemical reactions that start in the body's immune system, and answers to our questions require detailed knowledge of these reactions.

The recent explosive growth in immunology has provided some of the knowledge by uncovering many of the steps in allergic reactions. Lawrence Lichtenstein and Anne Kagey-Sobotka at the Johns Hopkins University School of Medicine and we at Los Alamos have been conducting a joint study of a step important in most allergies—the release of histamine, one of the chemicals that cause the allergy sufferer's discomfort. By coupling experimental measurements of histamine release (Johns Hopkins) with theoretical modeling of the pertinent reactions (Los Alamos), it was possible to map this step in detail. The study also provided important information about desensitization, that is, the turning off of the allergic response. It is hoped that this study will bring us closer to control of this common, but often debilitating and sometimes deadly, phenomenon.

The Immune Response

The human immune system is an amazing defense system capable of distinguishing what is foreign from what is self. This system responds to and destroys invading substances that can cause infection or disease, but in some people it also responds to many seemingly harmless foreign substances. Whether it be a virus, a bacterium, or only that troublesome pollen, any foreign material that triggers an immune response is called an antigen.

The immune system reacts to antigens with one or both of two responses, a cellular response and an antibody response. The cellular response involves production of special white blood cells (T lymphocytes) that are capable of binding to and destroying the antigen. A classic example of this response is the rejection of organ transplants. Material from an organ grafted to a genetically different recipient of the same species acts as an antigen and elicits a cellular immune response that destroys the graft.



A human basophil, a cell that plays an important role in human allergies. (Photomicrograph courtesy of Ann M. Dvorak.)

In the second response the immune system reacts to an antigen by producing and releasing into the blood special proteins called antibodies. A popular immunology experiment exemplifies this response. If a mouse is injected with red blood cells from a sheep, the mouse's immune system produces antibodies that can bind to the sheep's red blood cells. The bound antibodies act as tags that clearly advertise the foreign nature of the antigen. Once tagged, the antigen becomes subject to attack and destruction by other molecules and cells of the body.

An important characteristic of the reaction between an antigen and an antibody or T lymphocyte is its specificity: an antibody or a T lymphocyte binds to the antigen that triggered its formation but with rare exceptions does not bind to other antigens. For example, the antibody that binds to influenza A virus does not bind to influenza B virus. As shown in Fig. 1, this specificity occurs because the shape and charge distribution of a particular molecular structure on the antigen—the binding site—are matched by a complementary shape and charge distribution of a binding site on the antibody or T lymphocyte. The two structures mesh somewhat like a lock and key.

Both immune responses possess the property of memory. If a rejected graft is followed by another from the same donor, it will be rejected more rapidly than the first. If a mouse is injected a second time with red blood cells from a sheep, it will produce greater amounts of antibodies more rapidly, and these antibodies will bind more strongly to the sheep's red blood cells.

With these marvelous properties the immune system seems designed to protect us from disease and infection. However, as we have said, not all immune responses are directed against harmful substances and not all immune responses improve our well being, as those of us who have allergies well know.

Allergic Reactions, IgE, and Histamine

There are two types of hypersensitive, or allergic, reactions, each involving one of the two immune responses. Delayed hypersensitive reactions (so called because they evolve slowly, peaking in about 1 to 4 days) involve the cellular, or T lymphocyte, response. Examples are allergies to poison ivy and industrial chemicals. Immediate hypersensitive reactions, on the other hand, involve the antibody response. In the remainder of the article, we will discuss only immediate hypersensitive reactions, and there certainly are enough of them. Hay fever, hives, and asthma, as well as allergies

to grasses, dog and cat danders, certain foods, bee venom, and penicillin, are all examples of antibody-mediated hypersensitivity.

Human antibodies are grouped into five classes (immunoglobulin A, D, E, G, and M) according to their biological functions, and the immune system can produce all these classes in response to normally harmless foreign substances. But no allergic symptoms will result unless immunoglobulin E (IgE) antibodies are produced. An antigen that causes the immune system to produce IgE antibodies is called an allergen.

IgE antibodies were discovered in 1967 by Kimishige Ishizaka and Teruko Ishizaka when they were studying the blood serum of hay-fever patients at the Children's Asthma Research Institute and Hospital in Denver. This class of antibodies was the last to be discovered, in part because it is normally produced in very small amounts. The concentration of all antibodies in the serum of a nonallergic, healthy person is about 15 milligrams per milliliter, but the concentration of IgE antibodies is only about 0.0001 milligrams per milliliter. Allergic individuals, however, tend to have higher concentrations of IgE antibodies. (People suffering from certain parasitic infections also have elevated levels of IgE antibodies, but why this is so is still an open question.)

All antibodies, regardless of class, are made up of similar Y-shaped units. Each unit contains two identical heavy polypeptide chains whose molecular weight depends on the antibody class and varies from 55,000 to 75,000 daltons. These heavy chains are joined by one or more disulfide bonds along some portion of their lengths to form the base of the Y, called the Fc region. The remaining lengths of the heavy chains are free and form flexible arms of the Y. Each arm, or Fab region, also contains a light polypeptide chain joined to the heavy chain by a disulfide bond. The light chains in the two arms are identical and have a molecular weight of about 23,000 daltons. As shown in Fig. 2, IgE antibodies are monomers; that is, they consist of only one of these Y-shaped units. (IgD and IgG are also monomers, IgM is a pentamer, and IgA exists as a monomer, dimer, or trimer.)

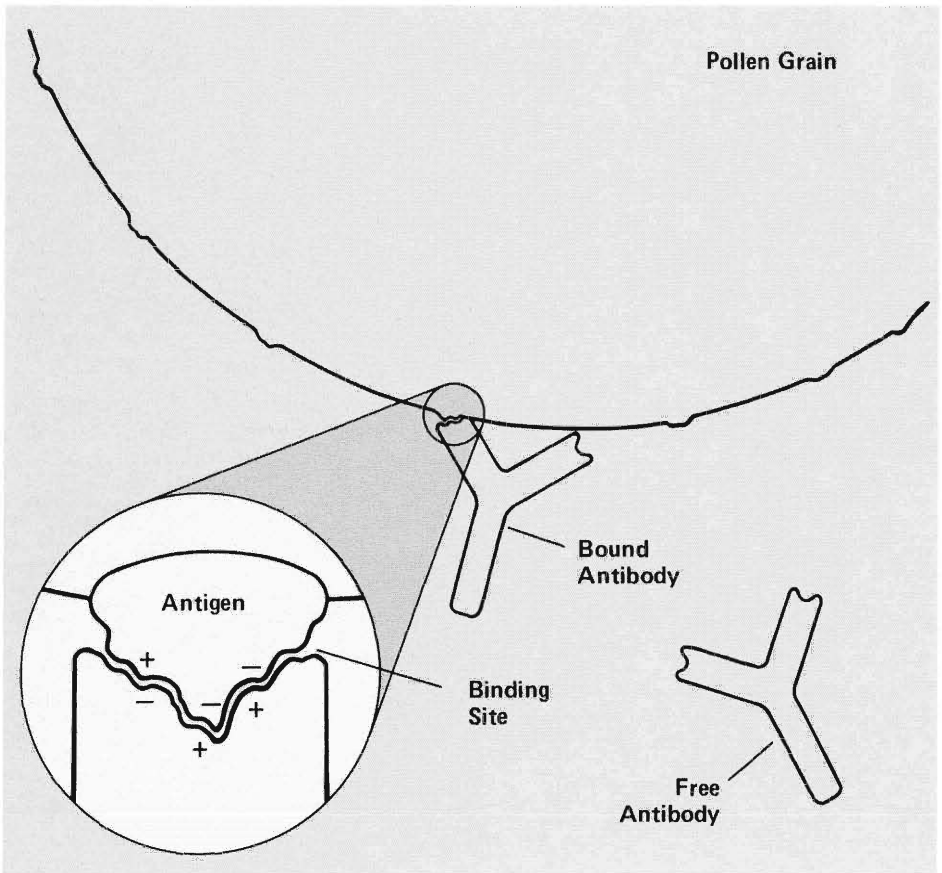


Fig. 1. Antibody-antigen binding. The "lock-and-key" model depicted here explains the high specificity of antibody-antigen binding. The bond between the two is not covalent, but is due to a combination of ionic and induced-dipole dispersion forces. As a result, a strong bond occurs when both the ionic charges and the shapes are complementary, the latter permitting the two surfaces to approach closely and thus maximize the dispersion force. The Y-shaped antibody molecule has two identical binding sites, either or both of which can participate in the binding.

It is the Fc region of an antibody that interacts with cells and molecules of the body and thus determines the antibody's biological functions. A class of antibodies includes all those antibodies with identical Fc regions. The name Fc arises from the fact that these identical fragments can be crystallized from a sample of, say, IgE antibodies. (Crystallization would not occur if these fragments were heterogeneous.)

Identical antigen binding sites are located near the ends of the Fab regions. (The name Fab stands for antigen binding fragment.) Here the antibody's specificity is determined: the binding sites have the correct three-dimensional structure to attach to the sites on the antigen that triggered the antibody's

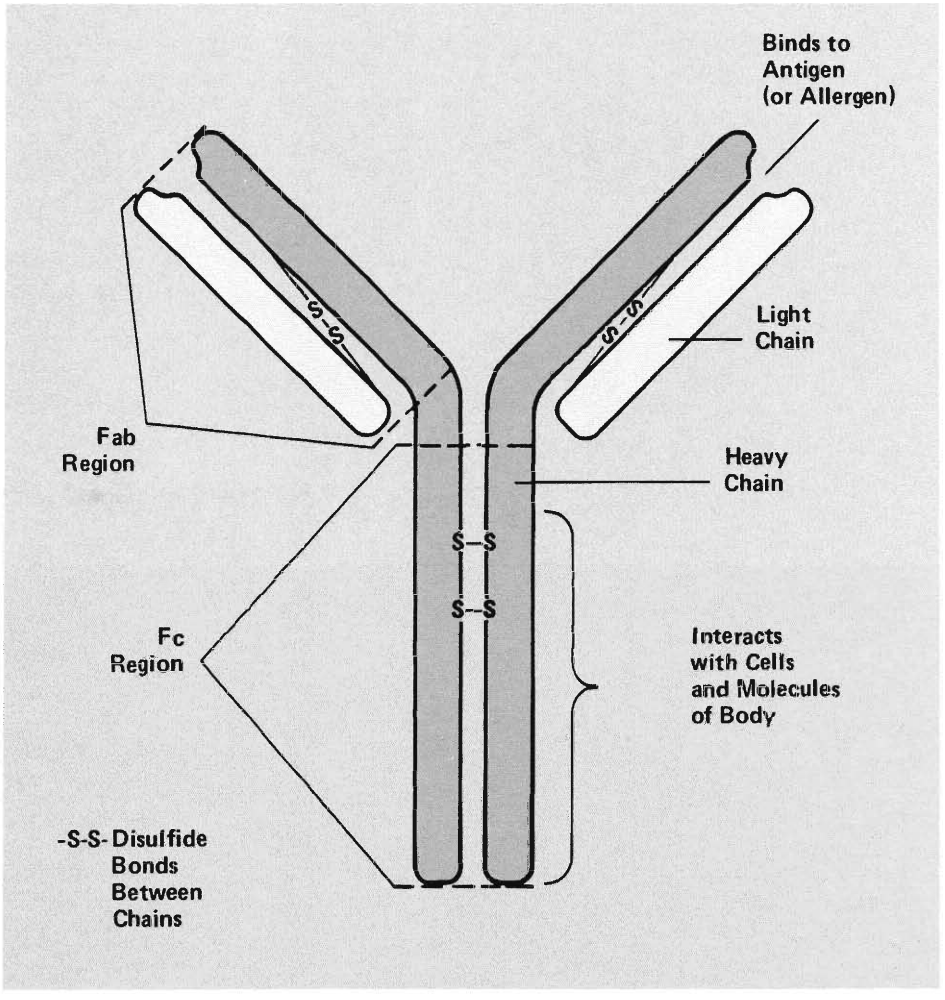


Fig. 2. Important features of an IgE antibody molecule. The molecule consists of two identical heavy polypeptide chains and two identical light polypeptide chains linked by disulfide bonds into an overall Y shape. The configuration depicted here is only symbolic of complicated three-dimensional chains that intertwine and fold back on themselves. Each arm is flexible and the angle between the arms of the Y can vary, possibly from 0 to 180 degrees. The Fc region is identical in all IgE antibodies. This region interacts with the cells and molecules of the body and thus determines the biological functions of the IgE class of antibodies. IgE antibodies specific to different antigens (or allergens) differ in their Fab regions. At the end of each arm are binding sites that match the binding site on the antigen (or allergen) that caused formation of the antibody. The binding sites are here depicted by concave half circles.

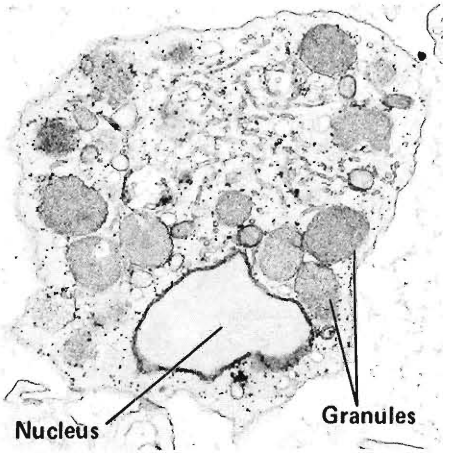


Fig. 3. Photomicrograph of a human basophil (7,000×). During an allergic reaction the granules of these cells release histamine and other potent chemicals into the blood stream. (This photo, kindly supplied by Ann M. Dvorak of the Beth Israel Hospital in Boston, is copyrighted by the United States-Canadian Division of the International Academy of Pathology and is reproduced with their permission.)

production. Thus, antibodies of the same class but specific to different antigens are identical in their Fc regions and differ in their Fab regions.

IgE antibodies interact with two types of cells—basophils and mast cells—that play a central role in allergic reactions. Mast cells are found in tissue and are most numerous in the linings of the respiratory and gastrointestinal tracts and under the skin. Basophils (Fig. 3) are found in blood serum and make up approximately 1 per cent of the white blood cells. The surfaces of both these types of cells are dotted with binding sites for the Fc portion of IgE antibodies. These binding sites are called Fc_ϵ receptors (the epsilon subscript indicates specificity for IgE antibodies). The number of receptors per cell varies widely. For example, on a human basophil there may be anywhere from 5000 to 500,000. Basophils of allergic people tend to have considerably more Fc_ϵ receptors than do those of nonallergic people.

Stored in granules within basophil and mast cells are biologically potent molecules, including histamine and the recently discovered leukotrienes. When these molecules are released from the cells they have several important effects. They cause contraction of smooth muscles such as those surrounding blood vessels and air passages in the lungs. It has been shown that histamine contracts the smooth muscles of the larger air passages in the lungs, and the leukotrienes contract the smaller peripheral airways. These molecules also affect the permeability of blood vessel walls and other membranes and cause glandular hypersecretion. At the site of a puncture wound, these effects result in blood flow changes, inflammation, fluid secretion, and the passage through various membranes of molecules and cells that attack and destroy harmful substances. In an allergic reaction a body-wide release of these same chemicals results in symptoms such as fluid secretion in the nose and throat and obstruction of air passages in the lungs.

There are several steps leading up to

histamine release; these are depicted schematically in Fig. 4. The first step is the initial response of a person's immune system to an allergen: the production of IgE antibodies specific to the allergen. Some of these IgE antibodies then bind through their Fc regions to Fc_ϵ receptors on mast cells or basophils; the two arms of the Y and the allergen-specific binding sites project outward from the cell surface. The cells are now "sensitized" to the allergen. Another exposure to the same allergen results in binding between the allergen molecules and the IgE antibodies that are bound to the surfaces of the cells. It is this interaction that

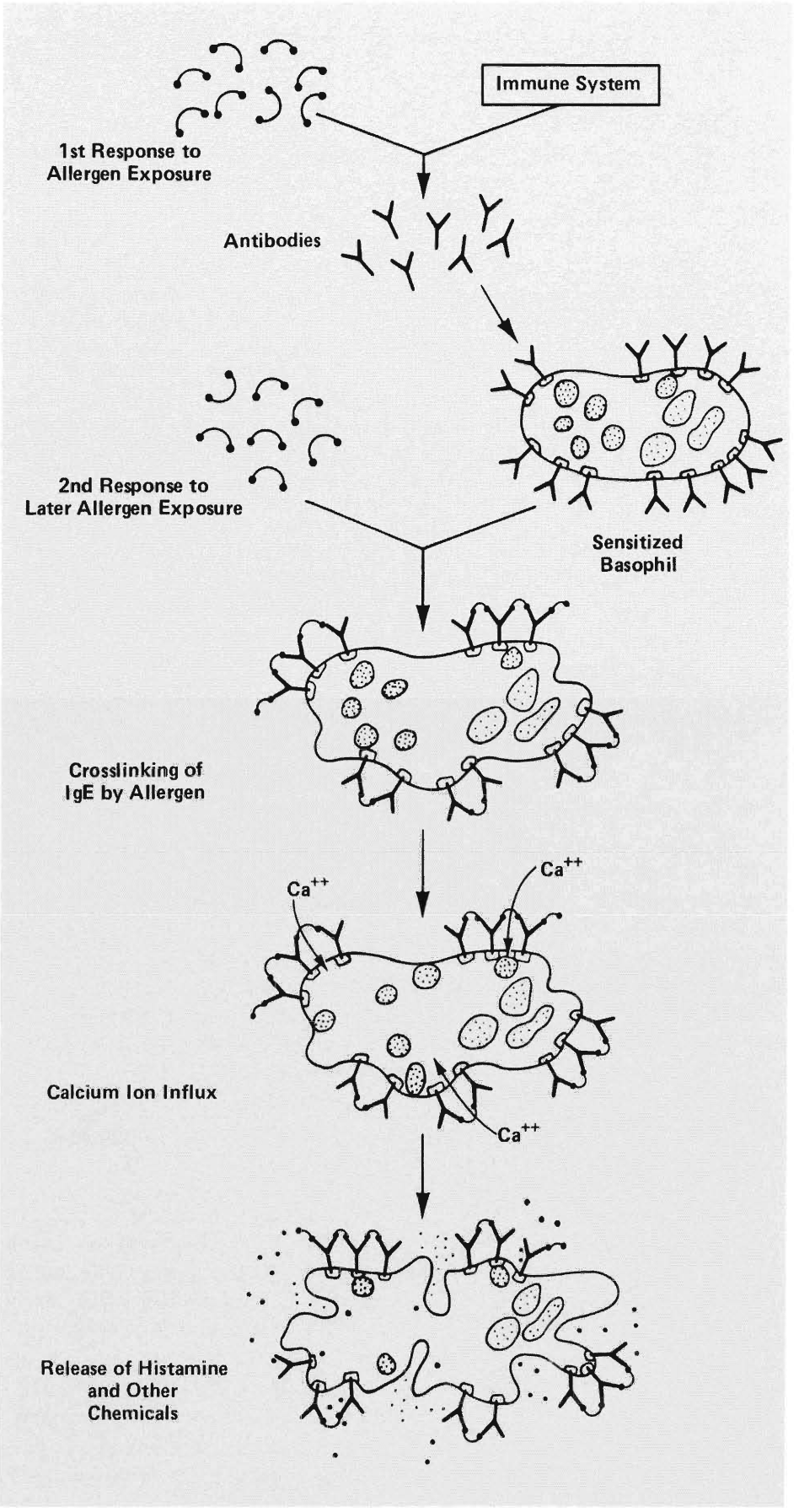


Fig. 4. The allergic reaction. The body's immune system may react to an allergen by producing IgE antibodies specific to that allergen. Many of these antibodies bind to the surfaces of basophils and mast cells through receptors that are specific to the Fc region of IgE antibodies; the arms of the antibodies containing the binding sites specific to the allergen point outward. A cell stippled with antibodies specific to an allergen is said to be sensitized to that allergen. Another exposure to the same allergen can lead to the formation of crosslinks, or bridges, between the IgE antibodies bound to the cell surfaces. These crosslinks form as sites on an allergen bind to complementary sites on adjacent antibodies. (Note that the allergen must have more than one binding site to form crosslinks. For simplicity the allergen shown here is divalent; that is, it has two binding sites.) Wherever a crosslink is formed a channel opens for transport of calcium ions (Ca^{++}) into the cell interior. This influx triggers a mechanism in which the granules release their contents of histamine and other chemicals into the blood serum.

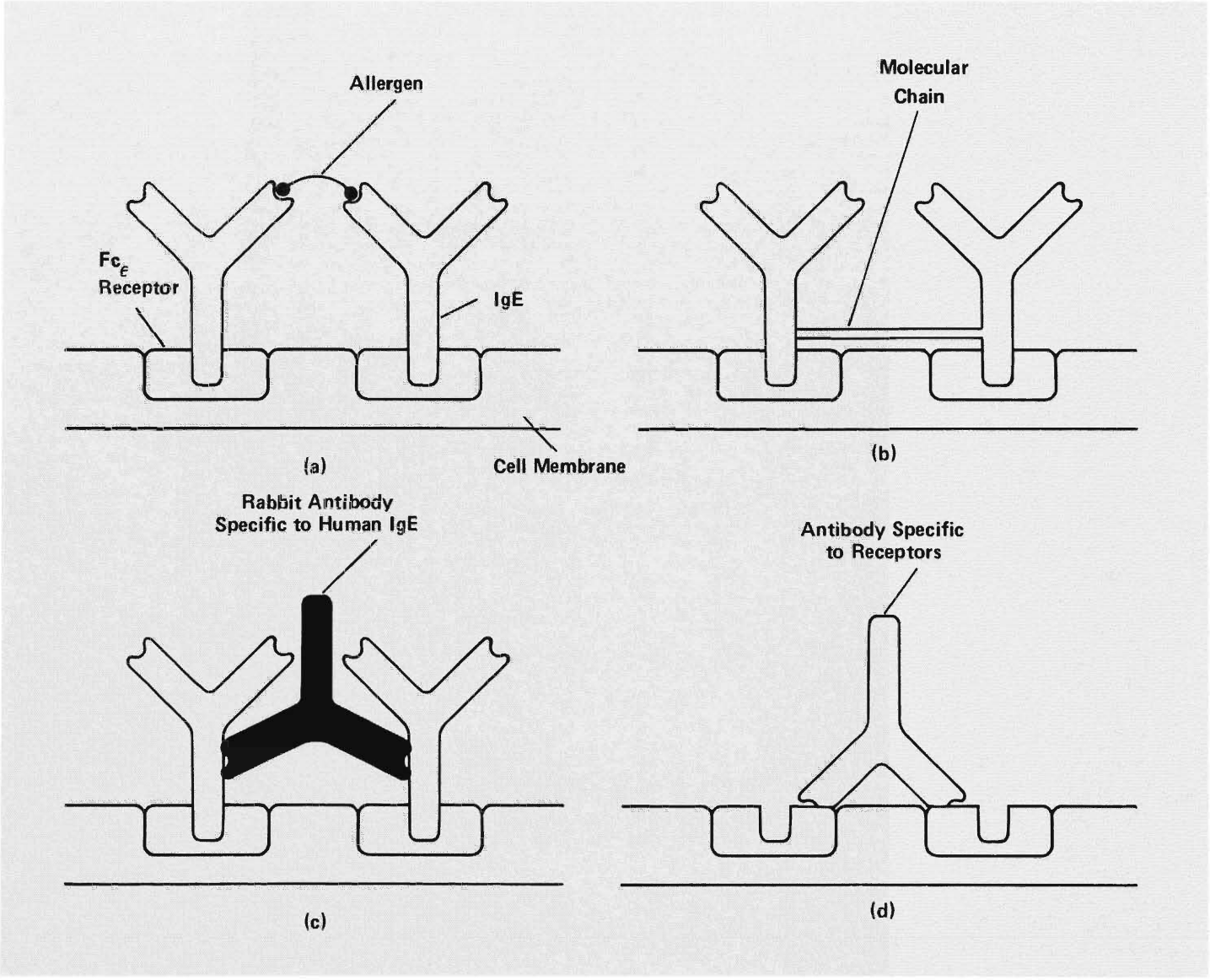


Fig. 5. A necessary condition for the release of histamine from a basophil is that two or more Fc_ϵ receptors on the cell surface be brought and held in proximity. This condition is achieved as any of the following processes takes place: (a) allergens form crosslinks between the arms of IgE antibodies bound to the receptors; (b) a molecular chain covalently binds two IgE antibodies bound to the receptors; (c) a rabbit antibody specific to human IgE antibodies binds to the Fc region of IgE antibodies bound to the receptors; and (d) a specially prepared antibody binds directly to the Fc_ϵ receptors.

can trigger the release of histamine from the granules of the cells. Here, then, is a critical step in an immediate hypersensitive allergic reaction.

In the remainder of the article we will discuss the histamine release mechanism in more detail, in particular what turns it on and what turns it off. We will concentrate on basophils, but almost everything we say about those cells holds true for mast cells as well.

The Basophil's "On" Signal

One condition necessary for histamine release from basophils is that the allergens must bind to the IgE antibodies in such a way that crosslinks, or bridges, are formed between the antibodies (Fig. 5a). Evidence that this crosslinking is necessary lies in the fact that allergens with only one binding site, which are physically incapable of crosslinking two IgE antibodies, do not trigger histamine release. Chains, rings, or other configurations of many crosslinked IgE antibodies may form on the basophil surfaces, but such large aggregates of crosslinked antibodies are not necessary for histamine release. Rather, the formation of crosslinked antibody pairs is sufficient. David Segal, Joel Taurog, and Henry Metzger of the National Institutes of Health demonstrated this sufficiency in 1978 by exposing basophils to molecules consisting of two covalently linked IgE antibodies (Fig. 5b). These permanently linked IgE antibody pairs caused histamine release in the absence of allergen.

Allergens crosslink IgE antibodies by binding to their Fab regions, but other molecules that crosslink IgE antibodies by binding along the Fc regions in the base of the Y can also cause histamine release. This phenomenon was originally demonstrated with molecules prepared by injecting rabbits with human IgE antibodies; the rabbits produced antibodies that bound specifically to the Fc regions of human IgE antibodies. (Other animal antibodies specific to human IgE antibodies are prepared in a similar

fashion.) When basophils sensitized with human IgE antibodies were exposed to these rabbit antibodies, histamine release occurred even though the antibodies were crosslinked through their Fc regions (Fig. 5c).

It is now clear that the requirement for histamine release of crosslinked IgE antibodies is really a requirement that Fc_ϵ receptors, which are mobile on basophil and mast cell surfaces, be brought and held in proximity. This was verified in 1978 by the Ishizakas and their collaborators at Johns Hopkins University, as well as a group at the National Institutes of Health, when they made an antibody that bound to Fc_ϵ receptors on mast cells and basophils of rats. This antibody triggered histamine release from the cells simply by bringing empty Fc_ϵ receptors close together (Fig. 5d).

Histamine release from basophils also demands another condition. If no calcium ions are present in the medium surrounding the basophils, no histamine will be released. The proximity of Fc_ϵ receptors brought about by crosslinked IgE antibodies somehow allows calcium ions to cross the cell membrane, and this influx of calcium ions is an essential signal for histamine release. If calcium ions can be introduced into the cell in some other way, crosslinked IgE antibodies are not needed. For example, calcium ionophores, substances that cause calcium ion channels to form in cell membranes, will induce histamine release in the absence of crosslinking if calcium ions are available. Injecting calcium ions directly into basophils also induces histamine release. In a test tube the calcium ion concentration can be manipulated, but in serum, the natural milieu of the basophil, calcium ions are always present at a concentration (2 to 5 millimolar) that is sufficient to insure histamine release.

In summary, an immediate hypersensitive reaction is turned on by the flow of calcium ions into sensitized basophils made possible by allergen-linked IgE antibodies on the basophil surfaces. What turns the reaction off?

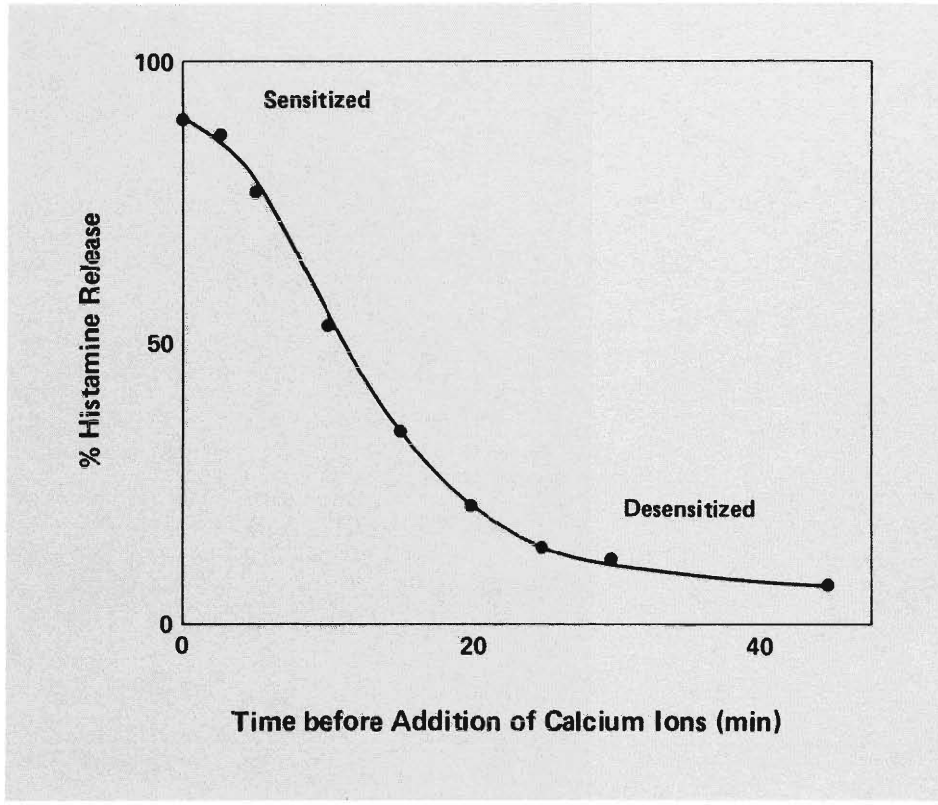


Fig. 6. Basophil desensitization. Basophils sensitized to the major ragweed allergen release a large percentage of their histamine when exposed to the allergen in the presence of calcium ions. (The basophils studied in the experiment depicted here released almost 90 per cent of their histamine under these circumstances.) However, if the basophils are exposed to the allergen in the absence of calcium ions, the cells gradually lose their ability to release histamine. The cells are said to be desensitized when very little histamine is released in the presence of calcium ions.

A Nonspecific "Off" Signal

As early as 1964 Lawrence Lichtenstein and Abraham Osler at Johns Hopkins University showed that the same agent that caused histamine release from basophils of hay-fever patients, ragweed pollen, could also desensitize these cells, that is, block their release of histamine. In 1971 Lichtenstein investigated desensitization further with experiments on white blood cells, including basophils, of hay-fever patients. The cells were exposed for various lengths of time to the major ragweed allergen (a 38,000-dalton protein isolated by T. P. King and his collaborators at Rockefeller University) in a medium containing no calcium ions. Calcium ions were then added, and the amount of histamine released by the cells during the following 30 minutes was measured. These experiments showed that the longer the cells were exposed to the ragweed allergen in the absence of calcium, the less histamine they were capable of releasing in the presence of calcium (Fig. 6). Lichtenstein obtained similar results using as the crosslinking agent

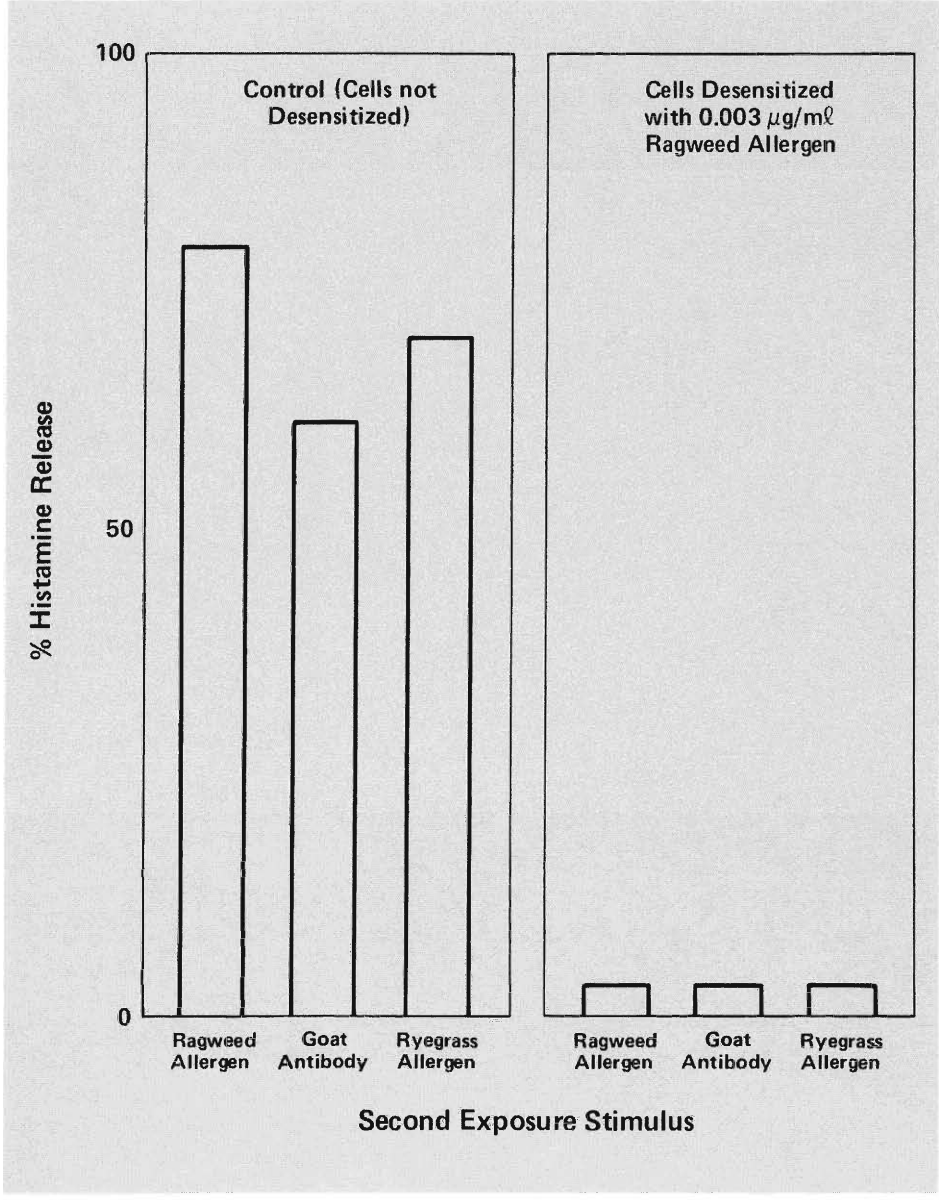


Fig. 7. Nonspecific desensitization. The basophils studied in the experiment illustrated here were from a donor who was allergic to both the major ragweed allergen and a ryegrass allergen. On the left are the percentages of histamine released by these doubly sensitized basophils when exposed to either of the two allergens or to a goat antibody specific to human IgE antibodies. On the right are the percentages of histamine released by these cells when exposed to the ragweed or ryegrass allergens or the goat antibody after *being desensitized to the ragweed allergen. The negligible release shows the cells to be nonspecifically desensitized to any allergen or crosslinking agent.*

a rabbit antibody specific to human IgE antibodies. It appears that the crosslinking of IgE antibodies on the basophil surface initiates two signals: a signal for histamine release in the presence of calcium ions and a signal for blockage of histamine release in the absence of calcium ions.

Lichtenstein performed similar experiments on white blood cells from a donor who was allergic to both the ragweed allergen and a ryegrass allergen. Therefore, some of the Fc_ϵ receptors on the donor's basophils were occupied by IgE antibodies specific to the ragweed allergen and some by IgE antibodies specific to the ryegrass allergen. As before, after sufficient exposure to the ragweed allergen in the absence of calcium ions, addition of calcium ions caused no significant histamine release. The cells, now desensitized to the ragweed allergen, were exposed in the presence of calcium ions either to the ryegrass allergen or to a goat antibody specific to human IgE antibodies. Neither substance caused significant release (Fig. 7). Desensitization to one allergen turned the basophils off to other allergens and crosslinking agents as well. This "nonspecific" desensitization occurred despite the fact that the desensitizing allergen interacted with only a fraction of the IgE antibodies on the cell surfaces, namely, those specific to the ragweed allergen.

Although nonspecifically desensitized basophils cannot be made to release histamine by crosslinked IgE antibodies, they do release histamine when exposed to calcium ionophores or injected with calcium ions. Nonspecific desensitization must therefore involve the shutdown of a calcium ion transport mechanism. Apparently, crosslinking of IgE antibodies in the presence of calcium ions at first activates the transport mechanism, but with time that activation somehow degrades.

Blocking histamine release by withholding calcium ions represents an artificial situation. Does nonspecific desensitization occur in the natural milieu of the basophil, that is, in the

presence of calcium ions? Evidence for a positive answer comes from numerous *in vitro* studies of histamine release from basophils as a function of allergen concentration. Data from such studies are generally displayed as plots, known as dose response curves, of the percentage of histamine released versus the logarithm of the allergen concentration. Figure 8 shows dose response curves for basophils from three allergic donors. The initial rise of the dose response curves reflects the increase in the number of IgE antibodies that are crosslinked by the allergen.The more interesting feature of the curves is their eventual fall, a phenomenon known as allergen excess inhibition.

Karl Becker, Henry Metzger, and Philip Grimley of the National Institutes of Health, in collaboration with the Ishizakas, showed that for basophils from allergic donors, which generally have large numbers of specific IgE antibodies on their surfaces, excess inhibition is accompanied by large numbers of crosslinks. They did so by studying the distribution of IgE antibodies crosslinked by a fluorescent form of a sheep antibody specific to human IgE antibodies (Fig. 9). At low concentrations of the sheep antibody the distribution of fluorescence was diffuse, but, at the concentrations at which excess inhibition occurs, the distribution of fluorescence became patchy. These observations suggest that at high concentrations of the crosslinking agent large aggregates of crosslinked IgE antibodies had formed.

Excess inhibition can be understood in terms of the effect of crosslinking on the calcium ion transport mechanism. Low numbers of crosslinks (low allergen concentrations) activate the transport mechanism and histamine release occurs. Degradation of the transport mechanism takes place slowly, and histamine release can be blocked only by withholding calcium ions until this gradual shutdown has been completed. As the number of crosslinks increases (higher allergen concentrations), the transport mechanism degrades more rapidly. Eventually, degradation dominates and histamine release is

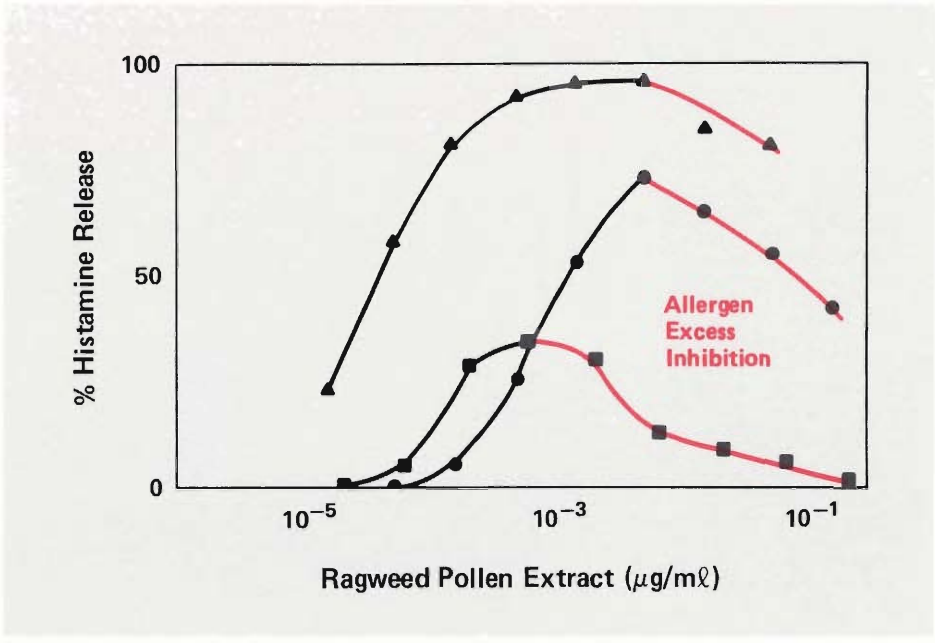


Fig. 8. Typical dose response curves. These curves show the percentages of histamine released by basophils from three ragweed-pollen-allergic donors when exposed to various concentrations of ragweed pollen extract. In all cases the percentage of histamine released first rises with increasing concentration but then peaks and declines. This decline at high allergen concentrations is called allergen excess inhibition. The ragweed pollen extract used in the experiment contained five ragweed allergens. From Lawrence M. Lichtenstein and Abraham G. Osler, Journal of Experimental Medicine 120, *507 (1964).*

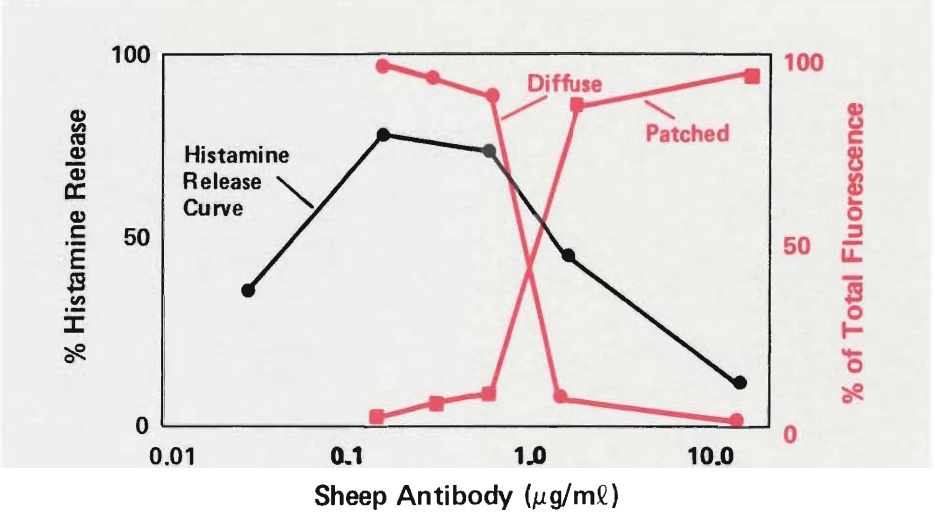


Fig. 9. At low concentrations of a sheep antibody specific to human IgE antibodies, the dose response curve (black) is rising and, as shown by fluorescence microscopy (red), the distribution of IgE antibodies on the basophil surface is mainly diffuse. At concentrations above that for maximum release of histamine, the distribution of IgE antibodies becomes mainly patched. The patches of fluorescence indicate the formation of large aggregates of crosslinked antibodies. Based on data from Karl E. Becker, T. Ishizaka, H. Metzger, K. Ishizaka, and Philip M. Grimley, Journal of Experimental Medicine 138, *394 (1973).*

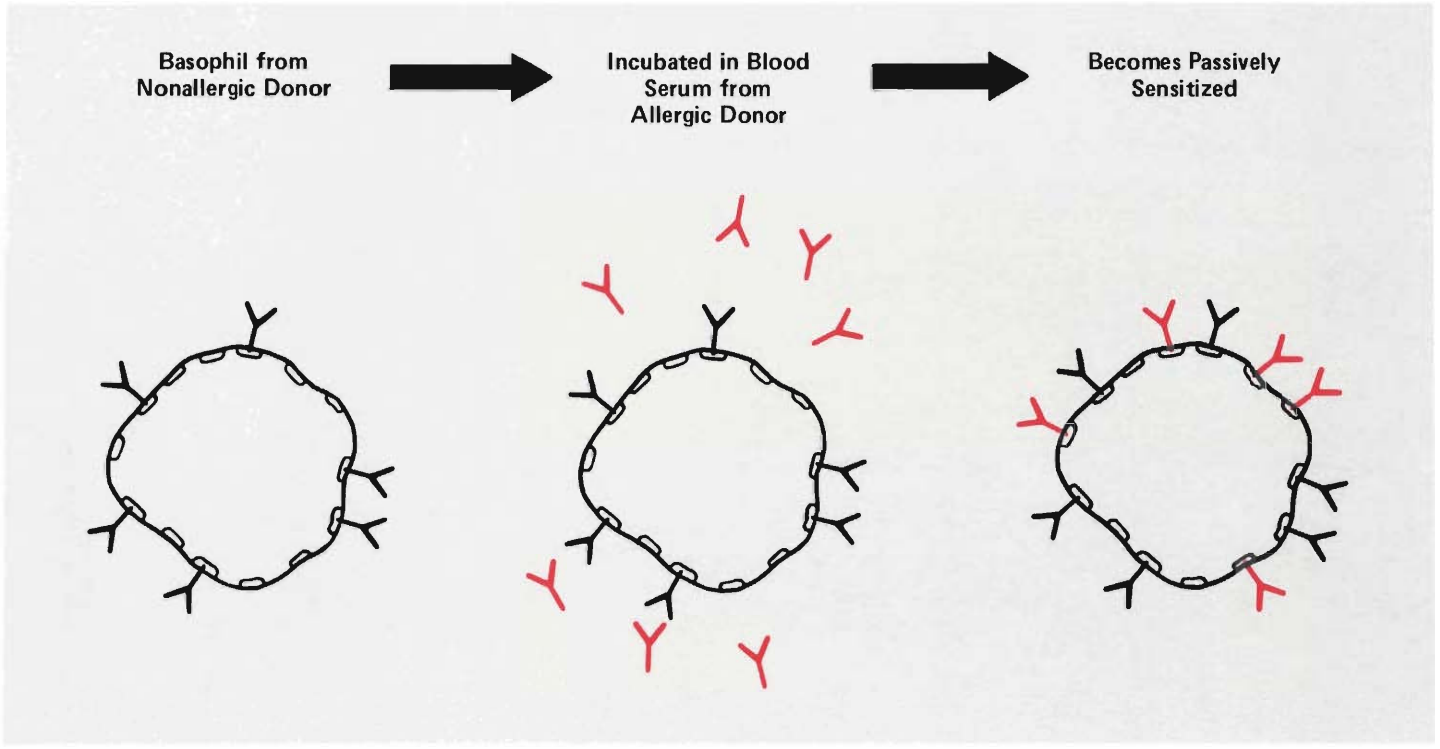


Fig. 10. Passive sensitization. Basophils with empty Fc_ϵ receptors can be sensitized to an allergen by incubating the cells in blood serum from a donor who is allergic to the allergen. This passive sensitization occurs because IgE antibodies specific to the allergen (red) fill some of the empty Fc_ϵ receptors.

blocked even in the presence of calcium ions.

Only on basophils with large numbers of the allergen-specific IgE antibodies bound to their surfaces can there be sufficient crosslinks for nonspecific desensitization to occur in the presence of calcium ions. Therefore, generally only basophils from allergic donors exhibit allergen excess inhibition because a substantial fraction of their large number of Fc_ϵ receptors is filled with the allergen-specific IgE antibodies.

What happens, however, when there are only small numbers of the allergen-specific IgE antibodies on the basophil surfaces? The answer to this question was provided by our joint research with Lichtenstein and Kagey-Sobotka. What happens proved to be not only something different but also quite interesting. In particular, we observed another mechanism for turning off the allergic reaction.

The Experiments

The goal of our initial collaboration with Lichtenstein and Kagey-Sobotka in 1977 was to test our theoretical predictions, in particular, predictions we had made about the dose response curves as the number of allergen-specific IgE antibodies on the basophil surfaces is increased. To eliminate extraneous variables that would blur a correlation between theory and experiment, we needed to use the same allergen, the same allergen-specific IgE antibody, and basophils from the same donor. Moreover, we needed basophils with free Fc_ϵ receptors that could be filled with different amounts of the IgE antibody of our choice. This requirement forced us to use basophils from nonallergic donors, because basophils from allergic donors tend to have most of their Fc_ϵ receptors filled. However, the total number of receptors is much smaller on basophils from nonallergic donors. As a result, our study was of basophils with relatively small numbers of the allergen-specific IgE antibodies on their surfaces, usually less than 10,000 per cell. Serendipitously, this circumstance led to our discovery of a second mechanism for desensitization.

Other experimental restraints resulted from the fact that no one had, or has yet, learned how to keep human basophils alive outside the body for longer than a day or so. The research thus depended on the availability of donors, both of basophils and blood serum. Another difficulty was the day-to-day variation in the basophil donor's exposure and immune system response that changed the number of empty Fc_ϵ receptors on the basophil surfaces.

To obtain basophils with different and well-characterized numbers of the allergen-specific IgE antibody on their surfaces, we incubated the cells with free Fc_ϵ receptors in serum from a donor who was extremely allergic to penicillin. (Ninety per cent of the penicillin-allergic donor's IgE antibodies were specific to the benzylpenicilloyl, or BPO, group.) The length of the incubation or the dilution of the penicillin-allergic donor's serum determined how many BPO-specific IgE antibodies filled Fc_ϵ receptors on the basophil surfaces.

This technique for sensitizing cells is called passive sensitization (Fig. 10) and has been known since 1921 when Carl Prausnitz injected into the skin of a nonallergic subject a small quantity of serum from Heinz Küstner, who was extremely allergic to fish. Twenty-four hours later fish extract was injected into the same area of the nonallergic subject's skin. Immediately a wheal appeared. We now know that Küstner's serum contained IgE antibodies specific to an allergen found in fish. When transferred to the skin of the nonallergic subject, these antibodies passively sensitized his mast cells. Then, when exposed to the allergen in the fish extract, the sensitized mast cells released

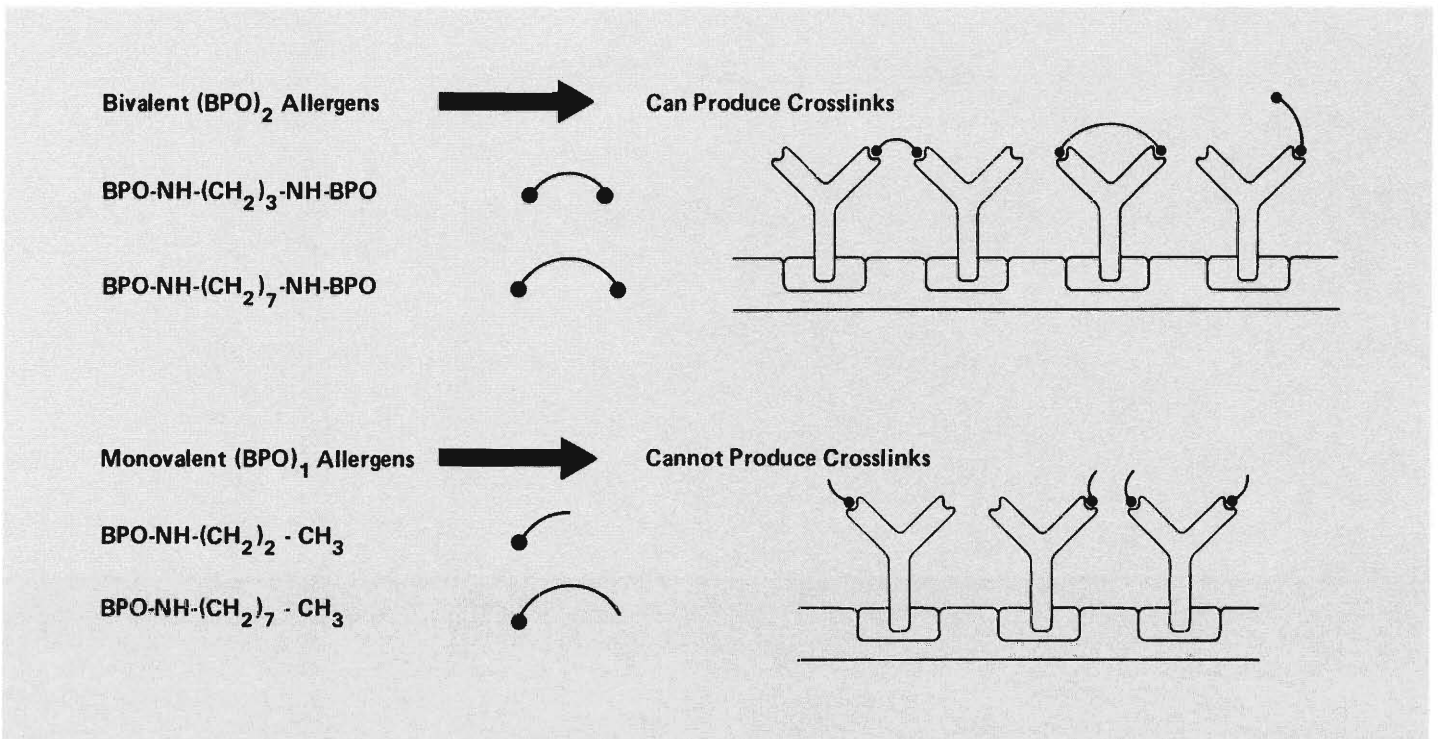


Fig. 11. Synthetic benzylpenicilloyl (BPO) allergens. The bivalent molecules, with a BPO group (circle) at each end of the chain, are able to form bridges between binding sites on the arms of the IgE antibodies. The monovalent molecules, with only one BPO group, cannot form such crosslinks.

histamine and other chemicals that produced the wheal. The PK test for allergy is named for these early investigators.

Our choice of using IgE antibodies from a penicillin-allergic person was based on the availability of simple, well-defined synthetic penicillin allergens. Bernard Levine at the New York University School of Medicine had first prepared a series of such allergens in 1967. Some of these are illustrated in Fig. 11. Each is a linear chain of different length with either a BPO group at each end (bivalent) or a BPO group at one end only (monovalent).

We resurrected the bivalent and monovalent synthetic penicillin allergens because they were ideal for our purpose: to start to build a mathematical model of histamine release from basophils. In $(BPO)_2$ we had the simplest possible crosslinking agent, a symmetric linear molecule with two identical binding sites. In $(BPO)_1$ we had a tool for testing our ideas about what happens when the number of crosslinks is reduced.

The final step of the experiments entailed determining the dose response curves for the passively sensitized basophils, that is, measuring the percentages of histamine released by the basophils when exposed to various concentrations of $(BPO)_2$ and $(BPO)_1$ allergens. We hoped that by comparing the dose response curves with our theoretical calculations about crosslinks we could learn something new about the role of crosslinks in histamine release.

We shall first present the results of our calculations and then a comparison of these results with those of the experiments.

Results of Crosslinking Calculations

We learned experimentally that the binding between the IgE antibodies used to passively sensitize the basophils and the $(BPO)_2$ or $(BPO)_1$ allergens came to equilibrium within seconds after the basophils were exposed to the allergens and well before any measurable histamine release or desensitization had occurred. Thus, histamine release and desensitization were governed by the equilibrium concentration of crosslinks. For these experiments we could neglect the details of the binding during the first few seconds of exposure and instead treat the basophils as if the crosslinks formed instantaneously. Thus, we calculated, for a given number of BPO-specific IgE antibodies per basophil, the equilibrium concentration of

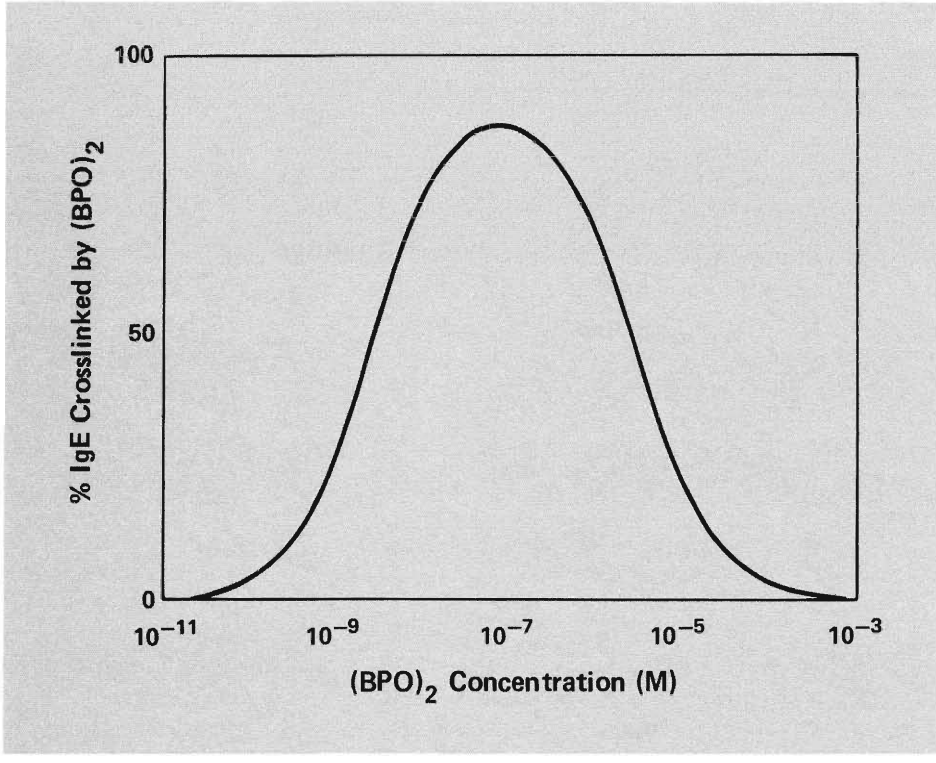


Fig. 12. According to our calculations, the fraction of IgE antibodies crosslinked by $(BPO)_2$ varies as shown with the concentration of $(BPO)_2$ to which the basophils are exposed. The curve has a single maximum and is symmetric about that maximum. The initial rise in the curve is expected; the decline is due to a saturation effect depicted in Fig. 13.

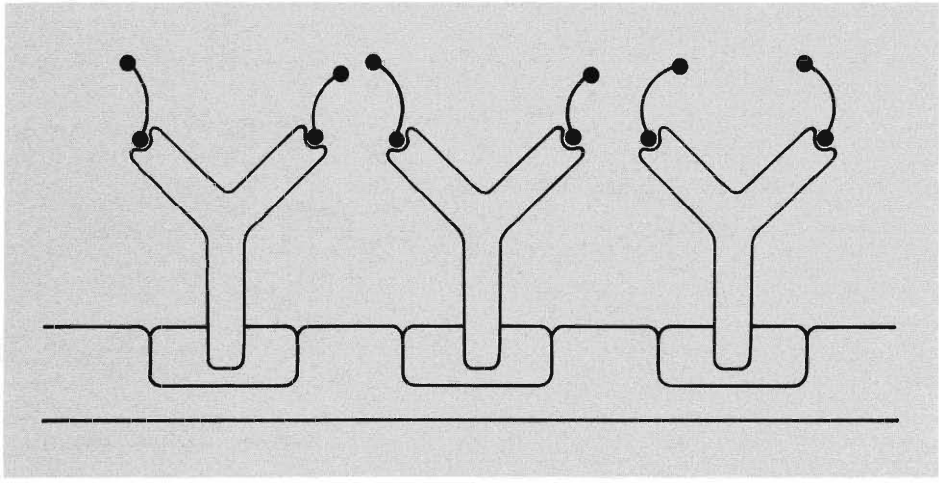


Fig. 13. At high $(BPO)_2$ concentrations most of the binding sites on the antibodies are filled with $(BPO)_2$ molecules. Because there are few empty sites available, these bivalent $(BPO)_2$ molecules are unable to form bridges between the antibodies.

crosslinks on the basophil surfaces for various concentrations of $(BPO)_2$ and $(BPO)_1$ allergens.

Our method of attacking this equilibrium problem is described in the sidebar "Crosslinking—A Theoretical Approach." The results of the calculations revealed interesting predictions about crosslinking. For example, Fig. 12 shows a plot, called a crosslinking curve, of the fraction of crosslinked antibodies versus the logarithm of the $(BPO)_2$ concentration. The curve rises from zero to a maximum and then decreases to zero. The curve falls because, as the $(BPO)_2$ concentration increases, more and more of the binding sites on the antibodies are occupied by $(BPO)_2$ allergens with one unbound BPO group (Fig. 13). With fewer potential sites available for that unbound BPO group the number of crosslinks decreases.

Another property of the crosslinking curves is symmetry about their maxima. That is, if the maximum number of crosslinks occurs at a $(BPO)_2$ concentration of 10^{-7} molar (as it does in Fig. 12), there will be just as many IgE antibodies crosslinked at 10^{-6} molar as at 10^{-8} molar.

The $(BPO)_2$ concentration at which the maximum of the crosslinking curve occurs is determined by only two parameters: K, the equilibrium constant for the binding between a BPO group and a binding site on an antibody, and the concentration of $(BPO)_1$ allergen. If $[(BPO)_2]_{max}$ is the $(BPO)_2$ concentration at which crosslinks are a maximum and $[(BPO)_1]$ is the $(BPO)_1$ concentration to which the basophils are exposed, then

$$[(BPO)_2]_{max} = \frac{1}{2}\left(\frac{1}{K} + [(BPO)_1]\right) \quad . \quad (1)$$

One of the most interesting predictions of our calculations is that $[(BPO)_2]_{max}$ does not depend on the total number of BPO-specific antibodies on the basophil surfaces. In other words, if the number of BPO-specific anti-

Sidebar

crosslinking — a theoretical approach

The problem at hand was to calculate the equilibrium concentration of crosslinks formed when BPO-specific IgE antibodies on basophil surfaces are exposed to the monovalent and bivalent synthetic penicillin allergens $(BPO)_1$ and $(BPO)_2$. We began by constructing a model consisting of all the binding reactions that can occur in this situation. Crucial to the calculations is knowledge of the equilibrium constants for these binding reactions. Although the model includes an infinite number of reactions, some reasonable assumptions reduce to a manageable number the equilibrium constants that must be known.

First, we let K be the equilibrium constant for the binding between a monovalent allergen and a "monovalent antibody." (A monovalent antibody does not, of course, exist but is useful as a theoretical construct because the equilibrium constant for its reaction with a monovalent allergen is indicative of the basic strength of the forces between the binding sites.) Consider now the simplest of the reactions depicted in the accompanying figure, those initial reactions in which a $(BPO)_1$ or $(BPO)_2$ allergen binds to one of the two sites on a free antibody (that is, an antibody bound through its Fc region to the cell surface but with each of the binding sites in its Fab regions free). The equilibrium constant for the reaction when $(BPO)_1$ is involved is $2K$ since the antibody offers two possible binding sites. Similarly, the equilibrium constant for the reaction involving $(BPO)_2$ is simply $4K$ because in this case the allergen also offers two possible binding sites. We assume that the equilibrium constants for these two reactions are unaffected if the free antibody is replaced by a chain of crosslinked antibodies with a free binding site on the antibody at each end of the chain.

Considering next the binding of a $(BPO)_1$ or $(BPO)_2$ allergen to the single available site on the products of the initial reactions, we assume that the equilibrium constants for these reactions are also related to K by appropriate statistical factors.

Next, we let K_x be the equilibrium constant for the basic crosslinking reaction, the binding of the complex containing one antibody and one $(BPO)_2$ allergen, each with a free binding site, to a free antibody. Again we assume that the equilibrium constant is unaffected if the free antibody is replaced by a chain of crosslinked antibodies with a free binding site on the antibody at each end.

Finally, a ring containing i antibodies is formed when a free BPO group of a $(BPO)_2$ allergen bound to one end of a chain of i crosslinked antibodies binds to the free site on the antibody at the other end of the chain. We assume that for $i \geq 2$ the equilibrium constant J_i for such a reaction is inversely proportional to i^2. Therefore, $J_i = 4J_2/i^2$ for $i \geq 2$. We let J_1 be the equilibrium constant for formation of the "ring" consisting of a single $(BPO)_2$ allergen spanning the sites on a single antibody.

Armed with the four equilibrium constants K, K_x, J_1, and J_2, we can calculate the equilibrium concentrations of all possible reaction products. (Reasonable estimates for the magnitudes of these constants can be obtained from various experimental data.) We will not present details of the calculations but rather the general concepts on which they are based.

The accompanying figure shows that seven complexes contain one antibody. The equilibrium concentrations of each of these complexes can be expressed as a function of K, J_1, and the $(BPO)_1$ and $(BPO)_2$ concentrations multiplied by the equilibrium concentration of free antibody. Therefore, W_1, the total equilibrium concentration of complexes containing one antibody, is obtained simply by adding together the equilibrium concentration of each of the complexes. We find that

$$W_1 = X f\left([(BPO)_1],[(BPO)_2],K,J_1\right) \quad ,$$

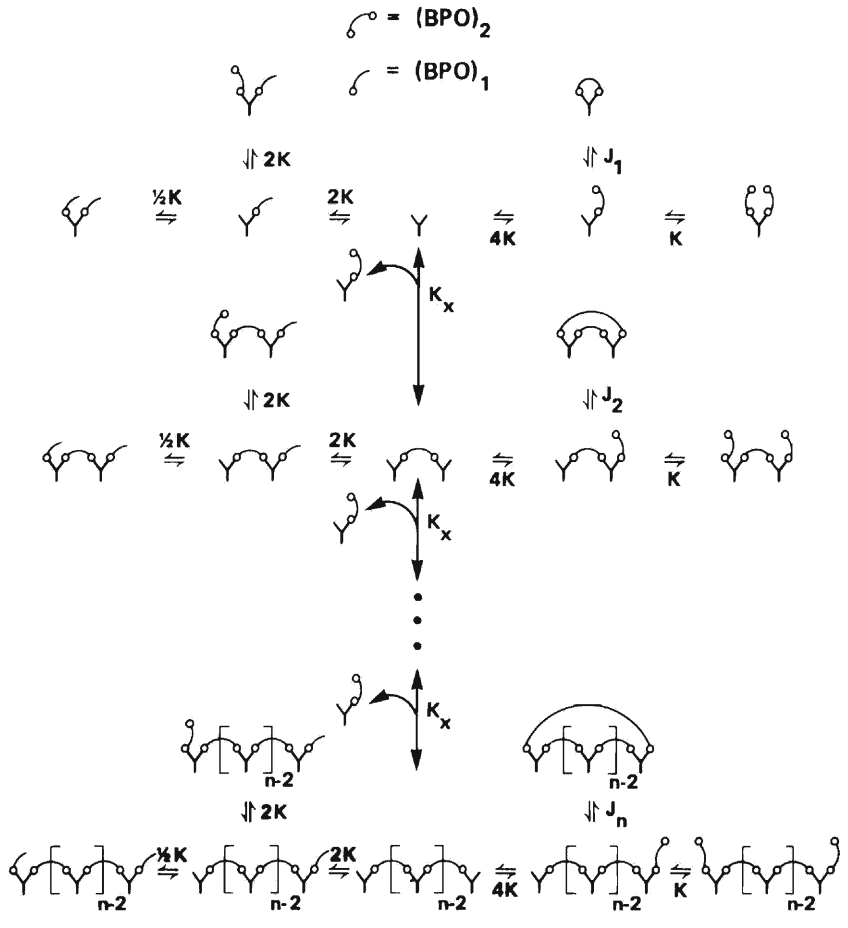
where

$$f = \left\{1 + K[(BPO)_1] + 2K[(BPO)_2]\right\}^2 + 4KJ_1[(BPO)_2] \quad .$$

In these equations $[(BPO)_1]$ and $[(BPO)_2]$ are the concentrations of $(BPO)_1$ and $(BPO)_2$, respectively, and X, the only unknown, is the equilibrium concentration of free antibody.

The accompanying figure also shows that seven complexes contain two antibodies. We can express the concentration of each of these complexes as a function of K, J_2, $[(BPO)_1]$, and $[(BPO)_2]$ multiplied by the concentration X_2 of the crosslinked chain containing two antibodies with a free binding

Sidebar



Shown here are all the binding reactions that can occur when BPO-specific IgE antibodies on basophils are exposed to the monovalent and bivalent synthetic penicillin allergens $(BPO)_1$ and $(BPO)_2$. The equilibrium constants for each reaction are also given.

site on the antibody at each end. But X_2 is in turn a function of K, $[(BPO)_2]$, and X, namely

$$X_2 = 4KK_x[(BPO)_2]X^2 \quad .$$

We can continue this process iteratively and develop a general expression for W_n, the equilibrium concentration of complexes containing n antibodies, as a function of $[(BPO)_1]$, $[(BPO)_2]$, K, K_x, J_n, and X:

$$W_n = X^n\{4KK_x[(BPO)_2]\}^{n-1}$$

$$\times f([(BPO)_1],[(BPO)_2],K,J_n) \quad .$$

The conservation law for total antibody concentration X_T leads to the equation

$$X_T = \sum_{n=1}^{\infty} nW_n \quad .$$

When we express all the W_n in this infinite series in terms of X, the infinite series can be summed, and we obtain an algebraic equation that can be solved for X. By substituting the solution for X into the expression for a particular W_n or for the equilibrium concentration of a particular complex, we can compute values for these expressions.

Because of the central role of crosslinks in the activation and desensitization of basophils, we are particularly interested in X_{poly}, the fraction of antibodies incorporated at equilibrium into complexes containing more than one antibody, that is, in the fraction of crosslinked antibodies. An expression for X_{poly} is easily derived since

$$X_{poly} \equiv \sum_{n=2}^{\infty} nW_n = (X_T - W_1)/X_T \quad .$$

Since this calculation was first presented, much theoretical work has been done on both the equilibrium and the kinetic theory of binding of multivalent antigens to antibodies on cell surfaces. As a result primarily of the work of Alan Perelson (Los Alamos), Charles DeLisi (National Institutes of Health), and Catherine Macken (Lincoln College, New Zealand), much progress has been made toward understanding the bonding of these more complicated antigens. ■

bodies is increased (by, for example, longer incubation), the $(BPO)_2$ concentration at which the maximum occurs does not change (Fig. 14a). On the other hand, increasing the concentration of $(BPO)_1$, which reduces the number of potential crosslinks, causes the maximum to shift to higher $(BPO)_2$ concentrations (Fig. 14b).

Histamine Release vs Crosslinking

How do the experimental dose response curves and the theoretical crosslinking curves compare? Figure 15 shows that the dose response curves exhibit the same properties as the crosslinking curves: the $(BPO)_2$ concentration at which the maximum occurs does not change with the number of BPO-specific antibodies on the basophil surfaces (Fig. 15a) but does change with the $(BPO)_1$ concentration (Fig. 15b). There is other agreement as well. For example, the properties of the basophil do not appear in Eq. 1; only the binding between the BPO group and the antibody is important. Therefore, we predict that basophils from different donors should exhibit maxima in their crosslinking curves at the same $(BPO)_2$ concentration. This prediction is in agreement with the observation that each of three dose response curves for basophils subjected to the same experimental treatment but from different nonallergic donors had its max-

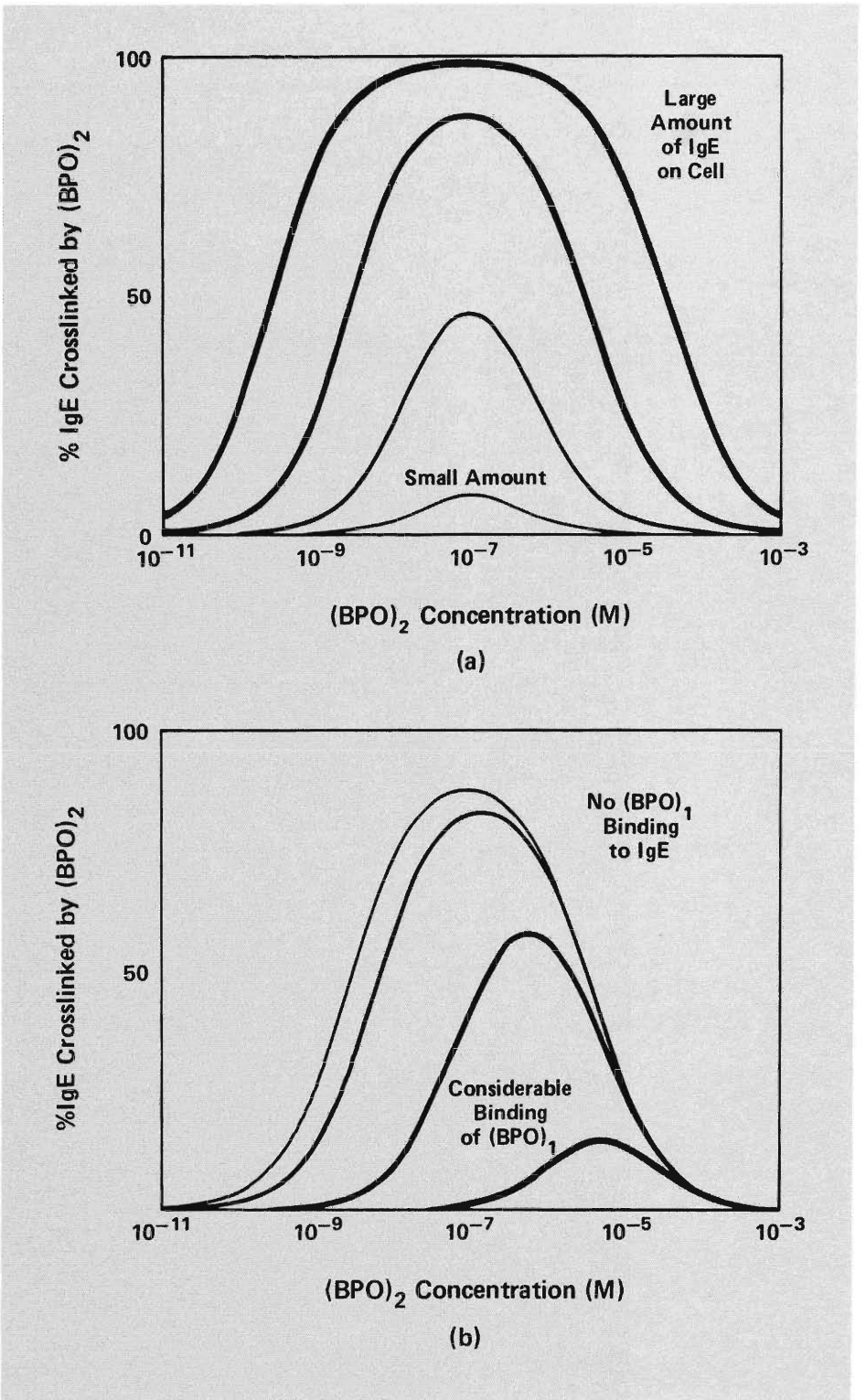
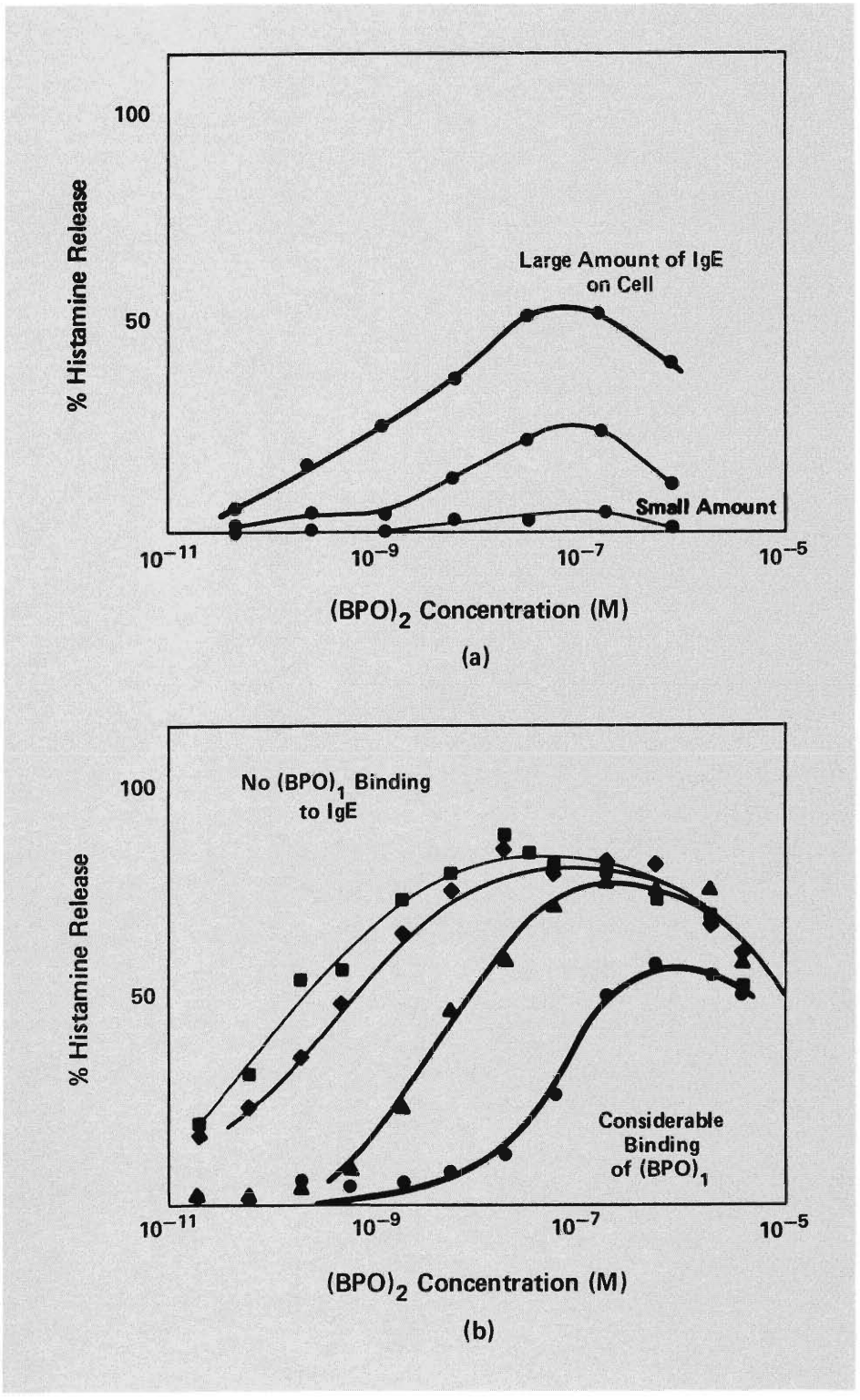


Fig. 14. Behavior of theoretical crosslinking curves. (a) As the amount of BPO-specific IgE antibodies on the basophil surfaces is increased, the crosslinking curves exhibit greater maxima, but the allergen concentration at which the maxima occur remains constant. (b) As the concentration of monovalent $(BPO)_1$ to which the BPO-specific IgE antibodies are exposed increases, the number of crosslinks decreases and the position of the maximum shifts to higher $(BPO)_2$ concentrations.



imum at the same $(BPO)_2$ concentration.

From these and other similar experiments on basophils from nonallergic donors, we concluded that the percentage of histamine released rises and falls directly with the rise and fall of the fraction of crosslinked antibodies. This behavior is quite unlike that of basophils from allergic donors. Recall that for these cells the fall in the dose response curve occurs because the number of crosslinks has increased to the point where nonspecific desensitization dominates over histamine release. In other words, too many crosslinks exist rather than too few. The difference in behavior is surely related to the fact that basophils from nonallergic donors have relatively small numbers of specific IgE antibodies bound to their surfaces, whereas cells from allergic donors have relatively large numbers.

Basophils from allergic and nonallergic donors should manifest other differences as well. In particular, can basophils from nonallergic donors be desensitized by withholding calcium ions?

Another "Off" Signal: Specific Desensitization

We found that passively sensitized

Fig. 15. Behavior of experimental dose response curves. Like the crosslinking curve shown in Fig. 12, the dose response curves for basophils from nonallergic donors exhibit maxima and are symmetric about those maxima. Further, the dependence of the dose response curves on the amount of BPO-specific antibodies on the basophil surfaces (a) and on the (BPO) concentration (b) is similar to that of the crosslinking curves (Fig. 14). The high concentration wings of the dose response curves are missing because of the difficulty in keeping high concentrations of $(BPO)_2$ in solution.

basophils from nonallergic donors do, in fact, desensitize. But they do so in a way that has not before been observed—they desensitize specifically. When exposed to $(BPO)_2$ in the absence of calcium ions, they behave in one respect just like basophils from allergic donors: they progressively lose their ability to respond to $(BPO)_2$ (by releasing histamine) when calcium ions are added. However, when exposed in the presence of calcium ions to a rabbit antibody specific to human IgE antibodies, they release histamine normally. Desensitizing with one allergen or crosslinking agent affected only the histamine release that is triggered by that allergen or crosslinking agent.

Our colleagues at Johns Hopkins performed an experiment showing that basophils from some *allergic* donors also desensitize specifically. After screening a number of individuals who were allergic to substances other than penicillin, they found one whose basophils had some fraction of their large numbers of Fc_ϵ receptors unfilled. They passively sensitized these basophils with BPO-specific IgE antibodies and then exposed them in the absence of calcium ions to $(BPO)_2$ at the concentration for maximum crosslinking of the antibodies. The cells were then exposed in the presence of calcium ions to either $(BPO)_2$, ryegrass allergen (the individual's natural allergen), or a goat antibody specific to IgE antibodies. The results (Fig. 16) show that the cells desensitized only to $(BPO)_2$. In this case, even though the number of crosslinks between BPO-specific IgE antibodies was at a maximum, the total number was too small for nonspecific desensitization to play a significant role, and another type of desensitization came into play.

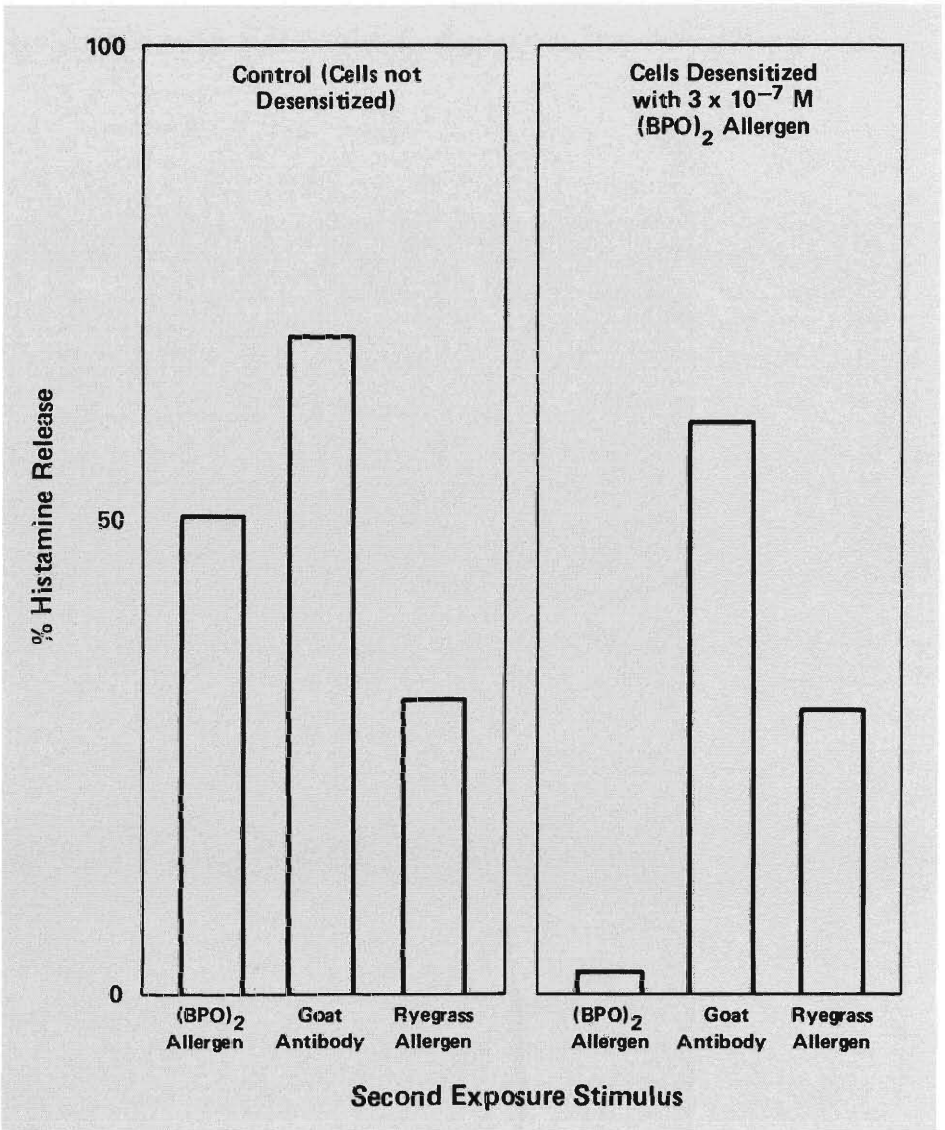


Fig. 16. Specific desensitization. The basophils studied in the experiment illustrated here were from a donor who was allergic to ryegrass. Although most of the Fc_ϵ receptors on the donor's basophils were filled with IgE antibodies specific to the ryegrass allergen, a sufficiently large number of empty Fc_ϵ receptors existed that the basophils could be passively sensitized to the BPO group. On the left are the percentages of histamine released by the passively sensitized basophils when exposed to either the synthetic $(BPO)_2$ allergen, a goat antibody specific to human IgE antibodies, or the ryegrass allergen. On the right are the percentages of histamine released by these cells when exposed to the allergens or the goat antibody after ***being desensitized to $(BPO)_2$. Only the histamine release triggered by the desensitizing agent is significantly affected.***

In a subsequent study we found that allergic individuals whose basophils desensitize specifically have much lower concentrations of IgE antibodies in their serum than do allergic individuals whose basophils desensitize nonspecifically. This fact is further support for the suggestion that specific desensitization is a phenomenon associated with small numbers of specific IgE antibodies on the basophil surfaces.

Because $(BPO)_1$, which can bind but not crosslink IgE antibodies, does not desensitize basophils, we know that specific desensitization is also triggered by crosslinks. We therefore predict that basophils from nonallergic donors will undergo the greatest specific desensitization at the allergen concentration producing the maximum number of crosslinks, that is, the concentration at which histamine release is a maximum. To test this prediction we performed the following experiment. Again, basophils from nonallergic donors were passively sensitized to $(BPO)_2$. The dose response curve for these

TABLE I

CHARACTERISTICS OF SPECIFIC AND NONSPECIFIC DESENSITIZATION

	Specific Desensitization	Nonspecific Desensitization
Substances to which the basophils desensitize	Only the desensitizing allergen	All allergens and crosslinking agents
Relation to number of specific IgE antibodies on the basophil surfaces	Occurs when the number is small	Occurs when the number is large
Relation to number of crosslinks on the basophil surfaces	Is maximum when the number, although small, is maximum	Is maximum when the number is maximum
Allergen concentration at which maximum histamine release occurs	Identical to concentration at which crosslinks and desensitization are maximum	Lower than the concentration at which crosslinks and desensitization are maximum
Explanation in model	Linked Fc_ϵ receptors are altered	Limited amount of gating factor is available

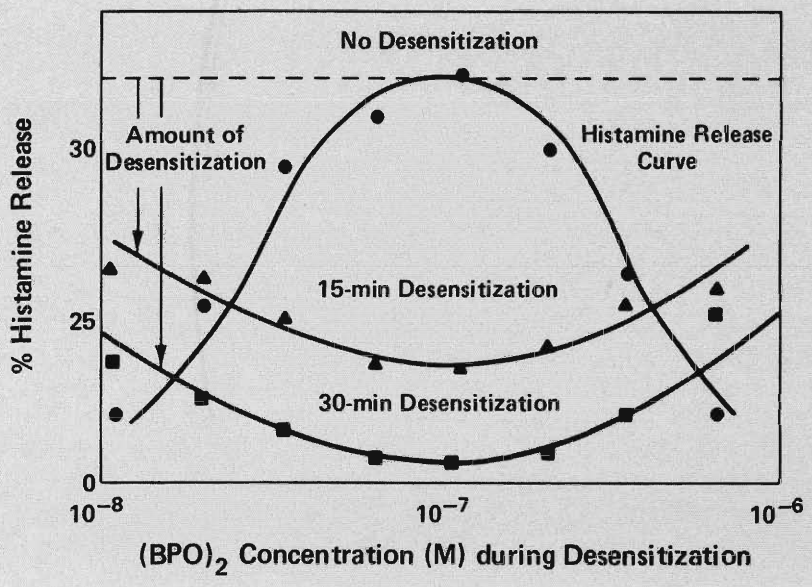


Fig. 17. The basophils studied in the experiment illustrated here were from a nonallergic donor and had been passively sensitized to the BPO group. The dose response curve for these cells exhibited a maximum at a $(BPO)_2$ concentration of 10^{-7} molar. The cells were then specifically desensitized to $(BPO)_2$ by exposing them in the absence of calcium ions to various $(BPO)_2$ concentrations for either 15 or 30 minutes. Finally, the percentages of histamine released by the cells when exposed in the presence of calcium ions to a $(BPO)_2$ concentration of 10^{-7} molar was measured. The difference between the maximum histamine released (dashed line) and the measured histamine release (lower curves) is a measure of the amount of specific desensitization undergone by the cells at that (BPO) concentration. Note that the desensitization is greatest for a desensitizing $(BPO)_2$ concentration of 10^{-7} molar, the $(BPO)_2$ concentration for maximum histamine release.

cells exhibited a maximum at a $(BPO)_2$ concentration of 10^{-7} molar. Groups of the sensitized cells were then desensitized by exposing them for a fixed time to $(BPO)_2$ in the absence of calcium ions. The $(BPO)_2$ concentration was varied from group to group. Finally, the desensitized cells were exposed in the presence of calcium ions to $(BPO)_2$ at the concentration for maximum histamine release, or 10^{-7} molar. If no desensitization had taken place, all the groups of cells would have released the maximum amount of histamine. But desensitization did, in fact, occur and was evidenced by the cells' release of less than the maximum amount of histamine. Further, the group of cells desensitized at a $(BPO)_2$ concentration of 10^{-7} molar released the least amount of histamine and thus underwent the greatest amount of desensitization (Fig. 17).

Recently, in collaboration with Henry Metzger of the National Institutes of Health, we used covalently linked pairs of IgE antibodies to show that large aggregates of crosslinked IgE antibodies are not required to induce specific desensitization. Just as for histamine release, the formation of linked pairs of IgE antibodies is the "unit" signal for specific desensitization.

The two modes of desensitization are compared in Table I. Are they related or are they independent mechanisms?

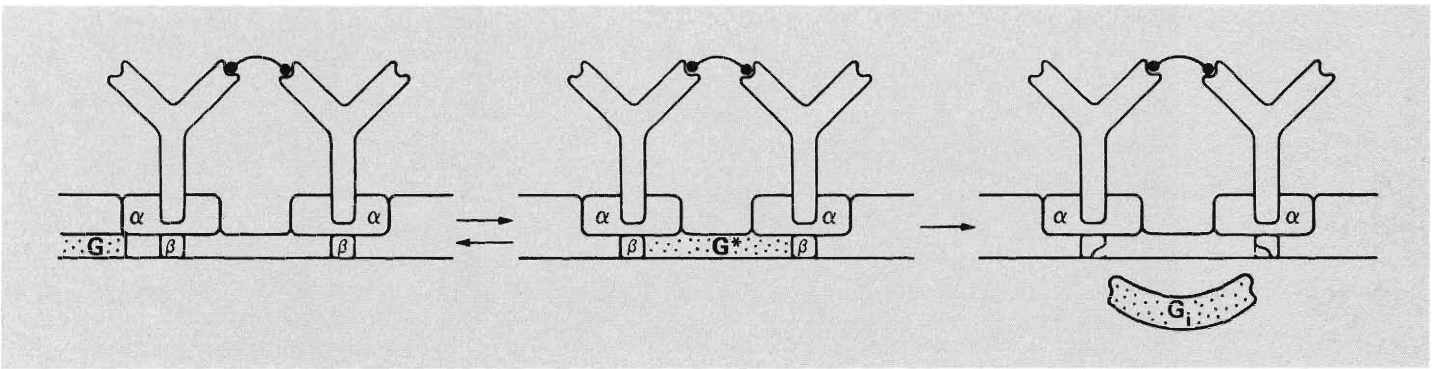


Fig. 18. Desensitization model. In this model an open calcium channel (G) is formed when an allergen crosslinks two IgE antibodies and a gating factor (G) combines with a part (β) of the Fc_ϵ receptors to which the antibodies are bound. Decay of the channel results in loss of the gating factor (G_i), which can account for nonspecific desensitization, and alterations in the β portion of the receptors, which can account for specific desensitization.*

The Transition from Specific to Nonspecific Desensitization

We have evidence that basophils with relatively small numbers of specific IgE antibodies on their surfaces desensitize specifically and that basophils with relatively large numbers of specific IgE antibodies on their surfaces desensitize nonspecifically. Perhaps if the number of specific IgE antibodies on the basophil surfaces could be varied over a large enough range, the type of desensitization exhibited by the basophils would undergo a smooth transition from specific through partially nonspecific to completely nonspecific.

In 1981 Donald MacGlashan and Lawrence Lichtenstein performed an experiment to test this suggestion. A major difficulty they faced was finding a donor whose basophils had enough free Fc_ϵ receptors, but eventually one was located whose basophils had close to 20,000 per cell. To take full advantage of the free sites available, they very carefully purified a BPO-specific antibody preparation before passively sensitizing the basophils. They estimated that in their experiments the number of BPO-specific IgE antibodies per basophil was varied from approximately 800 to 14,000. Over this range they observed a smooth transition in the type of desensitization from specific to nonspecific, although the nonspecific desensitization was not complete. Our expectation is that if more BPO-specific IgE antibodies could be placed on the basophils, complete nonspecific desensitization would be achieved.

A Speculative Model for Desensitization

Crosslinking of IgE antibodies on the surface of a basophil causes a flow of calcium ions into the cell that triggers the release of histamine-containing granules from the cell. But crosslinking also leads to desensitization of a type determined by the number of IgE antibodies on the basophil surface.

How does this all come about? We're not certain, but we have some ideas that we have formalized in a mathematical model. The model can explain what has been observed so far and can also make predictions that can be tested experimentally. The basic features of the model are sketched in Fig. 18. We propose that when two Fc_ϵ receptors are brought into proximity by crosslinked IgE antibodies, they combine with a "calcium gating factor" in the cell membrane. This reaction forms a channel through which calcium ions flow into the basophil. We have shown experimentally that channels formed by crosslinked IgE antibodies are short-lived. We have incorporated this observation into our model by assuming that a channel rapidly decays to an inactive form. We also assume that only a limited amount of gating factor is available.

This model explains nonspecific desensitization as follows. As basophils with large numbers of IgE antibodies specific to a particular allergen are exposed to the allergen in the absence of calcium ions, so many calcium channels are formed that the supply of gating factor is exhausted. When later exposed to *any* allergen in the presence of calcium ions, no further calcium channels can be formed and those formed during desensitization have decayed. Hence, no histamine is released.

Calcium ions act as stimulatory signals not only for basophils and mast cells but also for a variety of other cells, some of which undergo processes similar to desensitization. For example, cells in the blowfly's salivary gland can be stimulated to secrete by an influx of calcium ions triggered by the molecule 5-hydroxytryptamine. Continued

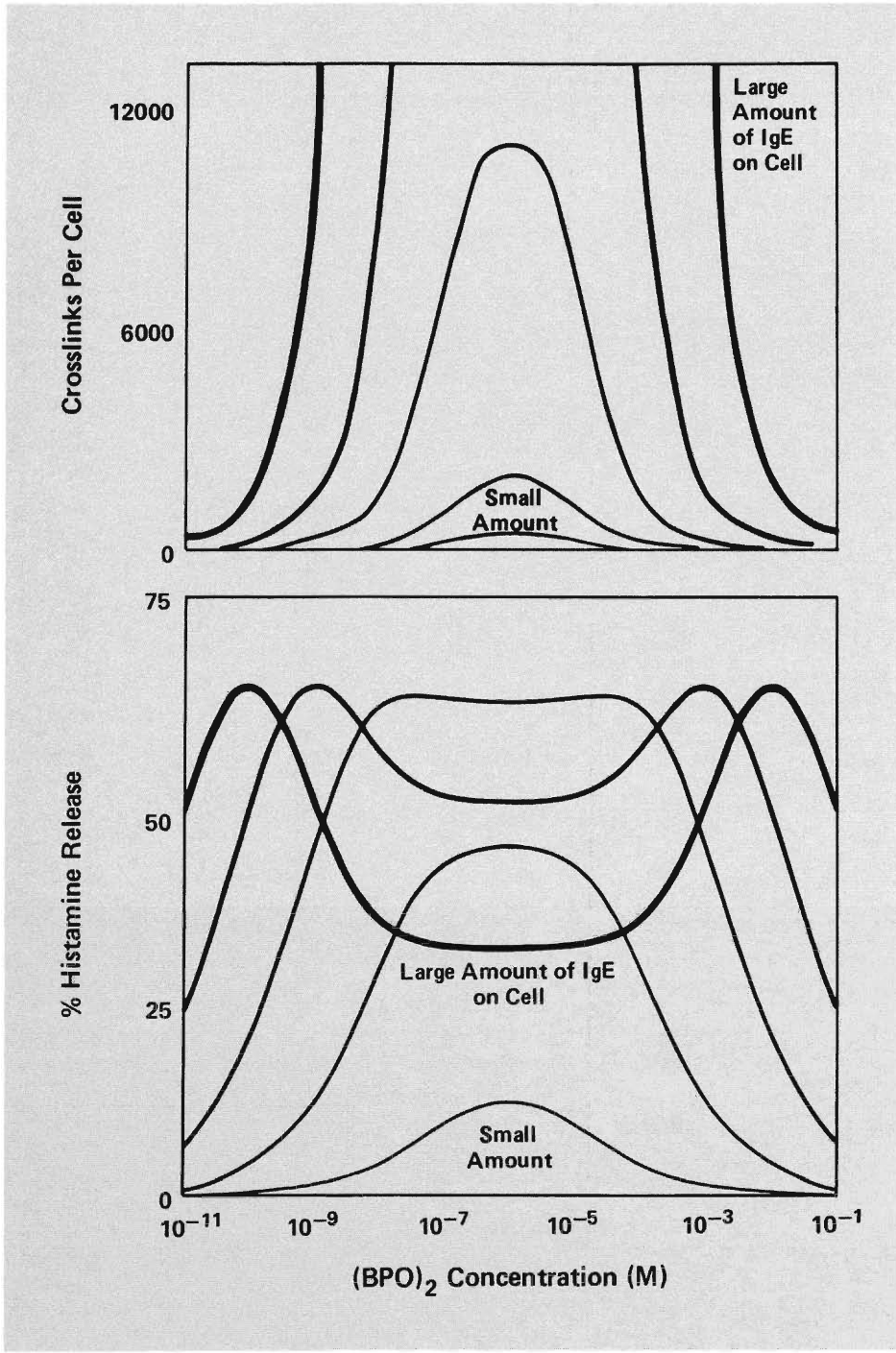


Fig. 19. These two graphs illustrate the relationship predicted by the model of Fig. 18 between crosslinks and histamine release. The upper graph shows that very large numbers of crosslinks can form on basophils with large numbers of BPO-specific IgE antibodies on their surfaces. These large numbers of crosslinks are accompanied, at $(BPO)_2$ concentrations around the optimum for release, by reductions in the percentage of histamine released (lower graph). This reduction represents the transition from specific to nonspecific desensitization.

treatment with 5-hydroxytryptamine leads to a shutdown of the mechanism for calcium ion transport and, hence, to a shutdown of secretion. The shutdown occurs because the calcium gating factor, which has been identified in this case as phosphatidylinositol, is only a minor constituent of the cell membranes and is eventually used up. When phosphatidylinositol is restored to the cells, desensitization is reversed, and the cells once again secrete normally.

What the calcium gating factor is in human basophils, if there is one, is not known. We have attempted to identify a gating factor by incorporating phosphatidylinositol and other likely candidates into nonspecifically desensitized basophils, but none of these substances caused their release of histamine to revert to normal.

Another feature of the model is invoked to explain specific desensitization. We propose that decay of a calcium channel is accompanied by inactivation of the Fc_ϵ receptors between which the calcium channel was formed. This inactivation affects the course of events as follows. As basophils with small numbers of IgE antibodies specific to a particular allergen are exposed to that allergen in the absence of calcium ions, only a small number of calcium channels are formed and the supply of gating factor is not exhausted. But as the calcium channels decay, so also do the Fc_ϵ receptors filled with these crosslinked antibodies. This inactivation of the receptors in effect inactivates all the IgE antibodies specific to the desensitizing allergen, and later exposure to that allergen in the presence of calcium ions does not cause histamine release. On the other hand, later exposure in the presence of calcium ions to a different allergen can cause histamine release because that allergen crosslinks IgE antibodies specific to itself. These crosslinked antibodies can then combine with the remaining gating factor to form calcium channels.

The transition between the two types of desensitization occurs when the number of crosslinks becomes significant compared to the amount of gating factor. These ideas also are summarized in Table I. Figure 19 shows the relationship between crosslinks and histamine release predicted by the model. The lower three curves in both graphs essentially duplicate the crosslinking and dose response curves for basophils with limited numbers of specific IgE antibodies on their surfaces (Figs. 14a and 15a). The upper two curves in both graphs show the effect of large numbers of specific IgE antibodies. In particular, the

onset of nonspecific desensitization is indicated at the middle $(BPO)_2$ concentrations by the continued rise (off scale) in the number of crosslinks accompanied by a decrease in the percentage of histamine released.

Our model couples specific and nonspecific desensitization. Because of this coupling the model makes strong predictions about the relationship between the time course of specific and nonspecific desensitization for experiments carried out with cells from the same donor. It also makes predictions about what happens when basophils desensitized with one allergen are later exposed to the same or different allergens. Experiments to test these predictions are in progress at Johns Hopkins.

Of course it could be that specific and nonspecific desensitization are not coupled at all. Indeed, our first guess was that nonspecific desensitization came about because some gating factor was used up whereas specific desensitization was the result of a totally independent process called receptor-mediated endocytosis. This process would literally transfer the allergen-linked IgE antibodies from the cell surface to the cell interior. This guess was based on a number of examples of endocytosis in which cells internalize a variety of their own surface receptors. In some cases the internalization is triggered by the binding of a molecule to the receptor and in other cases by the crosslinking of molecules bound to the receptors. However, Donald MacGlashan has recently shown that endocytosis does not occur in specifically desensitized basophils. He passively sensitized basophils with BPO-specific IgE antibodies and specifically desensitized these cells by exposing them to $(BPO)_2$ in the absence of calcium. He then washed off the $(BPO)_2$ and exposed the cells to radioactively labeled molecules containing many BPO groups. He observed that the basophils specifically bound the radioactive label. Therefore, the BPO-specific IgE antibodies still remained on the basophil surfaces after specific desensitization.

Conclusion

Allergic reactions of the immediate type arise because of a complicated chain of events: exposure to an allergen, recognition of the allergen by the immune system, production of IgE antibodies, sensitization of basophils or mast cells, re-exposure to the allergen, triggering of basophils or mast cells to release histamine and other chemicals, and response of the body to those chemicals. Treatments of allergies are designed to break this chain. Of course, the best thing to do is to avoid the allergen. This is straightforward if you are allergic to cod fish, but impossible if you are allergic to juniper pollen and insist on living in New Mexico.

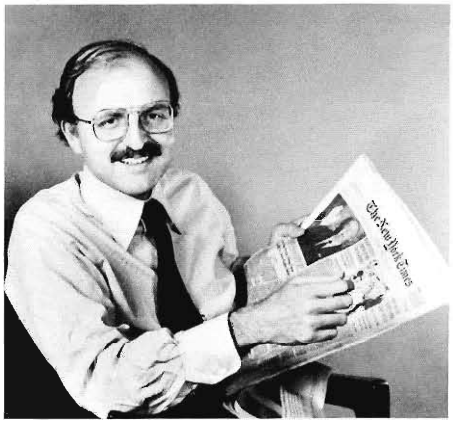
If you can't stay away from the allergen, you can try to break the chain by manipulating the immune response. Almost all allergy "shots" are directed toward this end; they attempt to affect, not the basophils or the mast cells, but the cells of the immune system that produce the antibodies. Exposures to low concentrations of an allergen (the shots) over long periods can sometimes desensitize these cells. The cells are then said to be tolerized because they no longer produce antibodies that bind to the allergen but rather tolerate its presence. If tolerance can be maintained so that the IgE response is constantly blocked, the chain is broken and the patient is free of symptoms. Unfortunately, a substantial fraction of those treated with low doses of allergen do not become tolerant to the allergen.

A second type of treatment involves increasing the immune response rather than decreasing it. The idea is to produce in an allergic individual such high concentrations of antibodies of classes other than IgE that these antibodies, flowing in the blood, bind to the allergen and prevent it from triggering basophils and mast cells. This raising of the "blocking" antibody concentration has been highly successful in the treatment of bee sting allergy. Allergic individuals receive regular injections of pure bee venom at concentrations that raise and maintain a high IgG response but do not trigger histamine release from basophils. If such an individual is stung by a bee, the IgG antibodies in solution bind to the bee venom and prevent it from crosslinking the venom-specific IgE antibodies on basophils. This treatment, however, is much less successful for allergens that are inhaled than for allergens that are injected.

An alternative to immunotherapy is drug therapy. Drugs such as the antihistamines have been used for some time in attempts to prevent the body from responding to the chemicals released during allergic reactions. Some of these drugs also inhibit, at least to some extent, the release of chemicals from the basophil and mast cell granules.

These are the major approaches to the treatment of immediate hypersensitivity. At the moment there is no therapy in use that is designed to bring about basophil and mast cell desensitization. Some cases have been reported in which immunotherapy produced desensitization of basophils, but this result was fortuitous rather than by design. Perhaps our expanding grasp of the desensitization process will alter this situation. ■

AUTHORS



Byron Goldstein was born and grew up in New York City. He received his Bachelor of Science in physics from the City College of New York in 1961 and his Ph.D. in theoretical physics from New York University in 1967. He began to work seriously on problems related to biology while he was a National Institutes of Health postdoctoral fellow at the University of California, San Diego. He joined the Theoretical Biology and Biophysics Group at the Laboratory in 1975. Before coming to Los Alamos he was a Professor of Physics and member of the Biophysics Group at Fairleigh Dickinson University.

Micah Dembo earned his Bachelor of Science in mathematics from Allegheny College in 1972 and his Ph.D. in biomathematics from Cornell University Medical College in 1977. After finishing graduate work, he came to Los Alamos as a postdoctoral fellow in the Theoretical Biology and Biophysics Group and remained as a staff member after the postdoctoral appointment ended. During his years in the group he has worked on a number of theoretical problems of importance in biology. In addition to developing mathematical models of cell activation and desensitization, he has worked on the modeling of cooperative interactions in proteins, on diffusion reaction problems, particularly with regard to membrane transport phenomena, and on fluid mechanical models of cell motility.

Further Reading

Kimishige Ishizaka and Teruko Ishizaka, "Mechanisms of Reaginic Hypersensitivity and IgE Antibody Response," Immunological Reviews **41**, 109-148 (1978).

Anne K. Sobotka, Floyd J. Malveaux, Gianni Marone, Larry L. Thomas, and Lawrence M. Lichtenstein, "IgE-Mediated Basophil Phenomena: Quantitation, Control, Inflammatory Interactions," Immunological Reviews **41**, 171-185 (1978).

Micah Dembo, Byron Goldstein, Anne K. Sobotka, and Lawrence M. Lichtenstein, "Histamine Release Due to Bivalent Penicilloyl Haptens: Control by the Number of Cross-Linked IgE Antibodies on the Basophil Plasma Membrane," Journal of Immunology **121**, 354-358 (1978).

Lawrence M. Lichtenstein, Martin D. Valentine, and Anne K. Sobotka, "Insect Allergy: The State of the Art," Journal of Allergy and Clinical Immunology **64**, 5-12 (1979).

Anne K. Sobotka, Micah Dembo, Byron Goldstein, and Lawrence M. Lichtenstein, "Antigen-Specific Desensitization of Human Basophils," Journal of Immunology **122**, 511-517 (1979).

Micah Dembo and Byron Goldstein, "A Model of Cell Activation and Desensitization by Surface Immunoglobin: The Case of Histamine Release from Human Basophils," Cell **22**, 59-67 (1980).

Ann M. Dvorak, H. H. Newball, H. F. Dvorak, and L. M. Lichtenstein, "Antigen-Induced IgE-Mediated Degranulation of Human Basophils," Laboratory Investigation **43**, 126-139 (1980).

Teruko Ishizaka, "Analysis of Triggering Events in Mast Cells for Immunoglobulin E-Mediated Histamine Release," Journal of Allergy and Clinical Immunology **67**, 90-96 (1981).

Donald W. MacGlashan, Jr., and Lawrence M. Lichtenstein, "The Transition from Specific to Nonspecific Desensitization in Human Basophils," Journal of Immunology **127**, 2410-2414 (1981).

Paul D. Buisseret, "Allergy," Scientific American Vol. 247, No. 2, 86-95 (1982).

B. Goldstein and M. Dembo, "The IgE-Mediated Activation and Desensitization of Human Basophils," in *Cell Surface Phenomena*, A. S. Perelson, C. DeLisi, and F. Wiegel, Eds. (Marcel Dekker, New York, in press).

Comments on The History of the H-Bomb

by Hans A. Bethe

Theoretical Division Leader, Los Alamos, 1943-45
Consultant, Los Alamos, 1946-Present

Back in 1954 I wrote an article on the history of the H-bomb, stimulated by a book by Shepley and Blair which gave an entirely distorted view of that history. It took until recently to have that article declassified. I had intended to put this article into the Laboratory's archives and not to publish it, in order not to stir up old controversies. However, now there has appeared the very popular book by Peter Goodchild, *J. Robert Oppenheimer: Shatterer of Worlds*. While this book is excellent in most respects, it gives among others a very wrong impression of the development of the H-bomb. Therefore, I am now publishing this article, and I have added a few remarks specifically correcting some of the mistakes in Goodchild's book. What follows is a (slightly edited) version of the 1954 article, which was written in some anger about certain events of 1953-54.

...The first of these events was an article by C. J. V. Murphy* in *Fortune* of May 1953 which presented a highly biased and inaccurate picture of the H-bomb development and of the efforts of many American scientists to establish a more adequate air defense system for this country. Next came the most important event, the Oppenheimer case. The hearings on this case, and their unexpected publication by the Atomic Energy Commission, have made the general public aware of the deep conflicts which, at various times, arose in connection with the thermonuclear development. Fortunately, the record of the Oppenheimer hearings contains testimony which enables anyone who takes the trouble to read through its 992 pages to form his own opinion on the issues.

Now, however, [that is, in 1954] a book has appeared which requires an immediate answer. It is written by James R. Shepley and Clay Blair, Jr., and purports to tell the American public the history of the hydrogen bomb. Apart from official public statements, which were in any case not particularly informative on the matters discussed so freely by the authors, the information and opinions presented in the book have obviously been obtained from persons holding extreme views on a number of matters. Whoever these persons may have been, they were extreme in their dislike and/or distrust of Oppenheimer, extreme in their certainty of the malfeasance of Los Alamos, extreme in their conviction that anyone who expressed misgivings or raised questions concerning the wisdom of committing ourselves to the H-bomb program was ipso facto subversive. As a result, the book is full of misstatements of fact, and so phenomenally biased as to retain little contact with the events that actually occurred.

Many of the readers of the book will be familiar, from other reports, with some of the political moves on the H-bomb project that went on in Washington. The book is made only more misleading because it reports a number of these moves outwardly accurately, as far as I can judge. Many readers may thereby be misled into believing that the progress of the technical work is also reported correctly by Shepley and Blair. With very few exceptions this is not so; and the fact that the technical history was different puts a completely different light on the reasons and justification for various "political" moves, e.g., on the agitation for the establishment of a second weapons laboratory.

In this article I will talk in the main about the technical history of the project since this is the only subject which I know first-hand. Unfortunately, any factual account of technical development must be incomplete because large parts of the subject remain classified. Many of the points in this article would become even more convincing if classified matters could be discussed.

I shall not attempt to give an exhaustive list of the misstatements of fact in the Shepley-Blair book. On many matters reported in the book I have no first-hand knowledge. Even where I do have such knowledge, I shall leave out much detail, as well as much that is still classified, and, finally, many of the points that were discussed by Dr. Bradbury in his excellent press statement and press conference which were published in *The New Mexican* of Santa Fe, New Mexico, on Friday, September 24, and Sunday, September 26, 1954.

At various points in this article, reference will be made to the book by Shepley and Blair, which will be quoted as SB with the page number. Reference will also be made to testimony in the Oppenheimer case, which will be quoted as OT with the page number in the official publication.

The historical material is arranged under three major headings: Wartime development, Postwar development of fission bombs, and Thermonuclear weapons. In these sections I try to follow the historical sequence and mention SB as I go along. In a fourth section

**This article and interviews of Mr. Murphy with persons concerned are quoted as one of their chief sources of information by Shepley and Blair. (Letter to the Editor of the* New York Herald-Tribune, *October 15, 1954.)*

I discuss the things which were required before success in a thermonuclear program could be achieved.

1. Los Alamos During Wartime

After the Los Alamos Laboratory was started in the Spring of 1943, it became clear that the development of a fission bomb was far more difficult than had been anticipated. If our work was to make any contribution to victory in World War II, it was essential that the whole Laboratory agree on one or a very few major lines of development and that all else be considered of low priority. Teller took an active part in the decision on what were to be the major lines. Before any specific work of an engineering or design nature could be taken up, it was necessary that theoretical investigations be brought to the stage where they could provide some detailed guidance. A distribution of work among the members of the theoretical division was agreed upon in a meeting of all scientists of the division, and Teller again had a major voice.

In the early Summer of 1944, the Laboratory adopted as its main line the development of the implosion, a method since described publicly, e.g., in the testimony in the Greenglass trial and in instructions to U.S. Customs and Postal Officials for the purpose of helping them to detect clandestine import of atomic bomb parts.

As soon as the implosion method was proposed by Neddermeyer, Teller advocated that the Laboratory should devote major effort to its development. In 1944 he was given the responsibility for all theoretical work on this problem. Teller made two important contributions. He was the first to suggest that the implosion would compress the fissile material to higher than normal density inside the bomb. Furthermore he calculated, with others, the equation of state of highly compressed materials, which might be expected to result from a successful implosion. However, he declined to take charge of the group which would perform the detailed calculations of the implosion. Since the theoretical division was very shorthanded, it was necessary to bring in new scientists to do the work that Teller declined to do. Partly for this reason, some members of the British Atomic Energy team, already working in the U.S. on other aspects of the Manhattan District project, were brought to Los Alamos and asked to help with this problem. The leader of the British theoretical group was Rudolf Peierls, and another very hardworking member was Klaus Fuchs.

With the pressure of work and lack of staff, the theoretical division could ill afford to dispense with the services of any of its members, let alone one of such brilliance and high standing as Teller. Only after two failures to accomplish the expected and necessary work, and only on Teller's own request, was he, together with his group, relieved of further responsibility for work on the wartime development of the atomic bomb. This was done by me, as the Leader of the Theoretical Division, not by Oppenheimer, the Director of the Laboratory.

About this same development Shepley and Blair have the following to say (page 40): "Edward Teller also worked at Los Alamos during the war. But because Oppenheimer did not like him personally—a fact that was perhaps traceable to their differing political views—Teller was denied a specific job in connection with the development of the atomic bomb." It is obvious that this is almost the exact opposite of the truth.

It is difficult to judge another man's personal feelings toward a third, even if you see both of them almost daily. But as far as I could see, the personal relations between Teller and Oppenheimer were very good at the beginning of Los Alamos. Later on, Teller's attitude toward his own work and toward the program of the Laboratory created a strain in his relations with Oppenheimer, and, to a lesser degree, in his relations with myself. At the start I had regarded Teller as one of my best friends and as the most valuable member of my division. Our relation cooled when Teller did not contribute much to the work of this division. More important perhaps for a disturbance of relations was his wish to spend long hours discussing alternative schemes which he had invented for assembling an atomic bomb or to argue about some remote possibilities why our chief design might fail. He wanted to see the project being run like a theoretical physics seminar and spent a great deal of time talking and very little time doing solid work on the main line of the Laboratory. To the rest of us who felt we had a vital job to do, this type of diversion was irksome. To come back to the relations between Teller and Oppenheimer, politics certainly played no role in them. Communism in particular was no issue at that time at Los Alamos.

The success of Los Alamos rested largely on its teamwork and the leadership of its director. Shepley and Blair do not wish to give credit to Oppenheimer because (footnote on page 28) "the technical contributions at wartime Los Alamos" were not made by him. It is not the primary function of the director of a laboratory to make technical contributions. What was called for from the Director of Los Alamos at that time was to get a lot of "prima donnas" to work together, to understand all the technical work that was going on, to make it fit together, and to make decisions between various possible lines of development. I have never met anyone who performed these functions as brilliantly as Oppenheimer, as Goodchild rightly emphasizes.

The individuals mentioned in the footnote on page 28 of SB as having made "the technical contributions at wartime Los Alamos" are an odd collection. Some, like von Neumann, really did contribute most important ideas. Other very important names like Kistiakowsky, Bradbury, Bacher, Rossi, Cyril Smith, R. R. Wilson, Feynman, et al., are omitted. Instead, the footnote mentions two persons who did not work significantly on the A-bomb at Los Alamos, but almost exclusively on the H-bomb.

The implosion, which has been mentioned as the main program of the Laboratory, consists of placing a large quantity of high explosive around the surface of a small sphere of uranium-235 or plutonium. This method was invented *during* the war, while SB, page 115, make it appear as if this method had been invented only in 1950. Also, the idea of using a fraction of a critical mass (fractional crit) for an atomic explosion orginated during the war; it was not "sparked by Teller's intuition" in 1950. Rather, it was common knowledge and

strongly advocated by the Los Alamos Laboratory, and by the Atomic Energy Commission, in 1948-49. The idea developed from the same implosion calculations which Teller had refused to perform. I believe in fact that I was the first to point out this possibility but it is true that Teller quickly supported it, all in 1944. However, it was not until the art of fission bombs had been thoroughly developed by the postwar Los Alamos Laboratory that the fractional crit became a practical possibility. In other words this scheme had long been on the Los Alamos books and was waiting only for the perfection of techniques. To give Teller and the year 1950 credit for this idea as SB do on page 115 is entirely false.

There are two interesting sidelights on the accuracy of SB's reporting. In the first place, the important development of the fractional crit weapon had no bearing on the thermonuclear work at all, contrary to SB's statement. Secondly, SB claim that the General Advisory Committee [a nine-man committee, established in 1947 and chaired by Oppenheimer until 1952, that advised the AEC on scientific and technical matters] was against fractional crit weapons. If they were, Oppenheimer must have had a badly split personality because the Vista report, with which Oppenheimer was prominently identified but which SB and their trusted colleague, C. J. V. Murphy, have criticized so much, recommended fractional crit weapons as a mainstay of our arsenal.

2. Postwar Development of Fission Bombs

It has been made amply clear in the Oppenheimer testimony and elsewhere that at the end of the war the number of scientists at Los Alamos declined severely and that this was especially true of the number of senior staff members. The theoretical division, which has the main responsibility for the conceptual design of weapons, was reduced from over thirty scientists to eight in 1946 (according to Bradbury's press statement); it has since increased again to over fifty [in 1954]. This decline was part of the general movement to "let the boys come home." We all felt that, like the soldiers, we had done our duty and that we deserved to return to the type of work that we had chosen as our life's career, the pursuit of pure science and teaching.

The older ones among us felt a heavy responsibility to our teaching. Wartime had shown that this country had a very short supply of competent scientists, and Los Alamos was one of the best examples. The young scientists whose careers had been interrupted by the war wanted to get training under the G.I. Bill of Rights. The largest graduate schools in physics before the war had about fifty graduate students; now this number jumped to a hundred and, in some universities, to over two hundred. The great effort which was made in training these young people has borne fruit in the meantime. Only because of it could laboratories like Los Alamos gather their large staff of highly competent scientists in the years since 1948. Only in this way could the Los Alamos theoretical division grow to its present [1954] 50-odd members, not to speak of the important work that other young scientists are doing in industry, in other governmental laboratories, and in the universities themselves.

For most of the scientists, young or old, who participated in the wartime work at Los Alamos, this was their first experience with work of a secret nature or work having immediate practical military significance. It is in no way surprising that most of them preferred the free interchange of ideas with their colleagues in this country and abroad which goes with pure, non-secret research. Moreover, it was not obvious in 1946 that there was any need for a large effort on atomic weapons in peacetime. All these factors help explain the exodus of scientists from Los Alamos and other wartime projects in 1946. The most effective cure for this attitude was the behavior of Russia in the first years after the war. For many scientists one of the most convincing points in the Russian behavior was their negative attitude toward our offer to make atomic power and atomic weapons an international rather than a national development, a plan to which Shepley and Blair (page 170) refer as the scientists wanting "to give the secrets of the A-bomb to the world". Most scientists soon recognized that the Russians were not willing to open the Iron Curtain to an International Atomic Authority and Oppenheimer was one of the first to recognize this, as has been demonstrated amply in the Oppenheimer testimony. The negotiations in the U.N. Atomic Energy Commission, as much as anything else, made many of the wartime members of the Los Alamos Laboratory willing to return to weapons work at least on a part-time basis.

The fact remains that in 1946 the Los Alamos Laboratory was very weak. To demand, as Teller did as a condition for his staying, that Los Alamos tackle the super-bomb on a large scale, or plan for twelve tests a year on fission bombs, was plainly unrealistic to say the least. Dr. Bradbury, in his statement of September 24, 1954, pointed out that only as late as 1951 could a schedule of twelve test shots be reached. In only one subsequent year, 1953, was the firing of such a large number again found necessary. It is hardly possible to give enough credit to the small group of scientists who decided to stay at Los Alamos in 1946 without making demands beyond the Laboratory's capacity.

The development laboratory at Los Alamos was not the only part of the atomic energy program which was hard hit immediately after the war. The very production of bombs of the existing models also declined severely. It has been reported, e.g., in SB page 53, that only a very small stockpile of atomic bombs existed when the AEC took over from the Manhattan District on January 1, 1947. Shepley and Blair, by being unclear about dates, find here one of their opportunities for conveying a false impression while not actually making a false statement. A casual reading of their remarks on page 53 gives the impression that Oppenheimer expressed himself as satisfied with the status of the weapons program as of January 1947. If you read carefully, however, you find that his satisfaction was expressed as of the Summer of 1949, a time when great strides had been made in the A-bomb program.

As soon as the AEC took over, it and the General Advisory Committee, under the chairmanship of Oppenheimer, considered the weapons program their most important task. This is amply shown by the testimony in the Oppenheimer case. SB, pages 114 and 115, state

that the GAC, and many other scientists, when they opposed the H-bomb advocated the improvement of atomic bombs, "though" (they had) "not" (done so) "before." Of course, this advocacy of better A-bombs was not made in public, but in the privacy of its reports the GAC recommended improved A-bombs from the beginning of its existence, which was shortly after the AEC took over from the military.

Already in the interim period of 1946, but especially when they received the full support of the AEC and GAC in 1947, Los Alamos set out to work on the improvement of A-bomb design. This work bore fruit as early as 1948 in the "Sandstone" tests. SB on page 100 quote a statement by Senator Johnson that the Sandstone bombs were already improved by a factor of 6 over the wartime A-bomb. I can neither confirm nor deny the accuracy of this figure or any other figures given in SB because such figures are classified. But, assuming the statement by SB to be correct, I submit that this was a tremendous achievement of the Los Alamos Laboratory in so short a span of time.

Immediately after the results of the Sandstone tests were known, the Los Alamos Laboratory began planning further improvements in fission bombs. It was also planned that these improved designs would be tested in another test series in the Pacific, and the approximate date of that series, known later as Greenhouse, was agreed upon. It must be realized that a long time is required between the first conceptual design* and the final test of an improved weapon.

First, theoretical calculations have to be done; then a great deal of experimentation, including non-nuclear explosions, is necessary to test the soundness of the theoretical concept; simultaneously fabrication techniques may have to be developed; then a final design must be made and fabricated; and finally elaborate preparations must be made for observing the performance of the weapon at the test and for the test itself. No such development can be accomplished in a few months as has often been implied in newspaper speculations on A- and H-bomb development. It is true that now with extensive experience and expanded resources such developments can be made much more rapidly than they used to; but planning in 1948 and 1949 for a major test series in Spring 1951 seemed then a fairly strenuous time scale.

Advanced designs of A-bombs, conceived at Los Alamos in 1948 and 1949 and tested in 1951, included weapons of small diameter. This idea was proposed by Los Alamos and most vigorously supported by the AEC and the GAC. There was little interest in it among the military at first, but now [1954] they are clamoring for more of these weapons. This throws some light on the remark of SB, page 10, that "The military was . . . uneasy about the development of weapons."

It also throws light on the charge that Los Alamos was "overcautious" (SB page 144) and therefore slow. The goal in technical development is usually reached faster if the development is methodical and sustained and if mistakes are avoided, than if novel schemes are pursued before the groundwork has been laid.

"Conceptual design" involves a general decision on the properties of a weapon to be developed, including its power and its approximate geometric arrangement.

3. The Development of the H-Bomb

The H-bomb was suggested by Teller in 1942. Active work on it was pursued in the summer of 1942 by Oppenheimer, Teller, myself, and others (see Oppenheimer's testimony). The idea did not develop from Teller's "quiet work" at Los Alamos during the war as claimed by SB, pages 40 and 45.

When Los Alamos was started in Spring 1943, several groups of scientists were included who did work on this problem specifically. However, it was realized that this was a long-range project and that the main efforts of Los Alamos must be concentrated on making A-bombs (see Section 1). Teller, working on the H-bomb at Los Alamos, discovered a major difficulty (testimony by Oppenheimer). This discovery made it clear that it would be a very hard problem to make a "classical super" work, as this type of H-bomb was called. I shall refer to the classical super as Method A.

It was decided to write down, at the end of the war, an extensive record of the technical knowledge of the entire Los Alamos project. In line with this effort, it seemed also desirable to record the status of the "Super" so that work on it could be resumed the better when more manpower and other requisites were available. A summary report on this subject was written by Teller's collaborators in 1946 which turned out to be very useful for later work. I believe (but I am not sure because I was not present at Los Alamos at that time) that the conference on the Super in April 1946 also was intended partly to provide a record for the future (particularly since almost all the persons who had been working on this program had made definite plans to return to academic or non-weapon work), and possibly in addition to get some physicists from outside Los Alamos who were attending the conference interested in the problems with the hope that they might continue to work on them, theoretically and rather quietly. SB on page 55 present this conference as "a last-minute effort ... to spur the government into proceeding further with the H-bomb."

The work on thermonuclear weapons at Los Alamos never stopped. At this stage of the development, the main requirements were for theoretical work and for a few experimental physics measurements. Both of these types of work went ahead. On the basis of the monthly reports of the theoretical division of Los Alamos, it has been estimated that between 1946 and 1949 the work of that division was about equally divided between fission weapon design and problems related to thermonuclear weapons. (In this respect I was mistaken when testifying in the Oppenheimer case. I said then, from memory, that a relatively small fraction of the scientists of the division, though consisting of especially able men, were working on thermonuclear problems. Actually, the fraction was large.)

Two new methods of designing a thermonuclear weapon were invented (Methods B and C). Both inventions were due to Teller. Method B was invented in 1946, Method C in 1947. Method B was actively worked on by Richtmyer, Nordheim, and others. However,

at the time, there seemed to be no way of putting Method B into practice, as Dr. Bradbury has mentioned in his statement to *The New Mexican.* Teller himself wrote a most pessimistic report on the feasibility of this method in September 1947.

Method C is different from all the others in that thermonuclear reactions are used only in a minor way, for weapons of relatively small yield. This method seemed quite promising from the start, and as early as the Summer of 1948 it was added to the devices to be tested in the Greenhouse tests.

Theoretical work on the "classical super," Method A, proceeded continually, since this method was considered the most important of all thermonuclear devices. New plans for calculations were made frequently, mostly by consultation between Teller and the senior staff of the theoretical division. However, as Teller stated in 1946, "The required scientific effort is clearly much larger than that needed for the first fission weapon." In particular, the theoretical computations required were of such complication that they could not be handled in any reasonable time by any of the computing machines then available. Some greatly simplified calculations *were* done but it was realized that they left out many important factors and were therefore quite unreliable. Work was therefore concentrated on preparing full-scale calculations "for the time when adequate fast computing machines become available"—a sentence which recurs in many of the theoretical reports of this period. The plans for such a calculation on Method A were laid in September 1948, and the mathematical work was virtually completed by December 1949—all *before* the directive of President Truman—but it was not until mid-1952 that adequate computing machines finally became available, and by that time the most capable of them were fully engaged on the new and more promising proposal (Method D) discussed below.

When Dr. Teller and Admiral Strauss proposed in the Fall of 1949 to start a full-scale development of H-bombs, the method in their minds, as well as in the minds of the opponents of the program, was Method A. To accomplish Method A, two major problems had to be solved which I shall call Part 1 and Part 2. Part 1 seemed to be reasonably well in hand according to calculations made by Teller's group from 1944 to 1946 although nobody had been able to perform a really convincing calculation, as discussed in the paragraph above. Teller now believed that he had a solution for Part 2. In principle, the accomplishment of Part 2 had never been seriously in doubt, although the question of whether or not any particular device would behave in the way required could not be settled without experiment.

The Greenhouse thermonuclear experiment mentioned in SB was designed to test Part 2. After President Truman made the decision to go ahead with a full-scale thermonuclear program, Los Alamos made plans to add to the Greenhouse test series an experiment intended to test a particular proposal relating to Part 2. Teller played a large part in the specification of this device, and as it turned out it behaved very well. However, as on previous occasions, Teller did not do so well in directing the detailed theoretical work of his group. Only as late as January 1951, a month or so before the test device had to be shipped to the Pacific, was the full theoretical prediction of the (probably successful) behavior of the device available. But even while complete theoretical proof was lacking, most of us connected with the work at Los Alamos were confident that the Greenhouse experiment would work. As far as I could make out, at a meeting at Los Alamos in October 1950 which I attended as a guest, this was also the opinion of the GAC including Dr. Oppenheimer. Shepley and Blair instead report on page 116 that Dr. Oppenheimer expected the test device to fail. (The correct story on Oppenheimer's attitude will be discussed below.)

A very large fraction of the members of the Los Alamos Laboratory, not just a "small handful of his" (Teller's) "associates" as SB say on page 115, were extremely busy from Spring 1950 to Spring 1951 with the preparation of Teller's thermonuclear experiment. They did this in addition to preparing the Nevada tests of early 1951. The hundreds of scientists and technicians who worked for months to get the Greenhouse test ready will not enjoy Shepley and Blair's reference (page 116) to the Laboratory's "unwillingness to get involved in Teller's work."

The major feature of the year 1950 was, however, the discovery that *Part 1 of Method A was by no means under control.* While Teller and most of the Los Alamos Laboratory were busy preparing the Greenhouse test, a number of persons in the theoretical division had continued to consider the various problems posed by Part 1. In particular, Dr. Ulam on his own initiative had decided to check the feasibility of aspects of Part 1 without the aid of high-speed computing equipment. He, and Dr. Everett who assisted him, soon found that the calculations of Teller's group of 1946 were wrong. Ulam's calculations showed that an extraordinarily large amount of tritium would be necessary, as correctly stated by SB on page 102. In the Summer of 1950 further calculations by Ulam and Fermi showed further difficulties with Part 1.

That Ulam's calculations had to be done at all was proof that the H-bomb project was not ready for a "crash" program when Teller first advocated such a program in the Fall of 1949. Nobody will blame Teller because the calculations of 1946 were wrong, especially because adequate computing machines were not then available. But he was blamed at Los Alamos for leading the Laboratory, and indeed the whole country, into an adventurous program on the basis of calculations which he himself must have known to have been very incomplete. The technical skepticism of the GAC on the other hand had turned out to be far more justified than the GAC itself had dreamed in October 1949.

We can now appreciate better the attitude of the GAC, and indeed of most of the members of Los Alamos, to the Greenhouse thermonuclear test. They did not expect it to fail, but they considered it as irrelevant because there appeared to be no solution to Part 1 of the problem. The correct description of this attitude is given by Oppenheimer in his own testimony, OT page 952.

The lack of a solid theoretical foundation was the only reason why the Los Alamos work might have seemed to some to have gotten off to a slow start in 1950 (SB page 114). Purely theoretical work may seem slow in a project intended to develop "hardware," but there was

simply no basis for building hardware until the theory had been clarified. As far as the mental attitude of Los Alamos in early 1950, it was almost the exact opposite of that described by Shepley and Blair. I visited Los Alamos around April 1, 1950 and tried to defend the point of view of the GAC in their decision of October 1949. I encountered almost universal hostility. The entire Laboratory seemed enthusiastic about the project and was working at high speed. That they continued to work with full energy on Teller's Greenhouse test, after Ulam's calculations had made the success of the whole program very doubtful, shows how far they were willing to go in following Teller's lead.

Teller himself was desperate between October 1950 and January 1951. He proposed a number of complicated schemes to save Method A, none of which seemed to show much promise. It was evident that he did not know of any solution. In spite of this, he urged that the Laboratory be put essentially at his disposal for another year or more after the Greenhouse test, at which time there should then be another test on some device or other. After the failure of the major part of his program in 1950, it would have been folly of the Los Alamos Laboratory to trust Teller's judgment, at least until he could present a definite idea which showed practical promise. This attitude was strongly held by most of those on the permanent staff of the Laboratory who were responsible for its operation. As might be expected, the many discussions of aspects of this situation bred considerable emotion.

Between January and May 1951, the "new concept" was developed. (This I shall call Method D.) SB, page 119, say of this period "Teller found it impossible to get the necessary help at Los Alamos to carry on with his 'new concept' at the pace he thought the idea and program deserved." It would not have been surprising if this had been the case and if, after the major effort the Laboratory had made to prepare the Greenhouse test on Part 2, which to everybody's understanding had lost the major part of its point before the test was fired, there might have been some hesitation about immediately becoming committed to a large-scale effort along a new line of inquiry. In addition, it should be remembered that between January and May both tests in Nevada and the Greenhouse series of tests took place, and this required many senior members of the Laboratory to be at the test sites for prolonged periods of time and the attention of many others was engaged on study of results of these tests.

But what are the actual facts about this alleged delay in work on the new concept? In January Teller obviously did not know how to save the thermonuclear program. On March 9, 1951, according to Bradbury's press statement, Teller and Ulam published a paper which contained one-half of the new concept. As Bradbury has pointed out, Ulam as well as Teller should be given credit for this. Ulam, by the way, made his discovery while studying some aspects of fission weapons. This shows once more how the important ideas may not come from a straightforward attack on the main problem.

Within a month, the very important second half of the new concept occurred to Teller, and was given preliminary checks by de Hoffman. This immediately became the main focus of attention of the thermonuclear design program.

It is worth noting that the entire new concept was developed before the thermonuclear Greenhouse test which took place on May 8, 1951. The literature is full of statements that the success of Greenhouse was the direct cause of the new concept. This is historically false. Teller may have been influenced by thinking about the Greenhouse design when developing the new concept, but the success of Greenhouse (which was anticipated) had no influence on either the creation of the new concept, or on its quick adoption by the Laboratory or later by the GAC. The new concept stood on its own.

As early as the end of May 1951, I received from the Associate Director of Los Alamos a detailed proposal for the future program of the Laboratory in which Teller's new concept figured most prominently. By early June, when I visited Los Alamos for two weeks, everybody in the theoretical division was talking about the new concept.

Not only was the acceptance of the new concept not slow; but the realization of the development was a sensationally rapid accomplishment, in the same class as the achievement of Los Alamos during the war.

The impression is given in SB, pages 119-21, that Los Alamos would not have put major effort on the new concept so quickly if it had not been directed to do so by Gordon Dean, then Chairman of the AEC. Actually, Teller's new concept was so convincing to any of the informed scientists that it was accepted very quickly anyway. Certainly the events of the year 1950 would hardly seem to have given Teller any justification to ask the AEC, in the Spring of 1951, to establish a second weapons laboratory to compete with Los Alamos, as he did according to SB, page 120. (I read for the first time in the book by Shepley and Blair that Teller had asked for the second laboratory as early as Spring 1951. I did not hear of this proposal until the end of that year, although Teller *was* arguing both at Los Alamos and in Washington through the Spring of 1951 that the requirements of the thermonuclear program could only be met if the Los Alamos Laboratory underwent a major reorganization.)

The immediate acceptance of Method D by the AEC and GAC has been described in the Oppenheimer testimony. This meeting is quite incorrectly described in SB on page 135. It was not a "mass meeting". Invitations were issued only to persons directly concerned with the program, not to "any . . . scientist who wished to attend." This would obviously have been against all security regulations. Many scientists besides Teller took part in explaining the method. The meeting by no means started out in gloom, because most participants (including some members of the GAC) had some advance knowledge of the new concept. It did not require much persuading to make the GAC accept the new concept. "If this had been the technical proposal in 1949," (they) "would never have opposed the development" (Oppenheimer testimony). Now at last there was a sound technical program, and now immediately the GAC and everybody else connected with the program agreed with it. The Oppenheimer testimony shows that the GAC went beyond the Los

Alamos recommendations in allocating money for the support of the new concept.

It is difficult to describe to a non-scientist the novelty of the new concept. It was entirely unexpected from the previous development. It was also not anticipated by Teller, as witness his despair immediately preceding the new concept. I believe that this very despair stimulated him to an invention that even he might not have made under calmer conditions. The new concept was to me, who had been rather closely associated with the program, about as surprising as the discovery of fission had been to physicists in 1939. Before 1939 scientists had a vague idea that it might be possible to release nuclear energy but nobody could think even remotely of a way to do it. If physicists had tried to discover a way to release nuclear energy before 1939, they would have worked on anything else rather than the field which finally led to the discovery of fission, namely radiochemistry. At that time, concentrated work on any "likely" way of releasing nuclear energy would have led nowhere. Similarly, concentrated work on Method A would never have led to Method D. The Greenhouse test had a vague connection with Method D but one that nobody, including Teller, could have foreseen or did forsee when that test was planned. By a misappraisal of the facts many persons not closely connected with the development have concluded that the scientists who had shown good judgment concerning the technical feasibility of Method A were now suddenly proved wrong, whereas Teller, who had been wrong in interpreting his own calculations was suddenly right. The fact was that the new concept had created an entirely new technical situation. Such miracles incidentally do happen occasionally in scientific history but it would be folly to count on their occurrence. One of the dangerous consequences of the H-bomb history may well be that government administrators, and perhaps some scientists, too, will imagine that similar miracles should be expected in other developments.

Before the end of the Summer of 1951, the Los Alamos Laboratory was putting full force behind attempts to realize the new concept. However, the continued friction of 1950 and early 1951 had strained a number of personal relations between Teller and others at Los Alamos. In addition, Teller insisted on an earlier test date than the Laboratory deemed possible. There was further disagreement between Teller and Bradbury on personalities, in particular on the person who was to direct the actual development of hardware. Bradbury had great experience in administrative matters like these. Teller had no experience and had in the past shown no talent for administration. He had given countless examples of not completing the work he had started; he was inclined to inject constantly new modifications into an already going program which becomes intolerable in an engineering development beyond a certain stage; and he had shown poor technical judgment. Everybody recognizes that Teller more than anyone else contributed ideas at every stage of the H-bomb program, and this fact should never be obscured. However, as an article in *Life* of September 6, 1954, clearly portrays: Nine out of ten of Teller's ideas are useless. He needs men with more judgment, even if they be less gifted, to select the tenth idea which often is a stroke of genius.

It has been loosely said that the people at Los Alamos couldn't "get along" with Teller and it might be worthwhile to clarify this point. Both during the difficulties of the wartime period and again in 1951, Teller was on excellent terms with the vast majority of the scientists at Los Alamos with whom he came in contact in the course of the technical work. On both occasions, however, friction arose between him and some of those responsible for the organization and operation of the Laboratory. In each case, Teller, who was essentially alone in his opinion, was convinced that things were hopelessly bad and that nothing would go right unless things were arranged quite differently. In each case, the Laboratory accomplished its mission with distinction. In September 1951, when the program for a specific test of the new concept was being planned, Teller was strongly urged to take the responsibility for directing the theoretical work on the design of Mike. But he felt sure the test date should be a few months earlier; he didn't like some of the people with whom he would have to work; he was convinced they weren't up to the job; the Laboratory was not organized properly and didn't have the right people. Teller decided to leave and left. The Mike shot went off exactly on schedule and was a full success.

It took much more than the idea of the new concept to design Mike. Major difficulties occurred in the theoretical design in early 1952, which happened to be a period when I was again at Los Alamos. They were all solved by the splendid group of scientists at Los Alamos.

At this time more than one-half of all the development work of the Los Alamos Laboratory went into thermonuclear weapons and into the preparation of the Mike test in particular. All but a small percentage of the theoretical division were thinking about this subject. In addition, there was a group of theorists working in Princeton under the direction of Professor John A. Wheeler in collaboration with the theoretical group at Los Alamos. Shepley and Blair, however, have to say of this period (on page 141) "Progress on the thermonuclear program still lagged."

Teller "helped" at this time by intensive agitation against Los Alamos and for a second laboratory. This agitation was very disturbing to the few leading scientists at Los Alamos who knew about it. Much precious time was spent in trying to counteract Teller's agitation by bringing the true picture to Washington. I myself wrote a history of the thermonuclear development to Chairman Dean of the AEC which was mentioned in the Oppenheimer testimony. This loss of time could be ill afforded at a time when the technical preparations for Mike were in a crisis.

Nevertheless, the theoretical design of Mike was completed by June 1952 in good time to make the device ready for test on November 1. Not only this, but, in the same period, much work was done leading to the conceptual design of the devices which were later tested in the Castle series in the Spring of 1954. The approximate date for the Castle tests was also set at that time, and it was planned then that it should lead to a deliverable H-bomb if the experimental Mike shot was successful. It is necessary always to plan approx-

imately two years ahead. Between Summer 1952 and Spring 1954, theoretical calculations on the proposed thermonuclear weapons proceeded; they were followed and in some cases paralleled by mechanical design of the actual device and finally followed by manufacture of the "hardware."

In July 1952, the new laboratory at Livermore was officially established by the AEC. Its existence did not, and in fact could not, accelerate the Los Alamos work because in all essentials the work for Castle had been planned before Livermore was established. In August 1952 an additional device was conceived at Los Alamos which might possibly have been slightly influenced by ideas then beginning to be considered at Livermore. In addition, Los Alamos decided to make a few experimental small-scale shots in Nevada in the Spring of 1953, and this program may have been slightly stimulated by the existence of Livermore. Livermore did assist in the observation of the performance of some of the devices tested at Castle.

Concerning the performance of Livermore's own designs, I will only quote the statement of Dr. Bradbury to the press which says, "Every successful thermonuclear weapon tested so far" [1954] "has been developed by the Los Alamos Laboratory."* This statement has not been contradicted.

(Note added in 1982: In the intervening 28 years, Livermore has contributed greatly to nuclear weapons development. Some weapons programs are assigned to Livermore, some to Los Alamos, and the talents of the two laboratories complement each other.)

4. Requisites for the Thermonuclear Program

The requirements for a successful thermonuclear program were four. First, there had to be an idea; second, there had to be many competent people who could work together in a team and could carry out this idea; third, there had to be well-developed, highly efficient fission bombs; fourth, there were needed high-speed computing machines.

The development of the idea has been dealt with in the last section. As far as people were concerned, Dr. Bradbury showed in his press conference that during 1950 the number of scientists in the theoretical division increased from 22 to 35. This is in striking contrast to the statement of Shepley and Blair (footnote on page 104), "The roster of theoreticians at the weapons laboratory actually declined during 1950, the year of President Truman's decision to build a hydrogen bomb." In the meantime [1954], this number has increased to over 50. That all this was possible was due to the extensive training program of graduate students in physics at our universities in the years following the war.

The third requirement, an excellent fission bomb, is perhaps the most important of all. It is well known that a fission bomb is needed to create the high temperatures necessary to ignite an H-bomb. Since in such a process there is an obvious need to adapt the fission bomb to the particular requirements of the situation, much more detailed understanding of the fission explosion process is required and much more flexibility in the design of the fission weapon itself than was needed to develop the first fission weapon. Not until 1950 or 1951 did we begin to have the sort of capability required for this important prerequisite to a real attack on the thermonuclear problem.

The obligation of Los Alamos and the AEC after the war was in the first place to develop advanced models of the fission bomb. I have tried to show in Section 2 that this was done with competence and speed. But even if our side aim had been to develop the H-bomb, we would probably not have proceeded along a very different path than we did. As far as experimental and hardware development was concerned, the fission bomb simply had to come first. It is therefore clear that the fission bomb requirement did not permit successful development of an H-bomb substantially earlier than we actually got it, even if Teller's new concept had been available much earlier. There simply are no three lost years from 1946 to 1949.

There was a great deal of theoretical exploration during those three years, as discussed in Section 3. One might have wished that still more theoretical work had been done, but this would have required more manpower, which perhaps was the scarcest item in the early postwar years. But even supposing the manpower had been available, the work would undoubtedly have been concentrated on Method A which proved futile. As far as one can imagine such a hypothetical history, we might then have known by the Fall of 1949 that Method A would most likely not work. Even had we reached that stage at that time there is no discernible argument to indicate that Method D would consequently have been uncovered earlier than it was. Of course, it might have been, since in principle there was nothing to prevent one from conceiving of this approach. But even if it had been invented somewhat earlier, the time from invention to realization would necessarily have been considerably longer than it was, the way things actually happened. The size of the Los Alamos Laboratory, the experience of its staff, and the sophistication of their control over fission bomb design were all enormously greater in 1951 than they had been a couple of years before. In addition, there is the matter of the revolutionary change in computing facilities and techniques between 1947 and the present time [1954], which was just beginning to take real effect about the beginning of 1951.

Immediately after the war at many places in the United States work was started to design and build high-speed computing machines. This work was pursued with great vigor and enthusiasm. The first machine of the modern type which was used in connection with the weapons program was the ENIAC, and from early in 1948 persons at Los Alamos had made considerable demands on this machine. It was, however, of very limited capacity by modern

**This shows that the GAC were right when they said in 1951 that the facilities of Los Alamos were quite adequate for both H-bomb and A-bomb development (SB page 121). SB reproached them for this because in 1949 they had said that H-bomb development would interfere with A-bomb program. However, the staff of the Los Alamos Theoretical Division had doubled between 1949 and 1951, much A-bomb progress had been achieved, and the new concept, as well as the advent of fast calculating machines, had made H-bomb development far easier than could be anticipated in 1949.*

standards. The IBM Company's SSEC in New York began to operate sometime in 1948 and although it had a very large capacity, it was very slow by modern standards. Against this situation one must judge the statement by SB, page 61, "Lawrence received assurance from Teller that Los Alamos and Princeton would begin the machine calculations immediately." No fast computing machine existed either at Los Alamos or Princeton at the time, and the two machines existing elsewhere were not adequate for the calculations which were to be performed.

The first major improvement in this situation occurred during 1951 when the SEAC began to operate at the Bureau of Standards in Washington. Not long after this machine was running, a large fraction of its time was taken over for calculations required in the thermonuclear program. Later in 1951 large blocks of time were taken over on various models of the UNIVAC. Early in 1952 the MANIAC at Los Alamos came into operation and was immediately put to work on the thermonuclear program. This machine had been built with thermonuclear calculations specifically in mind. In the program leading up to Mike and later to Castle, the resources of the new machines were taxed to the limit. This was true in spite of the fact that these machines could accomplish in days calculations which would have required weeks to handle on the ENIAC and months to handle with the means available at Los Alamos in 1947.

5. Was the H-Bomb Necessary?

Until now I have tried to give a factual history of the development of fission and H-bombs. The vast majority of the scientists connected with this development will agree with me on this history. What I have to say now is entirely my own responsibility, and my views may not be shared by many of my colleagues.

It seems to be taken as an axiom nowadays [1954] that the H-bomb simply *had* to be developed. Shepley and Blair, as well as the much more balanced accounts in *Life* (September 6, 1954) and in *Newsweek* (August 2, 1954) and even the dispassionate opinion rendered by the Gray Board [the Personnel Security Board convened in 1954 to deliberate on the charges against Oppenheimer], seem to take it for granted that a decision in favor of a full-scale H-bomb program was the only one possible in 1949. They seem to feel that a delay of even a few months would have endangered this country. Finally, SB say on page 228 that Oppenheimer's "tragically and frightfully wrong" recommendations of 1949 were "not criminal . . . only fatal." They imply, here and throughout their article, that we would be virtually defenseless, and therefore subject to any amount of Russian diplomatic pressure, if we had not developed the H-bomb and the Russians alone had done so. I do not agree with any of these axioms.

Let us first assume the worst case, namely that the Russians are where they are now, while we have no thermonuclear weapons at all, but only our fission weapons. In assessing this possibility, I shall use again the figures given by SB, whose accuracy I can again neither confirm nor deny.

According to them (page 230) the Russian bomb was one megaton, whereas we could "any time in the year 1954 . . . put 1,000 atomic bombs of 500 kilotons' force on Soviet targets." Five hundred kilotons is half a megaton, and this 500 kiloton bomb is, of course, the one which President Eisenhower mentioned in his speech to the United Nations in December 1953. Since the Russian H-bomb is a new development, it is not likely that they have many of them at present.

Even if the situation were as unfavorable as I have just pictured, it seems to me that we would still be in quite a good position. The "wrong decision" would have been by no means fatal.

It might be objected here that I am arguing by hindsight, that in 1949 we could not know whether the Russian bomb might not come much earlier or much bigger. But so are the partisans of Teller arguing by hindsight when they say that our H-bomb development was after all successful, contrary to what might reasonably have been expected in 1950.

Moreover, I think that in fact the shortest possible time scale of the H-bomb development, in Russia as well as here, was predictable, much more so than whether ultimate success would be achieved. Since good fission bombs have to come first, the Russians, just as we, could hardly have had their H-bomb much earlier than they did.

It is often held against reassuring predictions that General Groves and Dr. Bush predicted in 1945 that the Russians would need 15 or 20 years to build an atomic bomb. But this prediction was at the time strongly opposed by the majority of scientists. For instance, in the book *One World or None,* published in 1945, Professor F. Seitz and myself reasoned that it would take a determined nation about 5 years to build an A-bomb. None of us then knew about Fuchs' betrayal, which certainly helped the Russian effort.

In spite of all this, the possibility that the Russians might obtain an H-bomb was of course the most compelling argument for proceeding with our thermonuclear program. It was, in my opinion, the *only* valid argument. It is interesting in this connection to speculate whether the Russians were indeed already engaged in a thermonuclear program by 1949. Mr. Strauss has stated in a speech that the Soviet H-bomb test, coming as early as August 1953, indicated that they had started work on the thermonuclear bomb much in advance of the United States (SB page 156). I believe that the opposite conclusion is equally justified.

We have seen that even in the worst case, i.e., if the Russians had developed their H-bomb and we had not, our present situation would not be untenable. The best case on the other hand would have been if neither country had developed such a weapon, and if thereby the mortal peril in which the whole world now finds itself had been avoided. When I started participating in the thermonuclear work in Summer 1950, I was hoping to prove that thermonuclear weapons could not be made. If this could have been proved convincingly, this would of course have applied to both the Russians and ourselves and would have given greater security to both sides than we now can ever achieve. It was possible to entertain such a hope until the Spring of 1951 when it suddenly became clear that it was no longer tenable.

The GAC's minority plan of 1949 in which they proposed that we should try to reach an agreement between Russia and the United States so that neither side would proceed with the H-bomb development still does not seem to me utopian. This I will discuss later on.

After the worst and the best case, let us consider our actual situation at present [1954]. The balance of power is now much more in our favor than it would have been under the assumptions of the worst case. Clearly this is to be welcomed. However, it must always be kept in mind that the advantage we now enjoy through the greater power* of our H-bombs may not last. I will not venture a prediction of the time it will take for the Russians to catch up with us again.

While we have a temporary advantage in the armaments race, we now have the H-bomb with us for all time. In the words of SB, page 228, "it is inescapable that two atomic colossi are doomed for the time being 'to eye each other malevolently across a trembling world.' " We can now only rely on the sanity of the governments concerned to prevent an H-bomb holocaust.

In the course of time, the present conflict between Communism and Democracy, between East and West, is likely to pass just as the religious wars of the 16th and 17th century have passed. We can only hope that it will pass without an H-bomb war. But whichever way it goes, the H-bomb will remain with us and remain a perpetual danger to mankind. Some day, some desperate dictator like Hitler may have the bomb and use it regardless of consequences.

The U.S. atomic scientists foresaw in 1949 "The horror of this monstrous balance of potential annihilation", as SB themselves say at the end of their book (page 231). To anyone with such knowledge and with any imagination, the decision to start full-scale development of an H-bomb was a tremendous step to take, and one that must not be taken lightly. This was a decision for which the scientists, inside and outside the GAC, could not take the responsibility on themselves. It was also too big a responsibility for the AEC. One of the arguments of the GAC and of the majority of the AEC was that the decision had to be made at higher governmental levels. Furthermore, they felt it their duty to tell the President and his close advisors of the implications of this step, which they saw so clearly, while members of the government, not so familiar with the potential power of an H-bomb, could not visualize these consequences to the same extent.

I never could understand how anyone could feel any enthusiasm for going ahead. I could well understand that President Truman and his close advisors were forced to a positive decision by the potential threat of a Russian H-bomb development. But I am sure they came to this decision with a heavy heart, and that most of the scientists who went to work on this project also had heavy hearts. I certainly had the greatest misgivings when Teller first approached me in October 1949 to return to Los Alamos full-time to work on this project.

Yet there seemed to be some scientists who apparently had no scruples on this account. If we can believe SB, pages 88 and 89, or even the testimony of Alvarez in the Oppenheimer case, Lawrence, Alvarez, and others associated with them had only one concern, namely how to overcome the technical obstacles. This unquestioning enthusiasm for the thermonuclear program looks to me very much like the enthusiasm that many Germans felt in 1917 when the German Government declared unrestricted submarine warfare. This gave the Germans a temporary advantage in the war but later on was the main cause which brought the U.S. into the war against Germany and thus caused the German defeat.

To most of us the important question seemed not how to build an H-bomb, but whether one should be built. The conference which was to be called at Los Alamos for November 7, 1949 (SB page 68), was to discuss this problem at length as much as the technical problem. Nearly every scientist felt the way Oppenheimer did in his letter to Conant (SB page 70): "It would be folly to oppose the exploration of this weapon. We have already known it had to be done; and it does have to be done . . . But that we become committed to it as the way to save the country and the peace appears to me full of dangers." It is remarkable, by the way, that this letter could be quoted by anybody as evidence against Oppenheimer; it seems to me an excellent letter which is clear proof that Oppenheimer was only against a crash program, not against *exploration* of thermonuclear problems.

The GAC report concluded: "We all hope that by one means or another, the development of these weapons can be avoided. We are all reluctant to see the United States take the initiative in precipitating this development. We are all agreed that it would be wrong at the present moment to commit ourselves to all-out efforts towards its development." The report of the GAC might well be considered as a prayer for some solution to the dilemma, not as an answer. Scientists are not especially qualified to find a solution in the domain of statecraft. All they could do was to point out that here was a very major decision and it was worth every effort to avoid an irrevocable, and perhaps fatal, step. (An intermediate step which would have left time for careful consideration of the problem by the government and yet not have wasted time in the technical development, might have been to direct intensified theoretical work on the H-bomb at Los Alamos, but not to take any immediate steps toward any major "hardware" development.)

Although the GAC were seeking a solution rather than offering one, the proposal of its minority still seems worthwhile, even as seen from today's [1954] viewpoint. The proposal was to enter negotiations with Russia with the aim that both countries undertake an obligation not to develop the H-bomb. If such an agreement could have been reached and had been kept, it would have gone far to avoid the peril in which the world now stands. At that time neither we nor the Russians presumably knew whether an H-bomb could be made. In this blissful state of ignorance we might have remained for a long time to come. Since the technical program was a very difficult one, it could never be accomplished without a major effort. It is possible, perhaps likely, that the Russians would have refused to enter an agreement on this matter. If they had done so, this refusal would have been a great propaganda asset for us in the international field

**According to SB page 161, the largest of our test shots reached a force of 15 megatons, compared to the Russians' 1 megaton. As in the earlier cases, I cannot comment on the accuracy of the figures.*

and would in addition have gone far to persuade the scientists of this country to cooperate in the H-bomb program with enthusiasm.

Many people will argue that the Russians might have accepted such an agreement, but then broken it. I do not believe so. Thermonuclear weapons are so complicated that nobody will be confident that he has the correct solution before he has tested such a device. But it is well known that any test of a bomb of such high yield is immediately detected. Therefore, without any inspection, each side would know immediately if the other side had broken the agreement.

It is difficult to tell whether or not the Russians would have developed the H-bomb independently of us. I am not sure what would have happened if we had followed the recommendations of the GAC majority and had merely announced that for such and such reasons, we would refrain from developing the H-bomb. Once we announced that we would go ahead, the Russians clearly had no choice but to do the same. In the field of atomic weapons, we have called the tune since the end of the war, both in quality and in quantity. Russia has to follow the tune or be a second-class power.

In summary I still believe that the development of the H-bomb is a calamity. I still believe that it was necessary to make a pause before the decision and to consider this irrevocable step most carefully. I still believe that the possibility of an agreement with Russia not to develop the bomb should have been explored. But once the decision was made to go ahead with the program, and once there was a sound technical program, I cooperated with it to the best of my ability. I did and still do this because it seems to me that once one is engaged in a race, one clearly must endeavor to win it. But one can try to forestall the race itself.

This article, written in 1954, has now been declassified. In publishing it now, I wish to add a few remarks specifically correcting some of the mistakes in Peter Goodchild's book *J. Robert Oppenheimer: Shatterer of Worlds.*

The most important point concerns the meeting of the GAC in Princeton on June 16, 1951. The Goodchild book (page 210) states that "Teller was not included among those due to speak". This is incorrect. The whole meeting was held in order to discuss Teller's new concept for the design of an H-bomb. For this reason only, a number of scientists concerned with this concept were invited, namely Bradbury, Froman, and Mark representing Los Alamos and five more independent scientists, Teller, myself, Nordheim, von Neumann, and Wheeler. The most important part of the meeting was to be the presentation of Teller's new idea. Teller himself gave the main presentation, followed by me and the three others. I totally endorsed Teller's new idea. It was after this presentation that Oppenheimer warmly supported this new approach. So did Gordon Dean, the Chairman of the AEC.

Then, the meeting discussed the implementation of Teller's idea by the Los Alamos Laboratory. In this connection, the people directly involved with the Laboratory (Bradbury, Froman, and Mark), already well acquainted with Teller's ideas, presented their plans. As I remember it, Teller got impatient with these plans, and it was only then that he "could contain" (himself) "no longer" and "insisted on being heard" (page 210). He thought that the Los Alamos people were planning too slow a development, and he insisted on accelerating it. As it turned out, Los Alamos completed the development up to the Mike test in a mere 18 months.

The Goodchild book also gives the impression that Gordon Dean was unfavorable to Teller generally. This was by no means the case. Mr. Dean took me aside privately and asked how the breach between Teller and Oppenheimer could be healed. He wanted very much to have Teller's cooperation in weapons development.

Goodchild also quotes (page 214) a testimony of Teller to the FBI that I "had been sent by Oppenheimer to Los Alamos to see whether the H-Bomb was really feasible after all." (This refers to my visits to Los Alamos before Teller's invention, i.e., in 1950 and January 1951.) Nobody ever sent me to Los Alamos. I was a regular consultant to the Laboratory, and I was strongly urged by members of the Laboratory, particularly Bradbury and Mark, to come again after Truman's decision to develop the H-bomb. It is true that I would have much preferred the H-bomb to turn out impossible, and that I was happy at the calculation by Ulam in the early Summer of 1950 which made it appear that the H-bomb of the original design might not be feasible. But I had made up my mind myself with not the slightest influence by Oppenheimer.

The Goodchild book also repeats the statement that the Russians exploded an H-bomb in August 1953 (page 219). This was not a true H-bomb, as I know very well because I was the chairman of the committee analyzing the Russian results. This Russian test is well discussed in the book *The Advisors* by Herbert York. The first true H-bomb exploded by the Russians was in late 1955, three years after our Mike test.

The claim that the August 1953 test was a true and deliverable H-bomb was strongly maintained by Lewis Strauss to justify his contention that the United States' development of the H-bomb had been necessary and urgent. As far as I can tell, the Russians made the 1953 test essentially just to show that they could also develop such a device. But once more, it was not the real thing.

Still another claim (p. 209) is that the Russians in late 1950 tested some kind of thermonuclear device. This claim is a pure fabrication. Herbert York investigated the history of the Russian tests very carefully and concluded that there was no such test. ■

Reflections of the Polish Masters

I am what is called a mathematical physicist. I take this to mean the utilization of—and sometimes the attendant construction of—mathematics in a context posed by physical reality. Now I suppose that statement would fail to distinguish mathematical physics from mathematics or from physics; after all, numbers and geometry, the stuff at the core of all mathematics, have been abstracted from the context of the physical world. And physics is the *hard science, of necessity drawing sharp conclusions only from its mathematical language. Newton had to invent the calculus to extend Galileo's algebraic kinematics to a general framework, and yet Newton is always viewed as a physicist. Evidently the boundary between these disciplines is ultimately blurred, although at a given time in development, the attitudes of the various practitioners can be distinct.*

Having exposed my view of no hard distinction between physics and mathematics, I have also exposed a main thrust of the nature of the discussion I had in mind in the following interview. That is, I wanted to explore the (personal) "philosophical" views of just what connections are in the back of theorists' minds that drive the work they perform. It is hard, in understatement, to know a creator's internal vantage point from the technical products in print.

Los Alamos is fortunate in the presence—either on a temporary or permanent basis—of a number of great individuals. I count as one of my fortunes that being here has allowed my coming to know Mark Kac and Stan Ulam. A mutual interest in discussing these matters has, of course, allowed the possibility of this interview. Moreover, these gentlemen embody a tradition of technical education and a viewpoint toward science that, in starting some fifty years ago in a "different" world, are in ways at variance with the more "modern" tradition. Above all, I wanted to explore just what these differences might entail.

As a brief background—both will provide more detail themselves—Kac and Ulam are both internationally known and successful mathematicians. And as shall be evident from the interview, both also have a strong enthusiasm in science. Kac has been a pioneer in the development of mathematical probability as well as in its applications (largely to statistical physics). In particular, the modern method of quantization proceeds through a device often called the Feynman-Kac path integral. Similarly, Ulam has made diverse contributions to the various twentieth century branches of mathematics while simultaneously involving himself in a range of theoretical and technological scientific applications. In particular, his name has been associated with the development of the Monte Carlo method of numerical simulation.

A technically oriented reader will find himself disappointed if he expects to hear in any detail of the work they are known for. Rather, what is offered are the reflections of these men, toward the latter parts of their careers, on how they have seen education, mathematics, and science evolve in spirit over the course of their professional lives. Also, their attitudes toward the content and range of their subject will be viewed. It is a regrettable consequence of the medium of the written word that the rich inflection of voice and gesticulation of hand that so often color and amplify the words of these men are not available to the reader. Nonetheless, I hope some of their characteristic charm and humor is conveyed.

An Interview with Stan Ulam and Mark Kac

by Mitchell Feigenbaum

FEIGENBAUM: Would each of you give a brief biographical sketch? Stan, would you like to start?

ULAM: My name is Ulam, Stan Ulam. Stanislaw is the real first name. I was born in Poland. I received my doctorate in mathematics from the Polytechnic Institute in Lwów, ages ago. During the early thirties I visited some foreign centers of mathematics. In 1935 I received an invitation to come to Princeton for a few months, to the Institute for Advanced Study. I was not clever enough to see what was coming, really. Stupidity made me not even make such plans; but then I received an invitation from this famous, very world-famous mathematician, one of the great mathematicians of the century, John von Neumann, who was actually only about six or seven years older than I; and so I decided to visit the United States for three months. Of course, there were no planes. I had to go to some port in France to catch a boat to New York. I spent a few weeks in Princeton, and one day at a von Neumann tea, G. D. Birkhoff, who was the dean of American mathematics, was present. He knew a little about my work, apparently from his son, who was about my age, and he asked me when I would come to Harvard. Then I went back to Poland. But the next fall I returned to Cambridge as a member of the so-called Society of Fellows, a new Harvard institution. I was only twenty-six or so. I started teaching right away: first, elementary courses and then quite advanced courses. And then I became a lecturer at Harvard in 1940. But every year during that time I commuted between Poland and the United States. In the summers I visited my family and friends and mathematicians. In Poland the mathematical life was very intense. The mathematicians saw each other often in cafés such as the Scottish Café and the Roma Café. We sat there for hours and did mathematics. During the summers I did this again. And then in '39, I actually left Poland about a month before World War II started. It was very lucky in a sense. My mother had died the year before the war, and my brother, thirteen years younger, was more or less alone. My father, a lawyer, was busy; he thought it would be good for my brother to come to the United States, too, to study at the university. My brother was seventeen at the time and he came with me in 1939. I enrolled him at Brown University in Providence, which was not too far from Cambridge.

Then in 1940 I became an assistant professor at the University of Wisconsin in Madison. While there—it was in the spring or summer of 1943—I received an inquiry from John von Neumann whether I

Kac: I was, as Michael Cohen, one of our mutual friends, says, independently poor. And it did cost a little to visit in the Café.

INTERVIEW

would be interested in doing some very important war work in a place which he couldn't name, and I was to meet him in Chicago in some railroad station to learn a little bit more about it. I went there; and he couldn't tell me where he was going; and there were two guys, sort of guards, looking like gorillas, with him. He discussed with me some mathematics, some interesting physics, and the importance of this work. And that was Los Alamos at the very start. A few months later I came with my wife, but that is another story. I could talk for hours about the impressions of the trip, of arriving for the first time in a very strange place. But that is already in some books, including my own autobiography. What else would you like to know?

FEIGENBAUM: Why don't you quickly say something about your work?

ULAM: I have been publishing mathematics papers since I was eighteen. Though not very common, neither was it too unusual, because very often mathematicians start very early. I got my Ph.D., as I told you, in Poland. And in this country I published papers as a lecturer at Harvard and at Wisconsin, but the work here in Los Alamos was mainly physics, of course. I had always had some interest in physics, and I had read a lot of relativity, quantum theory, etc. It had been a platonic interest in the sense that most of my early papers were in pure mathematics.

FEIGENBAUM: Mark, would you now say something, as you put it, as Stan's younger colleague?

KAC: I was also born in Poland, although it was not clear that it was Poland. Because, in fact, where I was born, it was czarist Russia, and where Stan was born, it was Austria. In addition to other uncertainties connected with my birth is that my date of birth is not entirely right either, because under the czars they used the Julian calendar. So my birth certificate says I was born on August 3rd, and I maintain this fiction, but in reality I was born on the 16th. I was born 170 kilometers—that is 100 miles—almost directly east of where Stan was born. Nevertheless, within those 100 miles were two completely different worlds, because Poland had not existed as an independent country for 150 years. It was partitioned among Austria, Germany, and Russia, and the cultures of the occupying powers had made an enormous imprint. In my part of the world, nobody spoke Polish; my mother never learned to speak Polish. Anyway, I was born. After an evacuation in 1915 somewhat deeper into Russia, we returned to Poland in 1921, and then I went for my first formal schooling in Polish. Polish was actually the fourth language I learned. I first spoke Russian, because that was the language that everyone spoke; then, when we came back home after the evacuation, my parents engaged for me a French governess, a French lady who was a widow of a White Russian officer. For three years she came for half a day, and we'd conjugate French verbs, and I hated it. Then my father was briefly a principal of a lay Hebrew school. It was not a religious school, but all the subjects were taught in Hebrew, so I learned Hebrew, which I promptly forgot. Then, finally in 1925, at the age of eleven, I entered a Polish school, a very well-known Polish school, the Lycée of Krzemieniec. The town where I was born had a certain part in Polish history, one of the reasons being that one of the two great Polish romantic poets, Juliusz Slowacki, was born there (almost every Polish child would know the name). In addition, another very famous citizen of that town is Isaac Stern, whose parents were wise to take him out of Poland when he was only nine months old. After secondary school education I went to the university in the same town where Stan was born and where he studied, except he was in the Engineering school, which had, remarkably enough, a division that was devoted to pure science, that is to say, mathematics and physics.

I went to the regular university and I was, and still am, five years younger. At that time Stan was already a legend—and to me looked infinitely old. He was only twenty-two and I was seventeen. I met him for the first time, briefly, and it will be a fiftieth anniversary of that event next year, when he was awarded his doctorate in 1933. (Actually, I thought it was this year, but he corrected me, and he ought to know better when he got his doctorate.) I graduated, got my doctorate, in 1937, and unlike Stan I wanted to get out of Poland very badly. I did not know the disaster was going to be of the magnitude it turned out to be, but it was obvious that Europe, especially eastern Europe, was not the place to stay. But it was not very easy to get out in those days.

Ulam: In Poland the mathematical life was very intense. The mathematicians saw each other often in cafés . . . We sat there for hours and did mathematics.

INTERVIEW

Now, two episodes I have recalled because Mitchell and I have been tracing back the autobiographical part. In 1936, maybe '37, just before the time I got my doctorate, I was trying desperately to get out of Poland, and I would read *Nature,* because in *Nature* there would be ads of various positions. Most positions required being a British subject, but one of them (at that time, by the way, I knew not a word of English) was an ad for a junior lecturer in the Imperial College of Science and Technology at the salary of 150 pounds per annum, which in those days was about 750 dollars. Even then that was not very much money, and I thought that no self-respecting British subject would ever want to apply for a job like this. So I spoke to my teacher, Hugo Steinhaus, and asked whether it would be a good idea to apply, and he, partly in jest, partly seriously, said, "Well, let's estimate your chances of getting the job. I would say it is 1 in 5000. Let's multiply this by the annual salary. If this comes out to be more than the cost of the postage stamp, then you should not apply. If it is less than the cost of the stamp, you should." Well, it turned out to be a little bit less than the cost of the stamp, so I wrote. I got a letter from them later on saying that unfortunately the job was filled, so there had been after all a British subject who wanted the 150 pounds per annum. Many, many years later when I was in England, I was invited to give a lecture at the Imperial College of Science and Technology, and I said to them, "You know, you could have had me for 150 pounds per annum." I believe that they actually looked up and found the correspondence. This anecdote reminds me that, when I finally decided to come to the United States, it was very difficult to get visas, because already the German refugees were coming. It was a terrible time, and I managed to get only a visitor's visa for a six-month period. The Consul made me buy a round-trip ticket just to make sure that I would return. The return portion of the ticket I still have, and it was for a boat that was sunk in the early days of the second world war. A memento.

It was Hugo Steinhaus, my teacher and my friend, a very well-known Polish mathematician, who tried very hard to help me get out. And finally he succeeded in a very simple way by helping me get a small fellowship to go abroad to Johns Hopkins University. It is curious how small things change one's life, and in effect possibly save one's life. I applied for that scholarship in 1937, immediately after getting my doctorate and did not get it. I thought it was a tremendous injustice, but I got it a year later; that saved my life because if I had gotten it a year earlier, I would have been compelled to go back. This way the war caught me in this country and literally saved my life. I was at Johns Hopkins when the war started, and then I got an offer to Cornell, where I spent twenty-two very happy years. (Mitchell is going to be my successor there.) In fact, my whole family, that is, my acquired family in the United States, my wife and both my children, are native Ithacans. And I have actually lived in Ithaca longer than in any other place in the world.

ULAM: So it is the converse of Odysseus.

KAC: When I left Cornell I was forced to make a very brief speech, and I said, "Like Ulysses I, too, am leaving Ithaca, the only difference being I'm taking Penelope with me." That was how it was. I was then for twenty years at Rocky U, Rockefeller University, in New York City and then decided to spend my declining years, as it were, where there is more sun and less ice. So I am now at the University of Southern California, a little bit west of here.

FEIGENBAUM: I guess it's time to interrupt you from these reminiscences. Stan, perhaps you can say something about how you became interested in mathematics?

ULAM: As a young boy at the age of ten, I was very interested in astronomy and then in physics. I was reading popular books on astronomy; there weren't as many, and they were not as beautiful ones as now with incredible illustrations, but still, that was my passion. An uncle gave me a little telescope for my birthday when I was eleven or twelve. By then I was trying to understand the special theory of relativity of Einstein, and I think I had a pretty good qualitative idea of what it was all about. Then, later, I noticed that I needed to know some mathematics, so I went beyond what was given in the high school, gymnasium, as it was called. Students started gymnasium at age ten and went to age eighteen. When I was fourteen, I decided to learn more mathematics by myself, and I was sixteen when I really learned calculus all by myself from a book by Kowalevski, a German not to be confused with Sonia Kowaleska, a famous nineteenth century Russian woman mathematician. Then I read also about set theory in a book by Sierpinski, and I think I understood that. We had a good professor in high school, Zawirski, who was a lecturer in logic at the university. I talked to him about it then and when I entered the Polytechnic Institute.

FEIGENBAUM: He was teaching at the high school?

ULAM: Yes, he was teaching in high school to make money, because lecturers earned hardly any money at the university. When I entered the university, I attended a course by Kuratowski, a freshman professor who had just come from Warsaw. He was only thirty-one years old; I was eighteen. He gave an elementary course on set theory, and I asked some questions; then I talked to him after classes, and he became interested in a young student who evidently was interested in mathematics and had some ideas. I was lucky to solve an unsolved problem that he proposed.

FEIGENBAUM: Stan, did you feel at that point that your interests were changing from astronomy and physics and relativity toward mathematics?

ULAM: No, in fact, even now I don't think the interests have changed. I am interested in all three. Of course, I did much more work in pure mathematics than in applications or in theoretical physics, but my main interests remain. I have to make a confession: nowadays I don't read many technical mathematical jour-

Ulam: When somebody mentions the word pressure to me, I sort of see something, some kind of confined hot or turbulent material.
Kac: I cringe.

INTERVIEW

nals—rather, I read what is going on in astronomy and astrophysics or in technical physics in *Astrophysics Journal* and *Physics Today*. It always seems to me much more understandable. You know, this specialization in each science, especially in mathematics, has proceeded much apace the last few years. Mathematics is now terribly specialized, more so than, say, physics. In physics there are more clearly defined central problems than in mathematics itself. Of course, mathematics still has many important problems, fundamental ones.

FEIGENBAUM: You feel that this specialization is unfortunate?

ULAM: Oh, yes. Both of us have very similar views, it turns out, about science in general and about mathematics and physics in particular.

FEIGENBAUM: Mark, how did you begin in mathematics?

KAC: Stan and I are running in parallel. Actually my interest in mathematics also began very young, and probably I romanticize a little. (I was saying to Mitch that if you try to think of something that happened sixty years ago, it is not always infinitely reliable.) My father had a degree in philosophy from the University of Leipzig in Germany and knew mathematics. He also later got a degree from Moscow in history and philology, so he knew, among other things, all the ancient languages. Anyway, he earned a living during the war by giving private tutorials in a little one-room apartment, and among other things he tutored in elementary geometry. I heard all these incredible things: from a point outside a straight line you can drop a perpendicular and draw one and only one parallel, and such and such angles are equal. I was four years old, five maybe, and all these wonderful, ununderstandable sounds, in what seemed like ordinary language, impressed me. I would absolutely pester him to try to tell me what it was; in self-defense he began to teach me a little bit of elementary geometry, and somehow the structure, that there is such a fantastic tight structure of deduction, impressed me when I was a very young boy. In fact, at that time my father despaired because at the same time I was exceedingly bad learning multiplication tables. That one could know how to prove theorems of elementary geometry without knowing how much seven times nine was seemed more than slightly strange. That was the beginning of my interest in mathematics, but like Stan the interest in science came almost at the same time, primarily by reading popular books. One book, available in Russian translation, was called a *Short History of Science* and was by an English lady whose name was Arabella Buckley, or something of the sort. It was fascinating! I then later read Faraday's *Natural History of the Candle,* which is one of the great books. In school, when I finally went to the gymnasium, as it was called, I was equally interested and equally good in mathematics and physics, but finally decided on mathematics.

Actually, an event during the summer before my last year at the gymnasium, among other things, influenced my decision. Here's how it was. My mother had envisaged that I would pursue something sensible like engineering, but in the summer of 1930 I became obsessed with the problem of solving cubic equations. Now, I knew the answer, which Cardano had published in 1545, but what I could not find was a derivation that satisfied my need for understanding. When I announced that I was going to write my own derivation, my father offered me a reward of five Polish zlotys (a large sum and no doubt the measure of his skepticism). I spent the days, and some of the nights, of that summer feverishly filling reams of paper with formulas. Never have I worked harder. Well, one morning, there it was—Cardano's formula on the page. My father paid up without a word, and that fall my mathematics teacher submitted the manuscript to *Mlody Matematyk (The Young Mathematician).* Nothing was heard for months, but as it turned out, the delay was caused by a complete search of the literature to ascertain whether I had not in fact "rediscovered" a derivation. They found that my derivation was, after all, original, and so it was published. When my gymnasium principal, Mr. Rusiecki, heard that I was to study engineering, he said, "No, you must study mathematics; you have clearly a gift for it." So you see. I had very good advice.

At the university I actually thought of possibly starting physics, but physics in Lwów was very poor, theoretical physics especially. Mathematics was extremely good and very lively, so it was very easy to get involved in a tremendously exciting and energetically developing subject rather than struggle with a subject in which there was not really much activity. I took, naturally, courses in the physics department and took some exams in theoretical physics, but my interest, real interest, in physics was kindled considerably later.

FEIGENBAUM: I have the impression that somehow science and mathematics have similarly cross-fertilized in your minds and that you have—I think you have conveyed this feeling—some kind of intuition that is very important toward the way that you view mathematics.

KAC: Yes, this may be of interest to modern readers, and I am sure that Stan will confirm what I say. We belong to an academic generation that was only a little bit removed from the heroic times in the great centers of mathematics, Göttingen and Paris. There the distinction between mathematics and physics was not made as jurisdictionally sharp as it is now. The great mathematicians of that era, Poincaré and Hilbert, both made extremely important contributions to physics, Poincaré especially. Our teachers were taught physics and knew it. Banach, for instance, who is primarily known as the creator of the school of functional analysis and who is probably the greatest Polish mathematician of all times, taught mechanics. He wrote a very good textbook on it. The whole distinction of now you are a physicist, so you do this, now you are a mathematician, so you do that, was intellectually blurred. There were, of course, people who were more concrete, and others who were more abstract, and people

INTERVIEW

who were more interested in this or that. But there wasn't any of this kind of professionalism, nor the almost union card distinctions that are prevalent now, so that it was easy, not only because our makeups were conducive to do this, but also because nobody told me that I should not study physics because if I didn't study just mathematics, I'd never catch up. The idea of catching up, of something running away, never existed. Isn't that so?

ULAM: Absolutely. You are talking about a very long time ago, fifty years ago, and you know—some time ago I had this thought—my life, and Mark's too, occupies more than almost two per cent of the recorded history of mankind. You see, fifty or sixty years is that much. That it is a sizeable fraction of the whole history that we know about is a strange and very terrifying thought. Things have changed in many ways, not only in technology but in attitudes.

FEIGENBAUM: Here is a question. When you mention that there is something negative in your minds about specialization and that you have this connection in your minds between physics and mathematics, is there some kind of a special intuition that you think comes from these two things working together? Do you feel that's an important ingredient?

ULAM: You see, it depends very much on the person. Some mathematicians are more interested in the formal structure of things. Actually, for people in general there are two types of memory that are dominant, either visual memory or auditory memory, and seventy-five per cent (this Mendelian fraction) supposedly have visual memory. Anyway, some people have a very purely verbal memory, more toward the logic foundations and manipulation of symbols, rather than toward imagining physical phenomena. When somebody mentions the word pressure to me, I sort of see something, some kind of confined hot or turbulent material.

KAC: I cringe.

ULAM: Right, but other people, von Neumann for example, are more logically minded. To him pressure was, so to say, a term in an equation. I rather suppose that he did not visualize situations where pressure would do this or that, but he was also very, very good in physics. Certainly there are different attitudes in ways of thinking. Some mathematicians are more prone to the physical. Also, we don't really know too much about this. It could be a question of accidents in your childhood and in your youth or of the way you learned things.

FEIGENBAUM: Do you think that this kind of intuition that you have is more special to yourself? I mean by that, if you think back to when you started doing mathematics, were more people then like yourself rather than more formal.

ULAM: No, no. I don't think so. Many mathematicians that I knew at that time were different from Mark Kac and myself in their attitude toward physics. Even now in this country, I would say ninety per cent or more of mathematicians have less interest in physics than we do.

KAC: Partly, of course, it is educational. I think the education in this country has been, especially higher education, singularly bad. For instance, it is perfectly possible for a young man to get a doctorate in mathematics in a reputable school, like Harvard, without ever having heard of Newton's laws of motion.

ULAM: I was on a committee of the American Mathematical Society when I discovered that you could get a Ph.D. at Harvard and other places without knowing Newton's laws of motion, which were actually one of the central motives for the development of calculus, you might say. That is how it is now.

KAC: We were exposed to chemistry, to physics, to biology; there were no electives when you were in secondary school. Secondary schools in Europe, in Poland, in France were in a certain sense harder than the university because you had to learn a prescribed curriculum. There was no nonsense. If you were in a certain type of school, you had to take six years of Latin and four years of Greek and no nonsense about taking soul courses or folk music, or all that. I have nothing against taking such courses, except that it has become a substitute. You had to take physics, you had to learn a certain amount of chemistry, of biology, and if you didn't like it, so it was. But if there was some kind of resonating note in you, then you were introduced to it early. At the university you really specialized, although not entirely; every mathematician had, for example, to pass an exam in physics and even, God help me, go through a physics lab. That was one of my most expensive experiences because, being rather clumsy, I broke more Kundt's tubes than I could afford. Stan made an extremely important point to which I can bring a little extra light. I heard probably one of the last speeches by von Neumann. It was in May 1955. (In October of that year, while I was in Geneva on leave, it was discovered that he had incurable cancer, and he died

then sometime later in 1957.) He was the principal banquet speaker at the meeting, I believe, of the American Physical Society in Washington. I was there, and I went to the meeting, and after the speech we had a drink together. His speech was, "Why Am I Not a Physicist?" or something of the sort. He explained that he had contributed technical things to physics; for example, everybody knows what a density matrix is, and it was von Neumann who invented density matrices, as well as a hundred other things that are now, so to speak, textbook stuff for theoretical physicists. But he, nevertheless, gave a charming and also moving talk about why he was not really a physicist, and one thing he mentioned was that he thought in terms of symbols rather than of objects; I am reminded that his friend Eugene Wigner hit on it correctly by saying that he would gladly give a Ph.D. in physics to anyone who could really teach freshman physics. I know what he meant. I would attempt, I wouldn't be very good at it, but I would attempt to teach a first semester course in quantum mechanics, and I would probably teach it reasonably well. But I would not know how to teach a freshman course in physics, because mathematics is, in fact, a crutch. When you feel unsafe with something, with concepts, you say, "Well now, let's derive it." Correct? Here is the equation, and if you manipulate with it, you finally get it interpreted, and you're there. But if you have to tell it to people who don't know the symbols, you have to think in terms of concepts. That is in fact where the major breach between the two—how to say—the two lines of thought come in. You are either like von Neumann, and I am in that sense closer to him, or you are like Ulam, who when you say pressure, feels it. It is not the partial derivative of the free energy with respect to volume; it is really something you feel with your fingers, so to speak.

FEIGENBAUM: But isn't it nonetheless true that any good mathematician has a very strong conceptual understanding of the things he is working on? He isn't just doing some succession of little proofs.

KAC: Well, the really good ones, yes. But then, you see, there is a gamut, a continuum. In fact, let me put this in because I would like to record it for posterity. I think there are two acts in mathematics. There is the ability to prove and the ability to understand. Now the actions of understanding and of proving are not identical. In fact, it is quite often that you understand something without being able to prove it. Now, of course, the height of happiness is that you understand it and you can prove it. The next stage is that you don't understand it, but you can prove it. That happens over and over again, and mathematics journals are full of such stuff. Then there is the opposite, that is, where you understand it, but you can't prove it. Fortunately, it then may get into a physics journal. Finally comes the ultimate of dismalness, which is in fact the usual situation, when you neither understand it nor can you prove it. The way mathematics is taught now and the way it is practiced emphasize the logical and the formal rather than the intuitive, which goes with understanding. Now I think you would agree with me because, especially with things like geometry, of which Stan's a past master, seeing things—not always leading neatly to a proof, but certainly leading to the understanding—ultimately results in the correct conjecture. And then, of course, the ultimate has to be done also—because of union regulations, you also have to prove it.

ULAM: Let me tell you something. It so happens that I have written an article for a jubilee volume in honor of this gentleman here, Mark Kac, on his whatever anniversary, a volume which has not yet appeared. But the article is about analogy and the ways of thinking and reasoning in mathematics and in some other sciences. So it is sort of an attempt to throw a little light on what he was just talking about. These things are intertwined in a mysterious way, and one of the great hopes, to my mind, of progress, even in mathematics itself, will be more formalizing or at least understanding of the processes that lead both to intuition and to then working out not only the details but also the correct formulations of things. So there is a very, very deep problem and not enough thought has been really given to it, just cursory remarks made.

FEIGENBAUM: Do you have a hope that people will be able to formalize these things, the serious components?

ULAM: It is now premature, but some partial understanding of the functioning of the brain might appear in the next twenty years or even before—some inklings of it, more than is known at present. That is a marvelous prospect. You see, if I were a very young man,

Kac: There are two principles of pedagogy which have to be adhered to. One is, "Tell the truth, nothing but the truth, but not the whole truth."

INTERVIEW

maybe I would be working more in biology or neurology, that is to say the anatomy of the brain, and trying to understand its processes. Mark and I, driving to the Laboratory this morning from Santa Fe, were discussing how children learn to talk and use the phrases they hear—learn to use them correctly in different contexts with changed elements. It is really a mysterious thing.

FEIGENBAUM: Let's pick up on the last thing you said—that maybe there is a chance of understanding how the brain works. When you say that, what comes to my mind is that there are problems that *in principle* you can think of—for example, fully developed turbulence in a fluid and perhaps the brain. It might be that these problems really will rely on an immense number of details, and maybe there won't be any nice theories such as we've known how to write so far, and you really just have to put all these details on a computer. Do you have any thoughts about that and what it implies for the limitations of future mathematical effort?

ULAM: Well, actually, computers are a marvelous tool, and there is no reason to fear them. You might say that initially a mathematician should be afraid of pencil and paper because it is sort of a vulgar tool compared with pure thought. Indeed, say thirty years ago, professional mathematicians were a bit scared, as it were, of computers, but it seems to me that for experimentation and heuristic indications or suggestions, it is a marvelous tool. In fact, the meeting* that is going on right now, to a large extent, is possible because so much has been discovered experimentally.

FEIGENBAUM: That is absolutely true.

ULAM: So in physics, experiments lead finally to problems and to theories. Experimentation in mathematics could be purely mental, of course, and it was largely so over the centuries, but now there is an additional wonderful tool. So in answer to your question about understanding the brain, yes, it seems to me, indeed.

FEIGENBAUM: Certainly one has learned now, or is at the first stage of really learning, how to do experiments on computers that can begin to furnish intuition for problems that otherwise were impenetrable. The new intuition then enables you to write a more analytical theory. Do you think there are problems that are so complex that you won't be able to get that kind of a handle on them? For example, maybe memory in a brain has no global structure, but rather entails nothing more than a million different distinctly stored things, and then you wouldn't write any theory for it but rather only simulate such a system on a computer. Do you think there may be some limitation to what kinds of things you can analyze?

ULAM: It depends on what you call theory. I noticed you said the analytical method; it means that by habit and tradition you think that is the only way to make progress in pure mathematics. Well it isn't. There may be some eventual super effect from the use of computers. I was involved from the beginning in computers and in the first experiments done in Los Alamos. Even in pure number theory there were already tiny little amusements from the first. A time may come, especially because the overspecialization of mathematics is increasing so much that it is impossible now to know more than a small part of it, that there will be a different format of mathematical thinking in addition to the existing one and a different way of thinking about publications. Maybe instead of publishing theorems and listing them there will be a sort of larger outline of whole theories, and individual theorems will be left to computers or to students to work out. It is conceivable.

KAC: Slaves.

ULAM: Mathematics, which hadn't changed much in its formal aspect in the last 2000 years, is now undergoing some change. The great discoveries of this century, Gödel's, are of tremendous philosophical importance to the foundation of mathematics. Gödel proved there are statements that are meaningful but that are not demonstrably true or false in a given system of axioms. Hilbert, of course, was *the* great believer of the formal system for all mathematics. He said, "We will understand everything, but it all depends on what basis." That is no longer so. You see, the axiom systems themselves change as a result of what you learn by physical experimentation or by mental experimentation. I think Mark probably has a different perspective.

KAC: I don't want to step out too far because I am a believer in one of Wittgenstein's dicta: that about things one knows nothing, one should not speak. I wish more people followed this dictum. Well, computers play a multiple role: they are superb as tools, but they also offer a field for a new kind of experimentation. Mitchell should know. There are certain experiments you cannot perform in your mind. It is impossible. There are experiments that you can do in your mind, and there are others you simply can't, and then there is a third kind of experiment where you create your own reality. Let me give you a problem of simple physics: a gas of hard spheres. Now nature did not provide a gas of hard spheres. Argon comes close, but you can always argue that maybe, because of slight attractive tails, something is going to happen. There is no substance—nature was so mean to us that there is no gas of hard spheres. And it poses very many interesting problems. It is child's play on the computer to create a gas of hard spheres. True, the memories are limited, so that, as a result, we can't have 10^{23} hard spheres, but we can have thousands of them, and actually the sensitivity to Avogadro's number is not all that great. We can really learn something about reality by creating an imitation of reality, which only the computer can do. That is a completely new dimension in experimentation. Finally, I may be

"Order in Chaos," a conference on the mathematics of nonlinear phenomena. Sponsored by the Center for Nonlinear Studies at Los Alamos National Laboratory, May 24-28, 1982.

misquoting him, but a very famous contemporary biologist, Sidney Brenner, who gave a lecture at Rockefeller University while I was still there, said that perhaps theory in biology will not be like that of physics. Rather than being a straight deductive, purely mathematical analytical theory, it may be more like answering the following question. You have a computer, and you don't know the wiring diagram, but you are allowed to ask it all sorts of questions. Then you ask the questions, and the computer gives you answers. From this dialogue you are to discover its wiring diagram. In a certain sense, he felt that the area of computer science—languages, theories of programming, what have you—may be more of a model for theorizing in biology than writing down analytic equations and solving them.

FEIGENBAUM: A more synthetic notion.

KAC: Yes. In fact, I think we will go even farther in this direction if we introduce, somehow, the possibility of evolution in machines, because you cannot understand biology without evolution. In fact, my colleague Gerry Edelman, whom you know very well and who is a Nobel laureate in biochemistry, is now "into the brain" and is trying to build a computer that has the process of evolution built into it so that you evolve programs: you start with one program that evolves into another, etc. It is an attempt to get away from the static, all-purpose Cray, or whatever it is, and to endow the computer with that one extraordinary, important element of life, namely evolution. I also feel like Stan; if I were younger—Si la jeunesse savait; si la vieillesse pouvait,—as you say in French,* I'd also get into biology. Those are fantastically challenging problems, and they are problems that call for formulation, not only for solution. That's also exciting, to be present at the creation, to formulate the problem.

ULAM: I might add something to it. In fact, to some extent, the differences we talked about between mathematicians and physicists, or the bent of mind, is of that sort. I also wrote, a very crude picture, about the following system: mathematicians start with axioms and draw consequences, theorems. Physicists have theorems or facts, observed by experiment, and they are looking for axioms, that is to say, laws of physics, backwards. Just as you said, the idea is to deduce this system of laws or axioms from which the observed things would follow. Actually the so-called Monte Carlo approach is a little that way, even in problems of a very prosaic, very down-to-earth nature. You manufacture your own world, as you say, of hard spheres, or what have you.

FEIGENBAUM: Mark, I want to turn to something that you mentioned yesterday. You offered a quotation that "axiomatization is the obituary of a great idea." In context, you were talking about how sometimes you can sort of overkill the mathematics and leave it dead in some way, as opposed to letting it speak for itself and be alive. Will you amplify on the soul of mathematics?

KAC: I will try. There is, of course, axiomatization and axiomatization. If, indeed, we think of the process of natural sciences as the discovery of what we call laws of nature that you can say are its axioms, then, to the contrary, such a discovery is a birth announcement. But, for instance, take geometry: that's one of the oldest, best known parts of human knowledge and, in fact, one of the great achievements of the Greeks. Euclid is probably being given most of the credit, but it was a communal affair, this axiomatization (axiomatization in the sense that from a simple number of seemingly self-evident statements, one can deduce and create a whole world of facts). Then it turned out there were cracks in this edifice; suddenly there were certain concepts that were not fully axiomatized. The ultimate axiomatization of geometry came with Hilbert in 1895, 2000 years after Euclid. That was an obituary in a certain sense, because then it (axiomatization or geometry) could be relegated essentially to a computer. Once the subject becomes so well organized that every single thing can be reduced to a program, then there is nothing more to be done. In fact, Gödel gave hope by proving that reduction is impossible in the somewhat wider system of mathematics, that always, no matter how large, how complex a system is, there will be statements that you won't be able to prove or to disprove. That means there is always the possibility of creation, another axiom, or something or other. There is this tendency among mathematicians of trying to understand through axiomatization.

ULAM: And in physics this is nonsense.

KAC: There are people who still try to axiomatize thermodynamics. The very last thing anybody should be doing is axiomatizing thermodynamics. I mean, first of all, most physical theories, though thermodynamics, I must say, is one of the most durable ones, are only temporary. They change; they evolve. So why the heck should one axiomatize something that the next day is going to be obsolete? But, on the other hand, many mathematicians who are trained formally feel there is no other way to perceive a subject but by strict axiomatization. And worse yet, they try to teach little children in schools like that. To teach geometry through the complete systems of axioms is stupid. Teaching geometry is to tickle a young man's or a young woman's imagination in solving all the wonderful problems. It should not be work to prove that if A is between B and C, and D is between A and C, then D is between B and C. You'll just draw a picture, and it is trivially evident.

ULAM: Take the new math, for instance.

KAC: I could speak hours against new math.

ULAM: It's waning, isn't it?

KAC: Yes, that's flogging a dead horse.

FEIGENBAUM: Do you think that this idea of people's just being

** "If youth only knew; if age only could."*

Ulam: One motive for doing mathematics is that suddenly you feel the ability that you are good at something. Very human. Nothing wrong with that feeling.

INTERVIEW

trained from a purely axiomatic viewpoint is a growing phenomenon, or has it always been so amongst mathematicians and scientists?

KAC: I really don't know. I know only a very few people.

FEIGENBAUM: You alluded to that situation in saying it's now taught, for example, in terms of new math, although you say that the new math is dying.

KAC: It was true for a while because, somehow, a group of mathematicians sold this idea to poor high school teachers, who didn't even understand what it was all about and who then taught geometry and other things only through axioms. There are two principles of pedagogy which have to be adhered to. One is, "Tell the truth, nothing but the truth, but not the whole truth." That I had from a former colleague who is now unfortunately deceased. The other one is, "Never try to teach anyone how not to commit errors they are not likely to commit." Now, to give you an example. New math spends an awful lot of time in second grade, God forbid, in trying to tell the little kids that you write a little three and you write a big three, and yet the little three and the big three symbolize the same thing because it is the cardinal number of a set of three elements. Correct? That is sheer idiocy. If a kid is logically sophisticated and is bothered by it, then I would take him aside and give him special training, but to create confusion in the mind of a child who is perfectly willing for a while to know that this three and this three, even though one looks bigger than the other, represent the same thing—leave it be! I know it sounds a little funny, but I feel very strongly about it. The need for precision, for logic, must be not imposed from outside. It must be coming from within. If somebody really feels uncomfortable, then he or she has an enormously highly developed sensitivity to finer logical points.

ULAM: I try to make jokes about it. If you print a page of mathematics or anything else, it is not invariant, because if you look at it upside down, it looks different. So the idea in new math was to write in such a way that no matter what angle you look at it, it is the same. That's an ultramathematical point of view.

FEIGENBAUM: Another question I was thinking about was, in reminiscing back to the Scottish Café, what was the excitement for mathematics? Was there some feeling at that time that there was a scheme of understanding things that would continue into the future?

KAC: Stan, you are much more strongly connected with the Scottish Café.

ULAM: I don't think so really. People were so immersed in the actual problems. Occasionally there would be some kind of speculation about the more remote future. For example, in Lwów, my home town in Poland, Banach, this famous mathematician whom I think you mentioned earlier, decided to have a big notebook kept in the Scottish Café where we assembled every day. It was a book in which problems to be solved, remarks, and ideas were written down. It was

kept in the Café, and the waiter would bring it when we came in. A lot of interesting problems were written up. The book, by the way, is being published by Birkhäuser. I guess I started to say that occasionally there would be some speculation. The mathematician Mazur once said, for example, "There must be a way to produce automatic arrangements which will reproduce themselves." That was long before von Neumann actually went into this whole complex of problems and found one way to do it. Speculations of this sort appeared sporadically, but on the whole it was a more down-to-earth, mathematically defined collection of problems which interested us in various fields, such as functional analysis and set theory, fields which were in those days still young.

KAC: But aging already.

ULAM: Perhaps.

KAC: It is difficult to say. Functional analysis, of course, was Banach's creation, and partly Steinhaus's. Toward the end of my student career, it was Banach, himself, I felt, and also Mazur, who began to look for other worlds to conquer.

ULAM: The nonlinear program of studies.

KAC: That's right. Banach also was reading. I can remember because I was once in his office over some trivial matter, and he was reading Wiener's early papers on path integrals. I agree with Stan, though I was less of a habitué of the Scottish Café. First of all, my teacher, Steinhaus, frequented a more elegant establishment where there were special things to eat, and all that. Secondly, I was financially somewhat less affluent than Stan—I was, as Michael Cohen, one of our mutual friends, says, independently poor. And it

Kac: We can really learn something about reality by creating an imitation of reality, which only the computer can do. That is a completely new dimension in experimentation.

INTERVIEW

did cost a little to visit in the Café. What happened primarily was that people discussed problems of interest and then people thought about them. If, indeed, nothing immediate came out of the problem, nothing that appeared to be interesting and promising, then it would be recorded in the notebook. Actually, very few problems in the book proved to be completely trivial. Many of them had a very noble history. Papers were written on many of them, and some are still unsolved. In fact, I want to make a kind of a footnote here. It is so remarkable that the Poles did not publish this book; rather, it has been published in the United States through the efforts, really, of a very remarkable young friend of ours by the name of Dan Mauldin, who is a professor of mathematics at, of all the impossible places, North Texas State University in Denton, Texas. He is a first-rate mathematician, and he has the Polish soul with regard to mathematical problems. It would be interesting to interview him, because he was on his way to becoming an All-American linebacker on the famous Longhorn team, and he gave it up for mathematics.

ULAM: Yes, he was on the Texas football team and played in championship games.

KAC: And then to the disgust of his coach, in his senior year, when he would really do tremendous things, he gave up football and started worrying about set theory.

ULAM: He was offered a car and money.

KAC: A house and everything. It's rather interesting what passions mathematics can engender.

ULAM: One thing you forgot to say—one motive in mathematics is the feeling that you can do something by yourself. I think it is present in almost all mathematicians. One motive for doing mathematics is that suddenly you feel the ability that you are good at something. Very human. Nothing wrong with that feeling.

KAC: Very human, in fact. Actually, I don't think it is really either understood, or perhaps not even understandable at all, how some problems generate passion. Some of them, by the way, ultimately prove to be of relatively little importance. I remember one in connection with Stan. Stan generates problems and conjectures at probably the highest rate in the world. It is very difficult to find anybody in his class in that. Many of them we discuss. He came with one and said, "Look, I thought of the following modification of Fibonacci numbers." With ordinary Fibonacci numbers you start with 1 and 1 and add them, obtaining 2 as the third member of the sequence. Then you add 2 and 1, obtaining 3, then 3 and 2, which gives 5, etc. In other words, the $(n+1)$th member of the sequence is the sum of the nth member and the $(n-1)$th member. Symbolically, $a_{n+1} = a_n + a_{n-1}$ with $a_1 = a_2 = 1$. But in Stan's idea, the formula for a_{n+1} is now $a_{n+1} = a_n +$ either $a_1, a_2, \ldots, a_n$, each taken with probability $1/n$. My God, it is interesting as a coffee house conversation, but for some strange reason, it caught me, and I worked on it, and I even found the mean of a_n, and even the variance. And the variance is given by a tremendous formula with a square root of 17 in it. It even appeared as a little Los Alamos report. I probably spent, easily, a week of hard work on it. Why? I have no idea except I couldn't let the damned thing alone.

ULAM: What you did with the Fibonacci-like rule was beautiful work, and it has a certain simplicity, like the problem itself. And the solution was unexpected because a_n grows exponentially, not with respect to n, but with respect to the square root of n...

KAC: Square root of n, with a complicated constant. There is a point to it because in constructing the sequence, you need at every stage to know *all* the preceding terms—a highly non-Markovian affair. At the time when I was playing with it, it was almost like being an alcoholic. You know it isn't good for you.

ULAM: Another interesting problem is still unsolved—Fermat's. The sum of two squares can be a square, but the sum of two cubes cannot be a cube, and so on. Nobody can prove it for arbitrary powers. Of course, for cubes, quartics, and so forth, but in general, nobody has been able to do it. It seems like a silly little puzzle, and yet so many people worked on it that as a matter of fact some of the efforts to solve it gave rise to much of the modern algebra. This is a strange thing. The mathematical ideal theory and other algebraic theories came from efforts to solve this silly puzzle.

KAC: So you never can tell. You never can tell. Usually these puzzles, the good ones, generate some tremendous things later on, while others of them die. It is very much like survival of the fittest.

ULAM: Or some kind of mysterious thing about the problems that makes them important in the future. It is impossible to tell logically.

FEIGENBAUM: You are almost saying that the problems have a teleological spirit to them and that you don't necessarily realize their unique position at the time they're done.

ULAM: No, one shouldn't be completely mystical, but one day maybe a little will be understood. There must be some...

KAC: Oh, come on, let's be mystical! Why not?

ULAM: So far we are.

FEIGENBAUM: One last question. Have you ever had long-range hopes of finding a good way to analyze a problem and then seen these hopes realized over many years? I think in physics very often there are programs that are set out. Someone has an idea, there is a way you can do the problem, and a lot of people will work on it, perhaps over ten years; sometimes it pans out and sometimes it doesn't.

KAC: I think the best example of that is the recent solution of the classification of all simple groups, finite groups. That is really one of the few genuinely collective efforts in mathematics, including the computer by the way, and that was a program, too, because there were various breakthroughs, understandings came from various places. Well, when it became clear that the problem of classifying simple groups probably could be solved, then an enormous human

machinery was created to solve it. In general, mathematicians, even much more than theoretical physicists, tend to be loners. They are collaborative, but basically there are very few papers with, say, more than three coauthors. It would be interesting to plot a graph: by the time it is five authors, the graph hits zero.

ULAM: In mathematics it is zero. It is not uncommon in physics. In answer to your question, Mitch, Newton said something like—I have to paraphrase it, "If I have achieved something in my life in science, it is because I have thought so long and so much about these problems."

FEIGENBAUM: He also said that if he was able to see further than other people, it was because he was standing on the shoulders of giants.

KAC: Sidney Coleman paraphrased that with, "If I was able to see farther, it was because I was surrounded by midgets."

FEIGENBAUM: What are the things that you have done that you feel most warm towards?

KAC: To begin with, I was always interested in problems rather than in theories. In retrospect the thing which I am happiest about, and it was done in cooperation with Erdös, who also occasionally comes to Los Alamos, was the introduction of probabilistic methods in number theory. To put it poetically, primes play a game of chance. And also some of the work in mathematical physics. I am amused by things. Can one hear the shape of a drum? I also have a certain component of journalism in me, you see; I like a good headline, and why not? And I am pleased with the sort of thing I did in trying to understand a little bit deeper the theory of phase transitions. I am fascinated, also, with mathematical problems, and particularly, as you know as well or better than I, the role of dimensionality: why certain things happen in from three dimensions on and some others don't. I always feel that that is where the interface, will you pardon the expression, of nature and mathematics is deepest. To know why only certain things observed in nature can happen in the space of a certain dimensionality. Whatever helps understand this riddle is significant. I am pleased that I, in a small way, did something with it. And you, Professor?

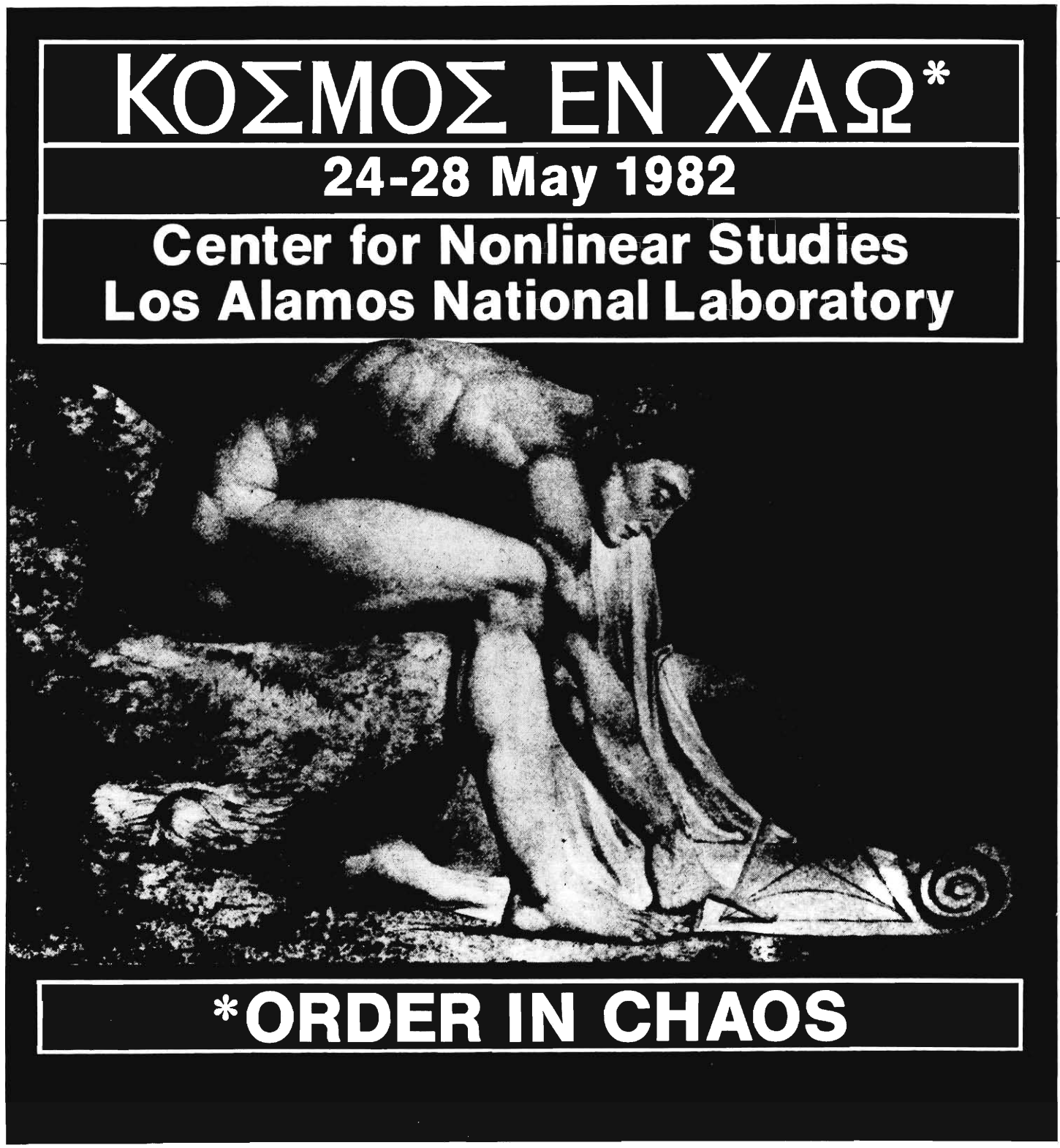
ULAM: I don't know. I think I was sort of lucky in a number of instances and not so clever. Dumb but lucky. Originally I worked in set theory and some of these problems are still being worked on intensively. It is too technical to describe: measurable cardinals, measure in set theory, abstract measure. Then in topology I had a few results. Some can be stated popularly, but we have no time for that. Then I worked a little in ergodic theory. Oxtoby and I solved an old problem and some other problems were solved in other fields later. In general I would say luck plays a part, at least in my case. Also I had luck with tremendously good collaborators in set theory, in group theory, in topology, in mathematical physics, and in other fields. Also some common sense approaches like the Monte Carlo method, which is not a tremendously intellectual achievement but is very useful, a few things like that.

KAC: I must interrupt because it's time for the afternoon session, but let me end by saying that it is the deserving ones who are also lucky. ■

Mitchell Feigenbaum, mathematical physicist and key contributor to the theory of chaos, proudly acknowledges that he, too, is half Polish. Born in New York City, he was, from an early age, deeply interested in understanding nature's puzzles. And, like his Polish seniors, Kac and Ulam, he has an abiding interest in both the nature of human experience and the nature of the human brain. One of his distant hopes is that his new approach to chaotic phenomena may provide a clue on how to model the complex processes of the brain. But speculation and fanciful notions notwithstanding, his work reflects his profound understanding of what makes for real progress rather than mere amusement in mathematical science.

Briefly, he discovered a universal quantitative solution characterized by specific measurable constants that describes the crossover from simple to chaotic behaviors in many complex systems. With the first experimental verification of these predictions for the onset of turbulence in fluids, it became clear that a new methodology had become available to treat previously intractable problems. The idea of the method is that a very low dimensional discrete nonlinear model that incorporates only the most basic qualitative features can, because of universality, correctly predict the precise quantitative details of a highly complex system. One is therefore directed to take very seriously—and not merely as a mathematically suggestive toy—the study of what had otherwise appeared to be a naive and oversimplified model. Indeed, these investigations of low dimensional discrete systems have by now blossomed into a large experimental and theoretical subdiscipline.

Thus, Feigenbaum is regarded as one of the founders of the modern subject of chaos and has several new mathematical/physical constants named after him. In 1980 he received a Los Alamos Distinguished Performance Award for this seminal work. A staff member at Los Alamos since 1974 and a Laboratory Fellow since 1981, he is currently on leave of absence as a Professor of Physics at Cornell University.



Conference poster by Gail Flower

Review of the CNLS Conference on Chaos in Deterministic Systems

by David Campbell, Doyne Farmer, and Harvey Rose

The phrase "order in chaos" seems self-contradictory: chaos is, after all, conventionally viewed as the complete absence of order. Yet precisely this title attracted two hundred and ten scientists from fourteen countries to Los Alamos from May 24-28, 1982, to attend the second annual international conference of the Center for Nonlinear Studies. The purposes of the conference were to survey the recent rapid developments and to anticipate the trends for future research in the area of "chaos in deterministic systems." The breadth of scientific interest in this topic was reflected in the variety of subjects discussed at the meeting. Presentations ranged from abstract mathematics through numerical simulations to experimental studies of fluid mechanics, chemistry, and biology. Even weather prediction made an appearance.

To appreciate the appeal of the conference title—and the importance of the field of research it describes—requires a closer look at the apparently contradictory components. The concepts of "order" and "determinism" in the natural sciences recall the predictability of the motion of simple physical systems obeying Newton's laws: the rigid plane pendulum, a block sliding down an inclined plane, or motion in the field of a central force are all examples familiar from elementary physics. In contrast, the concept of "chaos" recalls the erratic, unpredictable behavior of elements of a turbulent fluid or the "randomness" of Brownian motion as observed through a microscope. For such chaotic motions, knowing the state of the system at a given time does not permit one to predict it for all later times. In place of the determinism of the orderly systems, one has only probabilistic estimates and statistical averages.

Thus, in some sense, the possibility that chaos exists in deterministic systems runs directly counter to our intuition. To understand that this possibility is nonetheless real, we can refer to the deeper insight of Henri Poincaré, one of the founders of modern dynamical systems theory. Writing in the pre-quantum era of pure Newtonian determinism, Poincaré noted that

> *A very small cause which escapes our notice determines a considerable effect that we cannot fail to see, and then we say that the effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation* approximately. *If that enabled us to predict the succeeding situation* with the same approximation, *that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon.*

Hence, the crucial ingredient in deterministic chaos is a very sensitive dependence on initial conditions. Motions that start close to each other develop in time in dramatically different ways, and uncertainties in the initial values develop rapidly—exponentially, in fact—in time. Although the motion from instant to instant can be predicted, over macroscopic times it becomes no more predictable than a random sequence.

At first it might appear that the distinction between orderly and chaotic motions is merely one of the complexity of the system involved. In the parlance of dynamical systems theory, the orderly motions described above involve just one "degree of freedom," whereas the chaotic fluid involves many—in conventional hydrodynamics, infinitely many—degrees of freedom. It is thus tempting to associate simple systems with order and complicated ones with chaos.

In fact, this naive association is wrong for several fundamental reasons, some obvious and some subtle. First, everyday experience tells us that complicated systems with many degrees of freedom *can* undergo very orderly motion. For example, a fluid in smooth (laminar) flow moves in a regular, totally predictable manner.

Second, it is less familiar but nonetheless true that very simple physical systems can exhibit chaotic behavior, with all the associated randomness and unpredictability. Numerical experiments show that the motion of a rigid plane pendulum, if damped and driven, becomes truly chaotic. This result illustrates strikingly that a completely deterministic system can produce chaos without the addition of any external random noise. In other words, you don't have to put randomness in to get it out. The existence of deterministic motions that produce chaos is a clear example of order in chaos.

Third, it is now well established that, at least in some cases, the chaos observed in very complicated systems can be understood quantitatively in terms of simple models that involve very few degrees of freedom. This profound result, several examples of which were presented at the conference, is perhaps the most significant manifestation to date of order in chaos.

From this general motivation of the theme "order in chaos," we turn to a discussion of the specific results described at the conference. (The accompanying table lists the authors and titles of the talks presented.) Very roughly, the presentations divided into two major areas. First, there were attempts to identify the qualitative and quantitative essential features of deterministic chaos to describe and model it more accurately. Second, there were discussions of the transition from regular motion to chaos. Here the focus was on identifying various possible routes and establishing whether they had "universal" properties that were independent of the details of the mathematical model or physical system being studied.

1982 CNLS CONFERENCE TALKS[a]

Review

M. Feigenbaum Los Alamos National Laboratory	An Overview of Order in Chaos
H. Swinney University of Texas, Austin	Observations of Chaos

Theoretical

S. Aubry Laboratoire Léon Brillouin and Los Alamos National Laboratory	Devil's Staircase and Order Without Periodicity in Condensed Matter Physics
J. D. Farmer Los Alamos National Laboratory	Dimension, Fractal Measures, and Chaotic Dynamics
J. Ford Georgia Institute of Technology	Classical and Quantum Billiards: New Insights into Chaos
V. Franceschini University of Modena	Bifurcation Phenomena in Truncated Navier-Stokes Equations on a Two-Dimensional Torus
J. Guckenheimer University of California, Santa Cruz	Overview of Dynamical Systems Theory
M. Gutzwiller International Business Machines Corporation	Stochastic Behavior in Electron Scattering
E. Heller Los Alamos National Laboratory	Quantum Manifestations of Classical Chaos
P. Holmes Cornell University	Periodically Forced Nonlinear Oscillations of Dissipative Systems: Some Answers and Questions
P. Huerre University of Southern California	Long-Time Solutions to the Ginzburg-Landau Equation: A Numerical Study
C. Leith National Center for Atmospheric Research	Chaos and Order in Weather Prediction
R. MacKay Princeton University	A Renormalization Method for Orbits with Generalized Golden Ratio Rotation Number
B. Mandelbrot International Business Machines Corporation	Quadratic Chaos, Scaling, and Fractals
J. Marsden University of California, Berkeley	Fluids, Vortices, and Coadjoint Orbits
E. Ott University of Maryland	Strange Attractors in Crisis

A significant aspect of the conference was that in each of these areas there were important new developments both in theoretical modeling and prediction and in experimental observation.

These new developments were woven into the previous results in the conference's two introductory reviews surveying the field. In his "Observations of Chaos" Harry Swinney of the University of Texas, Austin, described experimental observations in electrical oscillators, chemical reactions (the Belousov-Zhabotinskii reaction), and fluid flows (Rayleigh-Bénard convection and circular Couette flow) that established the existence of deterministic chaos. He reviewed ex-

SHORT SUBJECT

N. Packard University of California, Santa Cruz	Measures of Chaos in the Presence of Noise
D. Ruelle Institut des Hautes Etudes Scientifiques	Unconventional Turbulent Structures
S. Shenker University of Chicago	Scaling Behavior in Maps of the Circle
E. Siggia Cornell University	A Universal Transition from Quasi-Periodicity to Chaos in Dissipative Systems
J. Yorke University of Maryland	The Dimension of Strange Attractors
A. Zisook University of Chicago	Universal Effects of Dissipation in Two-Dimensional Mappings

Experimental

I. Epstein Brandeis University	Oscillations and Chaos in Chemical Systems
L. Glass McGill University	Chaos in a Petri Dish: Nonlinear Dynamics of a Cardiac Oscillator
H. Haucke and Y. Maeno University of California, San Diego, and Los Alamos National Laboratory	Time-Dependent Convection in ^{3}He/Superfluid ^{4}He Solution
R. Keolian University of California, Los Angeles	Generation of Subharmonics and Chaotic Behavior in High-Amplitude, Shallow-Water Waves
O. Lanford, III University of California, Berkeley	Period Doubling in One and Several Dimensions
A. Libchaber Group de Physique des Solides de l'Ecole Normale Supérieure	Mercury in a Magnetic Field, A Rayleigh-Bénard Study
J. D. Roux Université de Bordeaux	Chaos in the Belousov-Zhabotinskii Reaction
R. Shaw University of California, Santa Cruz	The Dripping Faucet as a Model Chaotic System
C. W. Smith University of Maine	Bifurcation Universality for First-Sound Subharmonic Generation in Superfluid ^{4}He

[a]Proceedings of the conference will be published by North-Holland as a special issue of *Physica D* and also as a hardbound volume. The Center gratefully acknowledges support for the conference from the Applied Mathematical Sciences Program in the U.S. Department of Energy's Office of Basic Energy Sciences.

perimental data showing there are at least seven well-defined routes leading from smooth, regular motion to chaos. In his "Overview of Order in Chaos" Mitchell Feigenbaum of Los Alamos reviewed the theoretical description of deterministic chaos, recalling some of the essential ideas and methods and introducing simple model systems to illustrate these results. These introductory surveys set the stage for over a score of additional presentations, in which theory and experiment, abstraction and observation, were mingled in an appropriately chaotic manner.

To explain both the experimental and the theoretical results in

more detail, it is necessary to introduce some concepts from dynamical systems theory, which is the formal discipline underlying the study of all types of motion. These general concepts were discussed, in slightly differing contexts, in the survey by Feigenbaum and in the talks of Ford, Guckenheimer, Holmes, and Marsden. Dynamical systems can be divided into two broad categories—conservative and dissipative—depending on whether or not the energy is conserved. The Navier-Stokes equations of fluid mechanics are an important (infinite degrees of freedom) dissipative dynamical system, since the viscosity converts the energy of fluid motion into heat. Most of the research presented at the conference dealt with dissipative systems, where the long-time behavior is controlled by various kinds of "attractors." That is, different initial conditions evolve in time "toward" an attractor, and after initial transients die out the motion reduces, in a well-defined sense, to motion "on" the attractor.

A simple model description of a dissipative dynamical system, used in Feigenbaum's talk and familiar to the readers of *Los Alamos Science*,* is the discrete "logistic map," in which one point in the interval [0,1] is transformed to another according to

$$x_{n+1} = \lambda x_n (1 - x_n) \ . \qquad (1)$$

Since x_{n+1} follows uniquely from x_n, the map is deterministic. To view this map as a dynamical system, we need only think of the number of iterations of the map as "time" and the sequence of points $x_1, x_2, x_3, \ldots$ as the "motion." As a function of the parameter λ, the map has a variety of attractors. First, for $0 < \lambda < 1$, this map has a "fixed point" attractor at $x = 0$. As a simple exercise on a pocket calculator will demonstrate, for this range of λ initial points anywhere in the interval are eventually attracted, after many iterations (long time), toward the point $x = 0$. In real physical systems this type of attractor corresponds to motion that does not change in time. Thus, for example, when a pot is filled with water and placed on a flat surface, the initial sloshing dies out and the fluid comes to rest. For $1 < \lambda < 3$ the fixed point at the origin is unstable, and a new stable fixed point at $x = 1 - 1/\lambda$ appears. Analogous behavior is seen when a pot of water is heated and *steady* convection rolls form. Even though the fluid is moving, because the flow pattern is constant in time, the attractor is a fixed point.

A second type of attractor found in Eq. 1 is a periodic limit cycle, in which the sequence of values of x_n repeats itself periodically. As λ is increased in the range $3 < \lambda < 3.59...$, there is the famous sequence of periodic cycles with periods 2^n, $n = 1, 2, 3, \ldots$. In our analogy to a heated pot of water, a limit cycle corresponds to *oscillatory* convection rolls in which the flow pattern changes periodically in time.

The third type of attractor in Eq. 1 is much less familiar and, in fact, is called a strange attractor, a term first introduced by David Ruelle. These strange attractors (also called chaotic attractors) occur in Eq. 1 for certain values of $\lambda > 3.59...$ and describe chaotic "motion" in the map in the sense that the sequence of points $\{x_n\}$ is random. For some time, it has been thought that these strange attractors underlie the chaos observed in more complicated dynamical systems. Thus, for example, the turbulence seen in a pot of boiling water can be described by a strange attractor. One of the most exciting aspects of the conference, which we shall discuss in detail later, was the conclusive evidence from a variety of experiments that the chaos in several real systems can be described by low-dimensional strange attractors.

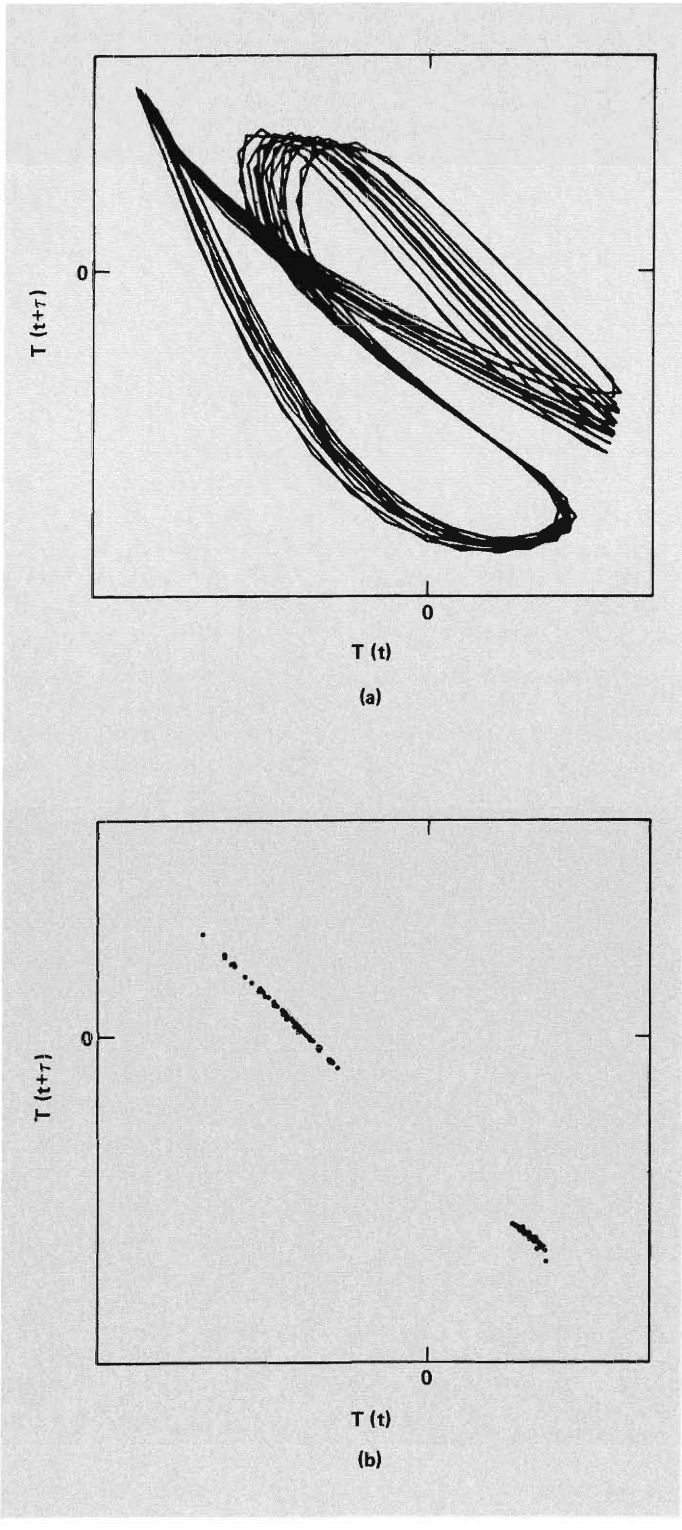
Chaotic or strange attractors are elegant incarnations of "order in chaos." Since many initial conditions collapse onto the attractor, the number of degrees of freedom "actively" participating in the chaos can be many fewer than in the full system. On the other hand, the chaos is real, because nearby points on the attractor separate initially at an exponential rate (determined by the positive Lyapunov exponents), causing small errors to amplify very fast and producing sensitive dependence on initial conditions.

Strange attractors are like bakers. Thinking of the whole space of possible initial conditions as dough, a strange attractor grabs the dough, stretches it, and then folds it back onto itself. Just as a small drop of vanilla will quickly get mixed throughout the dough by this process, a strange attractor rapidly mixes together all the initial conditions that it attracts, creating chaos.

The dimension of an attractor is, roughly speaking, the number of "active modes" that are left once all the transients have died out. It turns out, though, that the folding described above creates a complicated structure, something like the filo dough of Greek pastries. This structure, dubbed fractal by Mandelbrot and described in detail in his talk, makes dimension a difficult concept to define; in particular, the dimension need *not* be an integer. Relations among the dimension of a chaotic attractor and its other properties—such as the values of the Lyapunov exponents—were discussed in the presentations of Farmer and Yorke.

Given a chaotic dynamical system, one of the central problems is to ferret out the strange attractor—assuming it is present—and to estimate its dimension and other properties. This and related problems were analyzed in theoretical models in the talks by Farmer, Holmes, Huerre, Marsden, and Ott. In particular, Farmer's and Huerre's presentations underscored by example the possibility, mentioned in many other talks, that even in a dissipative dynamical system with infinitely many degrees of freedom, the chaotic attractor

*Mitchell J. Feigenbaum, "Univeral Behavior in Nonlinear systems," Los Alamos Science, Vol. 1, No. 1, 4-27 (1980).



may be low dimensional, involving perhaps only two or three active modes. In some of the most exciting reports to the meeting, this possibility was confirmed or at least supported in a variety of experimental systems. The experiments described included fluid flows (Haucke and Maeno, Keolian, and Libchaber), surface tension (Shaw), chemical reactions (Epstein and Roux), and physiological studies of heart beat irregularities (Glass). The accompanying figure, which is from the presentation of Haucke and Maeno, contains a portrait of the slightly more than two-dimensional strange attractor that evolved from convective flow in a Rayleigh-Bénard cell containing a mixture of helium-3 and superfluid helium-4 at 4 degrees kelvin.

Other experimental examples of chaos were much closer—some amusingly, some frighteningly—to our day-to-day reality. Have you ever been kept awake at night by a dripping faucet? If so, you might have noticed that some faucets drip periodically, while others drip in an unpredictable, apparently chaotic manner. In his talk Shaw analyzed a dripping faucet, and showed that, in some cases, the time intervals between drips are determined by a nearly two-dimensional chaotic attractor. Your frustration with the television weatherman might be moderated by considering some of the points raised by Leith. The extreme sensitivity of weather to initial conditions has long been recognized. Indeed, Leith recalled that one of the first simple dynamical systems known to lead to deterministic chaos was developed by meteorologist Edward Lorenz in the context of weather prediction. On the basis of model experiments, turbulence calculations, and studies of the differing evolution of similar atmospheric states, meteorologists estimate that the average doubling time for errors (which is related to the Lyapunov exponents mentioned earlier) is two and a half days. Yet another example, one with which none of us would want to be familiar, was discussed by Glass. The normal "motion" of a heart is a regular oscillation, which can be modeled by a periodic attractor. Glass presented experimental evidence that under certain stimuli the underlying attractor can

Illustrations from the presentation of Hauke and Maeno showing the (nearly) two-dimensional strange attractor underlying chaotic flow in a mixture of helium-3 and superfluid helium-4. (a) A "phase portrait" formed by plotting the temperature T ***at time*** t ***vs the temperature at a delayed time*** t + τ ***as measured on a probe. Since the attractor is nearly two-dimensional, it can be pictured as a ribbon (perhaps with folds and twists), and the plot is a projection of this ribbon onto a two-dimensional surface. (b) A "slice" through the attractor formed by making what is technically called a Poincaré section. The fact that each of the two disjoint parts of this slice looks one-dimensional (that is, like a line) demonstrates that the attractor is approximately two-dimensional.***

change from a limit cycle to a strange attractor, with the resulting chaos possibly being related to a heart attack.

Thus far we have focused on the *nature* of deterministic chaos, which was the first of the two major areas discussed at the conference. The second major area, routes to chaos, was also well represented both theoretically and experimentally. In his survey Swinney identified and gave experimental evidence for seven distinct routes to chaos, thus stressing the point that no single scenario exists to describe the transition to chaos in deterministic systems. On the theoretical side, the talk of Franceschini revealed the variety of routes to chaos in finite mode truncations of the Navier-Stokes equations. Since in this brief review we cannot possibly discuss all these routes, we shall focus on the two that arose most frequently in the presentations at the conference: the period doubling transition to chaos and the transition from periodic to quasi-periodic motion (which involves two or more frequencies that are not rational multiples of each other) to chaos.

The period doubling route to chaos has already been obliquely mentioned in our discussion of the logistic map. In fact, this route is known to exist in a wide class of dynamical systems, and, when it is observed, its essential properties are "universal" in the sense that they do not depend on the details of the specific system. This quantitative universality was first discovered and then extensively analyzed by Feigenbaum. Examples were discussed in his survey, in a presentation of two-dimensional maps by Zisook, and in a more general and abstract setting by Lanford. Experimentally, this route was observed in chemical reactions (Swinney and Roux), Rayleigh-Bénard convection experiments in mercury (Libchaber), first sound generation in superfluid helium-4 (Smith), heart beat irregularities (Glass), dripping faucets (Shaw), and water waves (Keolian).

One underlying mechanism for the periodic-quasi-periodic-chaotic transition to chaos was originally suggested by Ruelle and Takens in a general abstract mathematical framework. More recent work on a somewhat different, more explicit mechanism was discussed in the talks of Shenker and Siggia, and related calculations were described by MacKay. Experimental evidence for this route to chaos was discussed in chemical systems (Roux) and in Rayleigh-Bénard convection experiments in mercury (Libchaber). In several of the experimental talks, observations of some of the other routes to chaos mentioned by Swinney were also discussed.

Although most of the conference presentations fell into one of the two main areas already discussed, a number of talks addressed other topics related to chaos in deterministic systems. The nature of chaos in conservative systems, in which there cannot be attractors, was mentioned briefly in several talks and discussed more extensively by MacKay and Ford. Chaos in conservative systems has its historical roots in the fundamental questions of statistical mechanics. Why should a gas of interacting particles be described by the well-known statistical ensembles of Gibbs? Although we expect a large, isolated collection of interacting particles to be in thermal equilibrium, there is no generally applicable mathematical theorem that corroborates this expectation. In the specific context of the billiard ball problem, Ford discussed the possibility that the chaotic dynamics might lead to a state resembling thermal equilibrium.

Among the other topics discussed, several appeared to point the way to significant problems of the future. In emphasizing deterministic chaos, we have thus far explicitly excluded external noise or thermal fluctuations, which could add a separate, nondeterministic source of randomness to a dynamical system. Since in any experiment some level of noise can be anticipated, the response of chaotic deterministic systems to noise is a very important question. In particular, does external noise destroy the order in deterministic chaos? In his presentation Packard discussed this point and the scaling properties of information production rates for chaotic systems with external noise.

The possible role of chaos not in the time evolution of a dynamical system but in the spatial structure of condensed matter systems was discussed in the talks of Aubry and Ruelle.

Finally, there were presentations concerned with the manifestations of chaos in quantum mechanical systems. Gutzwiller displayed an example where the solutions to a particular Schrödinger equation depended sensitively on initial conditions. (Note that the sensitivity to initial conditions displayed in this example and in classical dynamics is quite different from the indeterminism in any measurement embodied by Heisenberg's uncertainty principle.) Heller illustrated the relations between the structure of quantum mechanical states and the orbits of the corresponding chaotic classical system.

In a very real sense, the Center for Nonlinear Studies' conference represented the "end of the beginning" of the field of deterministic chaos. Many of the fundamentals of low-dimensional chaos are theoretically modeled and experimentally verified, and a variety of intriguing questions seem ripe for answering. Given the panoply of models and the range of observed phenomena, it was no surprise that by the end of the conference most of the participants appeared ready to agree with the American poet Wallace Stevens, who, in his poem "Connoisseur of Chaos," asserted that

The squirming facts
exceed the squamous mind. ■

EDITIO POPULARIS

compiled by Barb Mulkin

With this issue we introduce "editio popularis"—popular edition. These brief accounts of ongoing work at the Laboratory are designed to provide a glimpse of the diversity of research conducted here. They are based on recent press releases sent from the Public Affairs Office to a national audience. Jim Breen is the Laboratory's Public Affairs Officer and Barbara Mulkin is the Deputy. Further information may be obtained by contacting the Public Affairs Office.

GEOCHEMICAL CONTROLS ON FISSION PRODUCT CONTAINMENT

Remarkably, remnants of natural fission reactors in the Oklo uranium mines in Africa still contain most of their nuclear wastes. David B. Curtis, Timothy M. Benjamin, and Alexander J. Gancarz have co-authored papers on migration of radionuclides from the site of several natural reactors discovered in thick layers of uranium ore at the Oklo Mines in equatorial Africa's Republic of Gabon.

The remains of the reactors were first identified in 1972 when French researchers found that uranium ore mined there was depleted in the fissionable isotope uranium-235. A complicated investigation revealed that uranium-235 was in short supply because it had been the fuel for natural fission reactors billions of years ago. The mere existence of these fossil reactors attests to their remarkable stability for about half the age of the earth. They also provide a unique opportunity to study the containment of waste from nuclear reactors in the earth's crust for geologic periods.

Los Alamos researchers are studying the migration of elements produced by neutron capture, nuclear fission, and radioactive decay. The latest papers report new data on the elements molybdenum, ruthenium, technetium, neodymium, and uranium and use these data to reconstruct the physical and chemical history of the reactors. The authors determined that fission began in the thick layers of uranium two billion years ago and that nuclear criticality was sustained for a few hundred thousand years. At some time within one million years after fission began, hot water circulating through the reactor cores removed certain elements and transported them into the surrounding rocks. Many of these mobile nuclear products were retained within a few meters of the reactor core. By comparing the chemistry of the mobile elements the authors were able to identify specific geochemical processes that may have controlled the loss of elements from the rocks of the reactor zones and led to their retention in peripheral rocks. Such information will be useful in evaluating the long term effectiveness of specific geologic sites in retaining commercially generated nuclear wastes.

The energy production of Oklo was about 15 times the energy that would be released in a nuclear-waste repository containing the entire 1980 U.S. inventory of spent commercial reactor fuel. The inventory in the Oklo reactors corresponded to 10 to 20 per cent of the uranium and 1 to 2 per cent of the nuclear products in the United States spent fuel inventory as of 1980.

Isotope and Nuclear Chemistry Division

SHUFFLER DESIGNED FOR FAST FACILITY

Los Alamos has designed and constructed a 15-ton nuclear material assay instrument for installation in the Department of Energy's new fuel reprocessing plant near Idaho Falls, Idaho. The instrument, nicknamed "Shuffler," is designed to measure the uranium-235 content of both spent fuel assemblies and the solid wastes generated by fuel reprocessing. It will be installed in the $200 million Fluorinel and Fuel Storage Facility (FAST) to be operated by EXXON Nuclear Idaho Company (ENICO).

George Eccleston, principal investigator and a member of the Laboratory's Energy Division Safeguards Assay Group, says the instrument is unique in that it was designed as an integral part of the facility and will be installed permanently. Usually, such instruments are portable or brought into a facility after it is complete.

In contrast to many foreign nations, the United States has no operating facilities for commercial reprocessing of spent reactor fuel, but federal facilities, such as that near Idaho Falls, will store and reprocess non-commercial spent uranium fuel from United States government test and research reactors and from the Navy's nuclear ship propulsion program.

Fuel assemblies will be brought to Idaho and stored in cooling ponds. Depending on their size, they may be cut into pieces before being reprocessed in an acid mixture. For materials accountability and criticality safety, the Shuffler will measure the uranium content in the fuel prior to dissolution of the fuel and separation of uranium.

The instrument irradiates, or interrogates, the fuel with a high flux of neutrons to produce uranium fissions and then counts the delayed neutrons resulting from the fissions. Given the high background of neutrons (17 million neutrons per second) and gamma rays (50,000 rads per hour), accurate measurements are difficult to make.

The californium-252 neutron source is manmade and emits up to 10 billion neutrons per second through spontaneous fission. This source is shuffled back and forth between a fuel interrogation position and a storage shield. The delayed neutrons are measured while the californium source is in the storage shield. The source may shuffle back and forth as many as 100 times while the instrument measures the uranium-235 in a spent fuel assembly.

The Shuffler's intricate functions were designed and evaluated using a Monte Carlo computer code developed by personnel of the Applied Theoretical Physics Division. The assay system will be remotely controlled and operated by a dedicated minicomputer system that can be easily operated by plant personnel. Design of the instrument as an integral part of the facility, rather than as a retrofit, enables both facility operations and assay measurements to be better coordinated. The instrument will be shipped and assembled at the FAST facility in the fall of 1982.

Energy Division

SPACE DETECTOR TESTED ON M-87

Researchers report a milestone in the effort to provide new tools for verifying activity in space that might violate the Limited Test Ban Treaty. A new system that is a marriage of two highly desirable features—good x-ray energy resolution and position sensitivity—has been developed and then tested successfully in space.

Richard Blake and Gordon Smith say the test 153 miles above the earth's atmosphere last summer imaged the galaxy M-87 in x rays and provided its x-ray energy distribution. M-87 is a powerful x-ray emitter and was chosen for this reason.

The Los Alamos equipment was lofted by a Black Brant rocket from White Sands Missile Range in New Mexico. In less than 150 seconds, the instrument package was on target. For almost 300 seconds, x rays streaming from the galaxy were focused by a telescope onto a detector that is a combination of a gas scintillator and an imaging proportional counter. X rays hitting the scintillator produced ultraviolet light that was passed through an optical window to the imaging counter. An electronic processor recorded the position and energy of each x-ray event. The data provided an image of the galaxy and its halo. The energy data can be transformed into information on the elements present in the galaxy and their physical properties, such as temperature.

Astrophysicists will be interested in data from the launch that may, for the first time, confirm a theory that elements in M-87 are not uniformly distributed throughout the galaxy, but "settle out" from the halo surrounding the galaxy's x-ray source according to their mass.

The experiment is one step in an effort to develop better capability to monitor nuclear explosions in space. More advanced versions may now be developed for longer space missions with higher sensitivity.

Because the new system is position sensitive, Blake also believes a modified version of this detector with suitable x-ray optics has potential for medical diagnosis, being able, in his opinion, to deliver high-resolution pictures with much lower x-ray doses to the patient.

Earth and Space Sciences Division

DATA "SANDWICHES" REVEAL NEW RESOURCES

Laboratory researchers have developed a rapid, efficient method of resource evaluation that turns reams of geologic and geophysical data and information from NASA satellites into photographic "sandwiches." The data-integration system is the latest tool in the battle to keep up with the information onslaught.

Thomas Weaver, principal investigator, says information is being "thrown at us so rapidly that we cannot keep up with the data." This system, which was developed as part of a DOE uranium resource evaluation program called NURE, was given a dry run at Talkeetna, Alaska. The success of the test led to development of a full-scale system that uses the Laboratory's Cray-1 supercomputers.

The Crays are programmed to accept data from aerial geophysical surveys, geologic maps, geochemical data, and Landsat imagery. The information is digitized, then transformed into photographic images that are overlaid, or "sandwiched," on film. The result is a spatially complete photo of all available information for a specific area.

After the success at Talkeetna, researchers chose to run a larger, more complex test on an area in southern Colorado's Montrose Quadrangle, which was also part of the earlier DOE uranium study. There are known mineral deposits and several types of mines in the quadrangle, and calibration data for the new system were available. Startling evidence of heretofore unknown concentrations of copper, lead, and zinc was shown. It would appear that much of the guesswork has been taken out of mineral exploration, and perhaps the wildcat methods of searching for mineral deposits may no longer be necessary.

The Department of Energy has contracted with the Laboratory to produce a geochemical atlas of the entire state of Alaska, using the sandwich format. Researchers believe the system could perhaps be adapted to include information on seismicity, gravity, and elevation, and so predict the probability of floods or earthquakes in various areas.

Earth and Space Sciences Division

SOVIETS COMMITTED TO NUCLEAR POWER

Short of a major policy switch, the Soviets will attempt a 10-fold increase in nuclear power by the year 2000, keeping the nuclear genie a factor in world negotiations regardless of public pressure in the West to dispense with this form of energy.

That's the conclusion of a Los Alamos report, "Soviet Nuclear Power," written by physicist William G. Davey.

Davey says although nuclear power is withering away in the United States—a victim of criticism and public pressure—the Soviet Union is demonstrating a commitment to nuclear power for electricity generation that is unmatched anywhere but France. Unless Soviet leadership changes its convictions, a 10-fold increase in generating capacity is likely in the next two decades. This 20-year projection is consistent with figures available for the period from 1971 to 1985. The figures show that the Soviet Union is doubling nuclear electrical generating capacity every five years. Official statements, which may be inflated, say that by 1985 the total electrical generating capacity will reach 1.6 trillion kilowatt-hours, with 14 per cent of that being nuclear and 15 per cent hydroelectric. Davey believes the 20-year projection is a conservative one in that it makes no allowance for special-purpose or fast breeder reactors, which were not included in the study because the detailed information was not available.

"The Soviets have an ingrained belief that large-scale high technology should support the national industrial base," Davey writes. "Also, they believe in political systems that allow centralized technical decisions to be implemented without regard to local or societal concerns. They recognize that nuclear fuel, because it is very compact, is especially valuable in a nation such as Russia, which encompasses great distances and harsh climates. And, finally, the Soviets maintain iron control over nuclear weapons' potential, when reactors are exported."

Office of the Assistant Director
for Planning and Analysis

REACTOR PROBE SHIPPED

A unique tool for diagnosing the physical events in loss-of-coolant experiments in nuclear reactor simulators has been shipped to Germany. The tool, a sophisticated videoprobe system, will be used at the PKL Facility at Erlangen, West Germany. The system is the third produced here; two others are in use at a similar facility in Japan under the Nuclear Regulatory Commission's multinational experimental and analytical research program known as 2D/3D.

Walter Kirchner says experiments in electrically heated reactor simulators reproduce the extremely harsh environment of a large reactor core. The video probe must withstand pressures varying from 600 pounds per square inch to atmospheric and rapid thermal shocks from 660 degrees Fahrenheit to ambient.

The probe looks a little like a fat Gatling gun. A miniature television camera, developed for military use and chosen for its extremely high resolution, is mounted in the barrel, and the barrel is inserted into the wall of the reactor simulator. Pictures are taken through a 1-inch-diameter window and lighted by miniature halogen lamps through four even smaller windows. Xenon strobe light is transmitted through bundles of fiber optics and synchronized at 30 frames per second to the TV camera.

A gold-plated annulus surrounds the jacket that holds the video equipment; it is filled with xenon gas to protect the delicate equipment and minimizes disturbance to the experimental reactor facility from the intrusion of the probe. Less than 300 watts of heat are removed from the simulator by the Los Alamos equipment—minimal disturbance.

Probes in use at the Japan Atomic Energy Research Institute at Tokai have worked well for more than a year. A similar probe, although not developed at Los Alamos, was recently used to examine the core of the crippled reactor on Three Mile Island.

Energy Division

BRITISH-AMERICAN PATENT ISSUED

Although joint patents between inventors from different countries are not rare, they are fairly uncommon in weapons-research establishments. A joint patent for a new detonator that is extremely safe and efficient has been issued to Los Alamos staff member Robert H. Dinegar and John Kirkham of the British Atomic Weapons Research Establishment. Their low-voltage, non-primary-explosive detonator has innumerable applications for defense and industry.

Most detonators, such as commercial blasting caps, use a sensitive primary explosive that propagates a shock wave to detonate an explosive charge. Commonly used detonator explosives are lead azide and lead styphnate, and both are touchy materials.

"Traditionally, detonators must be handled carefully and the best of them are rather dangerous," Dinegar says. "We have looked for years for a detonator that was efficient, small enough to be useful, but much safer than existing models. We believe the device we have patented meets these criteria."

The new detonator employs a configuration that provides confinement for a deflagration-to-detonation sequence (burning that builds up into a detonation) and uses a much less sensitive explosive. The device has obvious implications for safer weapons and also has potential for many industrial applications, especially in energy research. Dinegar says geothermal energy development and oil and gas exploration are obvious areas of interest. The detonator can also be used as a valve actuator to drive a piston. Research on this application is now underway.

Dynamic Testing Division

PLUTONIUM RECOVERY TECHNIQUE IMPROVED

A new technique for electrorefining of metal has allowed Los Alamos to almost double its recovery of pure plutonium metal. Electrorefining has been an effective tool for recycling metallic plutonium scrap since 1964, but the amount of plutonium processed at one time has been limited to 4 kilograms because of criticality considerations. The new procedure allows researchers to process up to 6 kilograms of plutonium without compromising safety, yield, or quality. The process costs about $45,000 and can yield $500,000 worth of plutonium. Principal investigator Lawrence Mullins claims the new method results in recovery of 82 per cent of the plutonium in scrap metal. The recovered plutonium has an average purity of 99.96 per cent.

Plutonium recycling at Los Alamos dates back to 1943. As greater amounts of scrap were generated in radiochemistry programs, reactors, and defense research, recovery became increasingly important. To meet its demands, the Laboratory switched from a slower, more complicated aqueous processing method to electrorefining. In aqueous processing the scrap metal was dissolved in acid and purified by solvent extraction. Additional steps were then necessary to convert the purified plutonium to metal. In electrorefining the impure metal is heated in a crucible to 750 degrees Celsius; then an electric current is applied. The plutonium, which is first oxidized and then reduced, drips to the bottom of the container. When it cools, the solidified pure metal is removed as a ring.

Los Alamos has also perfected a new process for direct reduction of plutonium oxide to plutonium metal. The processing of oxide by this method coupled with electrorefining uses less manpower and permits higher throughput than conventional aqueous methods.

Materials Science and Technology Division

SPACE BENEFITS IDENTIFIED

Consumers are reaping enormous benefits—more than $300 million a year—from the aggressive development of space, but the United States, which pioneered much of the research, may lose its competitive edge if the present research funding trend continues.

This view is expressed in a Laboratory report by Herbert "Bill" Lorber and Robert H. Drake. The authors conclude that NASA, the agency that put man on the moon, "finds itself trapped in the mundane activity of trying to reverse a negative cash flow."

Their report, "The Economic Benefits of Space Development," provides a glimpse of international space activity, summarizes the benefits of military and civilian space ventures, and characterizes areas of space research. Contrary to public perception, there is no single space program, but four separate efforts including military, intelligence, civilian-public, and commercial.

Most of the benefits to the consumer from the overall effort come from communications. A single Intelstat (international commercial communications) circuit has dropped from an annual cost of $64,000 in the 1960s to less than $6000 today. A rough estimate of the economic measure called "consumers' surplus" shows consumers are already enjoying a $300 million a year benefit in international telephone communications alone. While benefits are expected to increase as the industry grows, a

potential problem is the crowding at 23,000 miles from the earth, where communication satellites are placed in geosynchronous orbits. The Los Alamos economists say 97 satellite communication systems are in operation now, and because they must orbit at least two degrees apart, there is room for only 180.

The report describes other benefits that accrue from satellite observation systems. For example, the accuracy of routine weather forecasts has doubled since 1965 and improved monitoring of hurricanes and tornadoes has saved countless lives. Remote-sensing space systems such as Landsat bring direct economic benefits. A single, typical observation program of annual snowcover in the western United States, costing half a million dollars to perform, saves $50 million annually through improved water management in hydroelectric and irrigation systems.

Consumers can also look forward to savings from manufacturing in space. A 1984 launch is scheduled to manufacture high-yield semiconductors at dramatic cost reductions and gains in quality. The launch may also test low-gravity separation techniques for producing diabetic and antihemophilic drugs.

Analysis and Assessment Division

LASERS USED IN COAL GASIFIER

The original London Bridge was illuminated by coal-gas lamps. In the mid 1850s New York, Boston, and many other American cities were using town, or water, gas, as it was called. However, widespread use of coal gases fell victim to the development of cheap and abundant supplies of crude oil. Today there is growing interest in synthetic gas, but the coal gasification process must meet much more stringent requirements for cleanliness and environmental safety before it will be accepted as a leading technology.

To this end the Department of Energy has set up an experimental coal gasification system at its Morgantown Energy Technology Center in West Virginia. Now Los Alamos researchers have tested systems for on-line monitoring of gas composition and scrubber efficiency. The systems, involving laser-induced breakdown spectroscopy (LIBS) and coherent anti-Stokes Raman scattering (CARS), were set up in two sections of the experimental gasifier. Both worked well in the extremely harsh environmental conditions.

Lee Radziemski, David Cremers, and David Taylor describe LIBS as a straightforward method for determining the elemental composition of coal gasification streams—including products that are extremely corrosive. CARS is a sophisticated method for determining the presence and temperature of many molecules in the gasifier stream. The systems are complementary.

Both systems use readily available lasers. In LIBS pulses of laser light lasting 10 billionths of a second are focused down to a very small volume. The intense light creates a tiny fireball of hot plasma and reduces the material it strikes to its basic, elemental components. A multichannel analyzer can readily identify the resultant atomic spectra. Nitrogen, oxygen, hydrogen, and sulfur have been identified in real-time monitoring of the gasifier effluent, an improvement over the 10-minute turnaround time of existing monitoring equipment.

CARS relies on the fact that different molecules have different vibrational frequencies. By mixing two laser beams of different frequencies in the effluent gas, it is possible to stimulate the molecular species of interest to emit coherent radiation of a third frequency. The intensity of this third beam yields the species concentration. The CARS technique was successful in measuring concentrations of nitrogen, carbon monoxide, and hydrogen sulfide in a very "dirty" part of the gas stream, where the temperature is normally 1000 degrees Fahrenheit, the pressure is 200 pounds per square inch, and particle and tar-vapor loadings are high.

Chemistry Division

PROTON STORAGE RING PLANNED

By 1985 a major addition to the Weapons Neutron Facility will be operating at Los Alamos. Ground has been broken for a Proton Storage Ring, a $19 million facility designed to assist both weapons research and basic research in physics and materials science.

The ring will accept and store protons from the half-mile-long accelerator at the Los Alamos Meson Physics Facility (LAMPF). The stored protons will be released in short bursts to generate intense fluxes of neutrons at the Weapons Neutron Research target facility. The neutrons will then be employed in a wide range of basic nuclear research and practical applications. George Sawyer, construction manager, says the ring will make the Weapons Neutron Facility "the most powerful pulsed-neutron source anywhere, over a very broad neutron energy range." With this intense source researchers will be able, for the first time, to perform certain neutron experiments in nuclear physics, solid-state physics, and condensed matter science and to study high-current phenomena important to accelerator applications and fusion experiments.

LAMPF produces medium-energy protons at a higher intensity than any other accelerator in the world. It is a pulsed beam with an average current of 600 microamperes. When the proton storage ring is complete, it will receive every tenth pulse from the accelerator, accumulating the protons until the number of particles in the ring reaches 50 trillion.

Accelerator Technology Division

☆ U.S. GOVERNMENT PRINTING OFFICE: 1982 — 576-088/1