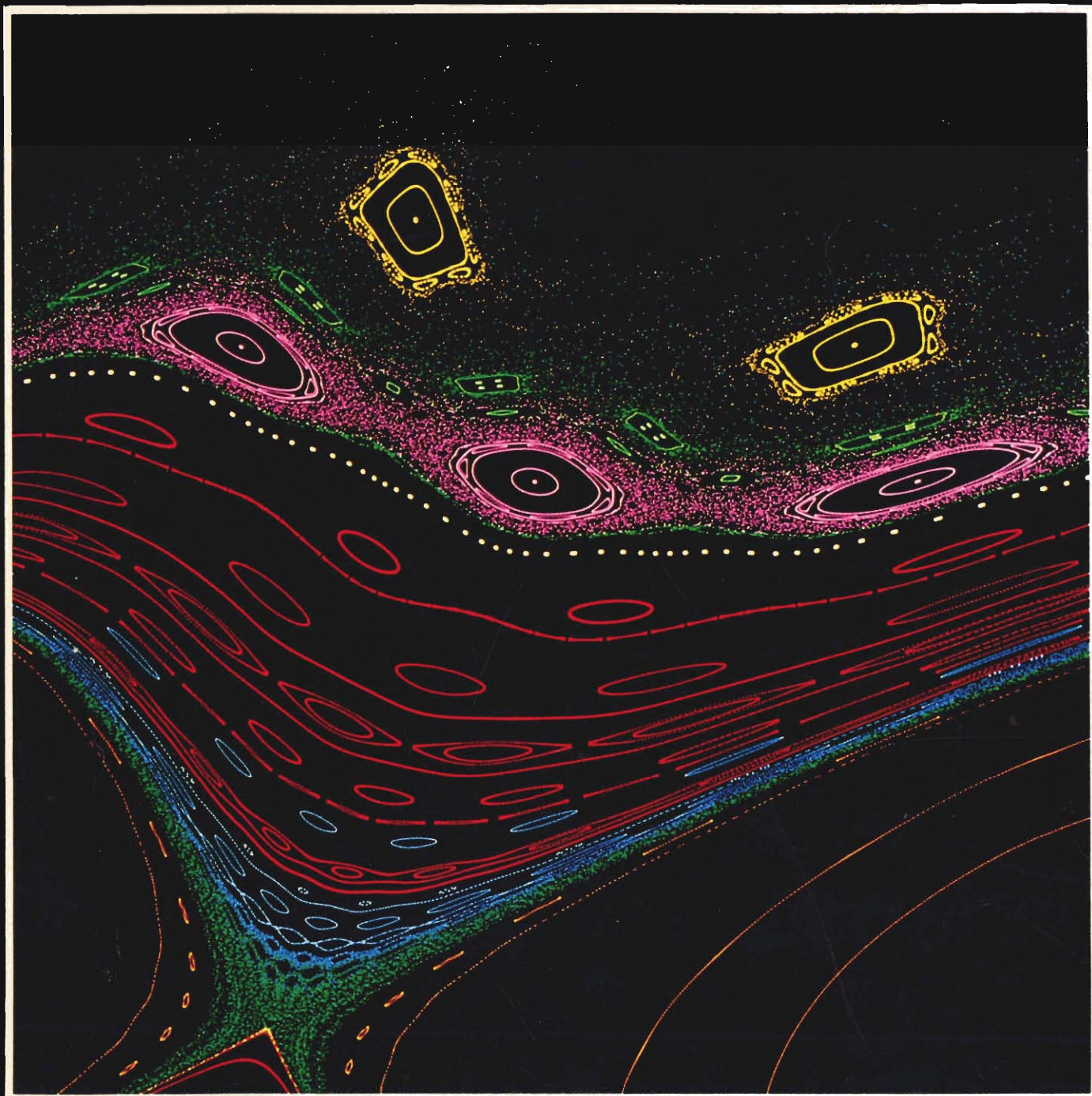


PREMIER ISSUE

SUMMER 1980

# Los Alamos Science

LOS ALAMOS SCIENTIFIC LABORATORY





**EDITOR**

Necia Grant Cooper

**ART DIRECTOR, DESIGN,  
PRODUCTION**

M. A. Garcia

**ASSOCIATE EDITOR**

Betty Leffler

**EDITORIAL STAFF**

DyAnne J. Short (Editorial Assistant)

Elaine Stanlick (Assistant Editor)

Dorothy Amsden (Contributing Editor)

**PHOTOGRAPHY**

LeRoy Sanchez (Portraits)

Group ISD-7

**ILLUSTRATION**

Anita Flores

Jim Mahan

Vicki Hartford

Gail Flower

Group ISD-3

Les Speir (Cartoonist)

**COVER GRAPHICS**

Bob Hotchkiss

**PRODUCTION**

Jo Ann Painter (Photocomposition)

Barbara Velarde (Photocomposition)

Judy Carlson

Gloria Riepe

William H. Regan

(Printing and Technical Coordination)

# Los Alamos Science

LOS ALAMOS SCIENTIFIC LABORATORY

Volume 1, Number 1

## CONTENTS

### RESEARCH AND REVIEW

- Universal Behavior in Nonlinear Systems** by Mitchell J. Feigenbaum 4
- High-Temperature Superconductivity: A Metallurgical Approach** 28  
by Angelo L. Giorgi, Gregory R. Stewart, James L. Smith, and Bernd T. Matthias
- Flow Cytometry: A New Tool for Quantitative Cell Biology** 40  
by L. Scott Cram, Dale M. Holm, and Paul F. Mullaney

### SHORT SUBJECTS

- Heavy Ion Fusion** by George Sawyer 59
- The Search for Muon-Number Violation at LAMPF** 62  
by Cy Hoffman and Minh Duong-Van

### PROGRAM FEATURE—NUCLEAR SAFEGUARDS

- Nuclear Safeguards—A Global Issue** by G. Robert Keepin 68
- Nondestructive Assay for Nuclear Safeguards** 88  
by Roddy B. Walton and Howard O. Menlove
- Dynamic Materials Accounting Systems** 116  
by Darryl B. Smith, Dante Stirpe, and James P. Shipley

### EDITORIAL

- Don Kerr on Nuclear Safeguards** 138

### SCIENCE IDEAS

- Plate Tectonics—Where the Action Is** by Bob Riecker 140

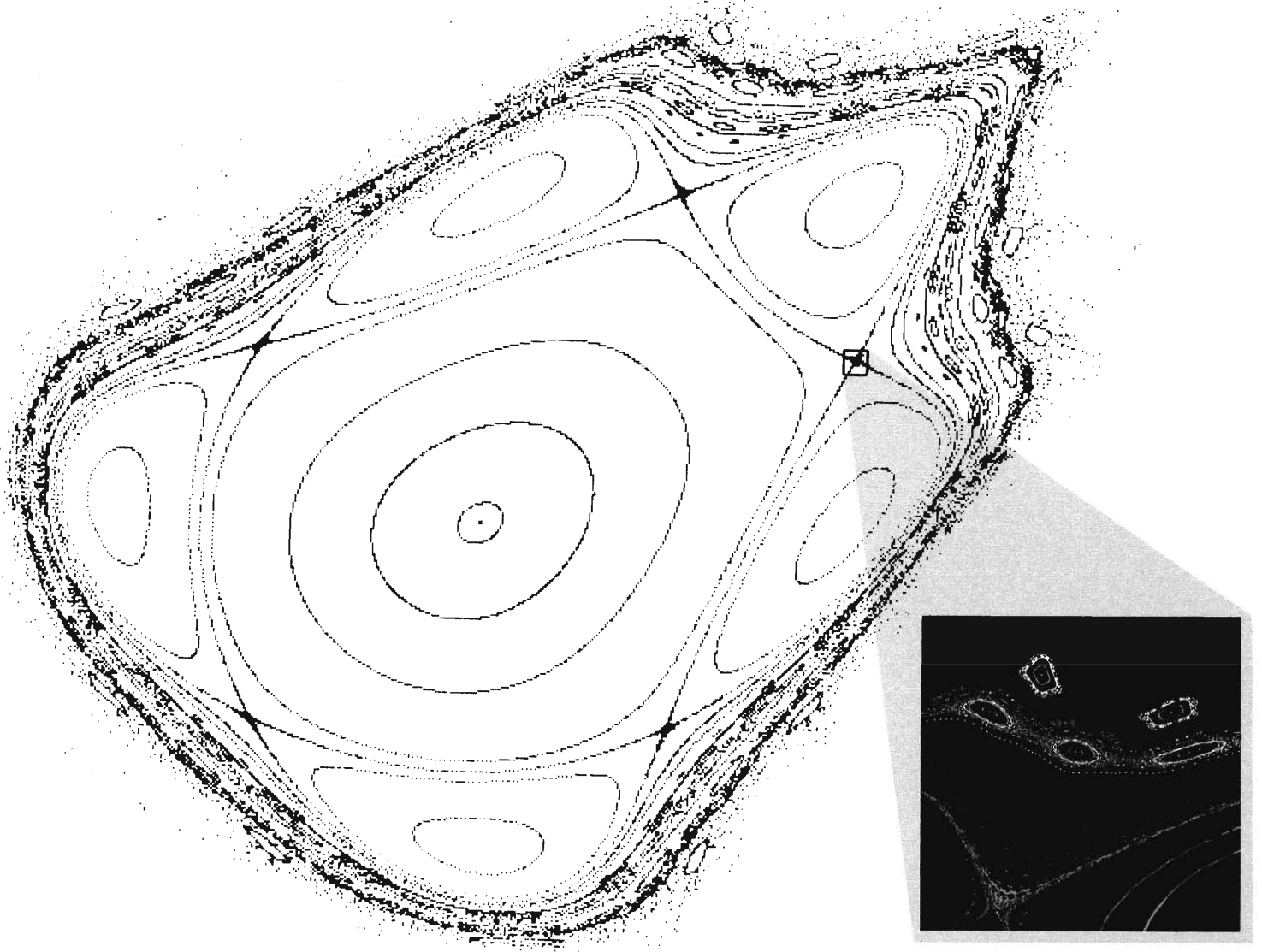
### PEOPLE

- Memorial to Professor (Fredrik) William H. Zachariasen** 146  
by Robert A. Penneman

Published 4 times a year by the University of California, Los Alamos Scientific Laboratory, Bldg. SM-43, Casa Grande Drive. Second class mailing permit pending approval at the Los

Alamos, NM Post Office.  
Address mail to MS-103, P. O. Box 1663, Los Alamos, NM 87545. Telephone (505)667-3905.  
Los Alamos Scientific Laboratory, an affirmative

action/equal opportunity employer, is operated by the University of California for the United States Department of Energy, under contract W-7405-ENG-36 Modification No. 187.



**ON THE COVER:** The cover figure was chosen to illustrate how an elementary two-dimensional transformation can exhibit unfathomably complex behavior — it is actually a blow-up of the boxed region in the figure shown alongside it.

The transition from simple to complex or pseudorandom behavior through period doubling in a large class of nonlinear dissipative systems can now be described quantitatively with the universality theory invented by Mitchell Feigenbaum and described in "Universal Behavior of Nonlinear Systems." The transformation depicted in the figure is an area-preserving map corresponding to an energy-preserving rather than a dissipative system. This type of system is known experimentally to proceed through period doubling, and Mitchell hopes to extend the universality theory to handle this Hamiltonian-type system.

The figure was constructed by taking first one and then another initial point  $(x,y)$  and iterating it under the transformation

$$x' = x^2 - y$$

$$y' = a + x$$

where  $a = -0.4224$ . This quadratic transformation preserves elements of area and can be viewed as a model Poincaré map of a Hamiltonian system; for example, a piston supported under gravity by a one-particle gas produces a similar Poincaré map.

The 50 or so initial points used to generate the cover figure can be decomposed into two sets. Those of the first set produce iterates that lie on ellipse-like curves. For example, three initial points produced, respectively, the central yellow

low dots, the smaller inside yellow ellipses, and the larger yellow ellipses inside the chain of nine yellow "island" ellipses. One initial point produced the central dots of both yellow clusters (as well as those of similar yellow clusters outside the viewing area). These points of the second set produce a "haze" of iterates, which wander randomly through certain regions of the figure. For example, the lavender haze surrounding the three lavender island clusters is produced by just one initial point. It is this haze that corresponds to the statistical behavior of the system. The complexity is far greater than illustrated here. In fact, a magnification of the lavender haze midway between two of the islands would show a similar constellation of islands and haze, and so forth, *ad infinitum*.

This behavior may be viewed as the consequence of an instability — such as in a plasma — that has led to an uncontrolled behavior of the system. In the Hamiltonian analogue, the whole plane is a surface of constant energy. A system obeying statistical mechanics should have wandered over the entire plane. Evidently the present system fails to do so, with a given point often found to circulate around one of the elliptical islands, or "invariant tori." The presence of these tori constitutes a violation of statistical mechanics, and their presence turns out to be insured by a famous theorem (KAM). A hope to establish the validity of statistical mechanics for a classical system is to determine how these tori may be destroyed with increasing dimensionality. Since period doubling is the route that successively destroys these tori, an extension of the present universality theory may describe this process quantitatively. The theory's extension is not yet worked out, but work is in progress. Apart from some understanding of how a Hamiltonian system comes to behave ergodically, there is the hope in the opposite direction that, having comprehended a general route to instability, an understanding of how to prevent this might come about. Such information is, of course, of crucial significance in the problem of magnetic fusion.

# INSIDE THIS ISSUE

## EDITOR'S NOTE

The impact of scientific discovery on the nature of our world has nowhere been felt more poignantly than by the first scientists at Los Alamos. Since the 1940s the Los Alamos Scientific Laboratory (LASL) has changed dramatically, enlarging and diversifying to the point where it is no longer just a weapons laboratory. Today it is a multipurpose institution concerned with the advancement of science and the solution of energy and defense problems. As a national laboratory, LASL seeks a complementary role vis-a-vis academic research, industrial development, and government policy, with the nature of that role still emerging. Clearly though, its many unique and significant contributions to science, to technology, and to society will continue to influence decisions on a broad range of national issues.

In this magazine, we hope to provide a forum for scientists and engineers at LASL to present their work to each other and to the wider community in a fashion that promotes understanding. This entails sharing personal insights into the broader implications of the work and the political issues that provide the context for it.

The exciting interplay between basic and applied research at LASL provides fertile ground for new ideas. The cover story, "Universal Behavior in Nonlinear Systems," is a good example. Upon coming to Los Alamos, Mitchell Feigenbaum was urged by Peter Carruthers, Theoretical Division Leader, to attempt an investigation into the nature of turbulence, a phenomenon that inhibits progress in many major technological programs. The result is the universality theory and the first quantitative understanding of how causal systems develop chaotic behavior. The ideas are new and fascinating and Mitchell has used his rare pedagogical skills to present them here. The presentation is by necessity much more technical than the others in this issue, but is well worth the reader's time.

"High-Temperature Superconductivity" is somewhat like a detective story, beginning with a mystery, following the false leads, and finally coming upon an intriguing solution. We began planning this article amidst conjecture and confusion as to the true interpretation, but the authors' enthusiasm ran high as they rapidly exploited a variety of measurement techniques available to them at LASL to piece together all the elements of the puzzle.

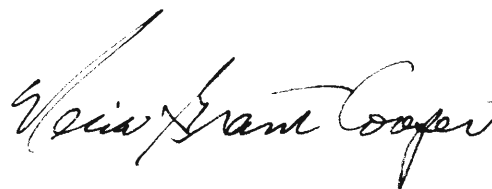
The focus on nuclear technology at LASL has always included studies of radiation effects on biological organisms. In the late 1960s, the biologists' need to understand radiation effects at the cellular level combined with the physicists' techniques for measurement and analysis to create a new research tool.

Known as the flow cytometer, this extraordinary instrument interrogates and sorts thousands of cells per second with high statistical accuracy. It is being used to investigate a host of problems in cell biology, including the morphology of individual chromosomes and the changes that accompany the transformation of a cell from normal to malignant. The story of its invention and its application to problems in cancer research are described in this issue.

In the midst of the raging antinuclear debate in this country and the expansion of the nuclear industry abroad, we present as our program feature the Nuclear Safeguards Program. Bob Keepin, considered by many to be the father of modern safeguards measurement technology, and a long-time participant in international negotiations on nuclear power and safeguards agreements, gives us a historical perspective on safeguards issues and a personal insight into LASL's contributions. The nuts and bolts of the LASL program are described in the two follow-up articles: an in-depth review of nondestructive measurement technology and its use by the nuclear industry, and a step-by-step description of how one designs a materials accounting system to deter a knowledgeable insider from diverting strategic nuclear material. An editorial by Don Kerr, LASL's Director, urges support for safeguards technology development and implementation in these troubled times of political terrorism and economic uncertainty.

The Laboratory's history is notable for the continued participation of great men of science. Willy Zachariassen was one, and from Bob Penneman's moving memorial to his long-time friend and collaborator we learn more about the man, his work, and his unique contribution to the Manhattan Project.

Few things are more stimulating than hearing one's implicit assumptions identified and challenged by new points of view. We encourage our readers to respond by contributing their own perspectives on the issues and ideas that appear in *Los Alamos Science*.



# *Universal Behavior in Nonlinear Systems*

by Mitchell J. Feigenbaum

Universal numbers,  
 $\delta = 4.6692016\dots$   
and  
 $\alpha = 2.502907875\dots$ ,  
determine quantitatively  
the transition from  
smooth to turbulent or  
erratic behavior  
for a large class of  
nonlinear systems.

There exist in nature processes that can be described as complex or chaotic and processes that are simple or orderly. Technology attempts to create devices of the simple variety: an idea is to be implemented, and various parts executing orderly motions are assembled. For example, cars, airplanes, radios, and clocks are all constructed from a variety of elementary parts each of which, ideally, implements one ordered aspect of the device. Technology also tries to control or minimize the impact of seemingly disordered processes, such as the complex weather patterns of the atmosphere, the myriad whorls of turmoil in a turbulent fluid, the erratic noise in an electronic signal, and other such phenomena. It is the complex phenomena that interest us here.

When a signal is noisy, its behavior from moment to moment is irregular and has no simple pattern of prediction. However, if we analyze a sufficiently long record of the signal, we may find that signal amplitudes occur within narrow ranges a definite fraction of the time. Analysis of another record of the signal may reveal the same fraction. In this case, the noise can be given a *statistical* description. This means that while it is impossible to say what amplitude will appear next in succession, it is possible to estimate the probability or likelihood that the signal will attain some specified range of values. Indeed, for the last hundred years disorderly processes have been taken to be statistical (one has

given up asking for a precise causal prediction), so that the goal of a description is to determine what the probabilities are, and from this information to determine various behaviors of interest—for example, how air turbulence modifies the drag on an airplane.

We know that perfectly definite causal and *simple* rules can have statistical (or random) behaviors. Thus, modern computers possess “random number generators” that provide the statistical ingredient in a simulation of an erratic process. However, this generator does nothing more than shift the decimal point in a rational number whose repeating block is suitably long. Accordingly, it is possible to predict what the *n*th generated number will be. Yet, in a list of successive generated numbers there is such a seeming lack of order that all statistical tests will confer upon the numbers a pedigree of randomness. Technically, the term “pseudorandom” is used to indicate this nature. One now may ask whether the various complex processes of nature themselves might not be merely pseudorandom, with the full import of randomness, which is untestable, a historic but misleading concept. Indeed our purpose here is to explore this possibility. What will prove altogether remarkable is that some very simple schemes to produce erratic numbers behave *identically* to some of the erratic aspects of natural phenomena. More specifically, there is now cogent evidence that the problem of how a fluid changes over from smooth to turbulent

flow can be solved through its relation to the simple scheme described in this article. Other natural problems that can be treated in the same way are the behavior of a population from generation to generation and the noisiness of a large variety of mechanical, electrical, and chemical oscillators. Also, there is now evidence that various Hamiltonian systems—those subscribing to classical mechanics, such as the solar system—can come under this discipline.

The feature common to these phenomena is that, as some external parameter (temperature, for example) is varied, the behavior of the system changes from simple to erratic. More precisely, for some range of parameter values, the system exhibits an orderly *periodic* behavior; that is, the system's behavior reproduces itself every *period* of time  $T$ . Beyond this range, the behavior fails to reproduce itself after  $T$  seconds; it almost does so, but in fact it requires *two* intervals of  $T$  to repeat itself. That is, the period has *doubled* to  $2T$ . This new periodicity remains over some range of parameter values until another critical parameter value is reached after which the behavior *almost* reproduces itself after  $2T$ , but in fact, it now requires  $4T$  for reproduction. This process of successive period doubling recurs continually (with the range of parameter values for which the period is  $2^n T$  becoming successively smaller as  $n$  increases) until, at a certain value of the parameter, it has doubled *ad infinitum*, so that the behavior is no longer periodic. Period doubling is then a characteristic route for a system to follow as it changes over from simple periodic to complex aperiodic motion. All the phenomena mentioned above exhibit period doubling. In the limit of aperiodic behavior, there is a unique and hence *universal* solution common to all systems undergoing period doubling. This fact implies remarkable consequences. For a given system, if we

denote by  $\Lambda_n$  the value of the parameter at which its period doubles for the  $n$ th time, we find that the values  $\Lambda_n$  converge to  $\Lambda_\infty$  (at which the motion is aperiodic) *geometrically* for large  $n$ . This means that

$$\Lambda_\infty - \Lambda_n \propto \delta^{-n} \quad (1)$$

for a fixed value of  $\delta$  (the *rate* of onset of complex behavior) as  $n$  becomes large. Put differently, if we define

$$\delta_n \equiv \frac{\Lambda_{n+1} - \Lambda_n}{\Lambda_{n+2} - \Lambda_{n+1}}, \quad (2)$$

$\delta_n$  (quickly) approaches the constant value  $\delta$ . (Typically,  $\delta_n$  will agree with  $\delta$  to several significant figures after just a few period doublings.) What is quite remarkable (beyond the fact that there is always a geometric convergence) is that, for all systems undergoing this period doubling, the value of  $\delta$  is *predetermined* at the universal value

$$\delta = 4.6692016 \dots \quad (3)$$

Thus, this definite number must appear as a natural rate in oscillators, populations, fluids, and all systems exhibiting a period-doubling route to turbulence! In fact, most measurable properties of *any* such system in this aperiodic limit now can be determined, in a way that essentially bypasses the details of the equations governing each specific system because the theory of this behavior is universal over such details. That is, so long as a system possesses certain *qualitative* properties that enable it to undergo this route to complexity, its *quantitative* properties are determined. (This result is analogous to the results of the modern theory of critical phenomena, where a few qualitative properties of the system undergoing a phase transition, notably the dimensionality, determine *universal* critical ex-

ponents. Indeed at a *formal* level the two theories are identical in that they are fixed-point theories, and the number  $\delta$ , for example, can be viewed as a critical exponent.) Accordingly, it is sufficient to study the simplest system exhibiting this phenomenon to comprehend the general case.

### Functional Iteration

A random number generator is an example of a simple iteration scheme that has complex behavior. Such a scheme generates the next pseudorandom number by a definite transformation upon the present pseudorandom number. In other words, a certain function is reevaluated successively to produce a sequence of such numbers. Thus, if  $f$  is the function and  $x_0$  is a starting number (or "seed"), then  $x_0, x_1, \dots, x_n, \dots$ , where

$$\begin{aligned} x_1 &= f(x_0) \\ x_2 &= f(x_1) \\ &\vdots \\ x_{n+1} &= f(x_n) \end{aligned} \quad (4)$$

is the sequence of generated pseudorandom numbers. That is, they are generated by *functional iteration*. The  $n$ th element in the sequence is

$$x_n = f(f(\dots f(f(x_0)) \dots)) \equiv f^n(x_0), \quad (5)$$

where  $n$  is the total number of applications of  $f$ . [ $f^n(x)$  is not the  $n$ th power of  $f(x)$ ; it is the  $n$ th *iterate* of  $f$ .] A property of iterates worthy of mention is

$$f^n(f^m(x)) = f^m(f^n(x)) = f^{m+n}(x), \quad (6)$$

since each expression is simply  $m + n$  applications of  $f$ . It is understood that

$$f^0(x) = x. \quad (7)$$

It is also useful to have a symbol,  $\circ$ , for functional iteration (or composition), so that

$$f^n \circ f^m = f^m \circ f^n = f^{m+n}. \quad (8)$$

Now  $f^n$  in Eq. (5) is itself a definite and computable function, so that  $x_n$  as a function of  $x_0$  is known in principle.

If the function  $f$  is *linear* as, for example,

$$f(x) = ax \quad (9)$$

for some constant  $a$ , it is easy to see that

$$f^n(x) = a^n x, \quad (10)$$

so that, for this  $f$ ,

$$x_n = a^n x_0 \quad (11)$$

is the solution of the *recurrence relation* defined in Eq. (4),

$$x_{n+1} = ax_n. \quad (12)$$

Should  $|a| < 1$ , then  $x_n$  geometrically converges to zero at the rate  $1/a$ . This example is special in that the linearity of  $f$  allows for the explicit computation of  $f^n$ .

We must choose a *nonlinear*  $f$  to generate a pseudorandom sequence of numbers. If we choose for our nonlinear  $f$

$$f(x) = a - x^2, \quad (13)$$

then it turns out that  $f^n$  is a polynomial in  $x$  of order  $2^n$ . This polynomial rapidly becomes unmanageably large; moreover, its coefficients are polynomials in  $a$  of order up to  $2^{n-1}$  and become equally difficult to compute. Thus even if  $x_0 = 0$ ,  $x_n$  is a polynomial in  $a$  of order  $2^{n-1}$ . These polynomials are nontrivial as can be surmised from the fact that for certain

values of  $a$ , the sequence of numbers generated for almost all starting points in the range  $(a - a^2, a)$  possess *all* the mathematical properties of a random sequence. To illustrate this, the figure on the cover depicts the iterates of a similar system in two dimensions:

$$\begin{aligned} x_{n+1} &= y_n - x_n^2 \\ y_{n+1} &= a - x_n. \end{aligned} \quad (14)$$

Analogous to Eq. (4), a starting coordinate pair  $(x_0, y_0)$  is used in Eq. (14) to determine the next coordinate  $(x_1, y_1)$ . Equation (14) is reapplied to determine  $(x_2, y_2)$  and so on. For some initial points, all iterates lie along a definite elliptic curve, whereas for others the iterates are distributed "randomly" over a certain region. It should be obvious that no explicit formula will account for the vastly rich behavior shown in the figure. That is, while the iteration scheme of Eq. (14) is trivial to specify, its  $n$ th iterate as a function of  $(x_0, y_0)$  is unavailable. Put differently, applying the simplest of *nonlinear* iteration schemes to itself sufficiently many times can create vastly complex behavior. Yet, precisely because the same operation is reapplied, it is conceivable that only a select few self-consistent patterns might emerge where the consistency is determined by the key notion of iteration and *not* by the particular function performing the iterates. These self-consistent patterns do occur in the limit of infinite period doubling and in a well-defined intricate organization that can be determined *a priori* amidst the immense complexity depicted in the cover figure.

### The Fixed-Point Behavior of Functional Iterations

Let us now make a direct onslaught against Eq. (13) to see what it possesses. We want to know the behavior of the system after many iterations. As we

already know, high iterates of  $f$  rapidly become very complicated. One way this growth can be prevented is to have the first iterate of  $x_0$  be precisely  $x_0$  itself. Generally, this is impossible. Rather this condition *determines* possible  $x_0$ 's. Such a self-reproducing point is called a *fixed point* of  $f$ . The sequence of iterates is then  $x_0, x_0, x_0, \dots$  so that the behavior is *static*, or if viewed as periodic, it has period 1.

It is elementary to determine the fixed points of Eq. (13). For future convenience we shall use a modified form of Eq. (13) obtained by a translation in  $x$  and some redefinitions:

$$f(x) = 4\lambda x(1 - x), \quad (15)$$

so that as  $\lambda$  is varied,  $x = 0$  is always a fixed point. Indeed, the fixed-point condition for Eq. (15),

$$x^* = f(x^*) = 4\lambda x^*(1 - x^*), \quad (16)$$

gives as the two fixed points

$$x^* = 0, x_0^* = 1 - 1/4\lambda. \quad (17)$$

The maximum value of  $f(x)$  in Eq. (15) is attained at  $x = 1/2$  and is equal to  $\lambda$ . Also, for  $\lambda > 0$  and  $x$  in the interval  $(0, 1)$ ,  $f(x)$  is always positive. Thus, if  $\lambda$  is anywhere in the range  $[0, 1]$ , then any iterate of any  $x$  in  $(0, 1)$  is also always in  $(0, 1)$ . Accordingly, in all that follows we shall consider only values of  $x$  and  $\lambda$  lying between 0 and 1. By Eq. (16) for  $0 \leq \lambda < 1/4$ , only  $x^* = 0$  is within range, whereas for  $1/4 \leq \lambda \leq 1$ , both fixed points are within the range. For example, if we set  $\lambda = 1/2$  and we start at the fixed point  $x_0^* = 1/2$  (that is, we set  $x_0 = 1/2$ ), then  $x_1 = x_2 = \dots = 1/2$ ; similarly if  $x_0 = 0$ ,  $x_1 = x_2 = \dots = 0$ , and the problem of computing the  $n$ th iterate is obviously trivial.

What if we choose an  $x_0$  *not* at a fixed point? The easiest way to see what happens is to perform a graphical analysis. We graph  $y = f(x)$  together with  $y = x$ .



Where the lines intersect we have  $x = f(x)$ , so that the intersections are precisely the fixed points. Now, if we choose an  $x_0$  and plot it on the x-axis, the ordinate of  $f(x)$  at  $x_0$  is  $x_1$ . To obtain  $x_2$ , we must transfer  $x_1$  to the x-axis before reapplying  $f$ . Reflection through the straight line  $y = x$  accomplishes precisely this operation. Altogether, to iterate an initial  $x_0$  successively,

1. move *vertically* to the graph of  $f(x)$ ,
2. move *horizontally* to the graph of  $y = x$ , and
3. repeat steps 1, 2, etc.

Figure 1 depicts this process for  $\lambda = \frac{1}{2}$ . The two fixed points are circled, and the first several iterates of an arbitrarily chosen point  $x_0$  are shown. What should be obvious is that if we start from any  $x_0$  in  $(0,1)$  ( $x = 0$  and  $x = 1$  excluded), upon continued iteration  $x_n$  will converge to the fixed point at  $x = \frac{1}{2}$ . No matter how close  $x_0$  is to the fixed point at  $x = 0$ , the iterates diverge away from it. Such a fixed point is termed *unstable*. Alternatively, for almost all  $x_0$  near enough to  $x = \frac{1}{2}$  [in this case, all  $x_0$  in  $(0,1)$ ], the iterates converge towards  $x = \frac{1}{2}$ . Such a fixed point is termed *stable* or is referred to as an *attractor* of period 1.

Now, if we don't care about the *transient* behavior of the iterates of  $x_0$ , but only about some regular behavior that will emerge eventually, then knowledge of the stable fixed point at  $x = \frac{1}{2}$  satisfies our concern for the *eventual* behavior of the iterates. In this restricted sense of eventual behavior, the existence of an attractor determines the solution *independently* of the initial condition  $x_0$  provided that  $x_0$  is within the *basin of attraction* of the attractor; that is, that it *is* attracted. The attractor satisfies Eq. (16), which is explicitly independent of  $x_0$ . This condition is the basic theme of universal behavior: if an attractor exists, the eventual behavior is independent of the starting point.

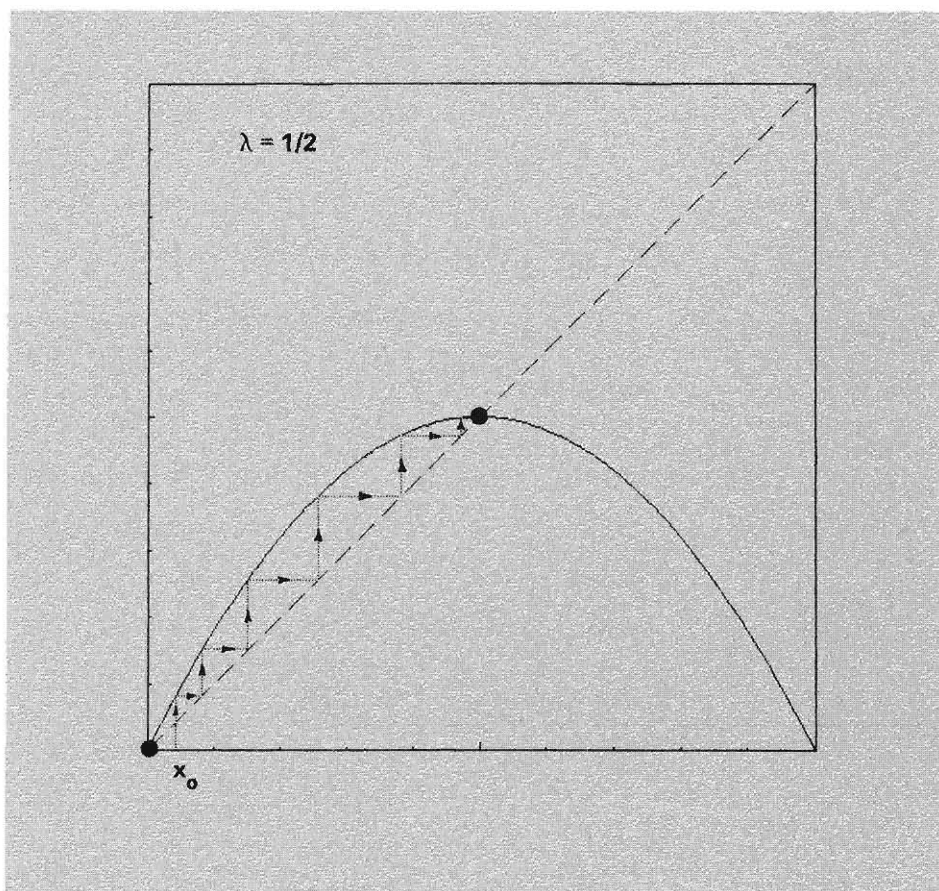


Fig. 1. Iterates of  $x_0$  at  $\lambda = 0.5$ .

What makes  $x = 0$  unstable, but  $x = \frac{1}{2}$  stable? The reader should be able to convince himself that  $x = 0$  is unstable because *the slope of  $f(x)$  at  $x = 0$  is greater than 1*. Indeed, if  $x^*$  is a fixed point of  $f$  and the derivative of  $f$  at  $x^*$ ,  $f'(x^*)$ , is smaller than 1 in absolute value, then  $x^*$  is stable. If  $|f'(x^*)|$  is greater than 1, then  $x^*$  is unstable. Also, only *stable* fixed points can account for the eventual behavior of the iterates of an arbitrary point.

We now must ask, "For what values of  $\lambda$  are the fixed points attracting?" By Eq. (15),  $f'(x) = 4\lambda(1 - 2x)$  so that

$$f'(0) = 4\lambda \quad (18)$$

and

$$f'(x_0^*) = 2 - 4\lambda. \quad (19)$$

For  $0 < \lambda < \frac{1}{4}$ , only  $x^* = 0$  is stable. At  $\lambda = \frac{1}{4}$ ,  $x_0^* = 0$  and  $f'(x_0^*) = 1$ . For  $\frac{1}{4} < \lambda < \frac{3}{4}$ ,  $x^*$  is unstable and  $x_0^*$  is stable, while at  $\lambda = \frac{3}{4}$ ,  $f'(x_0^*) = -1$  and  $x_0^*$  also has become unstable. Thus, for  $0 < \lambda < \frac{3}{4}$ , the eventual behavior is known.

## Period 2 from the Fixed Point

What happens to the system when  $\lambda$  is in the range  $\frac{3}{4} < \lambda < 1$ , where there are no attracting fixed points? We will see that as  $\lambda$  increases slightly beyond  $\lambda = \frac{3}{4}$ ,  $f$  undergoes period doubling. That is, instead of having a stable cycle of period 1 corresponding to one fixed point, the system has a stable cycle of period 2; that is, the cycle contains two points. Since these two points are fixed points of the function  $f^2$  ( $f$  applied twice) and since stability is determined by the slope of a function at its *fixed* points, we must now focus on  $f^2$ . First, we examine a graph of  $f^2$  at  $\lambda$  just below  $\frac{3}{4}$ . Figures 2a and b show  $f$  and  $f^2$ , respectively, at  $\lambda = 0.7$ .

To understand Fig. 2b, observe first that, since  $f$  is symmetric about its maximum at  $x = \frac{1}{2}$ ,  $f^2$  is also symmetric about  $x = \frac{1}{2}$ . Also,  $f^2$  must have a fixed point whenever  $f$  does because the second iterate of a fixed point is still that same point. The main ingredient that determines the period-doubling behavior of  $f$  as  $\lambda$  increases is the relationship of the slope of  $f^2$  to the slope of  $f$ . This relationship is a consequence of the chain rule. By definition

$$x_2 = f^2(x_0),$$

where

$$x_1 = f(x_0), \quad x_2 = f(x_1).$$

We leave it to the reader to verify by the chain rule that

$$f'^2(x_0) = f'(x_0)f'(x_1) \quad (20)$$

and

$$f^n(x_0) = f'(x_0)f'(x_1) \dots f'(x_{n-1}), \quad (21)$$

an elementary result that determines period doubling. If we start at a fixed point of  $f$  and apply Eq. (20) to  $x_0 = x^*$ , so that  $x_2 = x_1 = x^*$ , then

$$f'^2(x^*) = f'(x^*)f'(x^*) = [f'(x^*)]^2. \quad (22)$$

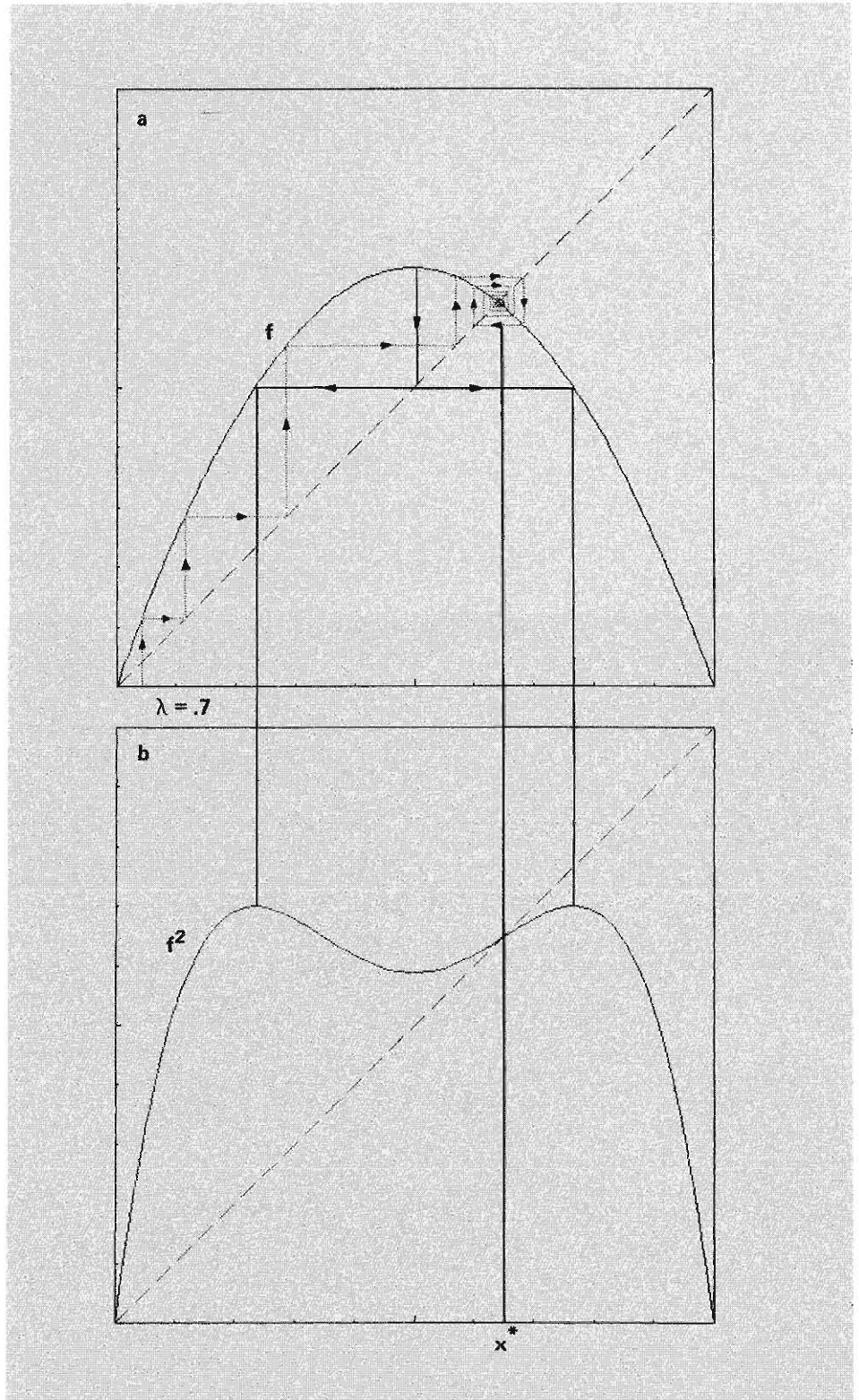


Fig. 2.  $\lambda = 0.7$ .  $x^*$  is the stable fixed point. The extrema of  $f^2$  are located in (a) by constructing the inverse iterates of  $x = 0.5$ .

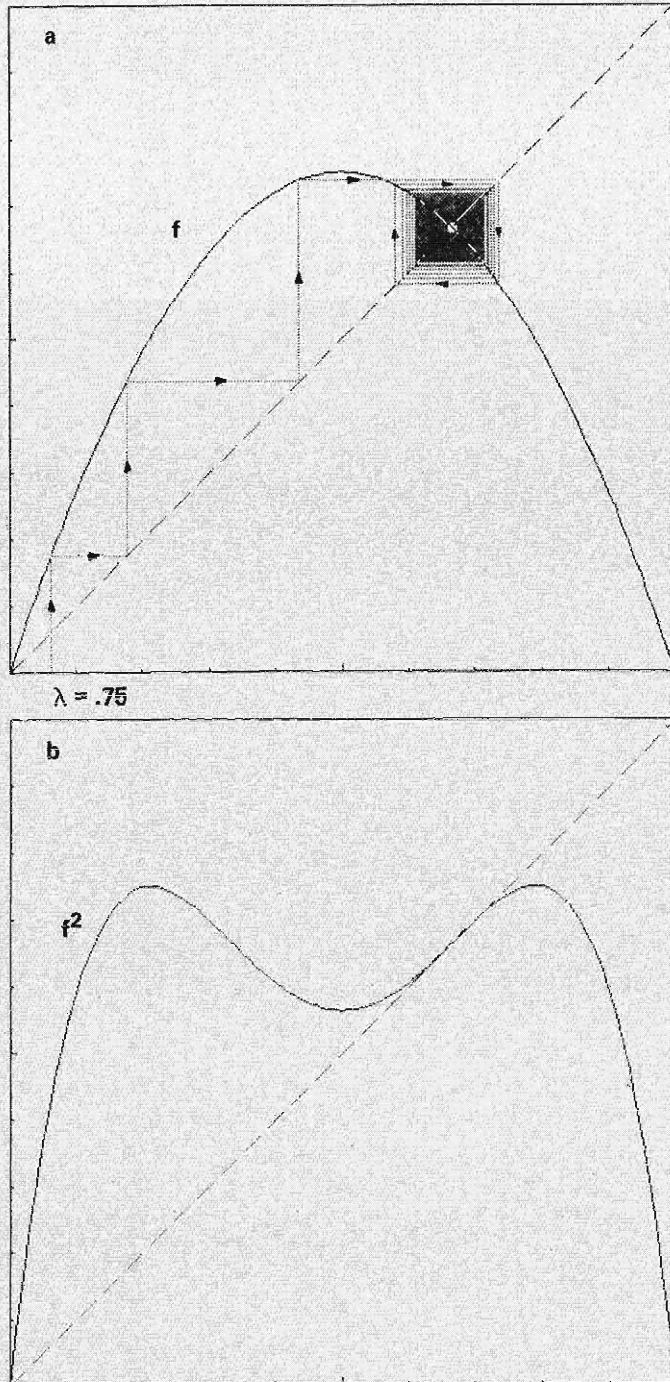


Fig. 3.  $\lambda = 0.75$ . (a) depicts the slow convergence to the fixed point.  $f^2$  osculates about the fixed point.

Since at  $\lambda = 0.7$ ,  $|f'(x^*)| < 1$ , it follows from Eq. (22) that

$$0 < f^{2n}(x^*) < 1.$$

Also, if we start at the extremum of  $f$ , so that  $x_0 = 1/2$  and  $f'(x_0) = 0$ , it follows from Eq. (21) that

$$f^n(1/2) = 0 \quad (23)$$

for all  $n$ . In particular,  $f^2$  is extreme (and a minimum) at  $1/2$ . Also, by Eq. (20),  $f^2$  will be extreme (and a maximum) at the  $x_0$  that will iterate under  $f$  to  $x = 1/2$ , since then  $x_1 = 1/2$  and  $f'(x_1) = 0$ . These points, the *inverses* of  $x = 1/2$ , are found by going *vertically* down along  $x = 1/2$  to  $y = f(x)$  and then *horizontally* to  $y = f(x)$ . (Reverse the arrows in Fig. 1, and see Fig. 2a.) Since  $f$  has a maximum, there are *two* horizontal intersections and, hence, the two maxima of Fig. 2b. *The ability of  $f$  to have complex behaviors is precisely the consequence of its double-valued inverse*, which is in turn a reflection of its possession of an extremum. A monotone  $f$ , one that always increases, *always* has simple behaviors, whether or not the behaviors are easy to compute. A *linear*  $f$  is always monotone. The  $f$ 's we care about always fold over and so are *strongly* nonlinear. This folding nonlinearity gives rise to universality. Just as linearity in any system implies a definite method of solution, folding nonlinearity in any system also implies a definite method of solution. In fact folding nonlinearity in the aperiodic limit of period doubling in any system is solvable, and many systems, such as various coupled nonlinear differential equations, possess this nonlinearity.

To return to Fig. 2b, as  $\lambda \rightarrow 3/4$  and the maximum value of  $f$  increases to  $3/4$ ,  $f'(x^*) \rightarrow -1$  and  $f^{2n}(x^*) \rightarrow +1$ . As  $\lambda$  increases beyond  $3/4$ ,  $|f'(x^*)| > 1$  and  $f^{2n}(x^*) > 1$ , so that  $f^2$  must develop two new fixed points beyond those of  $f$ ; that is,  $f^2$  will cross  $y = x$  at two more points. This transition is depicted in Figs. 3a and b for  $f$  and  $f^2$ , respectively, at  $\lambda =$

0.75, and similarly in Figs. 4a and b at  $\lambda = 0.785$ . (Observe the exceptionally slow convergence to  $x^*$  at  $\lambda = 0.75$ , where iterates approach the fixed point not geometrically, but rather with deviations from  $x^*$  inversely proportional to the square root of the number of iterations.) Since  $x_1^*$  and  $x_2^*$ , the new fixed points of  $f^2$ , are *not* fixed points of  $f$ , it must be that  $f$  sends one into the other:

$$x_1^* = f(x_2^*)$$

and

$$x_2^* = f(x_1^*) .$$

Such a *pair of points*, termed a *2-cycle*, is depicted by the limiting unwinding circulating square in Fig. 4a. Observe in Fig. 4b that the slope of  $f^2$  is in excess of 1 at the fixed point of  $f$  and so is an unstable fixed point of  $f^2$ , while the two new fixed points have slopes smaller than 1, and so are *stable*; that is, every two iterates of  $f$  will have a point attracted toward  $x_1^*$  if it is sufficiently close to  $x_1^*$  or toward  $x_2^*$  if it is sufficiently close to  $x_2^*$ . This means that the sequence under  $f$ ,

$$x_0, x_1, x_2, x_3, \dots ,$$

eventually becomes arbitrarily close to the sequence

$$x_1^*, x_2^*, x_1^*, x_2^*, \dots ,$$

so that this is a stable 2-cycle, or an *at-*

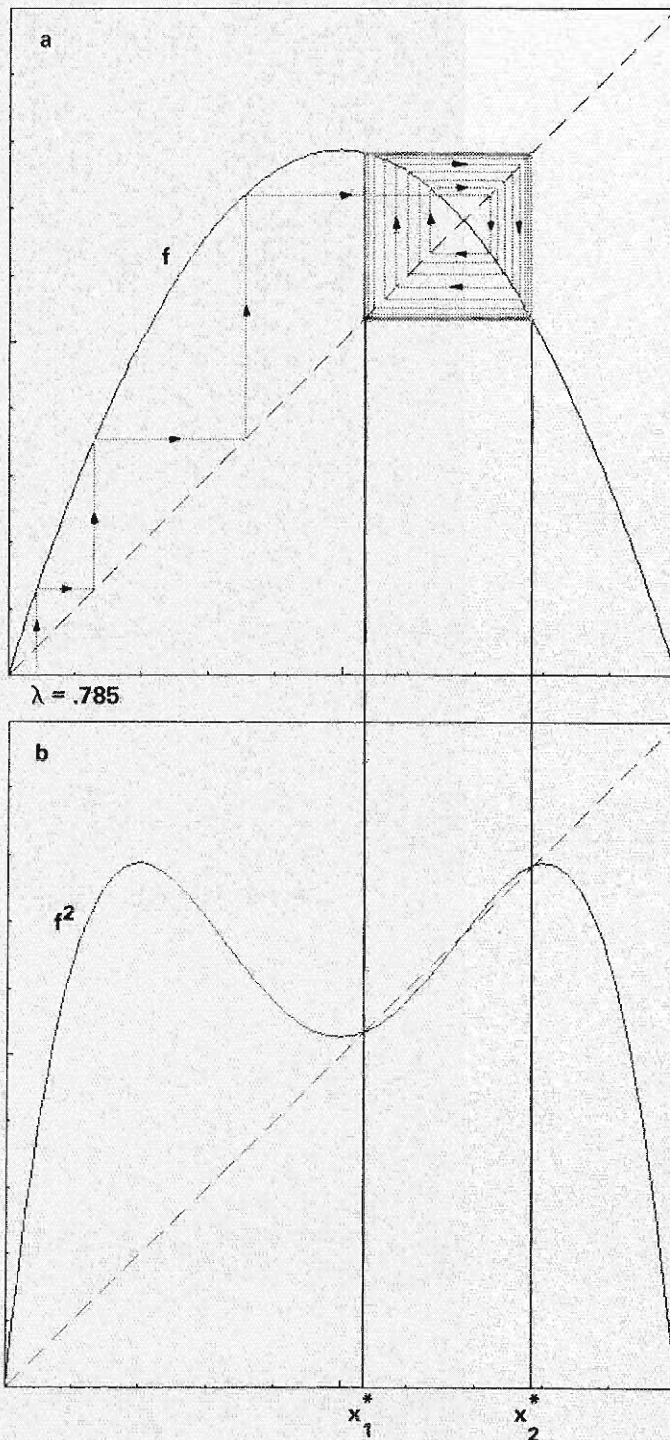


Fig. 4.  $\lambda = 0.785$ . (a) shows the outward spiralling to a stable 2-cycle. The elements of the 2-cycle,  $x_1^*$  and  $x_2^*$ , are located as fixed points in (b).

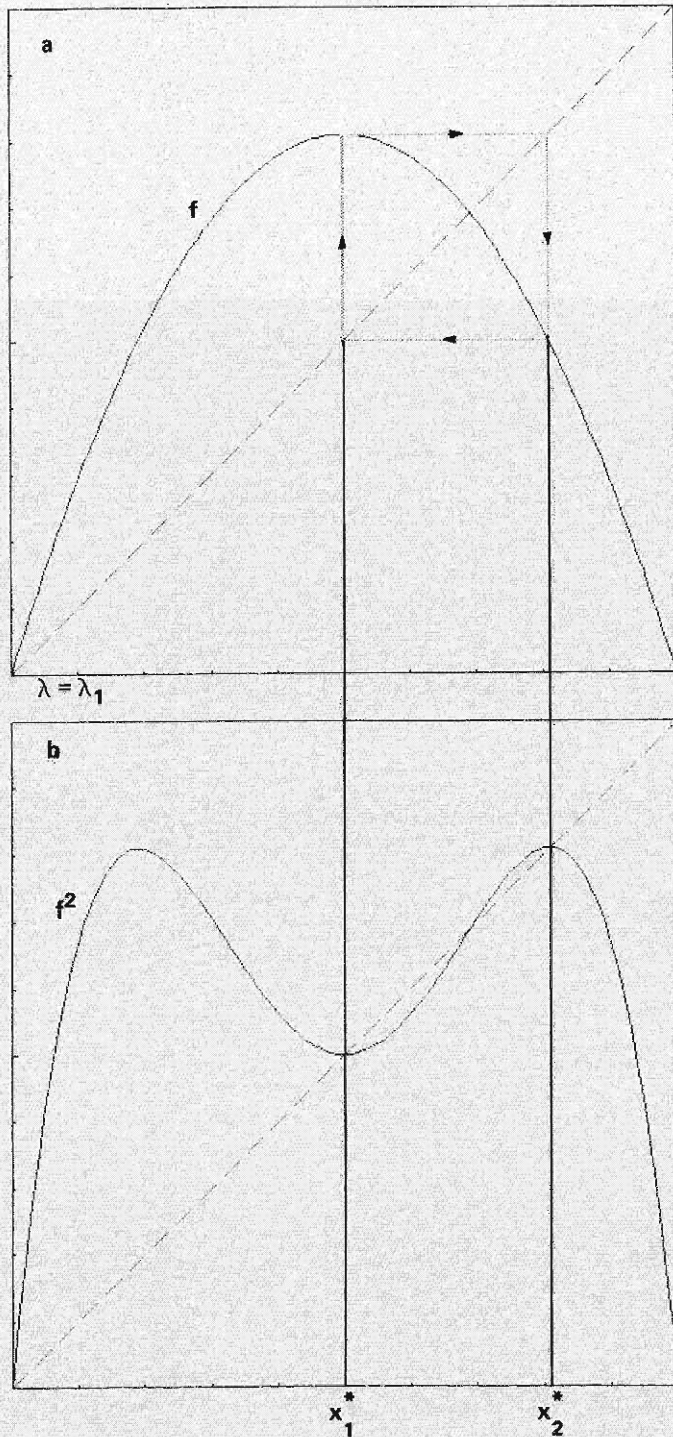


Fig. 5.  $\lambda = \lambda_1$ . A superstable 2-cycle.  $x_1^*$  and  $x_2^*$  are at extrema of  $f^2$ .

tractor of period 2. Thus, we have observed for Eq. (15) the first period doubling as the parameter  $\lambda$  has increased.

There is a point of paramount importance to be observed; namely,  $f^2$  has the same slope at  $x_1^*$  and at  $x_2^*$ . This point is a direct consequence of Eq. (20), since if  $x_0 = x_1^*$ , then  $x_1 = x_2^*$ , and vice versa, so that the product of the slopes is the same. More generally, if  $x_1^*, x_2^*, \dots, x_n^*$  is an  $n$ -cycle so that

$$x_{r+1}^* = f(x_r^*) \quad r = 1, 2, \dots, n - 1$$

and

$$x_1^* = f(x_n^*), \quad (24)$$

then each is a fixed point of  $f^n$  with identical slopes:

$$x_r^* = f^n(x_r^*) \quad r = 1, 2, \dots, n \quad (25)$$

and

$$f^{n'}(x_r^*) = f'(x_1^*) \dots f'(x_n^*). \quad (26)$$

From this observation will follow period doubling *ad infinitum*.

As  $\lambda$  is increased further, the minimum at  $x = 1/2$  will drop as the slope of  $f^2$  through the fixed point of  $f$  increases. At some value of  $\lambda$ , denoted by  $\lambda_1$ ,  $x = 1/2$  will become a fixed point of  $f^2$ . Simultaneously, the right-hand maximum will also become a fixed point of  $f^2$ . [By Eq. (26), both elements of the 2-cycle have slope 0.] Figures 5a and b depict the situation that occurs at  $\lambda = \lambda_1$ .

### Period Doubling *Ad Infinitum*

We are now close to the end of this story. As we increase  $\lambda$  further, the minimum drops still lower, so that both  $x_1^*$  and  $x_2^*$  have negative slopes. At some parameter value, denoted by  $\Lambda_2$ , the slope at *both*  $x_1^*$  and  $x_2^*$  becomes equal to  $-1$ . Thus at  $\Lambda_2$  the same situation has developed for  $f^2$  as developed for  $f$  at  $\Lambda_1 = 3/4$ . This transitional case is depicted in Figs. 6a and b. Accordingly, just as the fixed point of  $f$  at  $\Lambda_1$  issued into being a 2-cycle, so too does *each* fixed point of  $f^2$  at  $\Lambda_2$  create a 2-cycle, which in turn is a 4-cycle of  $f$ . That is, we have now encountered the second period doubling.

The manner in which we were able to follow the creation of the 2-cycle at  $\Lambda_1$  was to anticipate the presence of period 2, and so to consider  $f^2$ , which would resolve the cycle into a pair of fixed points. Similarly, to resolve period 4 into fixed points we now should consider  $f^4$ . Beyond being the fourth iterate of  $f$ , Eq. (8) tells us that  $f^4$  can be computed from  $f^2$ :

$$f^4 = f^2 \circ f^2 .$$

From this point, we can abandon  $f$  itself, and take  $f^2$  as the "fundamental" function. Then, just as  $f^2$  was constructed by iterating  $f$  with itself we now iterate  $f^2$  with itself. The manner in which  $f^2$  reveals itself as being an iterate of  $f$  is the slope equality at the fixed points of  $f^2$ , which we saw imposed by the chain rule. Since the operation of the chain rule is "automatic," we actually needed to consider only the fixed point of  $f^2$  nearest to  $x = 1/2$ ; the behavior of the other fixed point is slaved to it. Thus, at the level of  $f^4$ , we again need to focus on only the fixed point of  $f^4$  nearest to  $x = 1/2$ ; the other *three* fixed points are similarly slaved to it. Thus, a recursive scheme has been unearthed. We now increase  $\lambda$  to  $\lambda_2$ , so that the fixed point of  $f^4$  nearest to  $x = 1/2$  is again at  $x = 1/2$  with slope 0.

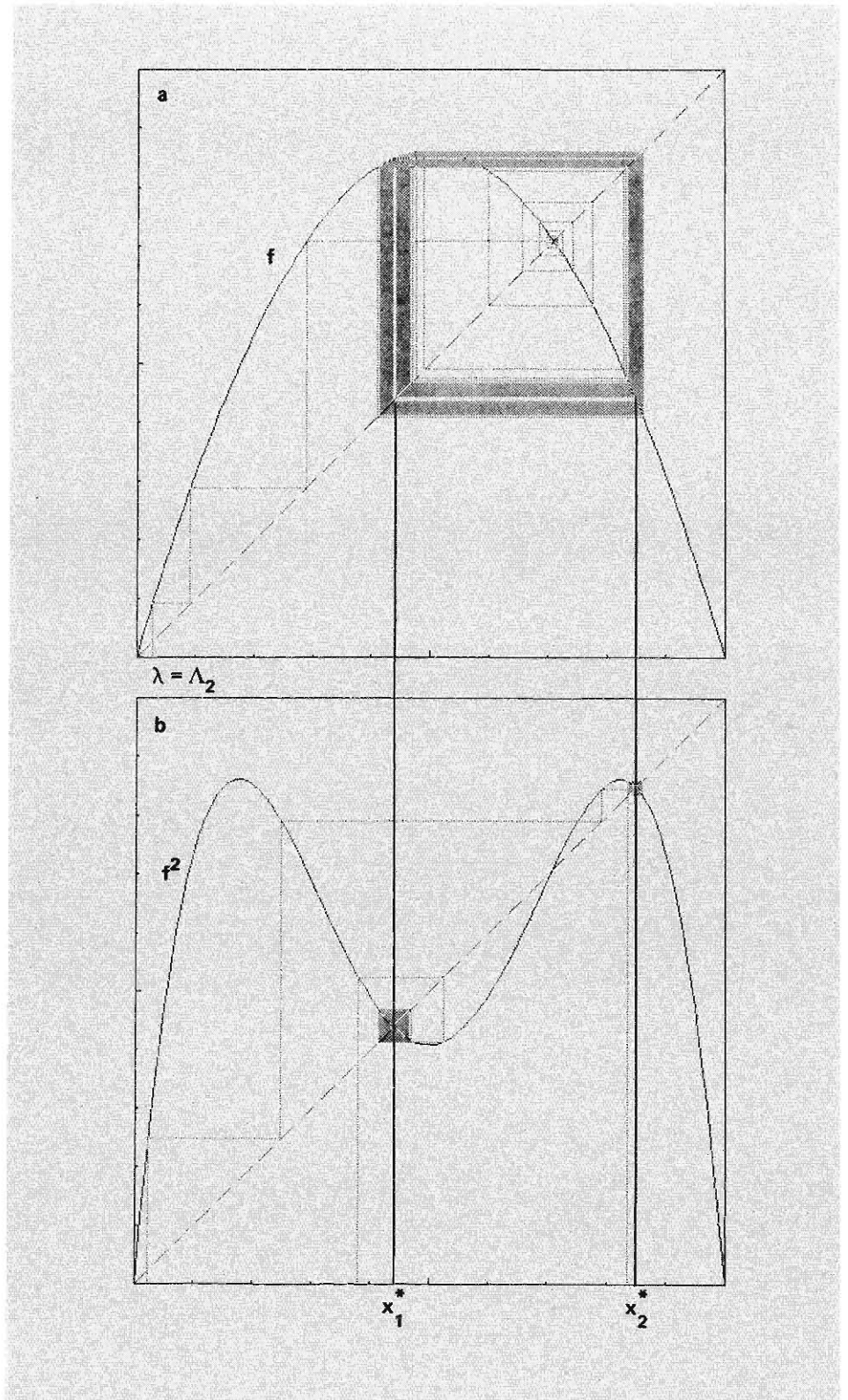


Fig. 6.  $\lambda = \Lambda_2$ .  $x_1^*$  and  $x_2^*$  in (b) have the same slow convergence as the fixed point in Fig. 3a.

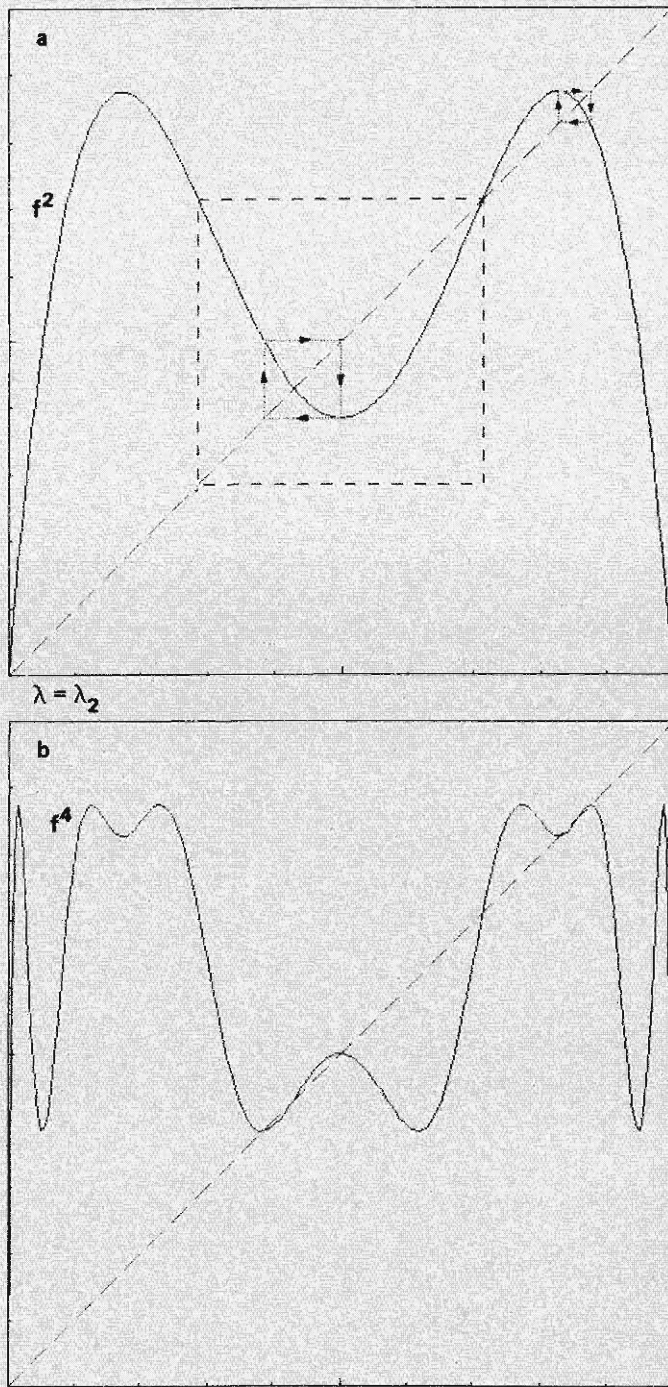


Fig. 7.  $\lambda = \lambda_2$ . A superstable 4-cycle. The region within the dashed square in (a) should be compared with all of Fig. 5a.

Figures 7a and b depict this situation for  $f^2$  and  $f^4$ , respectively. When  $\lambda$  increases further, the maximum of  $f^4$  at  $x = 1/2$  now moves up, developing a fixed point with negative slope. Finally, at  $\Lambda_3$  when the slope of this fixed point (as well as the other three) is again  $-1$ , each fixed point will split into a pair giving rise to an 8-cycle, which is now stable. Again,  $f^8 = f^4 \circ f^4$ , and  $f^4$  can be viewed as fundamental. We define  $\lambda_3$  so that  $x = 1/2$  again is a fixed point, this time of  $f^8$ . Then at  $\Lambda_4$  the slopes are  $-1$ , and another period doubling occurs. Always,

$$f^{2^{n+1}} = f^{2^n} \circ f^{2^n}. \quad (27)$$

Provided that a constraint on the range of  $\lambda$  does not prevent it from decreasing the slope at the appropriate fixed point past  $-1$ , this doubling must recur *ad infinitum*.

Basically, the mechanism that  $f^{2^n}$  uses to period double at  $\Lambda_{n+1}$  is the same mechanism that  $f^{2^{n+1}}$  will use to double at  $\Lambda_{n+2}$ . The function  $f^{2^{n+1}}$  is constructed from  $f^{2^n}$  by Eq. (27), and similarly  $f^{2^{n+2}}$  will be constructed from  $f^{2^{n+1}}$ . Thus, there is a definite operation that, by acting on functions, creates functions; in particular, the operation acting on  $f^{2^n}$  at  $\Lambda_{n+1}$ , (or better,  $f^{2^n}$  at  $\lambda_n$ ) will determine  $f^{2^{n+1}}$  at  $\lambda_{n+1}$ . Also, since we need to keep track of  $f^{2^n}$  only in the interval including the fixed point of  $f^{2^n}$  closest to  $x = 1/2$  and since this interval becomes increasingly small as  $\lambda$  increases, the part of  $f$  that generates this region is also the restriction of  $f$  to an increasingly small interval about  $x = 1/2$ . (Actually, slopes of  $f$  at points farther away also matter, but these merely set a "scale," which will be eliminated by a rescaling.) The behavior of  $f$  away from  $x = 1/2$  is immaterial to the period-doubling behavior, and in the limit of large  $n$  only the *nature of  $f$ 's maximum* can matter. This means that in the infinite period-doubling limit, all functions with a quadratic extremum will have identical behavior. [ $f'(1/2) \neq 0$  is the

generic circumstance.] Therefore, the operation on functions will have a *stable fixed point* in the space of functions, which will be the common universal limit of high iterates of any specific function. To determine this universal limit we must enlarge our scope vastly, so that the role of the starting point,  $x_0$ , will be played by an arbitrary *function*; the attracting fixed point will become a universal function obeying an equation implicating only itself. The role of the function in the equation  $x_0 = f(x_0)$  now must be played by an *operation* that yields a new function when it is performed upon a function. In fact, the heart of this operation is the functional composition of Eq. (27). If we can determine the exact operator and actually can solve its fixed-point problem, we shall understand why a special number, such as  $\delta$  of Eq. (3), has emerged independently of the specific system (the starting function) we have considered.

### The Universal Limit of High Iterates

In this section we sketch the solution to the fixed-point problem. In Fig. 7a, a dashed square encloses the part of  $f^2$  that we must focus on for all further period doublings. This square should be compared with the unit square that comprises all of Fig. 5a. If the Fig. 7a square is reflected through  $x = \frac{1}{2}$ ,  $y = \frac{1}{2}$  and then *magnified* so that the circulation squares of Figs. 4a and 5a are of equal size, we will have in each square a piece of a function that has the same kind of maximum at  $x = \frac{1}{2}$  and falls to zero at the right-hand lower corner of the circulation square. Just as  $f$  produced this second curve of  $f^2$  in the square as  $\lambda$  increased from  $\lambda_1$  to  $\lambda_2$ , so too will  $f^2$  produce another curve, which will be similar to the other two when it has been magnified suitably and reflected twice. Figure 8 shows this superposition for the first *five* such functions; at the resolution of the figure, observe that the last three

---

## A DISCOVERY

---

The inspiration for the universality theory came from two sources. First, in 1971 N. Metropolis, M. Stein, and P. Stein (all in the LASL Theoretical Division) discovered a curious property of iterations: as a parameter is varied, the behavior of iterates varies in a fashion independent of the particular function iterated. In particular for a large class of functions, if at some value of the parameter a certain cycle is stable, then as the parameter increases, the cycle is replaced successively by cycles of doubled periods. This period doubling continues until an infinite period, and hence erratic behavior, is attained.

Second, during the early 1970s, a scheme of mathematics called dynamical system theory was popularized, largely by D. Ruelle, with the notion of a "strange attractor." The underlying questions addressed were (1) how could a purely causal equation (for example, the Navier-Stokes equations that describe fluid flow) come to demonstrate highly erratic or statistical properties and (2) how could these statistical properties be computed. This line of thought merged with the iteration ideas, and the limiting infinite "cycles" of iteration systems came to be viewed as a possible means to comprehend turbulence. Indeed, I became inspired to study the iterates of functions by a talk on such matters by S. Smale, one of the creators of dynamical system theory, at Aspen in the summer of 1975.

My first effort at understanding this problem was through the complex analytic properties of the generating function of the iterates of the quadratic map

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

This study clarified the mechanism of period doubling and led to a rather different kind of equation to determine the values of  $\lambda$  at which the period doubling occurs. The new equations were intractable, although approximate solutions seemed possible. Accordingly, when I returned from Aspen, I numerically determined some parameter values with an eye toward discerning some patterns. At this time I had never used a large computer—in fact my sole computing power resided in a programmable pocket calculator. Now, such machines are very slow. A particular parameter value is obtained iteratively (by Newton's method) with each step of iteration requiring  $2^n$  iterates of the map. For a 64-cycle, this means 1 minute per step of Newton's method. At the same time as  $n$  increased, it became an increasingly more delicate matter to locate the desired solution. However, I immediately perceived the  $\lambda_n$ 's were converging geometrically. This enabled me to predict the next value with increasing accuracy as  $n$  increased, and so required just one step of Newton's method to obtain the desired value. To the best of my knowledge, this observation of geometric convergence has never been made independently, for the sim-



ple reason that the solutions have always been performed automatically on large and fast computers!

That a geometric convergence occurred was already a surprise. I was interested in this for two reasons: first, to gain insight into my theoretical work, as already mentioned, and second, because a convergence rate is a number invariant under all smooth transformations, and so of mathematical interest. Accordingly, I spent a part of a day trying to fit the convergence rate value, 4.669, to the mathematical constants I knew. The task was fruitless, save for the fact that it made the number memorable.

At this point I was reminded by Paul Stein that period doubling isn't a unique property of the quadratic map, but also occurs, for example, in

$$x_{n+1} = \lambda \sin \pi x_n .$$

However, my generating function theory rested heavily on the fact that the nonlinearity was simply quadratic and not transcendental. Accordingly, my interest in the problem waned.

Perhaps a month later I decided to determine the  $\lambda$ 's in the transcendental case numerically. This problem was even slower to compute than the quadratic one. Again, it became apparent that the  $\lambda$ 's converged geometrically, and altogether amazingly, the convergence rate was the same 4.669 that I remembered by virtue of my efforts to fit it.

Recall that the work of Metropolis, Stein, and Stein showed that precise qualitative features are independent of the specific iterative scheme. Now I learned that precise quantitative features also are independent of the specific function. This discovery represents a complete inversion of accustomed ritual. Usually one relies on the fact that similar equations will have qualitatively similar behavior, but quantitative predictions depend on the details of the equations. The universality theory shows that qualitatively similar equations have the identical *quantitative* behavior. For example, a system of differential equations naturally determines certain maps. The computation of the actual analytic form of the map is generally well beyond present mathematical methods. However, should the map exhibit period doubling, then precise quantitative results are available from the universality theory because the theory applies independently of which map it happens to be. In particular, certain fluid flows have now been experimentally observed to become turbulent through period doubling (subharmonic bifurcations). From this one fact we know that the universality theory applies—and indeed correctly determines the precise way in which the flow becomes turbulent, without any reference to the underlying Navier-Stokes equations.

curves are coincident. Moreover, the scale reduction that  $f^2$  will determine for  $f^4$  is based solely on the functional composition, so that if these curves for  $f^{2^n}$ ,  $f^{2^{n+1}}$ , converge (as they obviously do in Fig. 8), the scale reduction from level to level will converge to a definite constant. But the width of each circulation square is just the distance between  $x = 1/2$  when it is a fixed point of  $f^{2^n}$  and the fixed point of  $f^{2^n}$  next nearest to  $x = 1/2$  (Figs. 7a and b). That is, asymptotically, the separation of adjacent elements of period-doubled attractors is reduced by a constant value from one doubling to the next. Also from one doubling to the next, this next nearest element alternates from one side of  $x = 1/2$  to the other. Let  $d_n$  denote the algebraic distance from  $x = 1/2$  to the nearest element of the attractor cycle of period  $2^n$ , in the  $2^n$ -cycle at  $\lambda_n$ . A positive number  $\alpha$  scales this distance down in the  $2^{n+1}$ -cycle at  $\lambda_{n+1}$ :

$$\frac{d_n}{d_{n+1}} \sim -\alpha . \quad (28)$$

But since rescaling is determined only by functional composition, there is some function that composed with itself will reproduce itself reduced in scale by  $-\alpha$ . The function has a quadratic maximum at  $x = 1/2$ , is symmetric about  $x = 1/2$ , and can be scaled by hand to equal 1 at  $x = 1/2$ . Shifting coordinates so that  $x = 1/2 \rightarrow x = 0$ , we have

$$-\alpha g(g(x/\alpha)) = g(x) . \quad (29)$$

Substituting  $g(0) = 1$ , we have

$$g(1) = -\frac{1}{\alpha} . \quad (30)$$

Accordingly, Eq. (29) is a definite equation for a function  $g$  depending on  $x$  through  $x^2$  and having a maximum of 1 at  $x = 0$ . There is a unique smooth solution to Eq. (29), which determines

$$\alpha = 2.502907875 \dots \quad (31)$$

Knowing  $\alpha$ , we can predict through Eq. (28) a definite scaling law binding on the iterates of any scheme possessing period doubling. The law has, indeed, been amply verified experimentally. By Eq. (29), we see that the relevant operation upon functions that underlies period doubling is functional composition followed by magnification, where the magnification is determined by the fixed-point condition of Eq. (29) with the function  $g$  the fixed point in this space of functions. However, Eq. (29) does not describe a stable fixed point because we have not incorporated in it the parameter increase from  $\lambda_n$  to  $\lambda_{n+1}$ . Thus,  $g$  is not the limiting function of the curves in the circulation squares, although it is intimately related to that function. The full theory is described in the next section. Here we merely state that we can determine the limiting function and thereby can *determine the location of the actual elements of limiting  $2^n$ -cycles*. We also have established that  $g$  is an unstable fixed point of functional composition, where the rate of divergence away from  $g$  is precisely  $\delta$  of Eq. (3) and so is computable. Accordingly, there is a full theory that determines, in a precise quantitative way, the aperiodic limit of functional iterations with an *unspecified* function  $f$ .

### Some Details of the Full Theory

Returning to Eq. (28), we are in a position to describe theoretically the universal scaling of high-order cycles and the convergence to a universal limit. Since  $d_n$  is the distance between  $x = 1/2$  and the element of the  $2^n$ -cycle at  $\lambda_n$  nearest to  $x = 1/2$  and since this nearest element is the  $2^{n-1}$  iterate of  $x = 1/2$  (which is true because these two points were coincident before the  $n^{\text{th}}$  period doubling began to split them apart), we have

$$d_n = f^{2^{n-1}}(\lambda_n, 1/2) - 1/2. \quad (32)$$

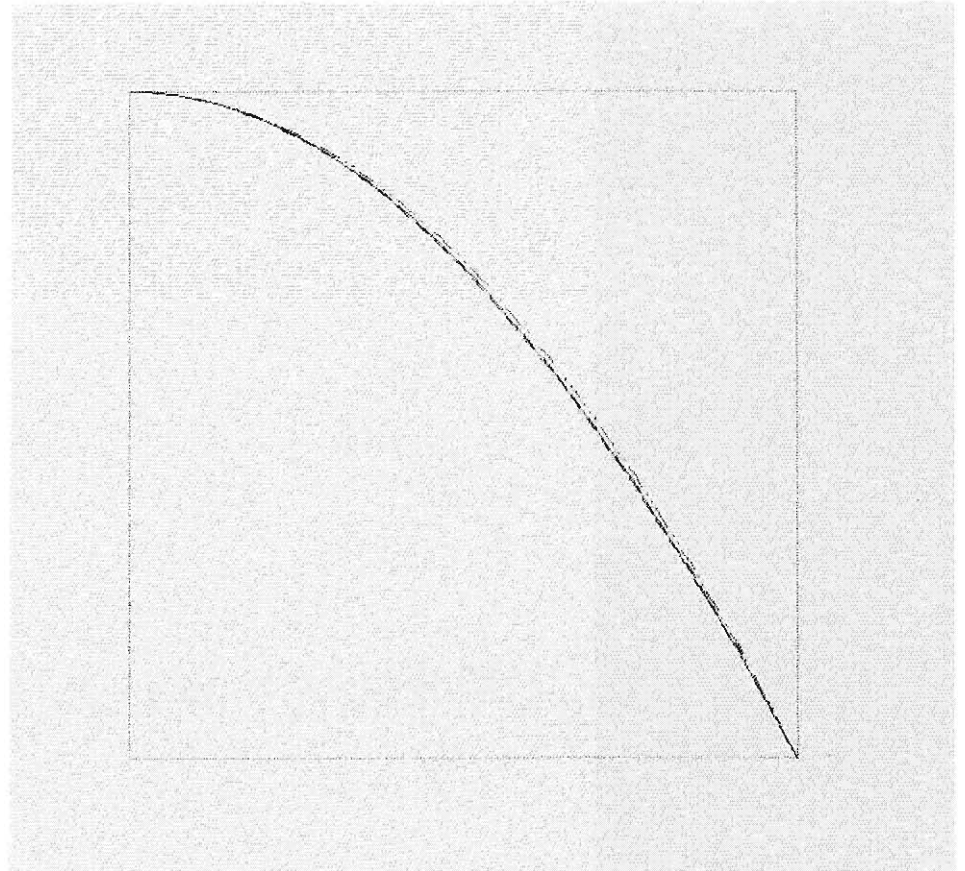


Fig. 8. The superposition of the suitably magnified dotted squares of  $f^{2^{n-1}}$  at  $\lambda_n$  (as in Figs. 5a, 7a, ...).

For future work it is expedient to perform a coordinate translation that moves  $x = 1/2$  to  $x = 0$ . Thus, Eq. (32) becomes

$$d_n = f^{2^{n-1}}(\lambda_n, 0). \quad (33)$$

Equation (28) now determines that the rescaled distances,

$$r_n \equiv (-\alpha)^n d_{n+1}.$$

will converge to a definite finite value as  $n \rightarrow \infty$ . That is,

$$\lim_{n \rightarrow \infty} (-\alpha)^n f^{2^n}(\lambda_{n+1}, 0) \quad (34)$$

must exist if Eq. (28) holds.

However, from Fig. 8 we know something stronger than Eq. (34). When the  $n^{\text{th}}$  iterated function is *magnified* by  $(-\alpha)^n$ , it converges to a definite function. Equation (34) is the value of this function at  $x = 0$ . After the magnification, the convergent functions are given by

$$(-\alpha)^n f^{2^n}(\lambda_{n+1}, x / (-\alpha)^n).$$

Thus,

$$g_1(x) \equiv \lim_{n \rightarrow \infty} (-\alpha)^n f^{2^n}(\lambda_{n+1}, x / (-\alpha)^n) \quad (35)$$

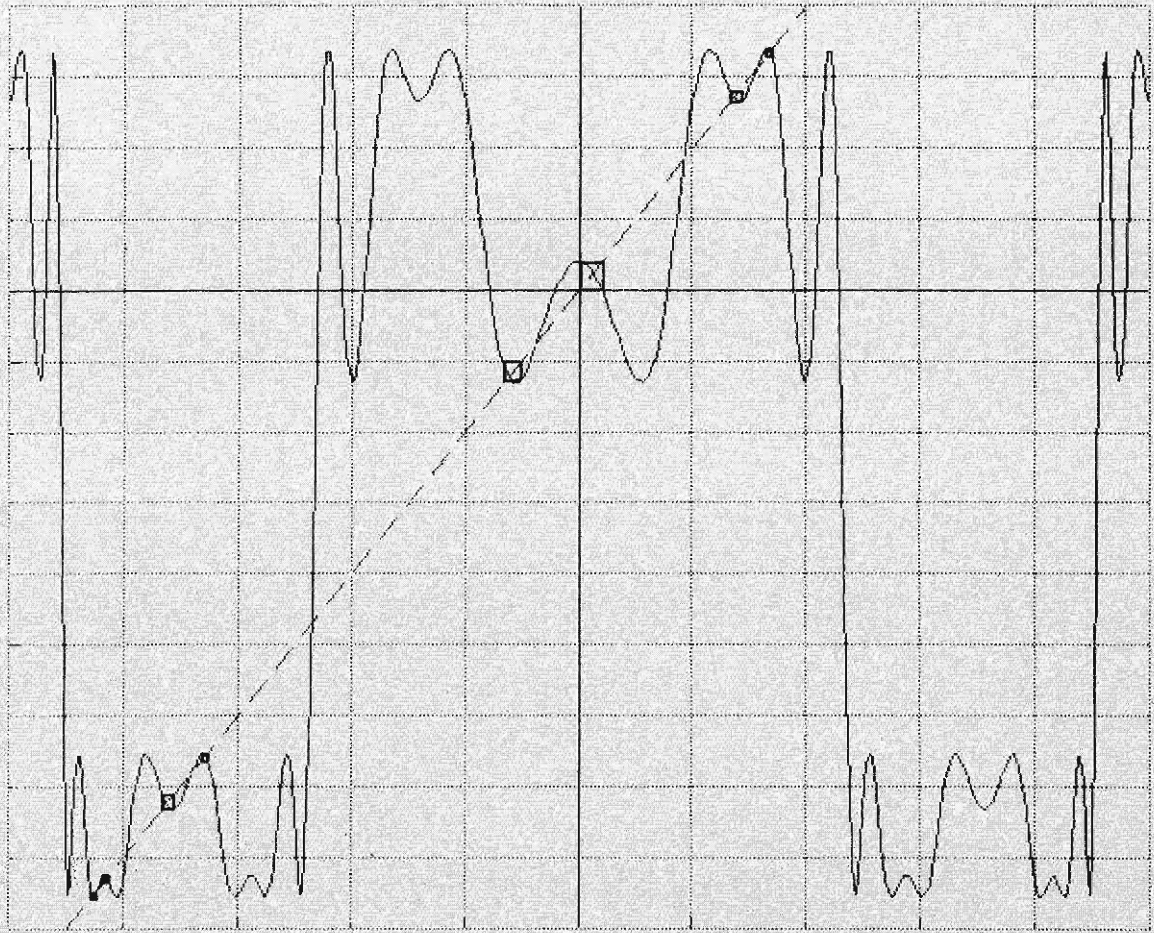


Fig. 9. The function  $g_1$ . The squares locate cycle elements.

is the limiting function inscribed in the square of Fig. 8. The function  $g_1(x)$  is, by the argument of the restriction of  $f$  to increasingly small intervals about its maximum, the *universal* limit of all iterates of all  $f$ 's with a quadratic extremum. Indeed, it is numerically easy to ascertain that  $g_1$  of Eq. (35) is always the same function independent of the  $f$  in Eq. (32).

What is this universal function good for? Figure 5a shows a crude approximation of  $g_1$  [ $n = 0$  in the limit of Eq. (35)], while Fig. 7a shows a better approximation ( $n = 1$ ). In fact, the extrema of  $g_1$  near the fixed points of  $g_1$  support circulation squares each of which contains two points of the cycle. (The two squares shown in Fig. 7a locate the four elements of the cycle.) That is,  $g_1$  determines the location of elements of high-

order  $2^n$ -cycles near  $x = 0$ . Since  $g_1$  is *universal*, we now have the amazing result that the location of the actual elements of highly doubled cycles is universal! The reader might guess this is a *very* powerful result. Figure 9 shows  $g_1$  out to  $x$  sufficiently large to have 8 circulation squares, and hence locates the 15 elements of a  $2^n$ -cycle nearest to  $x = 0$ . Also, the universal value of the scaling parameter  $\alpha$ , obtained numerically, is

$$\alpha = 2.502907875 \dots \quad (36)$$

Like  $\delta$ ,  $\alpha$  is a number that can be *measured* [through an experiment that observes the  $d_n$  of Eq. (28)] in any phenomenon exhibiting period doubling.

If  $g_1$  is universal, then of course its iterate  $g_1^2$  also is universal. Figure 7b

depicts an early approximation to this iterate. In fact, let us define a new universal function  $g_0$ , obtained by scaling  $g_1^2$ :

$$g_0(x) \equiv -\alpha g_1^2(-x/\alpha). \quad (37)$$

(Because  $g_1$  is universal and the iterates of our quadratic function are all symmetric in  $x$ , both  $g_1$  and  $g_0$  are symmetric functions. Accordingly, the minus sign within  $g_1^2$  can be dropped with impunity.) From Eq. (35), we now can write

$$g_0(x) = \lim_{n \rightarrow \infty} (-\alpha)^n f^{2^n}(\lambda_n x / (-\alpha)^n). \quad (38)$$

[We introduced the scaling of Eq. (37) to provide one power of  $\alpha$  per period doubling, since each successive iterate of  $f^{2^n}$

reduces the scale by  $\alpha$ ].

In fact, we can generalize Eqs. (35) and (38) to a *family* of universal functions  $g_r$ :

$$g_r(x) = \lim_{n \rightarrow \infty} (-\alpha)^n f^{2^n}(\lambda_{n+r}, x / (-\alpha)^n). \quad (39)$$

To understand this, observe that  $g_0$  locates the cycle elements as the fixed points of  $g_0$  at extrema;  $g_1$  locates the same elements by determining two elements per extremum. Similarly,  $g_r$  determines  $2^r$  elements about each extremum near a fixed point of  $g_r$ . Since each  $f^{2^n}$  is always magnified by  $(-\alpha)^n$  for each  $r$ , the scales of all  $g_r$  are the same. Indeed,  $g_r$  for  $r > 1$  looks like  $g_1$  of Fig. 9, except that each extremum is slightly higher, to accommodate a  $2^r$ -cycle. Since each extremum must grow by convergently small amounts to accommodate higher and higher  $2^r$ -cycles, we are led to conclude that

$$g(x) = \lim_{r \rightarrow \infty} g_r(x) \quad (40)$$

must exist. By Eq. (39),

$$g(x) = \lim_{n \rightarrow \infty} (-\alpha)^n f^{2^n}(\lambda_\infty, x / (-\alpha)^n). \quad (41)$$

Unlike the functions  $g_r$ ,  $g(x)$  is obtained as a limit of  $f^{2^n}$ 's at a *fixed value* of  $\lambda$ . Indeed, this is the special significance of  $\lambda_\infty$ ; it is an isolated value of  $\lambda$  at which repeated iteration and magnification lead to a convergent function.

We now can write the equation that  $g$  satisfies. Analogously to Eq. (37), it is easy to verify that all  $g_r$  are related by

$$g_{r-1}(x) = -\alpha g_r(g_r(-x/\alpha)). \quad (42)$$

By Eq. (40), it follows that  $g$  satisfies

$$g(x) = -\alpha g(g(x/\alpha)). \quad (43)$$

The reader can verify that Eq. (43) is in-

variant under a magnification of  $g$ . Thus, the theory has nothing to say about absolute scales. Accordingly, we must fix this by hand by setting

$$g(0) = 1. \quad (44)$$

Also, we must specify the nature of the maximum of  $g$  at  $x = 0$  (for example, quadratic). Finally, since  $g$  is to be built by iterating a  $-x^2$ , it must be both smooth and a function of  $x$  through  $x^2$ . With these specifications, Eq. (43) has a *unique* solution. By Eqs. (44) and (43),

$$g(0) = 1 = -\alpha g(g(0)) = -\alpha g(1),$$

so that

$$\alpha = -1/g(1). \quad (45)$$

Accordingly, Eq. (43) determines  $\alpha$  together with  $g$ .

Let us comment on the nature of Eq. (43), a so-called functional equation. Because  $g$  is smooth, if we know its value at a finite number of points, we know its value to some approximation on the interval containing these points by any sufficiently smooth interpolation. Thus, to some degree of accuracy, Eq. (43) can be replaced by a finite coupled system of nonlinear equations. Exactly then, Eq. (43) is an infinite-dimensional, nonlinear vector equation. Accordingly, we have obtained the solution to one-dimensional period doubling through our infinite-dimensional, explicitly universal problem. Equation (43) must be infinite-dimensional because it must keep track of the infinite number of cycle elements demanded of any attempt to solve the period-doubling problem. Rigorous mathematics for equations like Eq. (43) is just beyond the boundary of present mathematical knowledge.

At this point, we must determine two items. First, where is  $\delta$ ? Second, how do we obtain  $g_1$ , the real function of interest for locating cycle elements? The two

problems are part of one question. Equation (42) is itself an iteration scheme. However, unlike the elements in Eq. (4), the elements acted on in Eq. (42) are *functions*. The analogue of the function of  $f$  in Eq. (4) is the operation in function space of functional composition followed by a magnification. If we call this operation  $T$ , and an element of the function space  $\psi$ , Eq. (42) gives

$$T|\psi|(x) = -\alpha \psi^2(-x/\alpha). \quad (46)$$

In terms of  $T$ , Eq. (42) now reads

$$g_{r-1} = T|g_r|, \quad (47)$$

and Eq. (43) reads

$$g = T|g|. \quad (48)$$

Thus,  $g$  is precisely the fixed point of  $T$ . Since  $g$  is the limit of the sequence  $g_r$ , we can obtain  $g_r$  for large  $r$  by linearizing  $T$  about its fixed point  $g$ . Once we have  $g_r$  in the linear regime, the exact repeated application of  $T$  by Eq. (47) will provide  $g_1$ . Thus, we must investigate the stability of  $T$  at the fixed point  $g$ . However, it is obvious that  $T$  is *unstable* at  $g$ : for a large enough  $r$ ,  $g_r$  is a point arbitrarily close to the fixed point  $g$ ; by Eq. (47), successive iterates of  $g_r$  under  $T$  move away from  $g$ . How unstable is  $T$ ? Consider a one-parameter family of functions  $f_\lambda$ , which means a "line" in the function space. For each  $f$ , there is an isolated parameter value  $\lambda_\infty$ , for which repeated applications of  $T$  lead to convergence towards  $g$  [Eq. (41)]. Now, the function space can be "packed" with all the lines corresponding to the various  $f$ 's. The set of all the points on these lines specified by the respective  $\lambda_\infty$ 's determines a "surface" having the property that repeated applications of  $T$  to any point on it will converge to  $g$ . This is the surface of stability of  $T$  (the "stable manifold" of  $T$  through  $g$ ). But through each point of this surface issues out the corresponding line, which is one-

dimensional since it is parametrized by a single parameter,  $\lambda$ . Accordingly,  $T$  is *unstable* in only *one* direction in function space. Linearized about  $g$ , this line of instability can be written as the one-parameter family

$$f_\lambda(x) = g(x) - \lambda h(x), \quad (49)$$

which passes through  $g$  (at  $\lambda = 0$ ) and deviates from  $g$  along the unique direction  $h$ . But  $f_\lambda$  is just one of our transformations [Eq. (4)]! Thus, as we vary  $\lambda$ ,  $f_\lambda$  will undergo period doubling, doubling to a  $2^n$ -cycle at  $\lambda_n$ . By Eq. (41),  $\lambda_\infty$  for the family of functions  $f_\lambda$  in Eq. (49) is

$$\lambda_\infty = 0. \quad (50)$$

Thus, by Eq. (1)

$$\lambda_n \sim \delta^{-n}. \quad (51)$$

Since applications of  $T$  by Eq. (47) iterate in the opposite direction (diverge away from  $g$ ), it now follows that the rate of instability of  $T$  along  $h$  must be precisely  $\delta$ .

Accordingly, we find  $\delta$  and  $g_1$  in the following way. First, we must linearize the operation  $T$  about its fixed point  $g$ . Next, we must determine the stability directions of the linearized operator. Moreover, we expect there to be precisely one direction of instability. Indeed, it turns out that infinitesimal deformations (conjugacies) of  $g$  determine *stable* directions, while a unique unstable direction,  $h$ , emerges with a stability rate (eigenvalue) precisely the  $\delta$  of Eq. (3). Equation (49) at  $\lambda_r$  is precisely  $g_r$  for asymptotically large  $r$ . Thus  $g_r$  is known asymptotically, so that we have entered the sequence  $g_r$  and can now, by repeated use of Eq. (47), step down to  $g_1$ . All the ingredients of a full description of high-order  $2^n$ -cycles now are at hand and evidently are universal.

Although we have said that the function  $g_1$  universally locates cycle elements

near  $x = 0$ , we must understand that it doesn't locate all cycle elements. This is possible because a finite distance of the scale of  $g_1$  (for example, the location of the element nearest to  $x = 0$ ) has been magnified by  $\alpha^n$  for  $n$  diverging. Indeed, the distances from  $x = 0$  of all elements of a  $2^n$ -cycle, "accurately" located by  $g_1$ , are reduced by  $-\alpha$  in the  $2^{n+1}$ -cycle. However, it is obvious that some elements have no such scaling: because  $f(0) = a_n$  in Eq. (13), and  $a_n \rightarrow a_\infty$ , which is a definite nonzero number, the distance from the origin of the element of the  $2^n$ -cycle farthest to the right certainly has not been reduced by  $-\alpha$  at each period doubling. This suggests that we must measure locations of elements on the far right with respect to the farthest right point. If we do this, we can see that these distances scale by  $\alpha^2$ , since they are the images through the quadratic maximum of  $f$  at  $x = 0$  of elements close to  $x = 0$  scaling with  $-\alpha$ . In fact, if we image  $g_1$  through the maximum of  $f$  (through a quadratic conjugacy), then we shall indeed obtain a new universal function that locates cycle elements near the right-most element. The correct description of a highly doubled cycle now emerges as one of universal local clusters.

We can state the scope of universality for the location of cycle elements precisely. Since  $f(\lambda_1, x)$  exactly locates the two elements of the  $2^1$ -cycle, and since  $f(\lambda_1, x)$  is an approximation to  $g_1$  [ $n = 0$  in Eq. (35)], we evidently can locate both points exactly by appropriately scaling  $g_1$ . Next, near  $x = 0$ ,  $f^2(\lambda_2, x)$  is a better approximation to  $g_1$  (suitably scaled). However, in general, the more accurately we scale  $g_1$  to determine the smallest 2-cycle elements, the greater is the error in its determination of the right-most elements. Again, near  $x = 0$ ,  $f^r(\lambda_r, x)$  is a still better approximation to  $g_1$ . Indeed, the suitably scaled  $g_1$  now can determine several points about  $x = 0$  accurately, but determination of the right-

most elements is still worse. In this fashion, it follows that  $g_1$ , suitably scaled, can determine  $2^r$  points of the  $2^n$ -cycle near  $x = 0$  for  $r \ll n$ . If we focus on the neighborhood of one of these  $2^r$  points at some definite distance from  $x = 0$ , then by Eq. (35) the larger the  $n$ , the larger the *scaled* distance of this region from  $x = 0$ , and so, the poorer the approximation of the location of fixed points in it by  $g_1$ . However, just as we can construct the version of  $g_1$  that applies at the right-most cycle element, we also can construct the version of  $g_1$  that applies at this chosen neighborhood. Accordingly, the universal description is set through an acceptable tolerance: if we "measure"  $f^{2^n}$  at some definite  $n$ , then we can use the actual location of the elements as foci for  $2^n$  versions of  $g_1$ , each applicable at one such point. For all further period doubling, we determine the new cycle elements through the  $g_1$ 's. In summary, the *more accurately we care to know the locations* of arbitrarily high-order cycle elements, the *more parameters we must measure* (namely, the cycle elements at some chosen order of period doubling). This is the sense in which the universality theory is asymptotic. Its ability to have serious predictive power is the fortunate consequence of the high convergence rate  $\delta$  ( $\sim 4.67$ ). Thus, typically after the first two or three period doublings, this asymptotic theory is already accurate to within several percent. If a period-doubling system is *measured* in its 4- or 8-cycle, its behavior throughout and symmetrically beyond the period-doubling regime also is determined to within a few percent.

To make precise dynamical predictions, we do not have to construct all the local versions of  $g_1$ ; all we really need to know is the local *scaling* everywhere along the attractor. The scaling is  $-\alpha$  at  $x = 0$  and  $\alpha^2$  at the right-most element. But what is it at an arbitrary point? We can determine the scaling law if we order

elements not by their location on the  $x$ -axis, but rather by their order as iterates of  $x = 0$ . Because the time sequence in which a process evolves is precisely this ordering, the result will be of immediate and powerful predictive value. It is precisely this scaling law that allows us to compute the spectrum of the onset of turbulence in period-doubling systems.

What must we compute? First, just as the element in the  $2^n$ -cycle nearest to  $x = 0$  is the element halfway around the cycle from  $x = 0$ , the element nearest to an arbitrarily chosen element is precisely the one halfway around the cycle from it. Let us denote by  $d_n(m)$  the distance between the  $m^{\text{th}}$  cycle element ( $x_m$ ) and the element nearest to it in a  $2^n$ -cycle. [The  $d_n$  of Eq. (28) is  $d_n(0)$ ]. As just explained,

$$d_n(m) = x_m - f^{2^{n-1}}(\lambda_n, x_m). \quad (52)$$

However,  $x_m$  is the  $m^{\text{th}}$  iterate of  $x_0 = 0$ . Recalling from Eq. (6) that powers commute, we find

$$d_n(m) = f^m(\lambda_n, 0) - f^m(\lambda_n, f^{2^{n-1}}(\lambda_n, 0)). \quad (53)$$

Let us, for the moment, specialize to  $m$  of the form  $2^{n-r}$ , in which case

$$\begin{aligned} d_n(2^{n-r}) &= f^{2^{n-r}}(\lambda_n, 0) \\ &\quad - f^{2^{n-r}}(\lambda_n, f^{2^{n-1}}(\lambda_n, 0)) \\ &= f^{2^{n-r}}(\lambda_{(n-r)+r}, 0) \\ &\quad - f^{2^{n-r}}(\lambda_{(n-r)+r}, f^{2^{n-1}}(\lambda_n, 0)). \end{aligned} \quad (54)$$

For  $r \ll n$  (which can still allow  $r \gg 1$  for  $n$  large), we have, by Eq. (39),

$$d_n(2^{n-r}) \sim (-\alpha)^{-(n-r)} [g_r(0) - g_r((-\alpha)^{n-r} f^{2^{n-1}}(\lambda_n, 0))]$$

or

$$d_n(2^{n-r}) \sim (-\alpha)^{-(n-r)} [g_r(0) - g_r((-\alpha)^{-r+1} g_1(0))]. \quad (55)$$

The object we want to determine is the local scaling at the  $m^{\text{th}}$  element, that is, the ratio of nearest separations at the  $m^{\text{th}}$  iterate of  $x = 0$ , at successive values of  $n$ . That is, if the scaling is called  $\sigma$ ,

$$\sigma_n(m) \equiv \frac{d_{n+1}(m)}{d_n(m)}. \quad (56)$$

[Observe by Eq. (28), the definition of  $\alpha$ , that  $\sigma_n(0) \sim (-\alpha)^{-1}$ .] Specializing again to  $m = 2^{n-r}$ , where  $r \ll n$ , we have by Eq. (55)

$$\sigma(2^{n-r}) \sim \frac{g_{r+1}(0) - g_{r+1}((-\alpha)^{-r} g_1(0))}{g_r(0) - g_r((-\alpha)^{-r+1} g_1(0))}. \quad (57)$$

Finally, let us rescale the axis of iterates so that all  $2^{n+1}$  iterates are within a unit interval. Labelling this axis by  $t$ , the value of  $t$  of the  $m^{\text{th}}$  element in a  $2^n$ -cycle is

$$t_n(m) = m/2^n. \quad (58)$$

In particular, we have

$$t_n(2^{n-r}) = 2^{-r}. \quad (59)$$

Defining  $\sigma$  along the  $t$ -axis naturally as

$$\sigma(t_n(m)) \sim \sigma_n(m) \quad (\text{as } n \rightarrow \infty),$$

we have by Eqs. (57) and (59),

$$\sigma(2^{-r-1}) = \frac{g_{r+1}(0) - g_{r+1}((-\alpha)^{-r} g_1(0))}{g_r(0) - g_r((-\alpha)^{-r+1} g_1(0))}. \quad (60)$$

It is not much more difficult to obtain  $\sigma$  for all  $t$ . This is done first for rational  $t$  by writing  $t$  in its binary expansion:

$$t_{r_1 r_2 r_3 \dots} = 2^{-r_1} + 2^{-r_2} + \dots$$

In the  $2^n$ -cycle approximation we require  $\sigma_n$  at the  $2^{n-r_1} + 2^{n-r_2} + \dots$  iterate of the origin. But, by Eq. (8),

$$f^{2^{n-r_1} + 2^{n-r_2} + \dots} = f^{2^{n-r_1}} \circ f^{2^{n-r_2}} \circ \dots$$

It follows by manipulations identical to those that led from Eq. (54) to Eq. (60) that  $\sigma$  at such values of  $t$  is obtained by replacing the individual  $g_r$  terms in Eq. (60) by appropriate iterates of various  $g_r$ 's.

There is one last ingredient to the computation of  $\sigma$ . We know that  $\sigma(0) = -\alpha^{-1}$ . We also know that  $\sigma_n(1) \sim \alpha^{-2}$ . But, by Eq. (59),

$$t_n(1) = 2^{-n} \rightarrow 0.$$

Thus  $\sigma$  is discontinuous at  $t = 0$ , with  $\sigma(0 - \epsilon) = -\alpha^{-1}$  and  $\sigma(0 + \epsilon) = \alpha^{-2}(\epsilon \rightarrow 0^+)$ . Indeed, since  $x_{2^{n-r}}$  is always very close to the origin, each of these points is imaged quadratically. Thus Eq. (60) actually determines  $\sigma(2^{-r-1} - \epsilon)$ , while  $\sigma(2^{-r-1} + \epsilon)$  is obtained by replacing each numerator and denominator  $g_r$  by its square. The same replacement also is correct for each multi- $g_r$  term that figures into  $\sigma$  at the binary expanded rationals.

Altogether, we have the following results.  $\sigma(t)$  can be computed for all  $t$ , and it is *universal* since its explicit computation depends only upon the universal functions  $g_r$ .  $\sigma$  is *discontinuous* at all the rationals. However, it can be established that the *larger* the number of terms in the binary expansion of a rational  $t$ , the smaller the discontinuity of  $\sigma$ . Lastly, as a finite number of iterates leaves  $t$  unchanged as  $n \rightarrow \infty$ ,  $\sigma$  must be *continuous* except at the rationals. Figure 10 depicts  $1/\sigma(t)$ . Despite the pathological nature of  $\sigma$ , the reader will observe that basically it is constant half the time at  $\alpha^{-1}$  and half the time at  $\alpha^{-2}$  for  $0 < t < 1/2$ . In a succeeding approximation, it can be decomposed in each half into two slightly different quarters,

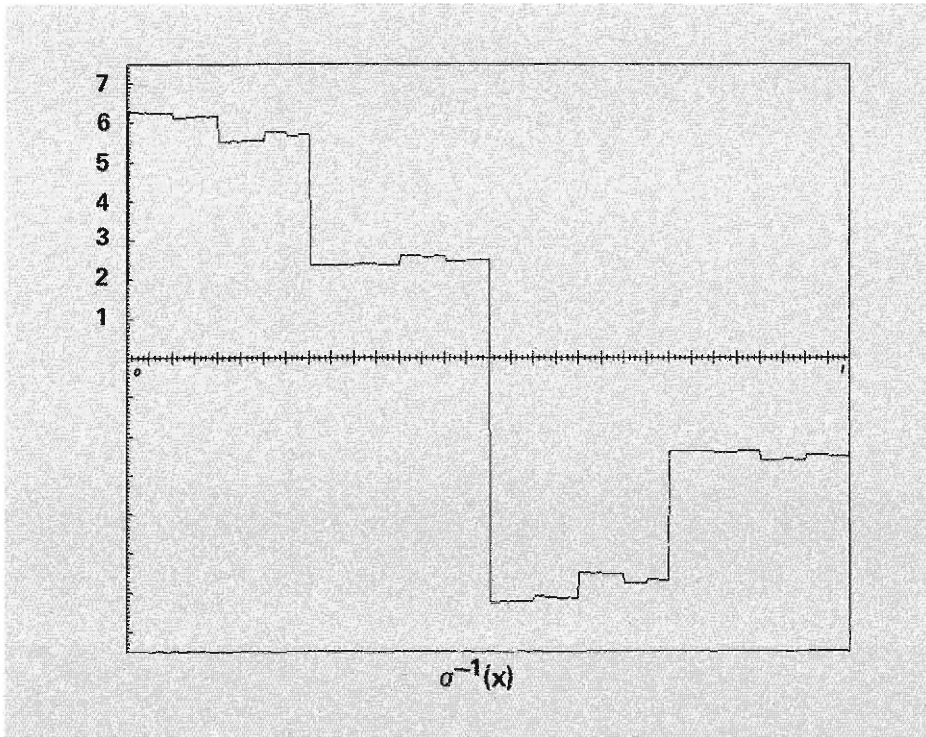


Fig. 10. The trajectory scaling function. Observe that  $\sigma(x + 1/2) = -\sigma(x)$ .

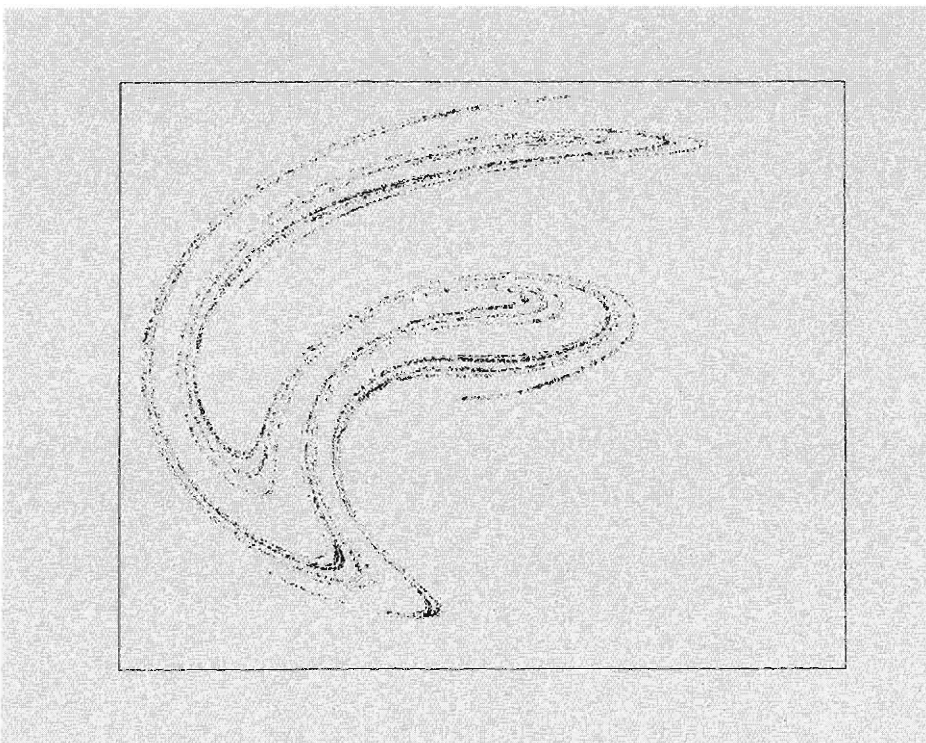


Fig. 11. The plotted points lie on the "strange attractor" of Duffing's equation.

and so forth. [It is easy to verify from Eq. (52) that  $\sigma$  is periodic in  $t$  of period 1, and has the symmetry

$$\sigma(t + 1/2) = -\sigma(t).$$

Accordingly, we have paid attention to its first half  $0 < t < 1/2$ .] With  $\sigma$  we are at last finished with one-dimensional iterates per se.

### Universal Behavior in Higher Dimensional Systems

So far we have discussed iteration in *one* variable; Eq. (15) is the prototype. Equation (14), an example of iteration in two dimensions, has the special property of preserving areas. A generalization of Eq. (14),

$$x_{n+1} = y_n - x_n^2$$

and

$$y_{n+1} = a + bx_n \tag{61}$$

with  $|b| < 1$ , contracts areas. Equation (61) is interesting because it possesses a so-called *strange attractor*. This means an attractor (as before) constructed by folding a curve repeatedly upon itself (Fig. 11) with the consequent property that two initial points very near to one another are, in fact, very far from each other when the distance is measured along the folded attractor, which is the path they follow upon iteration. This means that after some iteration, they will soon be far apart in actual distance as well as when measured along the attractor. This general mechanism gives a system highly sensitive dependence upon its initial conditions and a truly statistical character: since very small differences in initial conditions are magnified quickly, unless the initial conditions are known to *infinite precision*, all known knowledge is eroded rapidly to future ignorance. Now, Eq. (61) enters

into the early stages of statistical behavior through period doubling. Moreover,  $\delta$  of Eq. (3) is *again* the rate of onset of complexity, and  $\alpha$  of Eq. (31) is again the rate at which the spacing of adjacent attractor points is vanishing. Indeed, the one-dimensional theory determines all behavior of Eq. (61) in the onset regime.

In fact, dimensionality is irrelevant. The same theory, the same numbers, etc. also work for iterations in N dimensions, provided that the system goes through period doubling. The basic process, wherever period doubling occurs *ad infinitum*, is functional composition from one level to the next. Accordingly, a modification of Eq. (29) is at the heart of the process, with composition on functions from N dimensions to N dimensions. Should the specific iteration function contract N-dimensional volumes (a dissipative process), then in general there is one direction of slowest contraction, so that after a number of iterations the process is effectively one-dimensional. Put differently, the one-dimensional solution to Eq. (29) is always a solution to its N-dimensional analogue. It is the relevant fixed point of the analogue if the iteration function is contractive.

### Universal Behavior in Differential Systems

The next step of generalization is to include systems of differential equations. A prototypic equation is Duffing's oscillator, a driven damped anharmonic oscillator,

$$\ddot{x} + k\dot{x} + x^3 = b\sin 2\pi t. \quad (62)$$

The periodic drive of period 1 determines a natural time step. Figure 12a depicts a period 1 attractor, usually referred to as a *limit cycle*. It is an attractor because, for a range of initial conditions, the solution to Eq. (62) settles down to the cycle. It is period 1 because it repeats the same curve in every period of the drive.

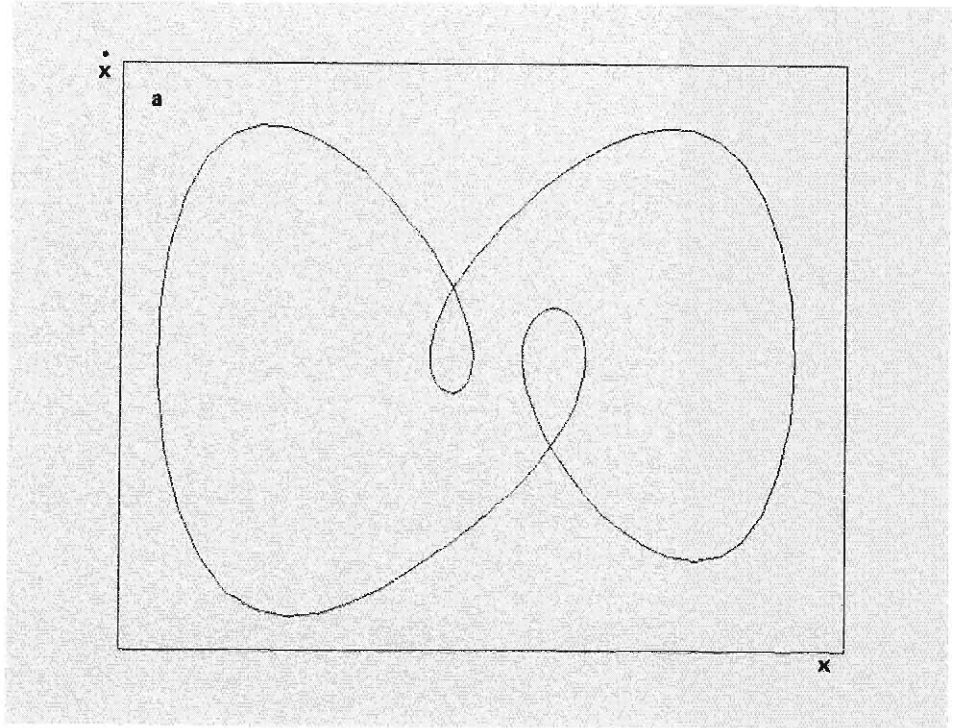


Fig. 12a. The most stable 1-cycle of Duffing's equation in phase space  $(x, \dot{x})$ .

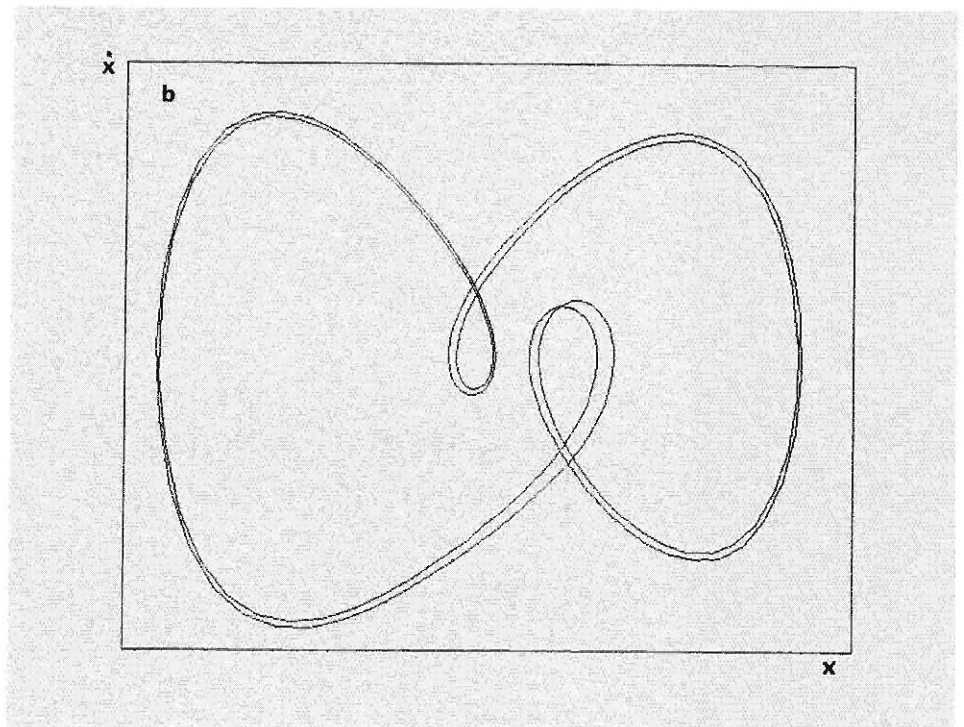


Fig. 12b. The most stable 2-cycle of Duffing's equation. Observe that it is two displaced copies of Fig. 12a.



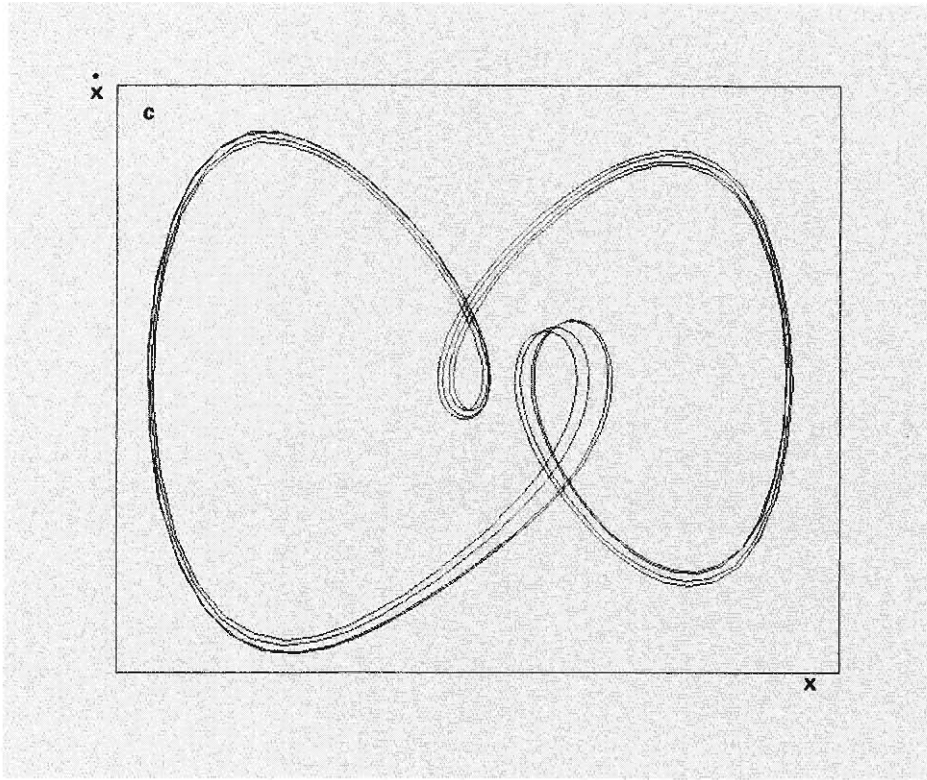


Fig. 12c. The most stable 4-cycle of Duffing's equation. Observe that the displaced copies of Fig. 12b have either a broad or a narrow separation.

Figures 12b and c depict attractors of periods 2 and 4 as the friction or damping constant  $k$  in Eq. (62) is reduced systematically. The parameter values  $k = \lambda_0, \lambda_1, \lambda_2, \dots$ , are the damping constants corresponding to the most stable  $2^n$ -cycle in analogy to the  $\lambda_n$  of the one-dimensional functional iteration. Indeed, this oscillator's period doubles (at least numerically!) *ad infinitum*. In fact, by  $k = \lambda_5$ , the  $\delta_3$  of Eq. (2) has converged to 4.69. Why is this? Instead of considering the entire trajectories as shown in Fig. 12, let us consider only where the trajectory point is located every 1 period of the drive. The 1-cycle then produces only one point, while the 2-cycle produces a pair of points, and so forth. This *time-one map* [if the trajectory point is  $(x, \dot{x})$  now, where is it one period later?] is by virtue of the differential equation a smooth and invertible func-

tion in two dimensions. Qualitatively, it looks like the map of Eq. (61). In the present state of mathematics, little can be said about the analytic behavior of time-one maps; however, since our theory is universal, it makes no difference that we don't know the explicit form. We still can determine the complete quantitative behavior of Eq. (62) in the onset regime where the motion tends to aperiodicity. If we already know, by measurement, the precise form of the trajectory after a few period doublings, we can compute the form of the trajectory as the friction is reduced throughout the region of onset of complexity by carefully using the full power of the universality theory to determine the spacings of elements of a cycle.

Let us see how this works in some detail. Consider the time-one map of the

Duffing's oscillator in the superstable  $2^n$ -cycle. In particular, let us focus on an element at which the scaling function  $\sigma$  (Fig. 10) has the value  $\sigma_0$ , and for which the next iterate of this element also has the scaling  $\sigma_0$ . (The element is not at a big discontinuity of  $\sigma$ .) It is then intuitive that if we had taken our time-one examination of the trajectory at values of time displaced from our first choice, we would have seen the same scaling  $\sigma_0$  for this part of the trajectory. That is, the differential equations will extend the map-scaling function continuously to a function along the entire trajectory so that, if two successive time-one elements have scaling  $\sigma_0$ , then the entire stretch of trajectory over this unit time interval has scaling  $\sigma_0$ . In the last section, we were motivated to construct  $\sigma$  as a function of  $t$  along an interval precisely towards this end.

To implement this idea, the first step is to define the analogue of  $d_n$ . We require the spacing between the trajectory at time  $t$  and at time  $T_n/2$  where the period of the system in the  $2^n$ -cycle is

$$T_n \cong 2^n T_0. \quad (63)$$

That is, we define

$$d_n(t) \equiv x_n(t) - x_n(t + T_n/2). \quad (64)$$

(There is a  $d$  for each of the  $N$  variables for a system of  $N$  differential equations.) Since  $\sigma$  was defined as periodic of period 1, we now have

$$d_{n+1}(t) \sim \sigma(t/T_{n+1})d_n(t). \quad (65)$$

The content of Eq. (65), based on the independence arising solely through the  $T_n$  in  $\sigma$ , and not on the detailed form of  $\sigma$ , already implies a strong scaling prediction, in that the ratio

$$\frac{d_{n+1}(t)}{d_n(t)},$$

when plotted with  $t$  scaled so that  $T_n =$

1, is a function *independent* of  $n$ . Thus if Eq. (65) is true for *some*  $\sigma$ , whatever it might be, then knowing  $x_n(t)$ , we can compute  $d_n(t)$  and from Eq. (65)  $d_{n+1}(t)$ . As a consequence of periodicity, Eq. (64) for  $n \rightarrow n + 1$  can be solved for  $x_{n+1}(t)$  (through a Fourier transform). That is, if we have measured any chosen coordinate of the system in its  $2^n$ -cycle, we can compute its time dependence in the  $2^{n+1}$ -cycle. Because this procedure is recursive, we can compute the coordinate's evolution for all higher cycles through the infinite period-doubling limit. If Eq. (65) is true and  $\sigma$  not known, then by measurement at a  $2^n$ -cycle and at a  $2^{n+1}$ -cycle,  $\sigma$  could be *constructed* from Eq. (65), and hence all higher order doublings would again be determined. Accordingly, Eq. (65) is a very powerful result. However, we know much more. The universality theory tells us that period doubling is universal and that there is a *unique* function  $\sigma$  which, indeed, we have computed in the previous section. Accordingly, by *measuring*  $x(t)$  in some chosen  $2^n$ -cycle (the higher the  $n$ , the more the number of effective parameters to be determined empirically, and the more precise are the predictions), we now can compute the entire evolution of the system on its route to turbulence.

How well does this work? The empirically determined  $\sigma$  [for Eq. (62)] of Eq. (65) is shown for  $n = 3$  in Fig. 13a and  $n = 4$  in Fig. 13b. The figures were constructed by plotting the ratios of  $d_{n+1}$  and  $d_n$  scaled respective to  $T = 16$  in Fig. 13a and  $T = 32$  in Fig. 13b. Evidently the scaling law Eq. (65) is being obeyed. Moreover, on the same graph Fig. 14 shows the empirical  $\sigma$  for  $n = 4$  and the recursion theoretical  $\sigma$  of Fig. 10. The reader should observe the detail-by-detail agreement of the two. In fact, if we use Eq. (65) and the theoretical  $\sigma$  with  $n = 2$  as empirical input, the  $n = 5$  frequency spectrum agrees with the empirical  $n = 5$  spectrum to

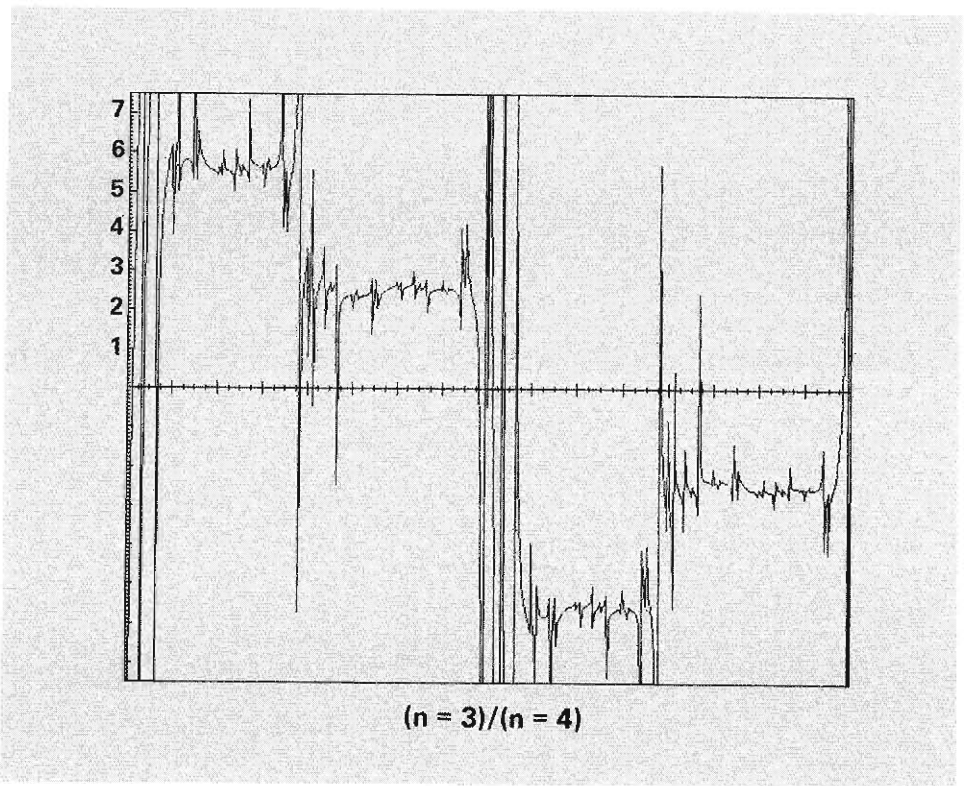


Fig. 13a. The ratio of nearest copy separations in the 8-cycle and 16-cycle for Duffing's equation.

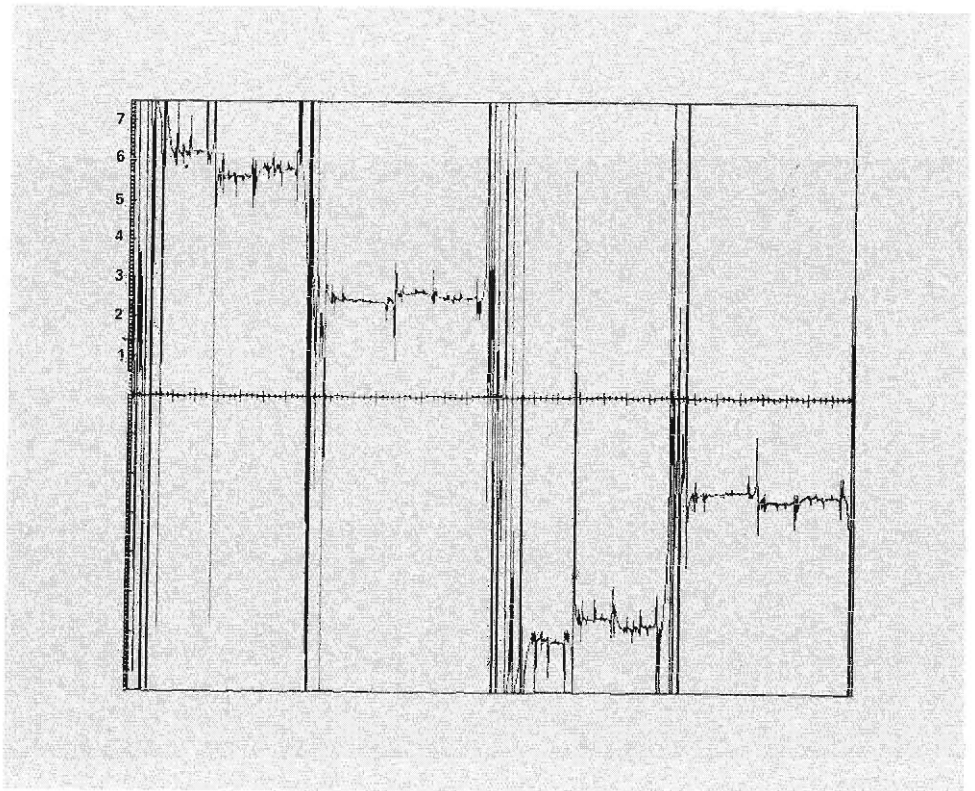


Fig. 13b. The same quantity as in Fig. 13a, but for the 16-cycle and 32-cycle. Here, the time axis is twice as compressed.

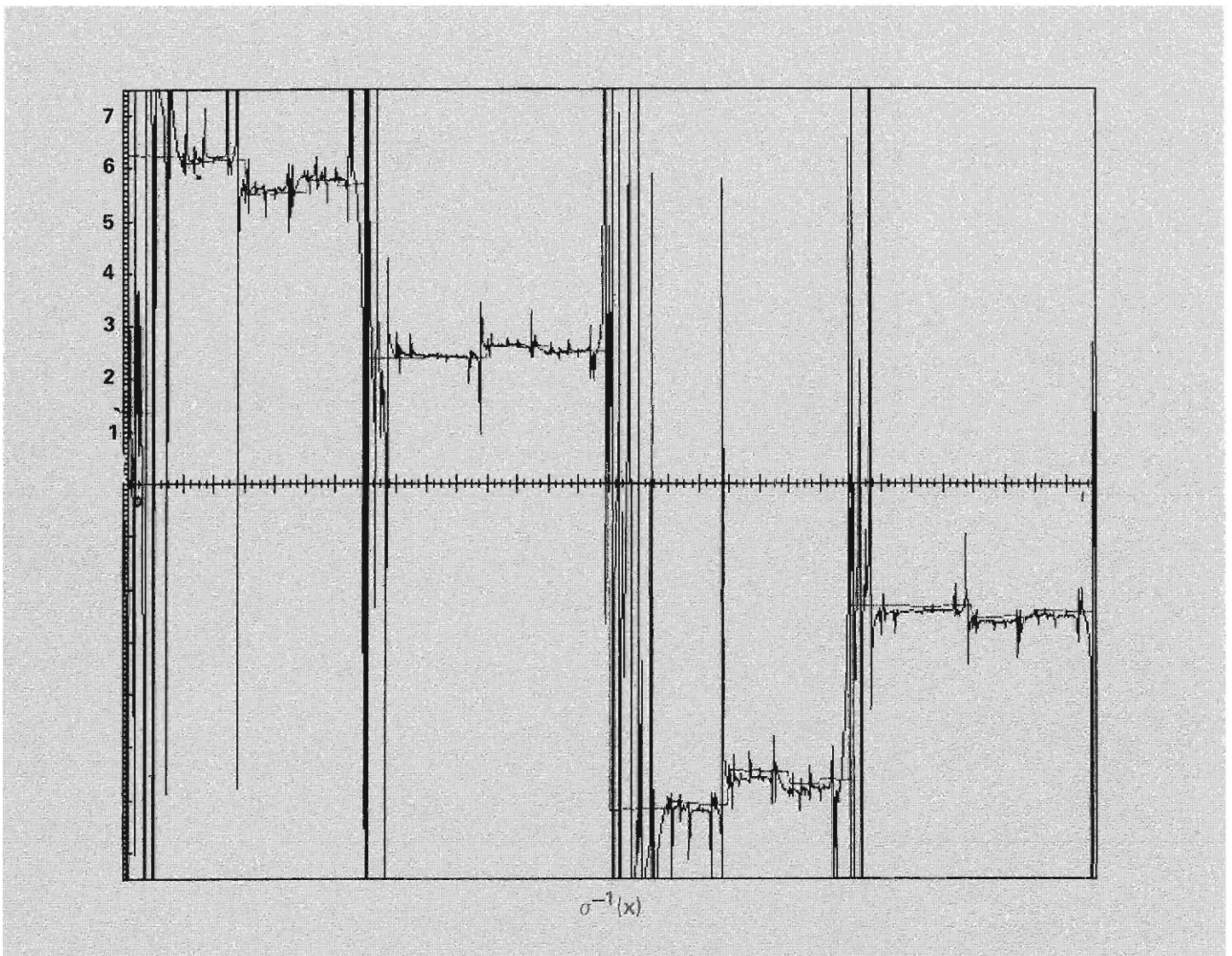


Fig. 14. Figure 13b overlaid with Fig. 10 compares the universal scaling function  $\sigma$  with the empirically determined scaling of nearest copy separations from the 16-cycle to the 32-cycle for Duffing's equation.

within 10%. (The  $n = 4$  determines  $n = 5$  to within 1%.) Thus the asymptotic universality theory is correct *and* is already well obeyed, even by  $n = 2$ !

Equations (64) and (65) are solved, as mentioned above, through Fourier transforming. The result is a recursive scheme that determines the Fourier coefficients of  $x_{n+1}(t)$  in terms of those of  $x_n(t)$  and the Fourier transform of the (known) function  $\sigma(t)$ . To employ the formula accurately requires knowledge of the entire spectrum of  $x_n$  (amplitude *and* phase) to determine each coefficient of  $x_{n+1}$ . However, the formula enjoys an

approximate local prediction, which roughly determines the amplitude of a coefficient of  $x_{n+1}$  in terms of the amplitudes (alone) of  $x_n$  near the desired frequency of  $x_{n+1}$ .

What does the spectrum of a period-doubling system look like? Each time the period doubles, the fundamental frequency halves; period doubling in the continuum version is termed half-subharmonic bifurcation, a typical behavior of coupled nonlinear differential equations. Since the motion *almost* reproduces itself every period of the drive, the amplitude at this original fre-

quency is high. At the first subharmonic halving, spectral components of the odd halves of the drive frequency come in. On the route to aperiodicity they saturate at a certain amplitude. Since the motion more nearly reproduces itself every two periods of drive, the next saturated subharmonics, at the odd fourths of the original frequency, are smaller still than the first ones, and so on, as each set of odd  $2^n$ ths comes into being. A crude approximate prediction of the theory is that whatever the system, the saturated amplitudes of each set of successively lower half-frequencies

define a smooth interpolation located 8.2 dB below the smooth interpolation of the previous half-frequencies. [This is shown in Fig. 15 for Eq. (62).] After subharmonic bifurcations *ad infinitum*, the system is now no longer periodic; it has developed a continuous broad spectrum down to zero frequency with a definite internal distribution of the energy. That is, the system emerges from this process having developed the beginnings of broad-band noise of a determined nature. This process also occurs in the onset of turbulence in a fluid.

### The Onset of Turbulence

The existing idea of the route to turbulence is Landau's 1941 theory. The idea is that a system becomes turbulent through a succession of instabilities, where each instability creates a new degree of freedom (through an indeterminate phase) of a time-periodic nature with the frequencies successively higher and incommensurate (*not* harmonics); because the resulting motion is the superposition of these modes, it is quasi-periodic.

In fact, it is experimentally clear that quasi-periodicity is incorrect. Rather, to produce the observed noise of rapidly decaying correlation the spectrum must become *continuous* (broad-band noise) down to zero frequency. The defect can be eliminated through the production of successive half-subharmonics, which then emerge as an allowable route to turbulence. If the general idea of a succession of instabilities is maintained, the new modes do *not* have indeterminate phases. However, only a small number of modes need be excited to produce the required spectrum. (The number of modes participating in the transition is, as of now, an open experimental question.) Indeed, knowledge of the phases of a small number of amplitudes at an early stage of period doubling suffices to determine the phases of the transition

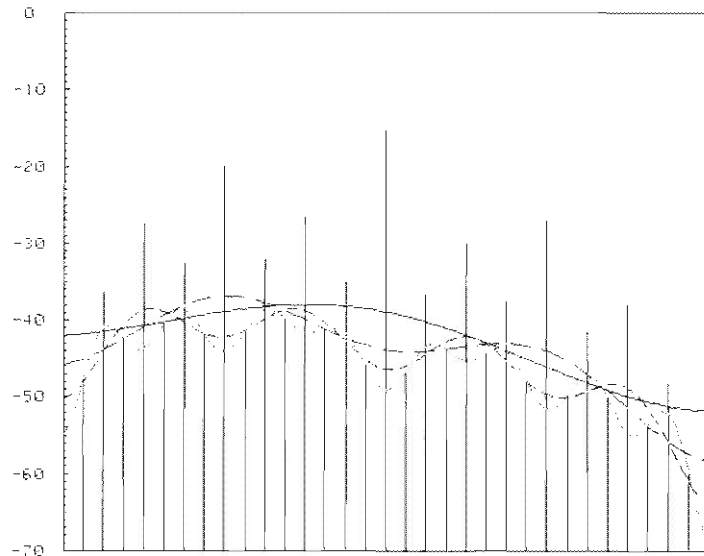


Fig. 15. The subharmonic spectrum of Duffing's equation in the 32-cycle. The dotted curve is an interpolation of the odd 32nd subharmonics. The shorter dashed curve is constructed similarly for the odd 16th subharmonics, but lowered by 8.2 dB. The longer dashed curve of the 8th subharmonics has been dropped by 16.4 dB, and the solid curve of the 4th subharmonics by 24.6 dB.

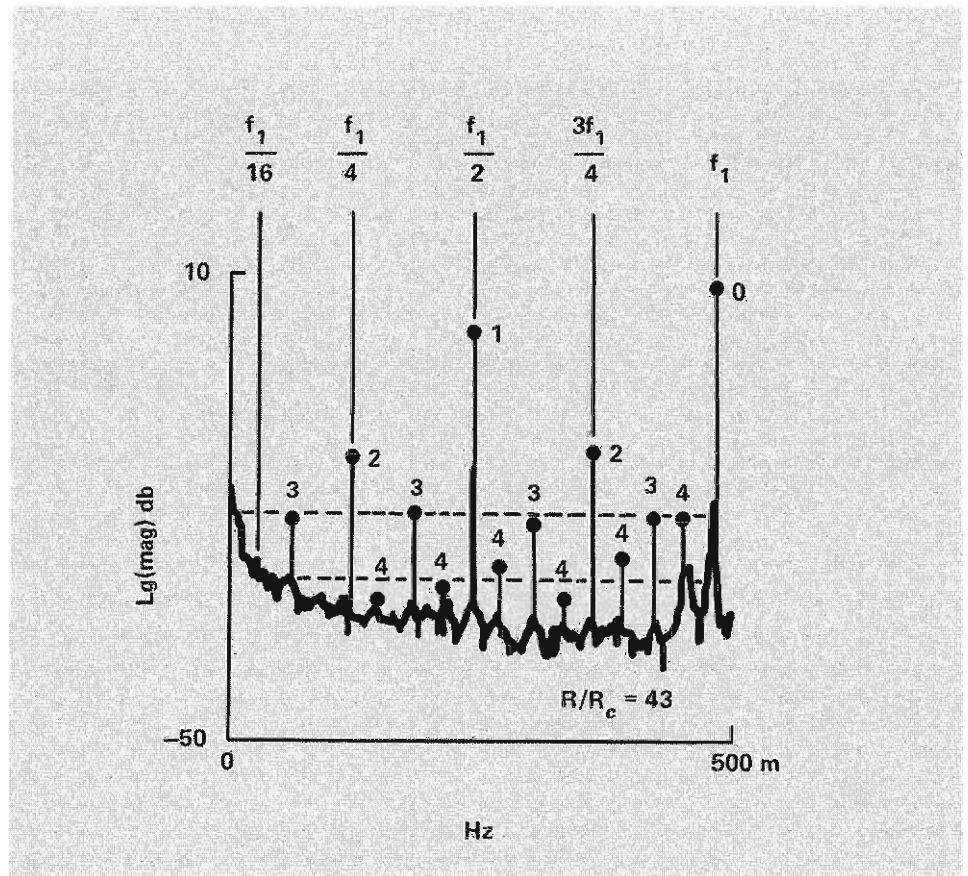


Fig. 16. The experimental spectrum (redrawn from Libchaber and Maurer) of a convecting fluid at its transition to turbulence. The dashed lines result from dropping a horizontal line down through the odd 4th subharmonics (labelled 2) by 8.2 and 16.4 dB.



Mitchell J. Feigenbaum is a mathematical physicist and inventor of the universality theory. He earned his bachelor of electrical engineering degree from the City College of New York in 1964 and his Ph.D. in elementary particle physics from Massachusetts Institute of Technology in 1970. He was a Research Associate at Cornell University and Virginia Polytechnic Institute before becoming a Staff Member in the Theoretical Division at LASL in 1974. In May, 1980, he received a LASL Distinguished Performance Award for his universality theory and development of the first quantitative understanding of the onset of turbulence. He is presently writing a monograph on this work for the series *Progress in Physics* published by Birkhauser Boston, Inc. His other technical interests include field theories and functional integrals, the renormalization group and phase transitions, and nonlinear dynamics generally. His peers at LASL often turn to him as the resident consultant on related technical matters.

spectrum. What is important is that a purely causal system can and does possess essentially statistical properties. Invoking *ad hoc* statistics is unnecessary and generally incompatible with the true dynamics.

A full theoretical computation of the onset demands the calculation of successive instabilities. The method used traditionally is perturbative. We start at the static solution and add a small time-dependent piece. The fluid equations are linearized about the static solution, and the stability of the perturbation is studied. To date, only the first instability has been computed analytically. Once we know the parameter value (for example, the Rayleigh number) for the onset of this first time-varying instability, we must determine the correct form of the solution after the perturbation has grown large *beyond* the linear regime. To this solution we add a new time-dependent perturbative mode, again linearized (now about a time-varying, nonanalytically available solution) to discover the new instability. To date, the second step of the analysis has been performed only numerically. This process, in principle, can be repeated again and again until a suitably turbulent flow has been obtained. At each successive stage, the computation grows successively more intractable.

However, it is just at this point that the universality theory solves the problem; it works only after enough in-

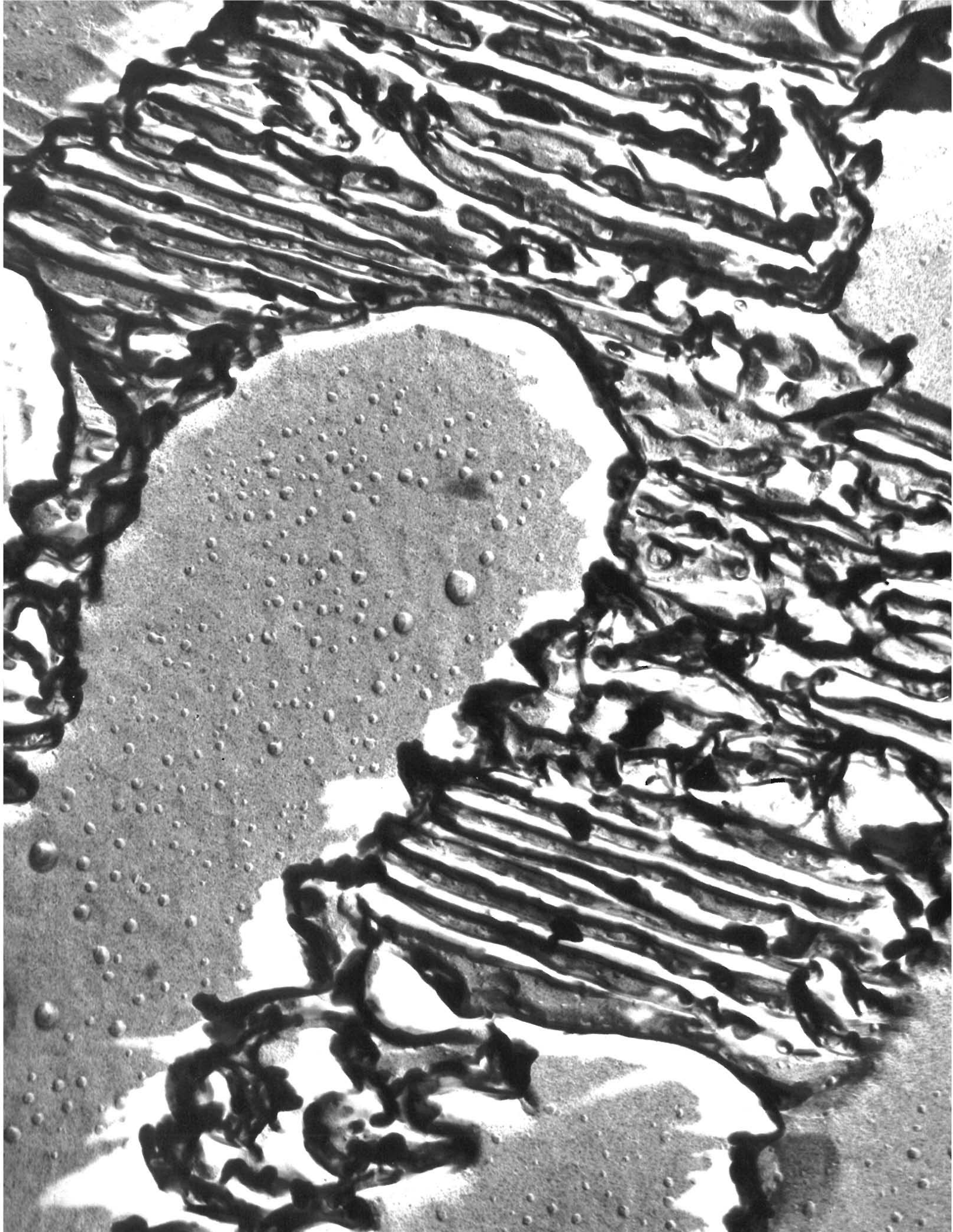
stabilities have entered to reach the asymptotic regime. Since just two such instabilities already serve as a good approximate starting point, we need only a few parameters for each flow to empower the theory to complete the hard part of the infinite cascade of more complex instabilities.

Why should the theory apply? The fluid equations make up a set of coupled field equations. They can be spatially Fourier-decomposed to an infinite set of coupled ordinary differential equations. Since a flow is viscous, there is some smallest spatial scale below which no significant excitation exists. Thus, the equations are effectively a finite coupled set of nonlinear differential equations. The number of equations in the set is completely irrelevant. The universality theory is generic for such a dissipative system of equations. Thus it is possible that the flow exhibits period doubling. If it does, then our theory applies. However, to prove that a given flow (or any flow) actually should exhibit doubling is well beyond present understanding. All we can do is experiment.

Figure 16 depicts the experimentally measured spectrum of a convecting liquid helium cell at the onset of turbulence. The system displays measurable period doubling through four or five levels; the spectral components at each set of odd half-subharmonics are labelled with the level. With  $n = 2$  taken as

asymptotic, the dotted lines show the crudest interpolations implied for the  $n = 3$ ,  $n = 4$  component. Given the small amount of *amplitude* data, the interpolations are perforce poor, while ignorance of higher odd multiples prevents construction of any significant interpolation at the right-hand side. Accordingly, to do the crudest test, the farthest right-hand amplitude was dropped, and the oscillations were smoothed away by averaging. The experimental results,  $-8.3$  dB and  $-8.4$  dB, are in surprisingly good agreement with the theoretical 8.2!

From this good experimental agreement and the many period doublings as the clincher, we can be confident that the measured flow has made its transition according to our theory. A measurement of  $\delta$  from its fundamental definition would, of course, be altogether convincing. (Experimental resolution is insufficient at present.) However, if we work backwards, we find that the several percent agreement in 8.2 dB is an *experimental observation* of  $a$  in the system to the same accuracy. Thus, the present method has provided a theoretical calculation of the actual dynamics in a field where such a feat has been impossible since the construction of the Navier-Stokes equations. In fact, the scaling law Eq. (65) transcends these equations, and applies to the *true* equations, whatever they may be.



# High-Temperature Superconductivity: A Metallurgical Approach

by Angelo L. Giorgi, Gregory R. Stewart, James L. Smith, and Bernd T. Matthias

A mysterious 30-fold enhancement of the critical temperature in the yttrium-iridium system has been traced to the formation of a eutectic structure and to a dramatic decrease in the stiffness of the crystal lattice.

Superconductivity is the sudden complete disappearance of electrical resistance in some materials when they are cooled below a critical temperature. This phenomenon (Fig. 1) has intrigued solid-state scientists, metallurgists, and engineers ever since its discovery by Kammerlingh Onnes in 1911. Visions of many possible applications of superconductivity to electrical power generation and distribution, fusion reactors, high-energy particle accelerators, and propulsion systems, and of the many new superconducting devices have provided a powerful stimulus to the search for superconducting materials with high superconducting critical temperature ( $T_c$ ). Development of a microscopic theory on superconductivity (the BCS theory) in 1957 brought hope that the theory would prove useful in predicting new high  $T_c$  materials. Unfortunately, none of the predictions have been successful. Many of the new high  $T_c$  superconductors shown in Table I, which form the basis for the present superconducting

technology, were discovered by application of Matthias' Rule, an empirical rule developed in the 1950s. This rule relates the  $T_c$  of elements, alloys, and compounds to their average number of valence electrons per atom ( $e/a$ ) as shown in Fig. 2.

In 1964, when workers at the Bell Laboratories were examining mixtures of the two metals yttrium and iridium as part of a study of the relationship between the  $T_c$  and the  $e/a$  ratio, they made a puzzling observation. When they added as little as 1 atomic per cent yttrium to iridium, the temperature at which the material became a superconductor quickly increased from 0.1 K to above 3 K. When they increased the amount of yttrium to 33 atomic per cent, the  $T_c$  remained at 3 K. What was causing this 30-fold increase in the superconducting critical temperature? The enhanced superconductivity was certainly not due to the small change in the  $e/a$  ratio produced by the addition of yttrium, nor could it be explained from the known materials in the system. The iridium  $T_c$  is

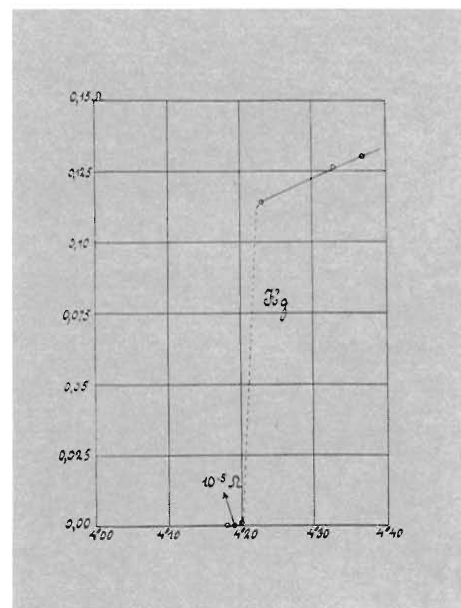


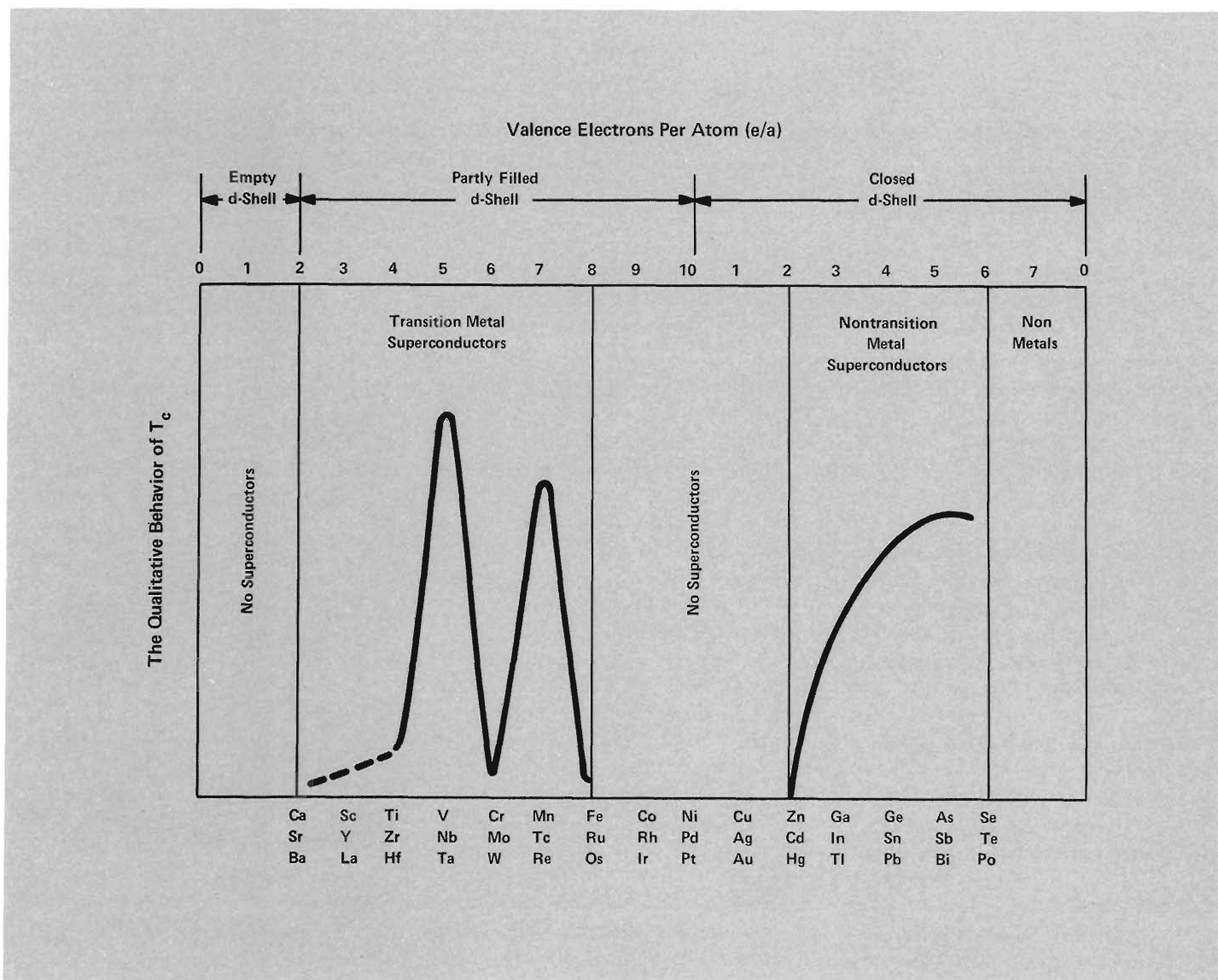
Fig. 1. A reproduction of Onnes' original data, which marked the discovery of superconductivity in 1911. The plot of resistance in ohms vs absolute temperature shows the complete disappearance of electrical resistance in mercury at 4.2 K.

only 0.1 K, and neither yttrium nor  $YIr_2$  (the only known compound in the system) become superconductors even when they are cooled to 0.3 K. The immediate conclusion was that another compound, probably with a composition close to  $YIr_4$ , must exist and that this hypothetical phase was the source of the superconductivity. Many samples, prepared with compositions varying between iridium and  $YIr_2$ , were heat treated and examined by x-ray diffraction. All attempts to discover a new phase were unsuccessful. The source of the enhanced superconductivity

TABLE I  
HIGH  $T_c$  SUPERCONDUCTORS

Compound	$T_c$ (K)	Compound	$T_c$ (K)
$Nb_3Ge$	23.2	$NbC_{.25}N_{.75}$	17.8
$Nb_3Ga$	20.3	$Nb_3Al$	18.8
$Nb_3Al_{.8}Ge_{.2}$	20.1	$V_3Ga$	16.5
$Nb_3Sn$	18.1	$V_3Si$	17.1

Fig. 2. Empirical behavior of  $T_c$  as a function of the average number of valence electrons per atom ( $e/a$ ) for pure elements and alloys. This empirical relationship was discovered by B. T. Matthias in 1957.





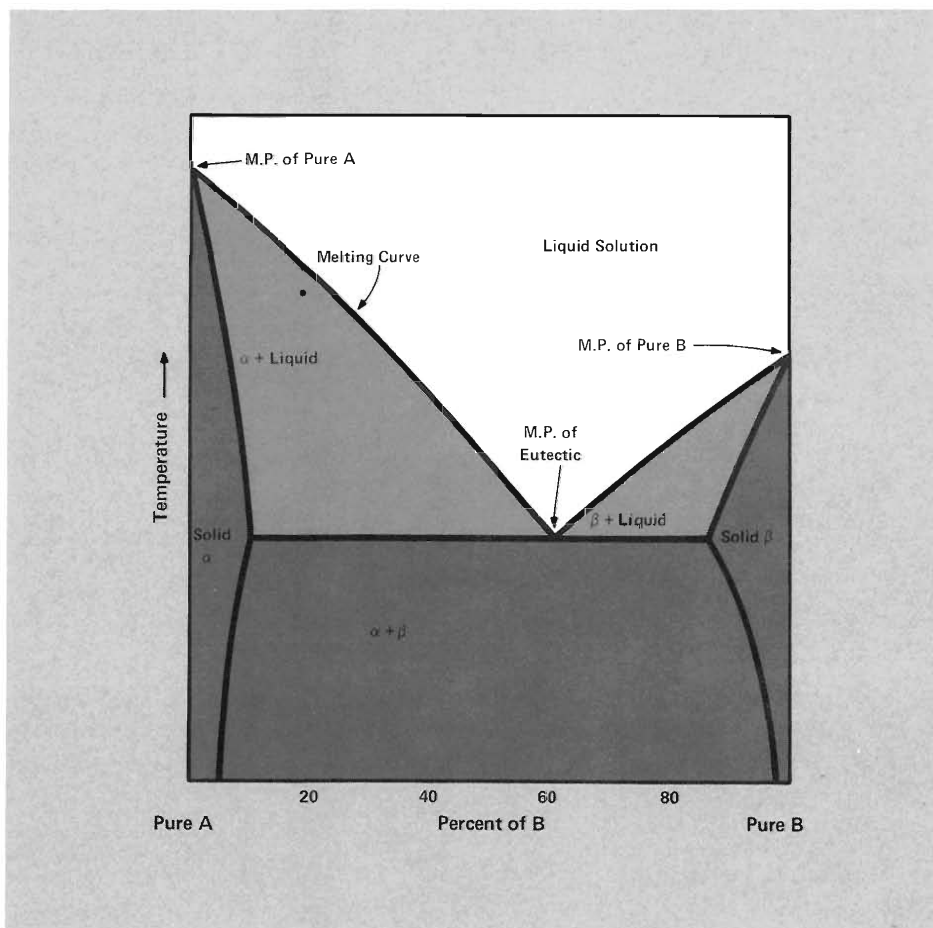


Fig. 3. A typical phase diagram for a mixture of metals showing melting point (M.P.) minimum at the eutectic composition. Solution of B into A lowers the M.P. of A; similarly, solution of A into B lowers the M.P. of B. Solid  $\alpha$  is a solid solution of B atoms dissolved in the crystal lattice of A. Solid  $\beta$  is a solid solution of A atoms in the crystal lattice of B.

remained a mystery. The investigation was finally abandoned and all but forgotten.

This year, when interest in the Y-Ir system was revived, a much more thorough study was conducted at the Los Alamos Scientific Laboratory (LASL) in collaboration with the University of California (UC) at La Jolla, California. The careful characterization of the various compositions was extended to include low-temperature specific heat measurements, metallographic examination, and transmission electron microscopy as well as the usual x-ray diffraction and magnetic susceptibility measurements. From this study, we have learned that the source of the enhanced superconductivity in the Y-Ir system is a eutectic structure consisting of a mixture of iridium and the neighboring phase,  $YIr_2$ .

A eutectic is the unique mixture of two constituents, usually metals, that has the lowest melting point. Eutectics have been known since the days of the Roman Empire and have been in continued use since then. Bronze and solder are common examples. A typical phase diagram for a eutectic mixture is shown in Fig. 3. The melting point for the eutectic composition is considerably lower than for either constituent.

The low-temperature specific heat measurements on the Y-Ir samples disclosed that, with the formation of the eutectic structure, the stiffness of the lattice decreases dramatically. The enhanced superconductivity probably is a result of the lattice softening. This intriguing solution to the mystery promises a new investigative approach to high-temperature superconductivity.

## Measurement Techniques

Reviewing some of the techniques used to measure the superconducting critical temperature will help explain the problems associated with a study of this type. At present there are over a thousand known superconducting materials. Their critical temperatures range from minimum values of a few millidegrees above absolute zero to the present maximum value of 23.2 K. Most of these materials are intermetallic alloys; that is, they are mixtures, compounds, or solid solutions of two or more metals.

The three methods used most commonly to determine superconductivity in such materials are based on three distinctive thermal and electromagnetic properties of the superconducting state:

1. The complete disappearance of electrical resistance.
2. The complete exclusion of magnetic fields, up to some critical value  $H_c$ , from the body of the superconductor.
3. A sharp increase in the electronic specific-heat of the superconductor at  $T_c$  owing to the marked decrease in the entropy.

In the method based on resistance, the sample is fashioned into a rod or wire and a fixed current is passed through it. As the temperature is lowered, the voltage across the sample is monitored by the four-probe technique illustrated in Fig. 4. When the material becomes

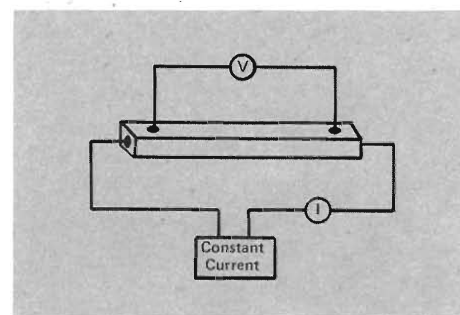


Fig. 4. Four-probe technique for measuring the change in electrical resistance as a sample becomes superconducting.

superconducting, the resistance and, therefore, the voltage suddenly drops to zero. The temperature at which the drop occurs is the  $T_c$  of the material. This method suffers from a demonstrated weakness. It requires a continuous superconducting path across the material, and it gives no indication of how much of the material is superconducting. Quite often the bulk of a material is not superconducting but instead contains microscopic filaments in a superconducting phase. A resistive measurement on such a sample gives results similar to those for a true bulk superconductor because the filaments, having zero resistance, short-circuit the sample and cause the voltage to drop to zero. Using x-ray diffraction to determine the phases present in a sample also can be misleading. If, as is often the case, the filaments in the superconducting phase represent only a small fraction of the total sample, their concentration may lie below the detection limit, and the x-ray diffraction pattern will indicate that the nonsuperconducting bulk material is the only phase present. An investigation limited to these techniques can lead to the erroneous conclusion that the phase representing the bulk material is the superconductor.

A much more widely used technique is the ac susceptibility method, based on the exclusion of magnetic fields (Fig. 5). A stable alternating current of less than 60 hertz is applied to the primary of a sensing coil. The output from two matched secondary coils, connected in opposition, is fed to the input of a frequency-locked amplifier. A small variable mutual inductance is connected between the primary and secondary circuits and adjusted to cancel out any residual imbalance in the secondary circuit. The sample, which can be a powder or solid, is placed in one of the secondary coils and cooled slowly. A sharp change in the sample's magnetic permeability at  $T_c$  causes an imbalance in

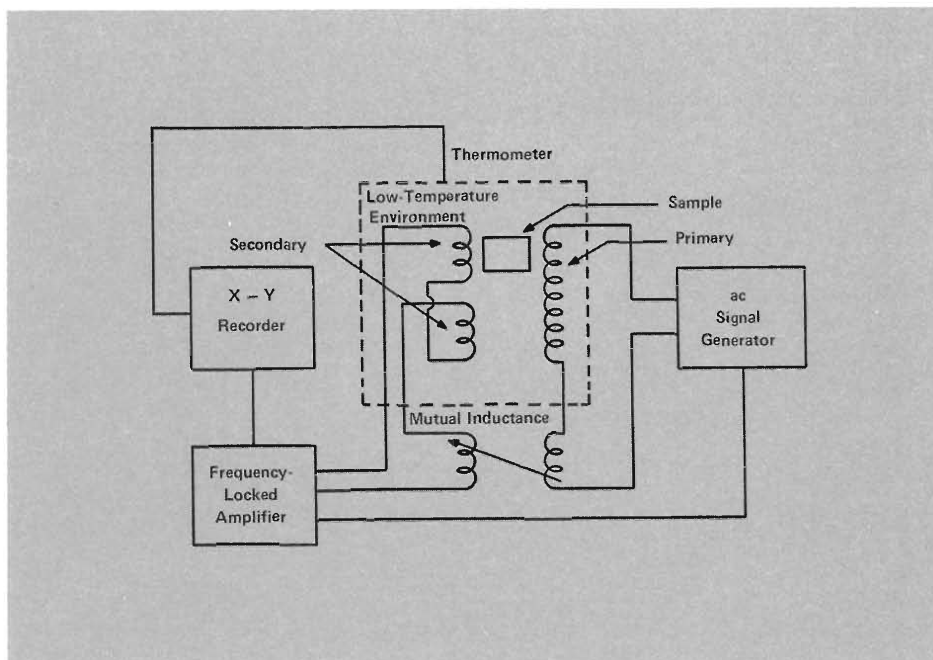


Fig. 5. Technique for measuring the change in magnetic permeability of a sample when it becomes superconducting. A rapid change in the electrical signal from the secondary coils caused by expulsion of the magnetic field from the sample indicates the onset of superconductivity.

the secondary circuit, which results in a signal at the output of the amplifier. Because the signal is proportional to the change in magnetic permeability, it gives a semiquantitative value of the amount of superconductor in the sample.

The ac susceptibility method is preferred over the resistive method because it detects superconductivity in the individual particles and does not depend on a continuous superconducting path. However, even this method can give misleading results. Occasionally, a superconducting phase is deposited as a thin film at the grain boundaries of a bulk material. When the film completely encloses the grains, it acts as a superconducting *can*; that is, it prevents any magnetic flux from penetrating into the *canned* bulk material. Under these conditions, the signal strength from the amplifier suggests that all the material is superconducting.

The method generally accepted as giving a true indication of both the presence and amount of superconducting material is measurement of the variation of the low-temperature specific heat with temperature (Fig. 6). The specific heat ( $C$ ) is the amount of energy ( $dQ$ ) required to raise the temperature of a unit mass (usually a mole) of material by a small increment ( $dT$ ).

$$C = dQ/dT. \quad (1)$$

The specific heat of a metal has two contributions, one from the thermal motion of the electrons surrounding the lattice and one from the lattice vibrations. For a normal metal, the two contributions to the specific heat  $C_n$  have a well-known temperature dependence.

$$C_n = \gamma T + \beta T^3, \quad (2)$$

where  $\gamma$  depends on the properties of the electrons and  $\beta$  depends on the crystal lattice. As shown in Fig. 7, these parameters are determined by plotting  $C_n/T$  vs  $T^2$ . The slope of the resulting straight-line-curve is equal to  $\beta$  and the intercept is  $\gamma$ . The information about the normal-state properties contained in these two parameters is discussed later.

When the metal becomes superconducting, the lattice contributions to the specific heat ( $\beta T^3$ ) remain the same, but the electronic contributions ( $\gamma T$ ) change dramatically. The specific heat ( $C_s$ ) for the superconducting phase is given by

$$C_s = ae^{-\Delta/kT} + \beta T^3. \quad (3)$$

The electronic term for a superconductor is clearly different in temperature dependence. The difference is due to the for-

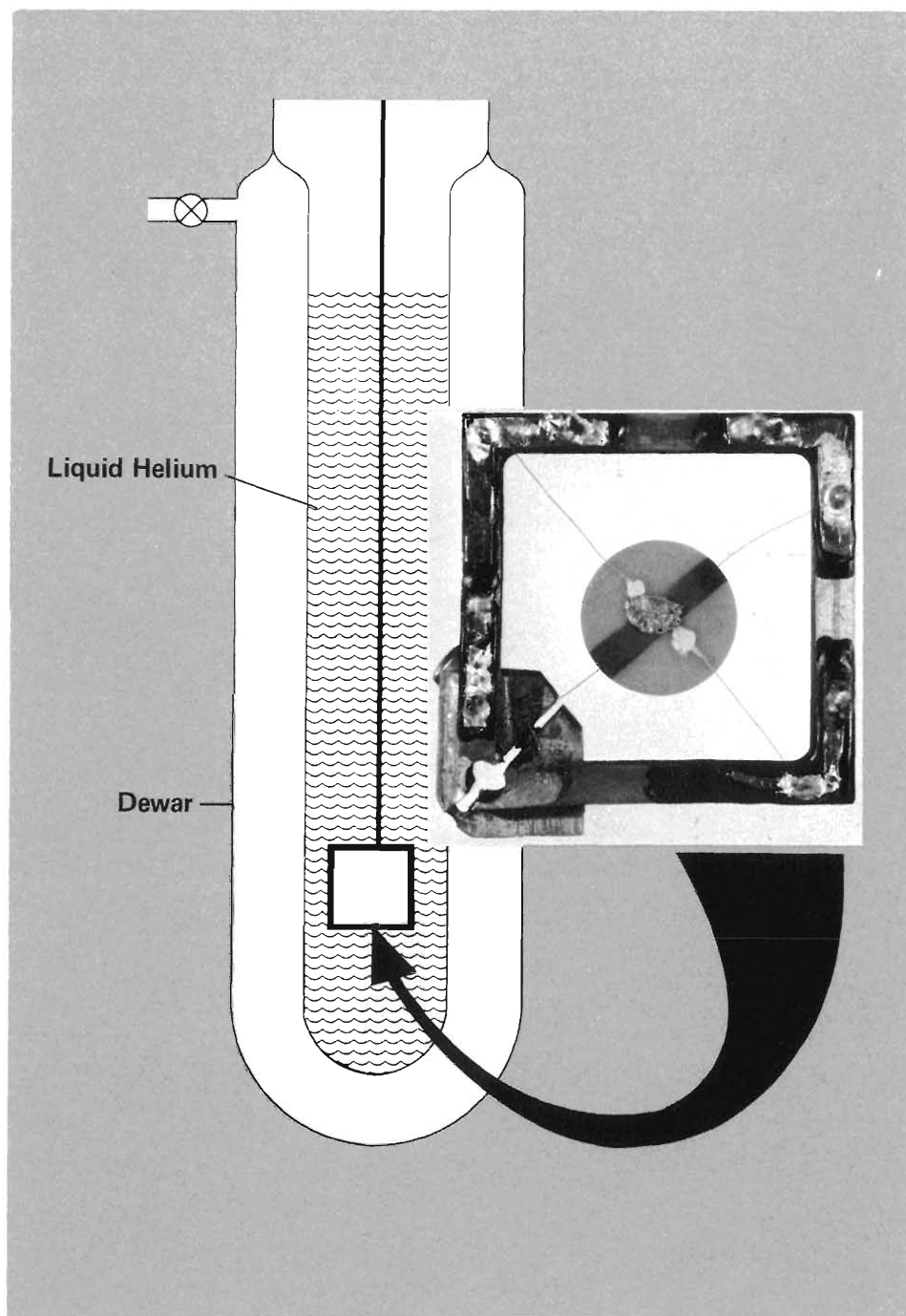
mation, at  $T_c$ , of an energy gap ( $\Delta$ ) in the allowable energy levels that the electrons can occupy in the material. Bound electron pairs of opposite spin and momentum (Cooper pairs) form at  $T_c$  and occupy energy levels below the energy gap. Energy is needed to break up the pairs and raise them to energy levels above the energy gap. Since  $ae^{-\Delta/kT}$  is usually more than twice as large as  $\gamma T$ , the change in the electronic contributions at  $T_c$  causes a marked discontinuity in the specific heat curve (Fig. 7). The discontinuity is the best indication of the presence of a superconducting state.

A generally accepted result of the BCS theory is that, for 100% bulk superconductivity,

$$ae^{-\Delta/kT_c}/\gamma T_c = 2.43. \quad (4)$$

Thus the ratio of the electronic specific heats at  $T_c$  can be used to estimate the amount of the sample that has become superconducting and to rule out ambiguities caused by either microscopic amounts of one-dimensional paths in the resistive method or two-dimensional grain boundary cans in the ac susceptibility method. Further, if x-ray diffraction indicates the presence of a second phase, the amount of second phase estimated from a comparison of the relative intensities of the diffraction patterns can also be compared to the ratio of the electronic contributions to the specific heat.

A much more accurate estimate of the amount of superconducting material is possible for materials with  $T_c$  values at least four times greater than the lowest temperature to which the specific heat can be measured (for example, a  $T_c \geq 4.8$  K and a lowest temperature of 1.2 K). The superconducting phase contributes significantly to the specific heat at  $T_c$ , but its contribution dies away exponentially as the temperature drops below  $T_c$ , when the only terms left in the specific heat behavior vs temperature expression are  $\beta T^3$  and any  $\gamma T$  contribution from nonsuperconducting material. The ratio of this remnant  $\gamma T$  to the value of  $\gamma T$  above  $T_c$ , where the sample is 100% nonsuperconducting, is the exact fraction of the material that is nonsuperconducting.



*Fig. 6. Technique for measuring low-temperature specific heat. The sample, mounted on a sapphire disc 25  $\mu\text{m}$  thick, is suspended in the center of a metal frame by 75- $\mu\text{m}$ -diameter gold wires. The gold wires isolate the sample and platform thermally while providing two separate electrical connections to the metal frame. One connection supplies heat to the sample and platform through a film heater (the dark strip across the platform). The other connection measures the change in temperature by monitoring the change in resistivity of a germanium thermometer. The heat is supplied to the sample and platform in fixed increments by pulses of electrical energy to the film heater. Since sapphire is an excellent thermal conductor, sample and platform equilibrate rapidly after each pulse. By measuring the changes in temperatures for fixed increments in energy, the specific heat for the platform and sample is measured as a function of temperature. The total specific heat minus the known specific heat of the sapphire platform is the specific heat of the sample.*

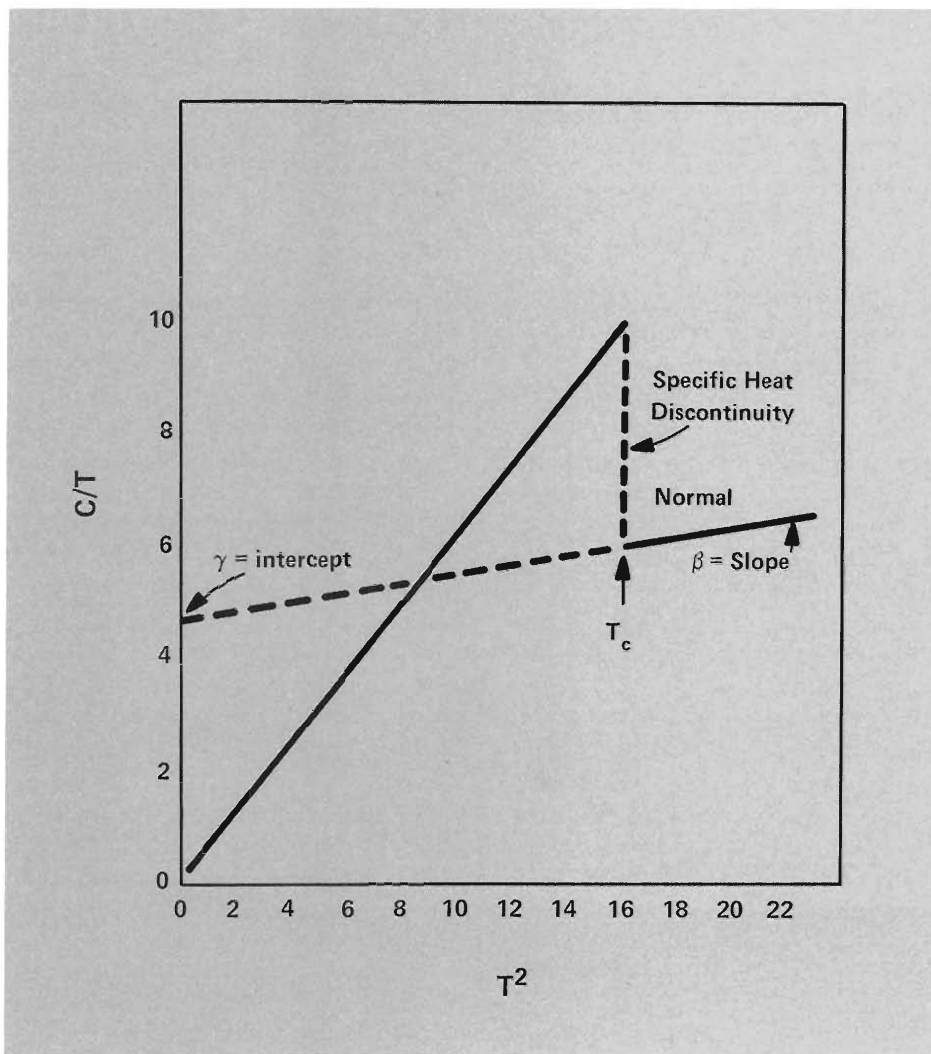


Fig. 7. Idealized specific heat curve for a superconductor with  $T_c \approx 4.0$  K. A significant jump occurs at  $T_c$ . In the absence of a superconducting transition, the specific heat curve follows the dashed line and intercepts the ordinate at  $\gamma$ . Near absolute zero, the lattice contributions go to zero rapidly, and the specific heat is due almost entirely to the electrons.

#### Importance of $\gamma$ and $\beta$

Obviously, certain properties of materials in their normal state must lead to the electrons' overcoming their repulsive Coulomb interaction to form Cooper pairs and create the energy gap  $\Delta$  at the superconducting transition. The purely theoretical considerations have been studied extensively and all known normal-state properties of various superconductors have been compared systematically. As a result, workers in the field are beginning to understand which properties are important for the superconducting state. The parameters  $\gamma$  and  $\beta$  determined by low-temperature specific heat measurement are found to be very important.

The parameter  $\gamma$  is proportional to  $N(0)$ , the number of electrons at the most highly populated energy level (the Fermi level) in the normal-state material, and to  $(1 + \lambda)$ , which depends on how

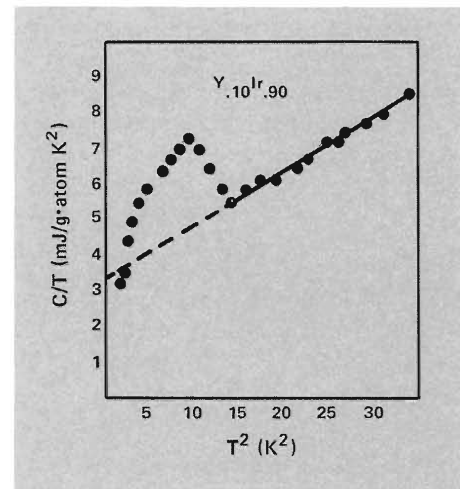


Fig. 8. Measured values of the specific heat for the  $Y_{10}Ir_{90}$  system. Compare these measurements with the idealized curve in Fig. 7.

strongly the material's electrons are tied to its lattice; it is given by

$$\gamma = \text{const.} \times N(0) \times (1 + \lambda). \quad (5)$$

The parameter  $\beta$  is a measure of how stiffly the ions are held in place in the lattice by the material's electronic forces; it is given by

$$\beta = \text{const.}/\theta_D^3, \quad (6)$$

where  $\theta_D$  is the characteristic Debye temperature. The term  $\theta_D$  was first introduced in a model of lattice vibrations proposed by Debye in 1912. The model describes the motion of the lattice ions by a superposition of quantized elastic waves propagating through the lattice. The quanta of elastic waves are called *phonons*. The Debye temperature, which is related to the cut-off frequency, can be thought of as the temperature at which the lattice forces responsible for the lattice stiffness are too weak to hold the ions in their places. Thus lead, a soft material, has a low  $\theta_D$  (108 K), whereas iron, a hard material, has a high  $\theta_D$  (464 K).

In 1957 Bardeen, Cooper, and Schrieffer developed the BCS microscopic theory of superconductivity and succeeded in relating  $T_c$  to  $\gamma$  and  $\beta$ .

$$kT_c = 1.14 \hbar \langle \omega \rangle \exp[-1/N(0)V], \quad (7)$$

where  $\hbar$  is Planck's constant,  $\langle \omega \rangle$  is some average of the lattice phonon frequency related to  $\theta_D$ ,  $N(0)$  is the electronic density of states at the Fermi surface, and  $V$  is the electron phonon interaction parameter. On this basis it was assumed that materials with high  $N(0)$  (high  $\gamma$ ) and high  $\theta_D$  (high  $\beta$ ) would be high  $T_c$  materials.

### Solution of the Y-Ir Mystery

Events leading to discovery of the source of the enhanced superconductivity in the Y-Ir system produced a number of false leads. They demonstrate why low-temperature specific heat and metallographic examinations are vital to any critical study of superconducting materials. We had long known that the addition of as little as 1 atomic per cent yttrium to iridium metal raises the inductive and resistive  $T_c$  from 0.1 K to over 3 K. More recent LASL-UC, La Jolla, studies showed that the behavior is similar when europium is substituted for the yttrium; that is, the addition of a very small amount of europium to iridium causes a dramatic increase in the  $T_c$ . This new result rekindled interest in both the Eu-Ir and Y-Ir systems. We chose the Y-Ir system for a complete study of this novel effect because of the apparent absence of significant second phase for compositions as rich in yttrium as 10 atomic per cent and the absence of superconductivity in  $YIr_2$ , above 0.3 K.

We remeasured Y-Ir compositions from  $Y_{.01}Ir_{.99}$  to  $YIr_2$  using the ac susceptibility method. The signal magnitude in these measurements first increased with increasing yttrium con-

centration, reaching a maximum value near the  $Y_{.20}Ir_{.80}$  composition and then decreased, becoming zero at the  $YIr_2$  composition. The same results had been obtained in the earlier investigation and no doubt led to the suggestion that the superconductivity is due to the presence of a hypothetical phase of approximate  $YIr_4$  composition. To investigate this possibility, we took new x-ray diffraction patterns of powdered samples of the same compositions using a Debye-Scherrer x-ray technique. The x-ray results indicated that no diffraction lines corresponding to a second phase were present in compositions up to 10 atomic per cent yttrium. The materials have a single-phase cubic structure with a relatively constant lattice parameter  $a_0 = 3.8395 \text{ \AA}$ , where  $a_0$  is the distance between lattice sites. This structure, which is identical to and slightly larger than pure iridium ( $a_0 = 3.8389 \text{ \AA}$ ), indicates that at most 1 atomic per cent yttrium is absorbed into the iridium lattice. The location of the remaining yttrium (up to 9 atomic per cent) remained a mystery.

We measured the low-temperature specific heats of the samples to determine how much of each sample was superconducting. Actual measurements for the  $Y_{.10}Ir_{.90}$  system are shown in Fig. 8. The size of the specific heat discontinuity at  $T_c$  indicated the following percentages.

Composition	Superconducting Phase (%)
$Y_{.01}Ir_{.99}$	8
$Y_{.025}Ir_{.90}$	20
$Y_{.10}Ir_{.975}$	56
$Y_{.17}Ir_{.83}$	70

These results are similar to the results observed in the ac susceptibility

measurements and indicate that the superconductivity is indeed due to a bulk superconductor whose concentration increases as the yttrium content increases. The low-temperature specific heat measurements also gave very valuable information on the parameters  $N(0)$  and  $\theta_D$ . As shown in Fig. 9, the value of  $\gamma$ , which is proportional to  $N(0)$ , remains almost unchanged through most of the range, while the value of  $\theta_D$ , which is inversely proportional to  $\beta$ , undergoes an enormous reduction.

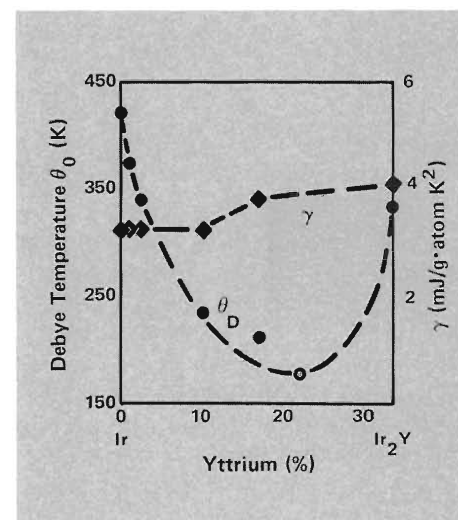
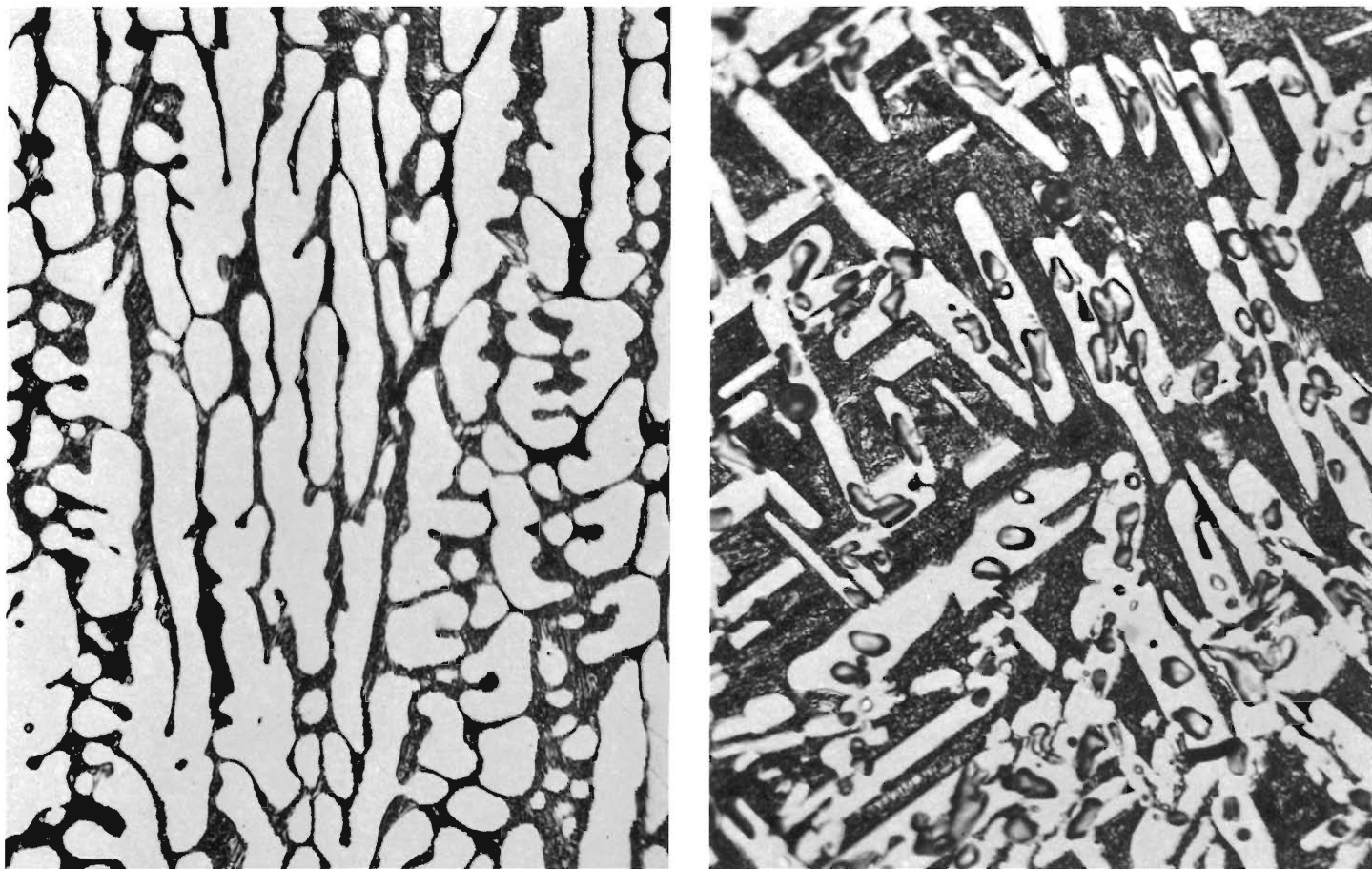


Fig. 9. Measured values of the Debye temperature  $\theta_D$  and the electronic specific heat parameter  $\gamma$  are shown for various compositions in the Y-Ir system. The electronic specific heat remains fairly constant, whereas the Debye temperature falls rapidly to a minimum and rises again as more yttrium is added to the iridium.



**Fig. 10. Metallographic photographs of the various Y-Ir system compositions. The eutectic is visible in all four photographs as the darker regions. At left, this page ( $Y_{.05}Ir_{.95}$ ), the light areas are pure iridium. As more yttrium is added to the system [at right ( $Y_{.20}Ir_{.80}$ )], the amount of eutectic increases; the light areas are probably pure iridium but the precise identification has not been made.**

The x-ray diffraction studies indicated that in compositions up to 10 atomic per cent yttrium only one phase appeared to be present, a cubic lattice of iridium containing less than 1 atomic per cent yttrium. At this point, we believed that the iridium cubic phase was the superconductor, but two questions remained unanswered. Why do increases in the yttrium concentration cause increases in the amount of superconducting phase, and where is the excess yttrium located? The excess yttrium neither enters the iridium lattice nor forms enough of a second phase to account for more than a small fraction of the yttrium present. The mystery of the missing yttrium was solved when we examined the microstructure of the samples using conventional optical metallography and transmission electron microscopy. Microscopic examination showed that a fine-grained eutectic (Fig. 10) was present in all samples and that the amount of the eutectic increased in direct propor-

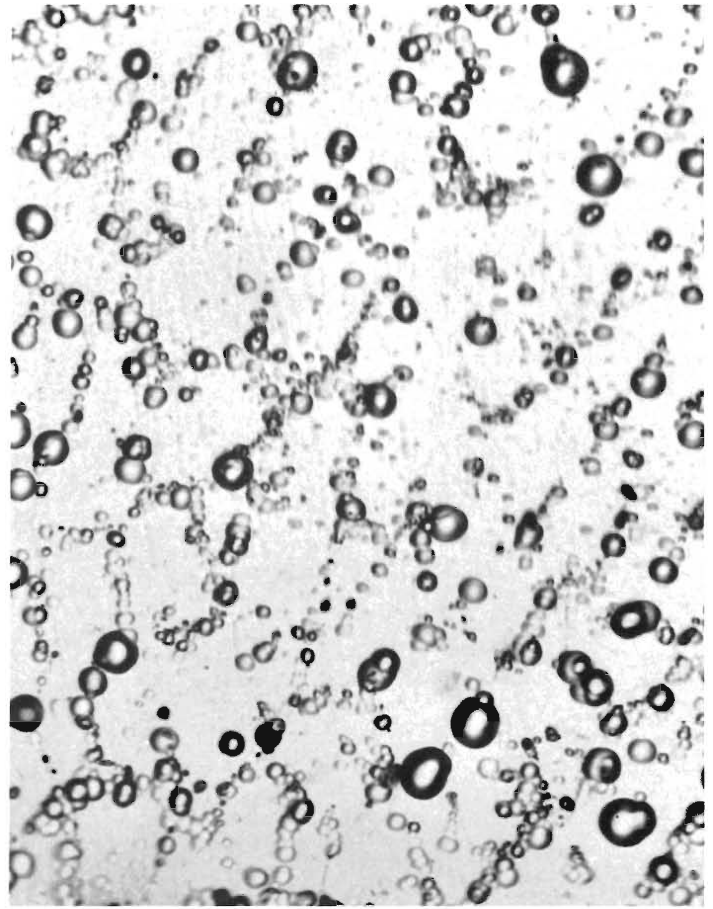
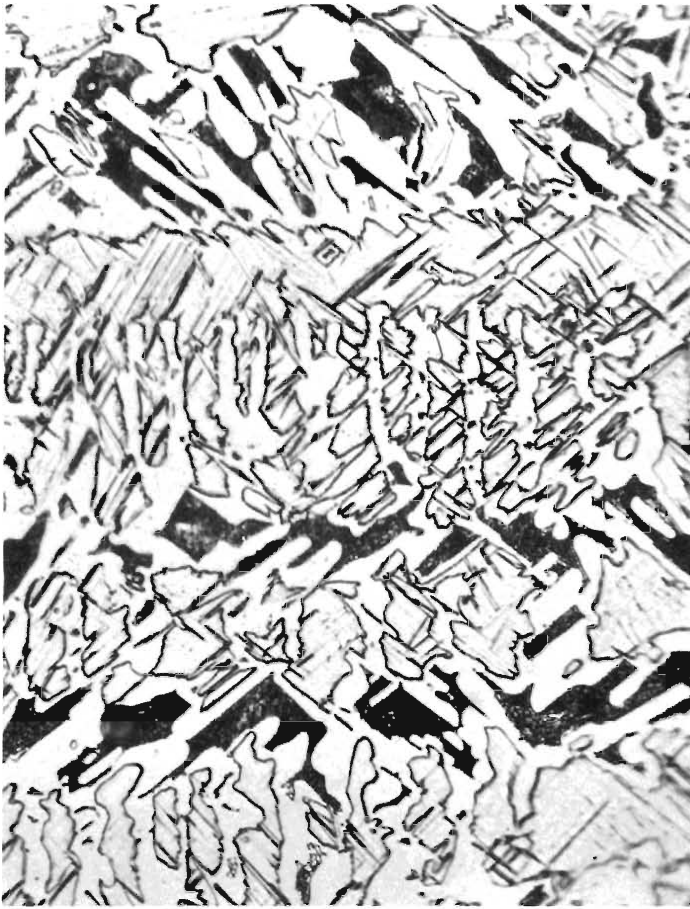
tion to the increase in the amount of bulk superconductor as determined by the low-temperature specific heat measurements. The x-ray results, the metallographic data, and the low-temperature specific heat measurements allow no other conclusion. The bulk enhanced superconductivity is associated with the eutectic and results from the extreme decrease in the lattice stiffness. The failure of the x-ray diffraction studies to detect the eutectic is surprising; it suggests that the structure must consist of extremely small crystallites.

#### **Search for High-Temperature Superconductors**

The search for methods to produce higher  $T_c$  materials has been the subject of innumerable investigations ever since superconductivity was first observed. It has taken such widely different approaches as the attempts of Little (1964)

to produce superconducting polymers, the studies of Abeles and coworkers (1966) on the effect of granular material on  $T_c$ , and the suggestion of Ginzburg (1968) on raising  $T_c$  by the formation of metal-dielectric sandwiches. Unfortunately, to date essentially no success has resulted from all these attempts.

With the development of the BCS theory,  $T_c$  was related to  $\theta_D$ ,  $N(0)$ , and  $V$ , the electron-phonon interaction parameter, as shown in Eq. (7). On this basis, a material with high  $\theta_D$ ,  $N(0)$ , and  $V$  values should have a high  $T_c$ . Unfortunately,  $V$  is not known for most materials and cannot be determined from other parameters. When  $V$  was ignored, it appeared that materials with high  $\theta_D$  and  $N(0)$  values would be high-temperature superconductors. Limited success was achieved in relating high  $N(0)$  values to high  $T_c$ . The dependence of  $T_c$  on  $\theta_D$  appears to be much more complex. We find many examples that support a correlation between high  $T_c$



*With further increase of yttrium [at left, this page ( $Y_{.25}Ir_{.75}$ )] the amount of eutectic begins to decrease; the light areas are probably  $YIr_2$ . At the right ( $Y_{.30}Ir_{.70}$ ), the eutectic, the dark bubble-like structures, has almost disappeared. The lighter areas are  $YIr_2$ . The amount of eutectic for each composition is proportional to the percentage of material that is superconducting as measured by specific heat measurements. (1000X) Metallography by Ramiro A. Pereysa.*

and high  $\theta_D$  and many that contradict it. Technetium ( $\theta_D = 454$  K) has a  $T_c$  of 8.9 K, whereas thallium ( $\theta_D = 78$  K) has a  $T_c$  of 2.4 K. On the other hand, osmium ( $\theta_D = 500$  K) has a  $T_c$  of only 0.66 K, whereas lead ( $\theta_D = 108$  K) has a  $T_c$  of 7.2 K. The present belief is that the electron-phonon interaction parameter  $V$  and the Debye temperature  $\theta_D$  are interrelated:  $V$  varies inversely with  $\theta_D$ . The dependence of  $V$  on  $\theta_D$  is not surprising since the electron-phonon interaction obviously varies with the lattice phonon frequency spectrum. Thus superconductivity should be favored in materials with high  $N(0)$  and/or low  $\theta_D$  values.

Other investigators have observed  $T_c$  enhancement, but never to the extent found in the Y-Ir system. Some enhancement occurs in dilute alloys of titanium, zirconium, or hafnium with the transition metals of groups V to VIII. Of particular interest is the investigation of the Ti-Mo system by Collings and Ho

(1969). They found that the  $T_c$  of the  $\alpha'$  phase increased gradually from 1.5 to 6 K as the molybdenum concentration was increased up to 4.5 atomic per cent. Low-temperature specific heat measurements indicated an increase in  $\gamma$  and a decrease in  $\theta_D$ , whereas metallographic examination showed that a hexagonal-close-packed martensitic structure was forming. Because the increase in  $\gamma$  and, therefore, in  $N(0)$  was not sufficient to explain the increase in  $T_c$ , part of the enhancement was attributed to the lattice softening shown by the drop in  $\theta_D$ . However, since both  $\gamma$  and  $\theta_D$  were changing, the cause of the  $T_c$  enhancement was unclear.

In the Y-Ir system the situation is more straightforward. Because  $\gamma$  remained essentially unchanged while  $\theta_D$  decreased, we can attribute the 30-fold increase in  $T_c$  directly to the modification of the lattice vibrations in the eutectic as compared to the vibrations in the iridium. This is the first documented case

of a system where the effects of  $\theta_D$  and  $N(0)$  can be separated clearly and the enhanced  $T_c$  can be linked directly to the change in the lattice vibrations. Changing the lattice vibrations has been anticipated theoretically as a means to enhance superconductivity. However, the experimental search involved making tiny spheres. The belief was that the vibration patterns of only a few hundred atoms would be severely altered compared to the vibration patterns of the same atoms located in an essentially infinite matrix of the same material. The alteration should result in a dramatic change in the Debye temperature, similar to the change observed in the eutectic. Although it is mere speculation at this time, the possibility that the eutectic in the Y-Ir system is made up of tiny regions of well-ordered atoms that are only slightly affected by the neighboring regions should be considered. If that is so, then the eutectic may well be a bulk form of an ensemble of small particles.

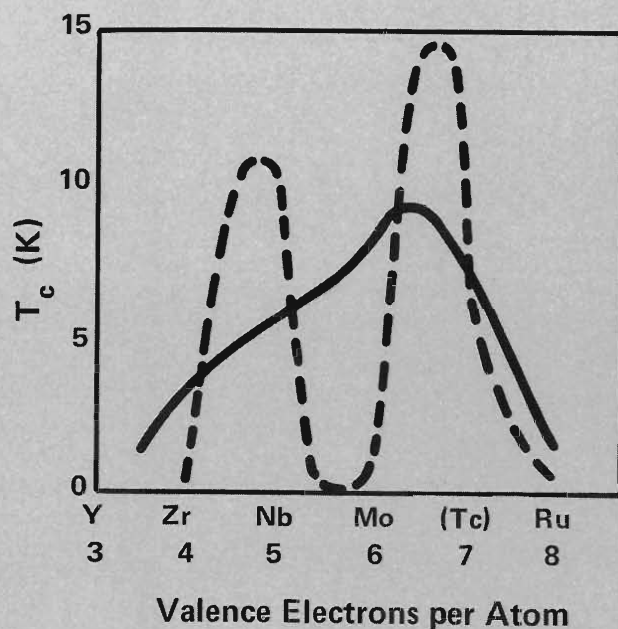


Fig. 11. The variation in  $T_c$  as a function of  $e/a$  for amorphous 4d transition metal alloys is shown by the solid line. The corresponding data for the same materials prepared in crystalline form are shown by the dashed curve. The distinct peaks for the crystalline materials have been broadened by preparing the materials in an amorphous state. The electronic levels appear to be blurred by the breakup of the crystalline structure.

Many implications now can be investigated. Within the field of superconductivity, the most obvious impact may be on amorphous materials. By definition, amorphous materials are disordered and noncrystalline; therefore their properties should be unaltered by radiation damage. Amorphous superconductors have been studied extensively as candidate materials for superconducting magnets in magnetic fusion reactors. Unfortunately, amorphous materials have low  $T_c$ . However, eutectics with high  $T_c$  and extremely small crystalline structures should exhibit similar resistance to radiation damage and should be useful in magnetic fusion applications.

Our work on eutectics suggests why amorphous materials have low  $T_c$ . Figure 2 shows the relationship between  $T_c$  and the  $e/a$  ratio for materials in crystalline form. When Matthias searched among the known superconductors for a relationship between  $T_c$  and the  $e/a$  ratio, he found two sharp peaks in the  $T_c$  curve at  $e/a$  ratios of 4.7 and 6.5 with sharp minimums between the peaks. These variations reflect the variation of  $N(0)$  in the elements and alloys across the periodic table. When Colver and Hammond (1973) prepared the same materials in the amorphous state, they found the peaks simply broadened. That is, the high  $T_c$  values were lowered and the low values were increased (Fig. 11). The simplest explanation of this result is that the sharp maxima in  $N(0)$  at certain compositions are blurred by the microscopic, liquid-like randomness of the atoms. To achieve high  $T_c$  in alloys, the  $N(0)$  blurring must be avoided. The eutectic may be the method of accomplishing this. The microscopic structure of many eutectics can be changed dramatically by varying their solidification conditions. Their particle size varies strongly with cooling



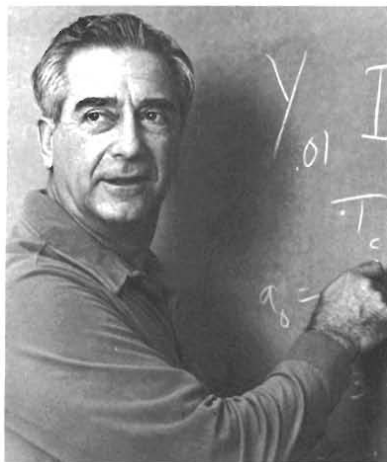
## THE AUTHORS

rates and approaches the amorphous state with very rapid quenches. These changes in structure cause changes in many of the magnetic, mechanical, optical, and thermal properties. As has been demonstrated in the Y-Ir system, it should be possible to prepare eutectics with regions of crystallinity large enough to keep the  $N(0)$  large and sharp, but small enough to ensure that the Debye temperatures are lowered dramatically. Such preparations could lead to much higher  $T_c$  materials and provide a tremendous impetus to the widespread application of superconductivity.

### Understanding Eutectics

The results of the present study may prove to be even more important outside the field of superconductivity. In metallurgy, eutectics have been known and studied for a long time without real progress in understanding when they occur and why they have such altered properties. The lowering of a melting point by mixing one substance with another has been known for a long time—but not understood. No theory or even hypothesis will predict the occurrence of eutectics. Now, the enhancement of  $T_c$  for the first time opens the way to a basic understanding of eutectics. We now realize that lattice softening goes hand in hand with the melting point minimum in the eutectics. We now have an important new piece in the puzzle of eutectic behavior.

The world has relied on eutectics for a long time, from getting rid of ice with salt to making soft solder, and from welding, to alkaline liquid coolants in reactors. Finally, some understanding is on the way; with it comes the possibility of progress toward higher  $T_c$ , technologically more useful superconductors, and perhaps eventually a better understanding of many metallic alloys.



Angelo L. Giorgi earned his bachelor of science degree in chemistry in 1939, came to LASL in 1946, and received his Ph.D. in physical chemistry from the University of New Mexico in 1956. He has over 30 years of varied experience in research, development, and instrumentation in the fields of electrochemistry, radiochemistry, high-temperature phase studies, and cryogenics. His continuing investigations of superconducting and magnetic properties of various carbides and intermetallic compounds have resulted in discoveries of several new superconductors, a new itinerant electron ferromagnet, and the first itinerant antiferromagnet.

Gregory R. Stewart earned his bachelor of science degree cum laude from California Institute of Technology in 1971, and received his Ph.D. from Stanford University in 1975. He did postdoctoral work at Stanford and at the University of Konstanz, FRG. He joined LASL's CMB Division as a Staff Member in 1977, where he developed a new type of small-sample calorimeter and constructed a 7-T superconducting magnet. He is involved in characterizing new materials that range from high  $T_c$  superconductors of technological interest to lower  $T_c$  superconductors, where the interest is in understanding what makes superconductivity occur.

James L. Smith is best known for his definitive work on three of the five known superconducting actinides. He also has studied high-pressure phase transitions of americium. As an offshoot of the work described in the article, he has made interesting discoveries about dilute magnetism; for example, he has discovered magnets that are 100 times more dilute than any found before. He earned his bachelor of science degree at Wayne State University in 1965 and his Ph.D. in physics from Brown University in 1974. He has been a Staff Member in LASL's CMB Division since 1973.

Bernd T. Matthias is Associate Director of the Institute for Pure and Applied Physical Sciences and Professor of Physics at the University of California at San Diego, and part-time member of the technical staff of Bell Telephone Laboratories, Inc. He has been a consultant for LASL since 1957 and was named a Fellow of the Los Alamos Scientific Laboratory in 1971. He is a world authority on superconducting materials, having been involved in the discovery of more of these than any other individual. He is also well known for his discoveries of ferroelectrics and ferromagnets, and for his work on the questions of the coexistence of magnetism and superconductivity. He is a member of the National Academy of Sciences, the American Academy of Arts and Sciences, and the Swiss Physical Society and a Fellow of both the American Physical Society and the American Association for the Advancement of Science. Matthias has received many scientific honors and awards, the two most recent awards being the Oliver E. Buckley Solid State Physics Prize in 1970, and American Physical Society International Prize for New Materials in 1979. The latter was for discovering unusually high temperature superconducting intermetallic compounds and alloys and demonstrating their usefulness in producing high magnetic fields for electric power technology and magnetic confinement of plasma. Originally from Frankfurt, Germany, he earned his Ph.D. from the Swiss Federal Institute of Technology in 1943.

*Fig. 1. A LASL flow cytometer chamber showing a liquid jet emerging from the bottom of the chamber and breaking into droplets. The chamber is illuminated by blue (488 nm) and red (647 nm) lasers. A piezoelectric crystal vibrates the flow chamber so that the jet breaks up into small uniform droplets that are illuminated by a 3-mm-diameter strobed yellow light. The charging collar silhouetted by the yellow light makes it possible to place a positive or negative electrical charge on a few droplets. An individual cancer cell located within a droplet can be sorted by charging the droplet and passing it through a static electric field. White reflections indicate where the laser beam passes through the flow cell.*

# Flow Cytometry

## A New Tool for Quantitative Cell Biology

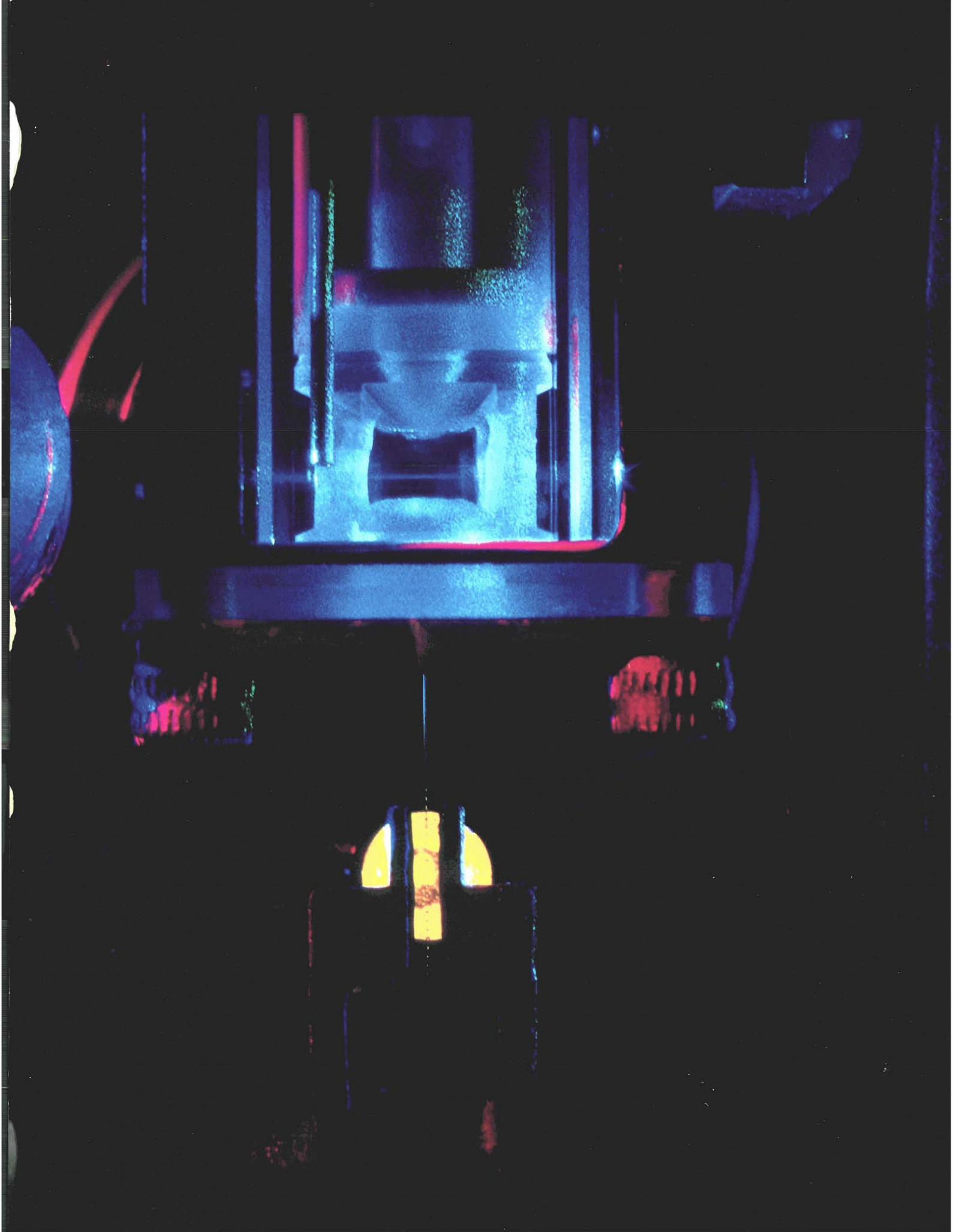
by L. Scott Cram,  
Dale M. Holm,  
and Paul F. Mullaney

The intense light of a laser beam illuminates a single cell stained with fluorescent dye. The cell fluoresces and if the signal is large, it indicates the presence of an elevated complement of DNA. Perhaps the cell is malignant. It is sorted automatically into a separate container for later study. In less than a millisecond, another cell passes through the laser beam to be analyzed quickly and effortlessly. The *flow cytometer* can examine and sort these cells as they flow single file through its narrow passageway at an average rate of 3000 cells per second.

This revolutionary tool (Fig. 1) was developed because biologists needed to analyze and sort individual cells according to specific characteristics.

The flow cytometer allows us to measure cellular properties and the dynamics of changes in those properties accurately in large populations of cells. At present, we can measure cell size, DNA content, the presence of specific antibodies, the permeability of cell membranes to particular molecules, the

migration of specific receptors on a cell surface, certain chemical reaction rates within cells, and the shapes and sizes of individual chromosomes—and the list grows longer every year. We can detect rare events occurring at frequencies of 1 in 1000 or 1 in 10,000 cells and can determine small differences in cell size, DNA content, or other properties among different sample populations of cells. Moreover, the sophisticated instrumentation of flow cytometry allows all these measurements to be made with great precision and high statistical accuracy. Most of these applications depend on tagging specific biological molecules in a cell with a fluorescent dye and measuring the fluorescence signal generated as the cell passes through the flow cytometer. A combination of hydrodynamic, optical, and electronic design ensures that cells are measured one at a time. The cells are illuminated uniformly by the laser beam so that the intensity and duration of fluorescence signals reflect the concentration and, in some instruments, the location of stained



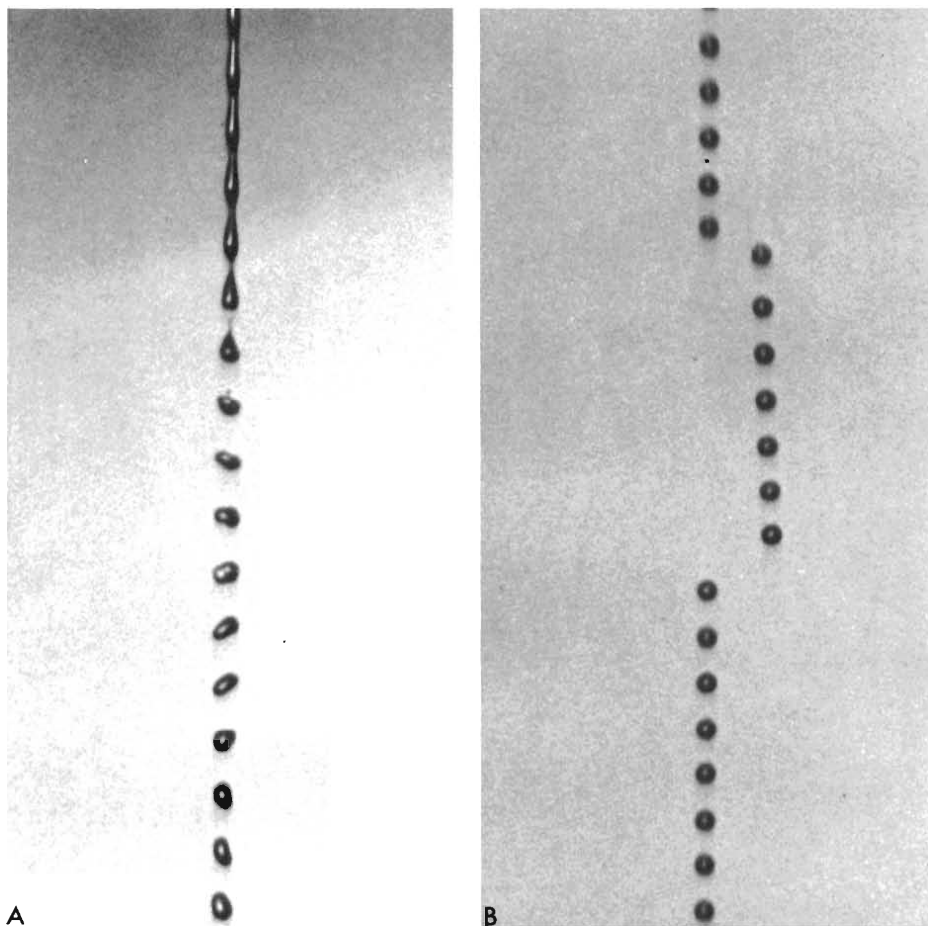
molecules within the cell.

Just as molecular biologists must be able to isolate and purify different biochemicals from the complex mixtures collected from disrupted cells, cell biologists must be able to obtain pure populations of cells from a heterogenous tissue or organ. Current techniques for separating viable cells include electrophoresis, centrifugation, and flow sorting. The first two are bulk isolation techniques. The third, when coupled to flow cytometry, can sort individual cells based on the variables measured on a particular cell. Thus flow sorting (Fig. 2) is a more precise method of separating closely related but functionally distinct cell types than either electrophoresis or centrifugation.

Flow cytometers were developed in the 1960s at Los Alamos Scientific Laboratory (LASL) and independently by Gohde in Germany. Although the first instruments lacked resolution, scientists soon recognized their potential for monitoring the growth pattern of cells, the transformation of cells from normal to malignant, and the function of the immune system. Several groups pioneered the early development of flow cytometry and its application to major problems in biomedicine. This extraordinary technique has been applied to problems in cancer diagnosis and treatment and to studies of basic cellular processes in normal and abnormal cells. Among early expectations was the possibility that this technique could be used for automated cancer detection and thus perhaps for mass cancer screening. This possibility still exists, but we must find new measurement variables that more clearly differentiate normal from malignant cells before it can be realized.

#### Early Staining and Measurement Techniques

The ability to stain DNA and other specific biochemical constituents of cells, the cornerstone of flow cytometry, dates



**Fig. 2.** A stroboscopic light appears to stop the movement of droplets as they jet out of the flow cytometer chamber. (A) Experimenters can program the flow cytometer electronics to separate a particular biological cell from the rest of the sample by charging a group of droplets as they break off from the solid stream of electrically conducting fluid. An electric field separates the charged droplets containing the cell of interest. (B) A group of 7 droplets is separated from the main stream. Since the average concentration of cells is about 1 cell per 50 drops, only one cell will normally be in the 7 droplets.

back more than 50 years to the work of Feulgen and Rossenbeck, who developed chemical procedures that allowed stoichiometric staining of DNA, the genetic material in cells. For the first time, the presence of DNA and its localization in the cell nucleus could be seen through a microscope. The procedure, called the Feulgen reaction, has been used widely to locate DNA in its various configurations including its condensation into chromosomes. A modification of the Feulgen reaction with fluorescent staining, developed in the 1960s, still serves as a standard for other staining procedures.

The first attempt to determine the quantity of DNA in a cell nucleus by optical means was made by Caspersen in 1936. He developed a microscope-

photometer to measure the amount of light absorbed by DNA. More absorption corresponded to more DNA. The correspondence was not exact, but it was useful nevertheless.

In the 1950s, Barder, Atkins, Mellors, Tolles, and others, using the early microscopic techniques, observed that elevated DNA levels are characteristic of cells derived from a large number and variety of human tumors. Thus the detection of malignant cells and, hence, clinical diagnosis might be based on the recognition of populations of cells with abnormal DNA distributions. However, microscope techniques are very slow and painstaking. Atkins, a prodigious worker in this field, spent years gathering data that can be acquired in minutes with flow cytometry.

## The Beginning of Flow Cytometry

When biophysicists at LASL developed the first flow cytometer, they were studying the effects of radiation on cells. The biophysics group had been concerned with monitoring the effects of radiation on whole organisms. How large a dose, they were asking, is required to affect life span measurably or to change tumor incidence? In 1965, the Atomic Energy Commission changed the direction of the biology program to the cellular level. At that time, the primary analytical tool available at LASL for monitoring cells was the Coulter counter, an electronic device that counts cells by measuring changes in electrical resistance. To use the Coulter counter, cells, immersed in a conducting medium, are passed through an insulating orifice. Because biological cells are quite good insulators, they decrease the conductivity across the orifice as they pass through it. The Coulter counter converts the decrease in conductivity to a voltage pulse for counting the number of cells per unit volume.

At that time, physicists who had transferred from nuclear reactor work to biophysics were using the techniques of gamma-ray spectroscopy to detect the presence and character of radioactive materials in humans and animals by counting and measuring the number of gamma rays emitted from an organism. By adapting the techniques of pulse-height analysis to the analysis of voltage pulses from the Coulter counter, they converted the Coulter counter into a device that quantitated cell volume. Now, volume distributions of large populations of cells could be measured. The desire to examine the cells corresponding to a specific volume led to another important development, the design by Fulwyler of the automatic *cell sorter*. The group used this device to sort individual cells with a specific volume into a separate container.

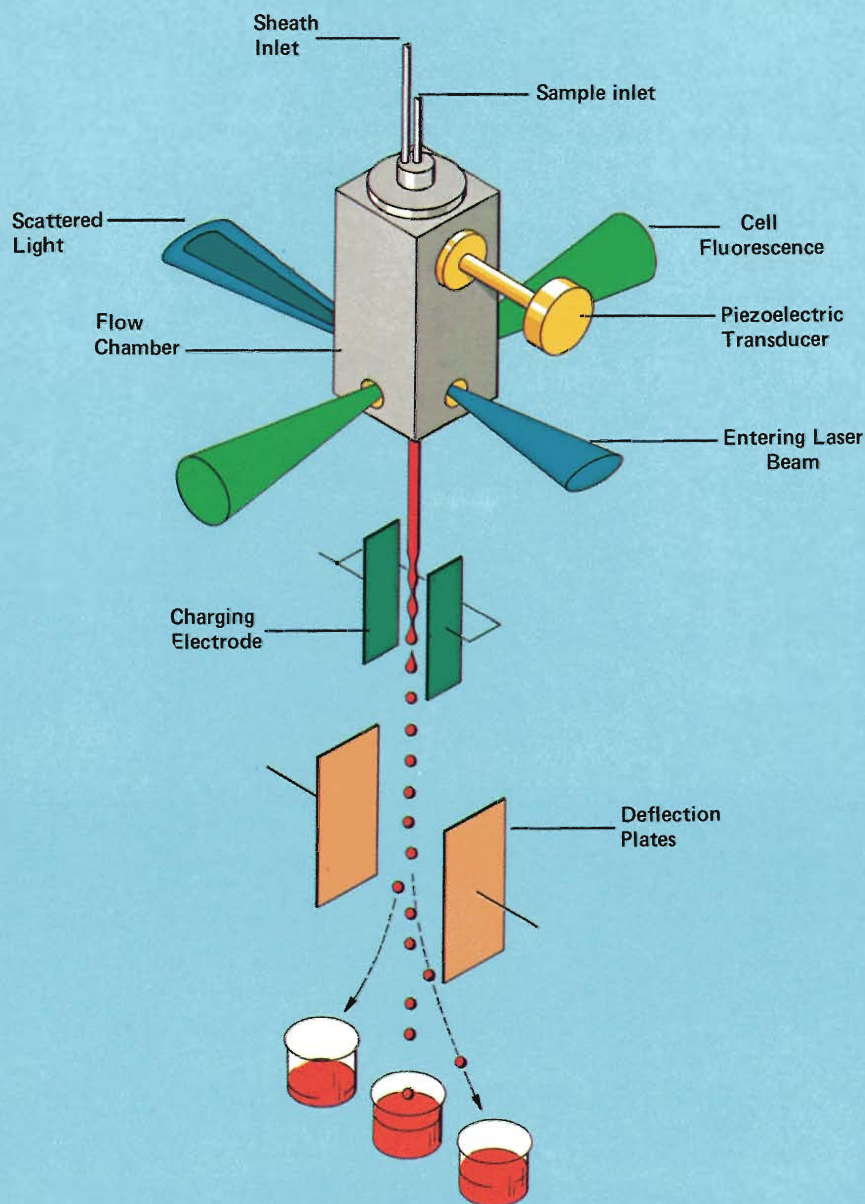
The volume-sorting instrument soon was applied to monitoring the life cycle of multiplying cells. LASL scientists considered cell volume a useful parameter to measure because a cell's DNA content doubles during the life cycle to insure proper transfer of genetic information when the cell divides, and an

increase in cell volume must accompany the increase in DNA content. Detailed studies led to the conclusion that cell volume is not a unique marker to differentiate cells at different stages of the life cycle. Fortunately, another parameter, DNA content, is unique. In 1966-67 Mullaney and Van Dilla constructed the progenitor of the LASL flow cytometers, an instrument that measures the fluorescence of a single cell as it passes through a laser beam. This device allows us to measure the DNA content of each cell in a population, if the cells have been stained with a dye chemically specific for DNA. The measurement allows us to follow the normal growth of cells or their abnormal growth caused by a perturbation of the cell's environment or as occurs in diseases.

Almost immediately, the biologists in the group were interested in using the flow cytometer to analyze the life cycle of exponentially growing cell populations by measuring the DNA distribution in cells exposed to various experimental conditions. The National Cancer Institute saw flow cytometry with its high accuracy and precision in measuring large populations of cells as a possible tool for early diagnosis of cancer, when the frequency of malignant cells is very low. Several groups including the LASL scientists, Wheelless and his group at the University of Rochester, Sweet and Bonner at Stanford University, and others improved the instrumentation and pioneered the application of flow cytometry to cancer diagnosis and to broader studies of cancer and the immune system.

## How the Instrument Works

All flow cytometers have three basic components: (1) a flow chamber in which cells are aligned for measurement; (2) a system for optical measurements consisting of a light source (usually a laser), beam-shaping and collection optics, and a light-detection device; and (3) electronics for signal acquisition, analysis, and display. The entire instrument is shown in Fig. 3 and details of the flow chamber, the optics, and the electronics are shown in Figs. 4, 5, and 6, respectively.



*Fig. 3. A schematic diagram illustrates how a flow cytometer works. The sample consisting of cells stained to identify a particular cellular property, such as the amount of DNA, enters at the top. The cells are dispersed into a single-cell suspension in the conducting medium (normally a saline solution). An electrically conducting sheath fluid is added at the top to ensure precise sample location in the flow cell (see Fig. 4). The laser beam enters the chamber from the right and is focused into a elliptically shaped slit of light to excite each fluorescently stained cell as it passes through the laser beam. The fluorescent light is analyzed with sophisticated electronics to quantitate the amount of fluorescent dye in each cell. As each cell passes through the laser beam, it also causes a scattering of the laser light that can provide additional information on cellular properties. The piezoelectric transducer is coupled mechanically to the flow chamber and tuned to about 40,000 hertz to vibrate the chamber and break the emerging stream into uniform droplets at a rate of about 40,000 droplets per second. The electrode can be charged rapidly to 75 volts so that droplets can be electrically charged as they break off from the main stream. The charged droplets are deflected by an electrical field supplied to the deflection plates. Thus, a group of droplets can be charged either positively or negatively and separated from the uncharged stream.*

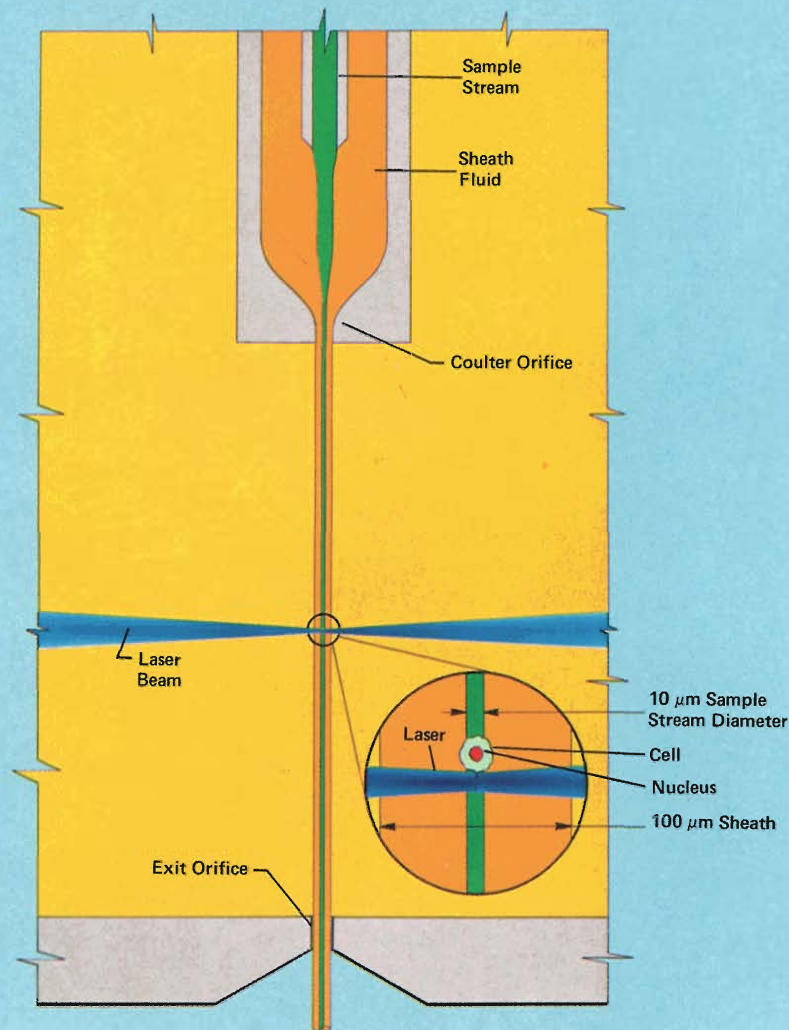
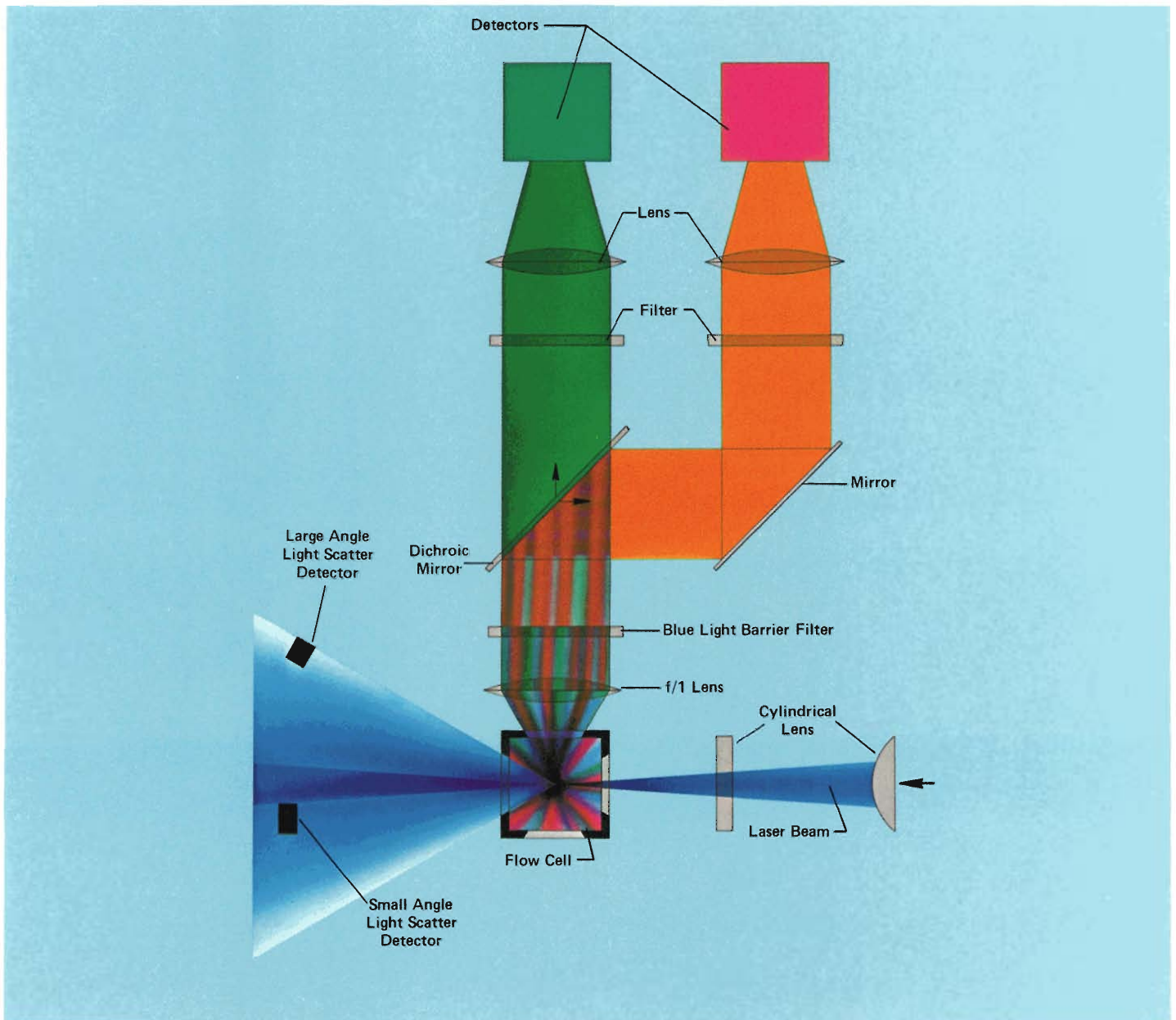


Fig. 4. The hydrodynamic design of the flow chamber is critical because cells must flow single file in a narrow path through the center of the Coulter orifice and then through the focused region of the laser beam. Most chambers currently in use are variations of a flow chamber originally designed by Crosslyn-Taylor to count particles in air. The three fluids are colored in this diagram to show their boundaries and to aid discussion. The sample inlet tube is located concentric with the sheath flow stream to ensure accurate cell stream positioning. The sheath container is shaped with a smooth transition region tapering from a few millimeters down to 100 microns, the size of the Coulter volume orifice. The shape causes a large increase in sample velocity to about 10 meters per second. The dc voltage across the Coulter orifice is applied to the sample inlet tube and to another electrode in the quiescent liquid. Because the sample stream is small, all cells pass through the same electrical field.

As the combined sample stream and sheath flow jet from the Coulter orifice across to the 100-micron-diameter exit orifice, the cells are illuminated by the highly focused laser beam (shown in the blowup of the inner section). With typical flow rates, the sheath volume is more than 100 times the sample stream flow rate, and the average diameter of the sample stream is only 10 microns. Thus, a 15-micron-diameter cell appears as a bulge in the sample stream. Cells are lined up much like beads on a string as they pass through the flow chamber. At a typical cell concentration of about 500,000 cells per milliliter, the average separation between cells is about 2 milliliters, so the likelihood of two cells passing through the laser beam simultaneously is small.

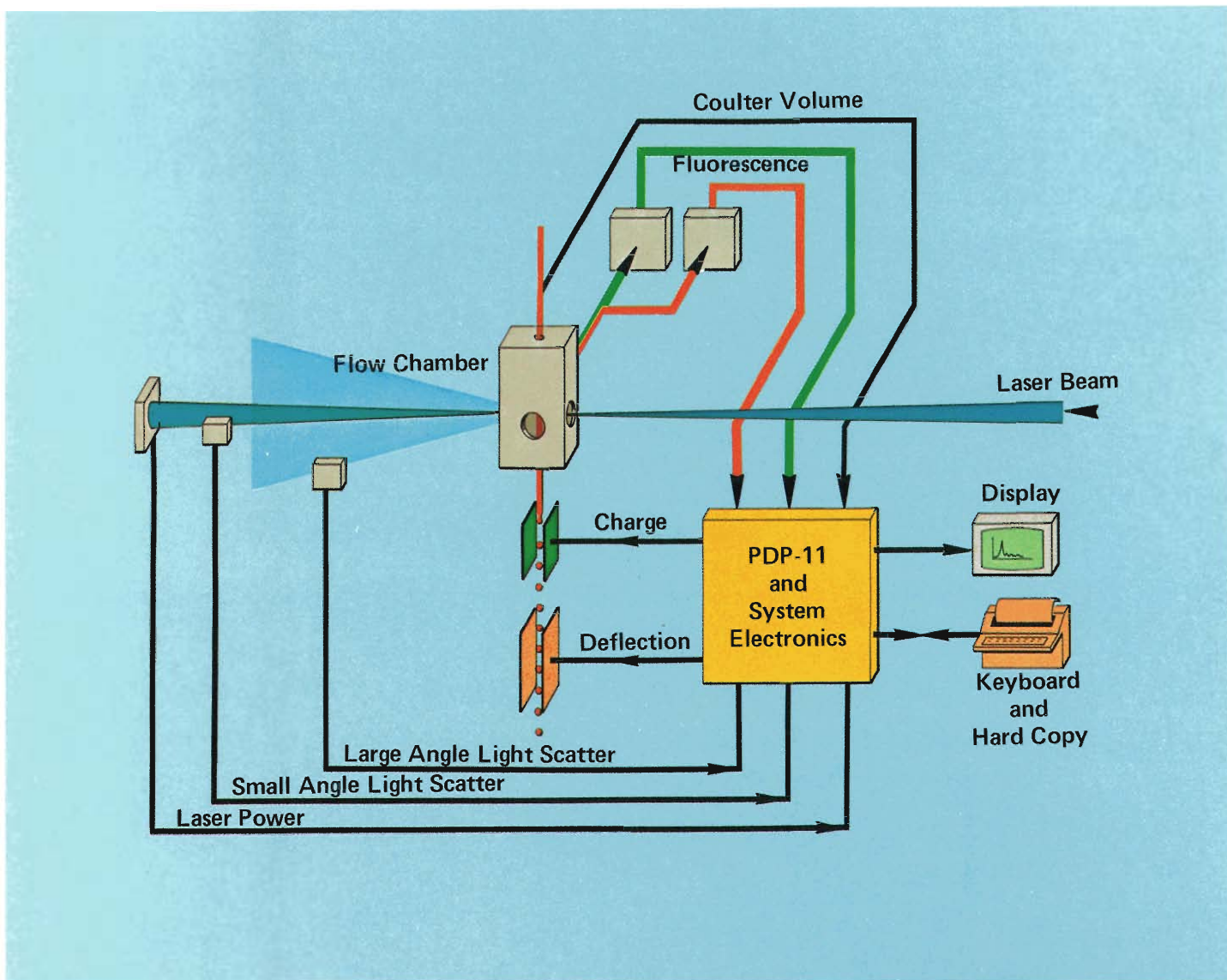


*Fig. 5. Flow cytometer optics. Cylindrical lenses of different focal lengths focus the laser beam into an ellipse at the center of the flow cell. As a stained cell passes through the laser beam, a pulse of fluorescent light is generated and a portion of it is collected by  $f/1$  light-collection optics. An interference "barrier filter" prevents scattered laser light from entering the highly efficient photomultiplier detectors. In the model shown, there are two photodetectors to measure green and red fluorescence originating from different parts of the cell. A dichroic mirror reflects the longer wavelength red fluorescence and transmits the shorter wavelength green fluorescence.*

*Concurrent with the emission of cell fluorescence, the blue laser light is scattered. The scattered light provides a means of obtaining information concerning the structure (morphology) of the scattering object. Since the size, shape, and mass (hence the refractive index) of a cell are the morphological features usually desired, an analysis of the light-scatter pattern produced by a cell may permit cell identification by providing a "signature" related to these physical properties. If light scatter signals from two detectors at different angles are compared, a signal can be generated that will discriminate fluorescent signals from debris and thereby enhance the purity of the signals from cells.*

*Two general types of laser illuminating light beams have been used. Earlier flow cytometers used spherical lenses to focus the laser beam into a spot of about 75-micron diameter at the cell stream intersection point to give light pulses about 40 microseconds long. The system shown here uses cylindrical lenses with different focal lengths to form an elliptical beam about 5-7 microns across its minor axis and about 100 microns across its major axis. Because the intensity distribution of the laser beam is Gaussian, these dimensions correspond to the points where the intensity of the laser beam is about 1/10th the intensity in the center of the laser beam. Cells traverse this laser beam in about 3-4 microseconds, so standard electronic circuitry for gamma ray spectroscopy can be used.*





*Fig. 6. This diagram represents the various signals that are handled by the electronics and the PDP-11 computer. The time between signals is calibrated and the signals are held so that all appropriate signals from the same cell can be compared and decisions can be made based on their values. This is all done before the cell arrives at the droplet break-off point. An induced charge can be placed on the few droplets most likely to contain the cell of interest, and these can be sorted. With this system purities exceeding 90% have been achieved. The operator interacts with the PDP-11 through a keyboard terminal and a cathode ray display.*

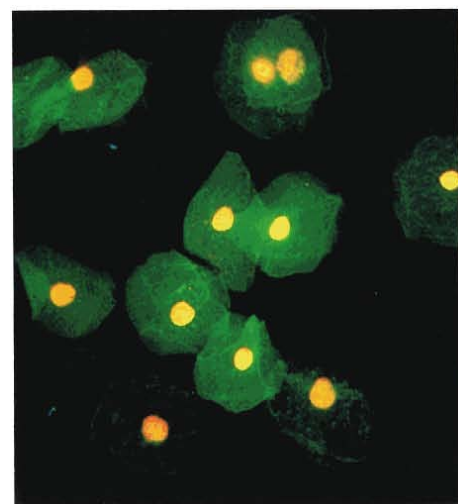
### Basic Fluorescence Measurements

Most applications of flow cytometry involve measurement of the fluorescence induced in a cell stained with a fluorescent dye, as the cell passes through a laser beam. Analysis of the light pulse determines the concentration and location of the stained biological molecules within the cell.

To illustrate, we consider a cell stained with two fluorescent dyes, a

yellow stain specific for the DNA in the cell nucleus and a green stain specific for the protein in the cytoplasm (Fig. 7). Exciting the cells with the 448-nm line of an argon laser generates two fluorescent signals, a red/yellow signal from the nucleus and a green signal from the cytoplasm.

The flow chamber ensures that each cell passes through the center of the laser beam and is illuminated uniformly by the laser light. In most LASL flow



*Fig. 7. Normal squamous cells stained with fluorescein isothiocyanate (green) and mithramycin (yellow). The cellular cytoplasm is green and the nucleus is yellow/red. The picture was taken through a fluorescent microscope with dark-field illumination.*

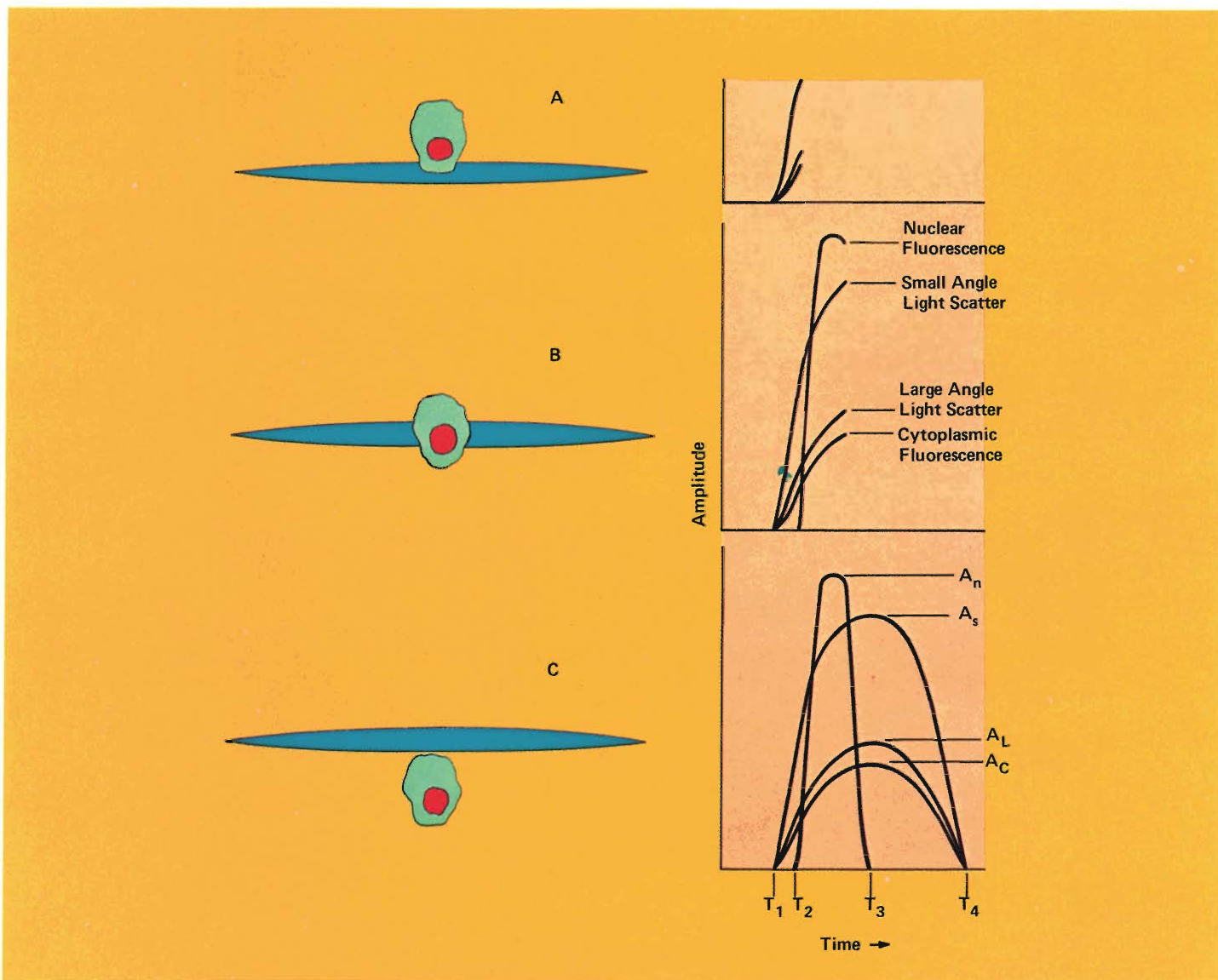


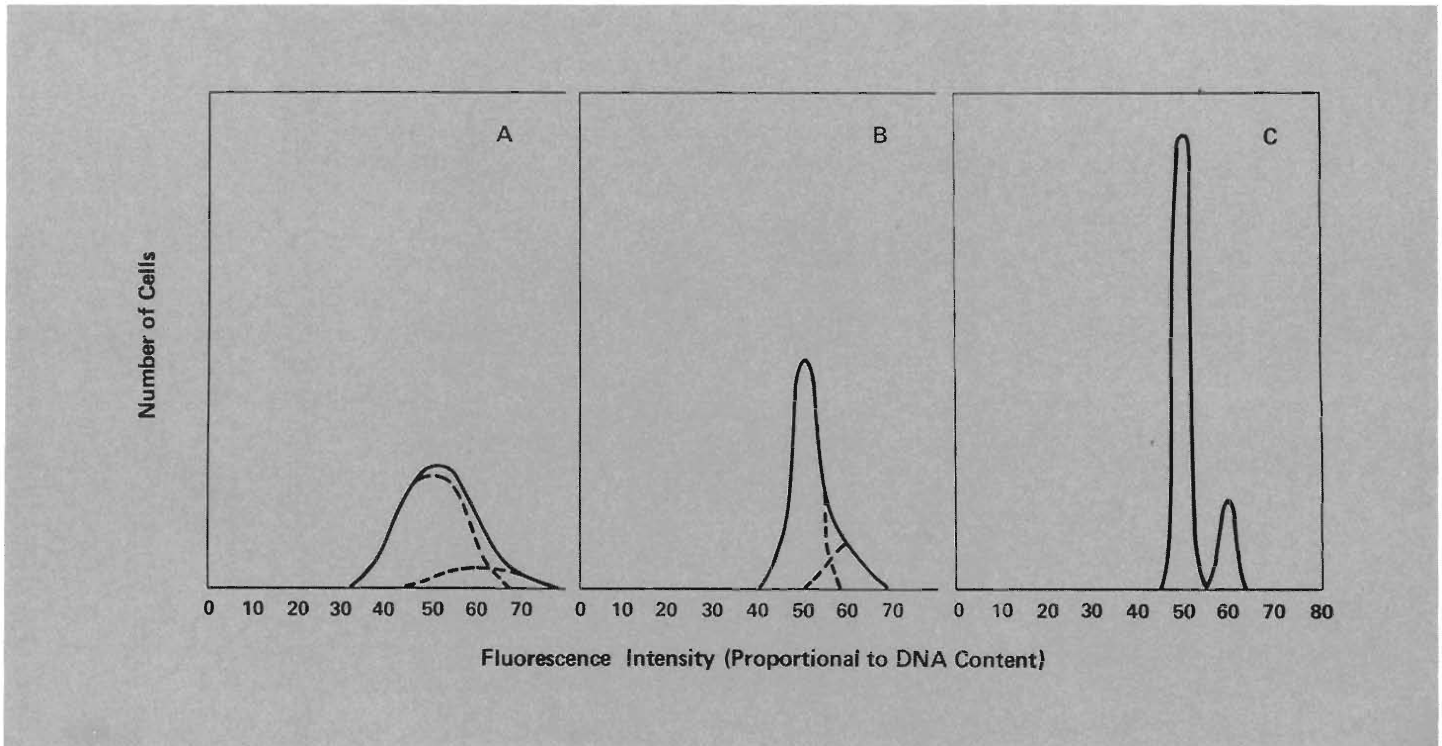
Fig. 8. The origin of electrical pulses that are generated as a two-color stained cell passes through the laser beam is illustrated here. In A, as the cell enters the laser beam, three signals are initiated: small- and large-angle light scatter and the excitation of green cytoplasm fluorescence. In B, the yellow nucleus fluorescence has been excited. In C, the complete pulse shapes from the four detectors can be seen. Several useful signals can be obtained from these voltages pulses. The amplitude of each signal ( $A_L$ ,  $A_S$ ,  $A_C$ ,  $A_N$ ) gives a measure of the density of the stain within the cell and the amount of light scattered. The length of the pulses above the threshold setting ( $T_4 - T_1$  and  $T_3 - T_2$ ) gives a measure of the cell and nuclear diameters, and the area under each curve gives a quantitative measure of the total amount of fluorochrome and the total amount of light scattered. The experimenter can decide which signals are best for analysis. In reality, the light-scatter signals are considerably larger than the fluorescence signals. Adjustment of amplifier gains allows one to display the signals overlaid on the same scale. Nuclear and cytoplasmic diameters can be obtained from the length of the pulses

cytometers the laser beam is shaped to an elliptical cross section at the intersection with the sample stream. The slit of laser excitation light provides a low-resolution scan of each cell as it passes through the laser beam. The fluorescent signals (Fig. 8) are measured at  $90^\circ$  from the laser beam optic axis to minimize the background light in the fluorescent light detectors (that is, to give dark-field illumination). The

fluorescence signal intensities are proportional to the amount of yellow and green fluorescent stain and thus to the amount of DNA and protein. The use of interference filters to separate the colors of light into the light detectors produces separate signals. The signals are analyzed individually by a two-dimensional pulse-height analyzer. The integrated intensity of a fluorescent pulse yields the DNA content (or protein con-

tent) of the cell, and the duration of the pulse yields the nuclear diameter (or cytoplasmic diameter).

A precise determination of DNA content was not possible with the first flow cytometers because of their poor resolution. For example, all normal, non-replicating cells should have the same DNA content. However, the DNA distribution for a population of such cells measured with early flow cytometers



*Fig. 9. Pulse-height distributions that might be obtained from two populations of cells stained for DNA content. The two populations are present in a ratio of 5:1 and their pulse-height distributions measured before mixing are shown by dashed and dotted curves. Cells in the less prevalent population have 20% more DNA. The solid lines in A, B, and C represent the pulse-height distribution of the populations mixed together. The distributions in A were measured by a low-resolution flow cytometer (14% cv) and in B by one having twice the resolution (7% cv). The resolution obtained with current instruments, shown in C, was four times the resolution in A. Here the two populations can be resolved completely and measured easily.*

had a coefficient of variation (cv) of 14%. A number of possible causes for the poor resolution were examined: non-uniform laser output, low laser beam intensity, low-efficiency photomultipliers, noisy electronics, poor staining procedures, and low-brightness dyes. We concluded that more powerful and stable lasers, brighter dyes, and better staining techniques were needed. With these improvements we routinely obtained a 3-4% cv, and by using two DNA fluorochromes of differing specificity we obtained 1.5-3% cv.

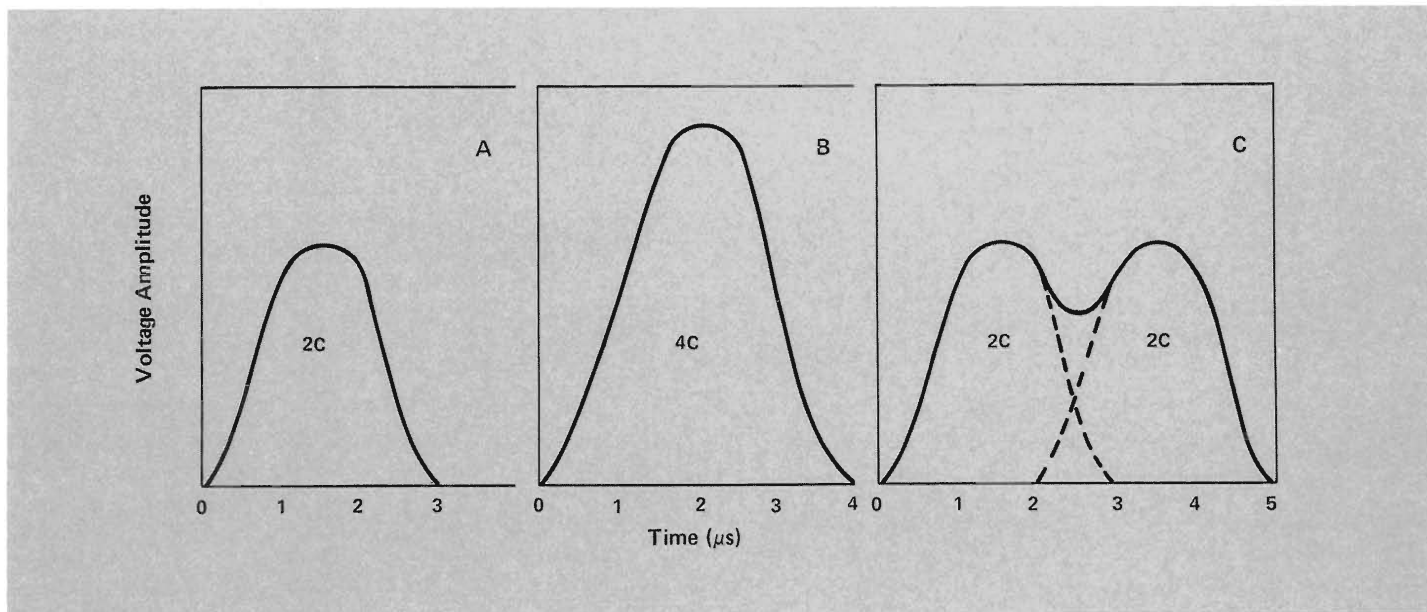
The curves in Fig. 9, representing histograms obtained on a pulse-height analyzer, illustrate how instrument

resolution influences data quality and the ability to see small differences in the DNA content of mixed cell populations. The histograms show the number of cells vs DNA content per cell.

The motivation for improving instrument resolution came from the desire to monitor the growth of replicating cells and from the need to produce safe human vaccines from cell lines maintained in the laboratory. At the time, it was proposed that only those mammalian cells that faithfully maintain their normal numbers of chromosomes should be used for vaccine production. By assumption, any increase in the number of chromosomes would be reflected by a

corresponding increase in DNA content. Consequently, if flow cytometers could measure DNA content more accurately, they could detect abnormal cells with increased DNA content and thus monitor the early stages of chromosome instability. We improved the resolution of our instruments for this purpose and found that a change in the number of chromosomes *does not necessarily* correspond to a change in DNA content. This highly significant result indicates that in some instances a cell can conserve its DNA while repackaging it into a different number of chromosomes.

The need to distinguish cell doublets (two cells stuck together) from a single



*Fig. 10. Pulse shape profiles. A. Pulse from a nucleus of a single cell ( $2c$  DNA content) as it passed through a laser beam that is small (7 microns) compared to the diameter of the nucleus. By definition,  $1c$  is the DNA content of a germ cell. B. Pulse from the nucleus of a single cell having twice the amount of DNA ( $4c$ ). C. Pulse from two cells (each having  $2c$  DNA content) that are stuck together. Since the electronics normally integrates the area of the pulses to get the total amount of fluorescent light, both the one cell in B and the two cells in C would be recorded as a single cell with twice the normal amount of DNA. In reality, the profile in C arises from two cells. The narrow width of the laser beam causes the pulses from the two cells to overlap (dashed part of curve C) to produce a saddle in a single pulse. Electronic circuitry has been designed to detect the saddle between the two peaks so that pulses of this shape can be discriminated from pulses of equal area such as shown in B. Thus, a single cell having twice the amount of DNA can be distinguished from a doublet.*

cell that has doubled its DNA content before cell division or contains double the normal amount of DNA led to another important improvement in instrument resolution. The original spherical optics that produced an illumination area much larger than a single cell were changed to optics that shaped the laser beam into an elliptical slit (as described above) much smaller than the cell nucleus. Figure 10 illustrates the difference in pulse shape obtained using a 7-micron laser beam for a single cell with double the DNA content of a nonproliferating cell and two nonproliferating cells stuck together. Electronics were incorporated to discriminate between the two pulse shapes.

The ability of flow cytometers to measure DNA content with high

statistical accuracy has been crucial for many biological studies, in particular, those related to the nature of cancer and its diagnosis and treatment.

### Nature of Cancer

Cancer is a general term used to describe what is probably a number of diseases. The common manifestation is growth of the cancer cell beyond the condition of simple repair and replacement to a condition that leads to interference with normal biological processes and eventual death. It appears certain that several steps are required for a cell to become malignant. Some of the steps involve growth characteristics, and

others involve immunological alterations. Although most normal cells in a mature organism are not growing (that is, dividing to produce new cells) there is some continuing amount of cell proliferation. Therefore, finding a very slight increase in cell proliferation in inappropriate cells (for early cancer detection) is a very difficult problem. However, prevalent opinion holds that there must be at least one, and perhaps more, unique and measureable properties of cells that are specific in cancer. A great deal of effort has been made to find such properties.

Because tumors can arise from a single cancer cell, all of the information that defines a particular cancer probably is contained within a single cell. This is one rationale for performing cellular

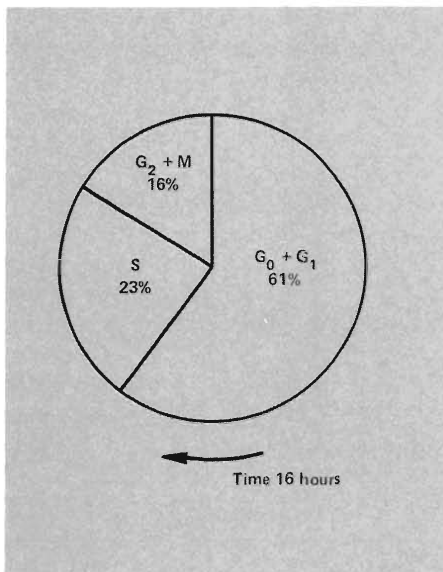


Fig. 11. A conceptual diagram of how Chinese hamster ovary cells (CHO), in exponential growth, are distributed throughout their life cycle.

research to find an unequivocal marker for a cancer cell. A suitable marker could be any of the following: the presence of specific antibodies or enzymes, an unknown chemical substance, a change in electrical properties, a change in the DNA, a loss of chromosome stability, or some presently unimagined feature. Experiments with DNA have high priority because the central cancer problem is thought to be one of altered gene expressions.

### The Cell Cycle

To define an abnormal or cancerous condition, we must understand the range of normal conditions. Therefore it is not truly meaningful to separate cancer research from basic research directed to the understanding and quantification of normal biological processes. Undoubtedly a detailed understanding of cellular growth throughout the life cycle of a cell holds the key to an understanding of where the process goes astray and produces uncontrolled tumor growth. One gross change that can be monitored during the life cycle is DNA content. Each phase of a cell's life cycle (Fig. 11) is characterized by the amount of DNA present in the cell nucleus. DNA is measured in units of  $c$ , the number of picograms of DNA in the sperm or egg cell of a particular species. All other cells in that species contain double this

amount, or  $2c$ . In Figs. 11 and 12,  $G_0$  represents the stage of a mature cell that is not multiplying (for example, a lymphocyte). If the cell receives a signal to replicate (the presence of a foreign material might initiate such a signal), it goes into stage  $G_1$ , in which several biochemical activities including the production of protein in the cytoplasm take place. Upon completion of these activities, the cell goes into stage S, in which the DNA replicates itself and the amount of DNA increases from  $2c$  to  $4c$ . Stage S is followed by  $G_2$ , a resting stage, and then by M, the mitosis stage, in which the DNA condenses into chromosomes and the cell divides.

Figure 12 shows the distribution of DNA content for a population of normal multiplying cells measured with a flow cytometer. The large peak, containing

most of the cells at  $2c$  DNA content, corresponds to stage  $G_1$ . The smaller peak around  $4c$  corresponds to cells in stages  $G_2$  and M. The region between the peaks contains the cells in stage S, undergoing DNA synthesis. In the normal life cycle, cells remain between  $2c$  and  $4c$ . However, when things go astray as in malignant growth, abnormal amounts of DNA are sometimes observed. As an organism ages or is subjected to certain chemicals or ultraviolet light and other environmental stresses, both the amount of DNA and the way it is packaged into individual chromosomes may change.

The measurement of DNA distributions therefore is extremely useful for tracking cells through the life cycle and for assaying the effects on the cell cycle of radiation, environmental conditions, and chemical drugs. The technique also

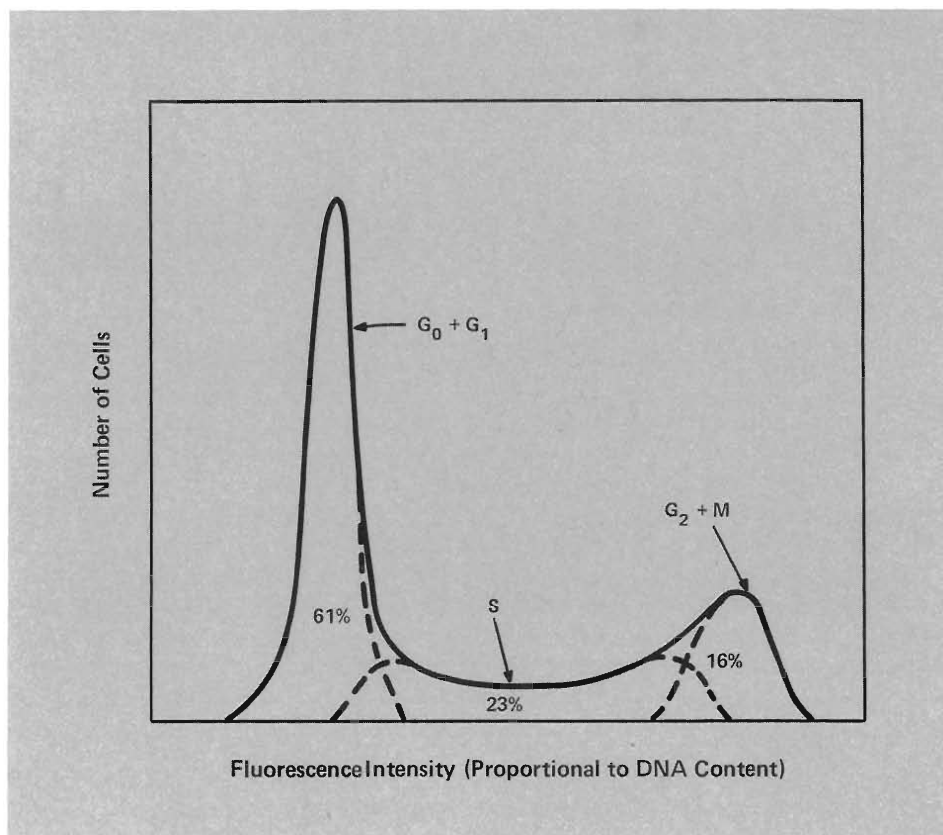
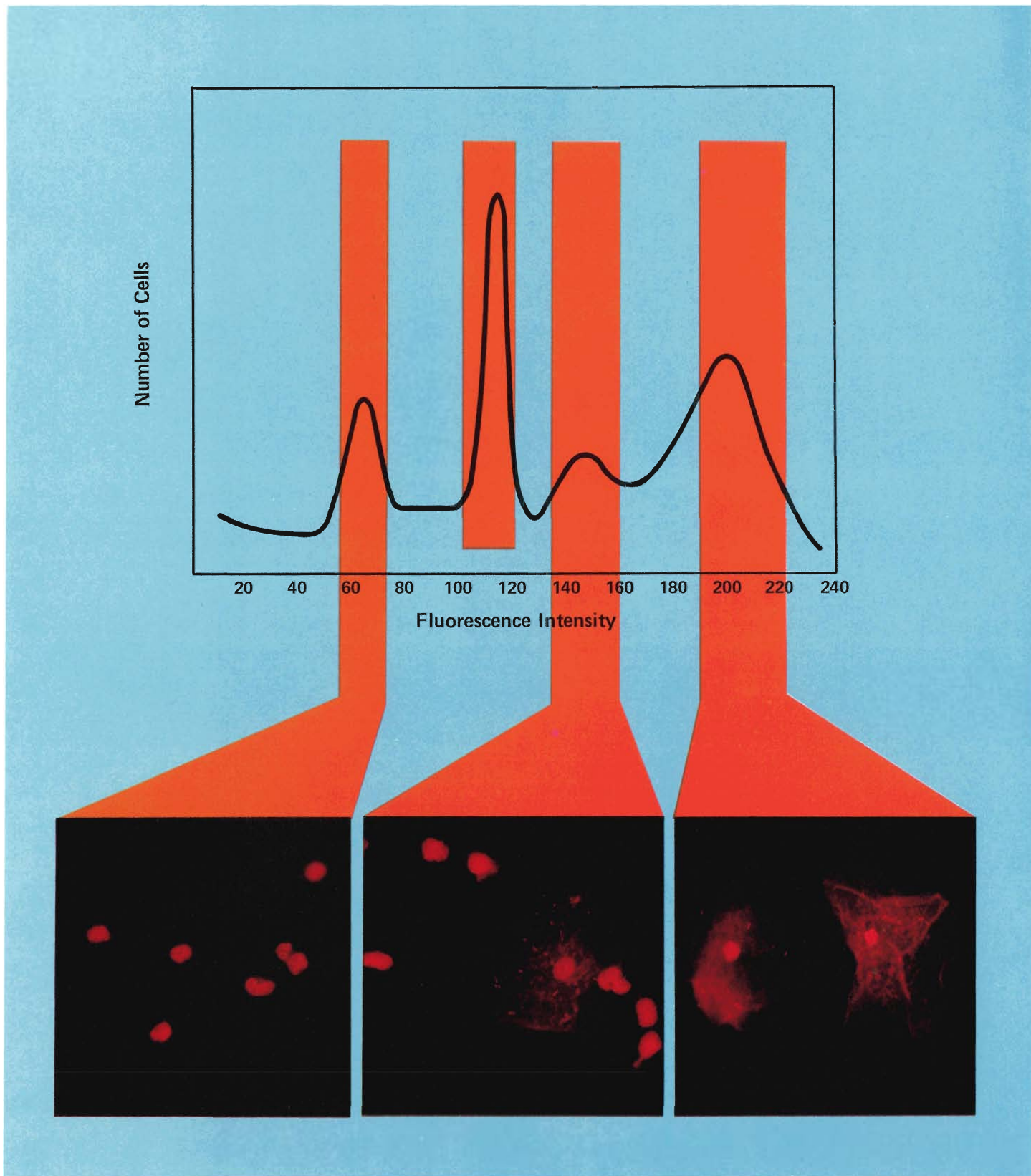


Fig. 12. The pulse-height distribution from exponential growing cells showing how the integral curve can be unfolded into its component parts to determine the percentage of cells in  $G_0 + G_1$ , S, and  $G_2 + M$ .



*Fig. 13. Intensity distribution of three cell types artificially mixed before fixation in 70% ethanol and stained using a nonspecific propidium iodide procedure. Fluorescence analysis was made using 488-nm laser illumination. Photomicrographs of cells sorted from each of the peaks are shown in the lower part of the figure. Based on cell morphology the cells from the indicated sorting windows were identified as: sort region 1, human white blood cells, magnification of 438 X; sort region 2, cells derived from a methylcolanthrene-induced skin tumor of mice (not shown); sort region 3, methylcolanthrene tumor cells with a few contaminating squamous cells from region 4, magnification of 276 X; region 4, normal human buccal cells (squamous cells) obtained from the lining of the oral cavity, magnification of 276 X. (Figure courtesy of Z. Svitra and John Steinkamp, LASL.)*

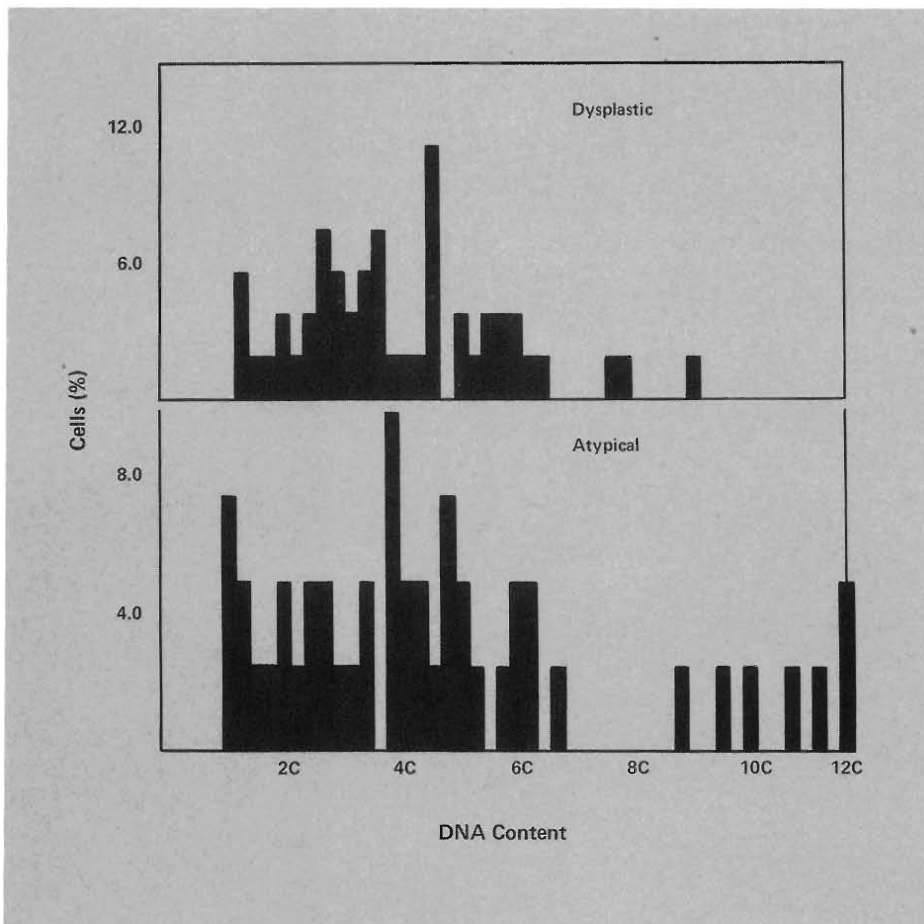


Fig. 14. DNA distributions obtained from abnormal clinical cervical samples —(A) Dysplasia, a recognized precancerous, but temporarily benign condition, and (B) frank cancer (carcinoma in situ). Both conditions have a similar characteristic, that is, cell proliferation as indicated by DNA densities at 4c, 6c, 8c, etc. Resting or mature cells display a DNA value of 2c.

has been used in attempts at cancer diagnosis and in monitoring the effects of chemotherapy on malignant cells. In one application, the cell sorter unscrambled a complex mixture of malignant cells and two types of normal cells as illustrated in Fig. 13. In this instance, the fluorescent staining procedure was only partially specific for the DNA of each cell type. Although the procedure is empirical, it produced clear resolution of the three cell types in the intensity distribution. If a stoichiometric staining procedure had been used, this resolution would not have been achieved. To verify which cells constitute each peak, several thousand cells were sorted from each peak; the results are illustrated in Fig. 13.

### Attempts to Automate Cancer Diagnosis

To the cytopathologist, malignant cells have certain distinguishing features that can be observed through the microscope. Because some of these features can be measured with a flow cytometer, we can hope that this instrument may one day be used as an inex-

pensive automated method for mass screening for cancer. As an example, consider the case of cervical cancer. This disease represents a major worldwide public health concern. In the United States alone, there are approximately 90 million women in the "at risk" age group. Moreover, clear evidence based on case histories tells us that early and accurate detection is the key to a good prognosis. For women who are screened regularly with the "Pap test," survival from this disease is greater than 95% if the disease is detected. However in the United States, mainly for socioeconomic reasons, the majority of the female population never undergoes this screening.

Thus for the general, infrequently screened population, survival is more like 50% if the disease is detected. For these reasons, cervical cancer detection has long been a prime candidate for automation.

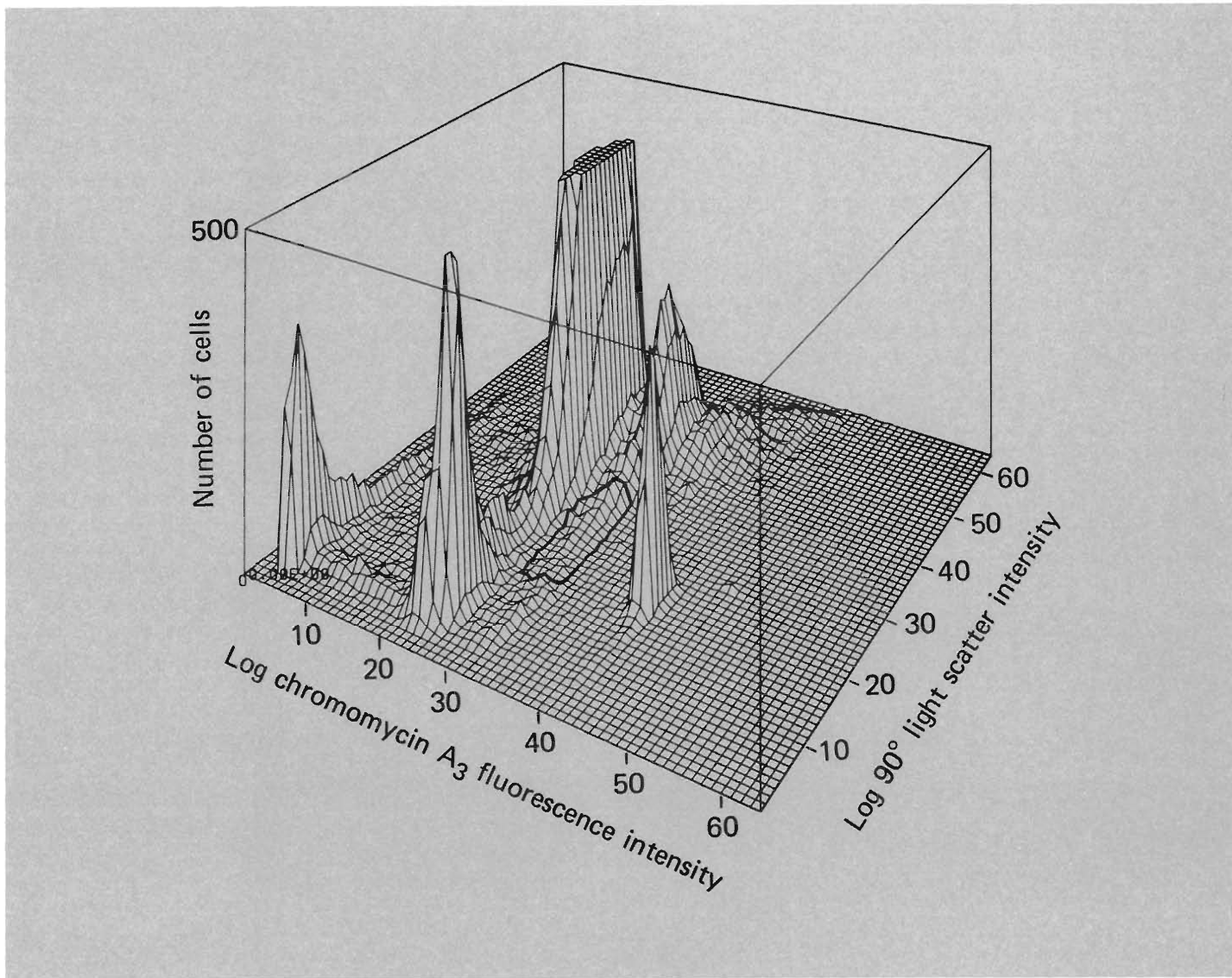
Under a microscope, the stained nucleus of a malignant cell looks irregular compared to the nucleus of a normal cell. Frequently the nucleus is larger than normal, and it may be less symmetrical. Sometimes the cell too is larger and more irregular and may have a granular appearance. These qualitative differences between normal and malignant cells are fairly easily detected by a trained eye.

The question is, can an instrument do as well? Detailed pattern recognition is slow and expensive, but a flow cytometer can measure several relevant parameters with good accuracy. These are DNA content, nuclear diameter, protein content, and cytoplasmic diameter. On the average these parameters change from normal to malignant cell populations as shown in Table I. The size features (nuclear and cytoplasmic diameters) overlap considerably between normal and malignant cell types. DNA content may increase significantly in malignant cells and the ratio of nuclear diameter to cytoplasmic diameter also tends to increase. The DNA distribution for cells from a frank cervical tumor is shown in Fig. 14. These cells can have DNA

TABLE I

#### Morphological Characteristics of Normal and Malignant Cells

Feature	Noncycling Cells	Malignant Cells
DNA ( $G_1$ )	2c	>2c
Nuclear diameter	<12 microns	>12 microns
Nuc/Cyto diameter	<0.5	>0.5
Cytoplasmic diameter	10-80 microns	20-40 microns



*Fig. 15. A two-parameter histogram of cervical and vaginal cells measured by flow cytometry. The histogram is generated by determining chromomycin A<sub>3</sub> fluorescence intensity as a measure of cellular DNA content (abscissa in the plane) and 90-degree light-scatter intensity as a measure of cell size (ordinate in the plane) simultaneously on each cell. Thus the location on the plane shows DNA content and cell size, while the height above the plane shows the relative frequency of occurrence. The intensities of the fluorescence and scatter signals are measured and presented logarithmically. An increase in seven channels in either direction corresponds to a doubling of intensity. With the cell sorter, it is possible to identify the cell types that give rise to histogram features. The peak and shoulder at the far left corner of the histogram (the lowest values of fluorescence and light-scatter) are due to cell debris and bacteria. The sharp peak in the foreground with a higher fluorescence intensity and low light-scatter is due to white blood cells; the large peak at equal fluorescence intensity but higher scatter values is due to intermediate and superficial epithelial cells. The small shoulders to the right of the white cell and epithelial cell peaks are due to cell aggregates. The sharp peak at even higher fluorescence is due to fluorescent microspheres that were added for machine calibration. The identities of the cells responsible for each feature have been confirmed by cell sorting and morphological analysis. (Figure courtesy of R. Jensen, Lawrence Livermore Laboratory)*

values as high as 12c, or 6 times the normal noncycling value. Since the geometric features have been shown clinically to have diagnostic value, sets of parameters such as DNA and size have been investigated for use in automated diagnosis of cervical cancer by flow cytometry.

At the Lawrence Livermore

Laboratory, clinical material has been examined for its DNA content and cell size. The DNA was measured by fluorescence, and the cell size was determined from light-scattering measurements. The data result in three-dimensional histograms of the two variables, as shown in Fig. 15 from an abnormal specimen.

Malignant and premalignant cells contain elevated amounts of DNA and tend to be intermediate in size. A large fraction of the abnormal cells in Fig. 15 show fluorescence and light-scatter signals that are localized to the right of the main peaks (in the heavily outlined area). A 20- to 30-fold enrichment of abnormal cells can be obtained for



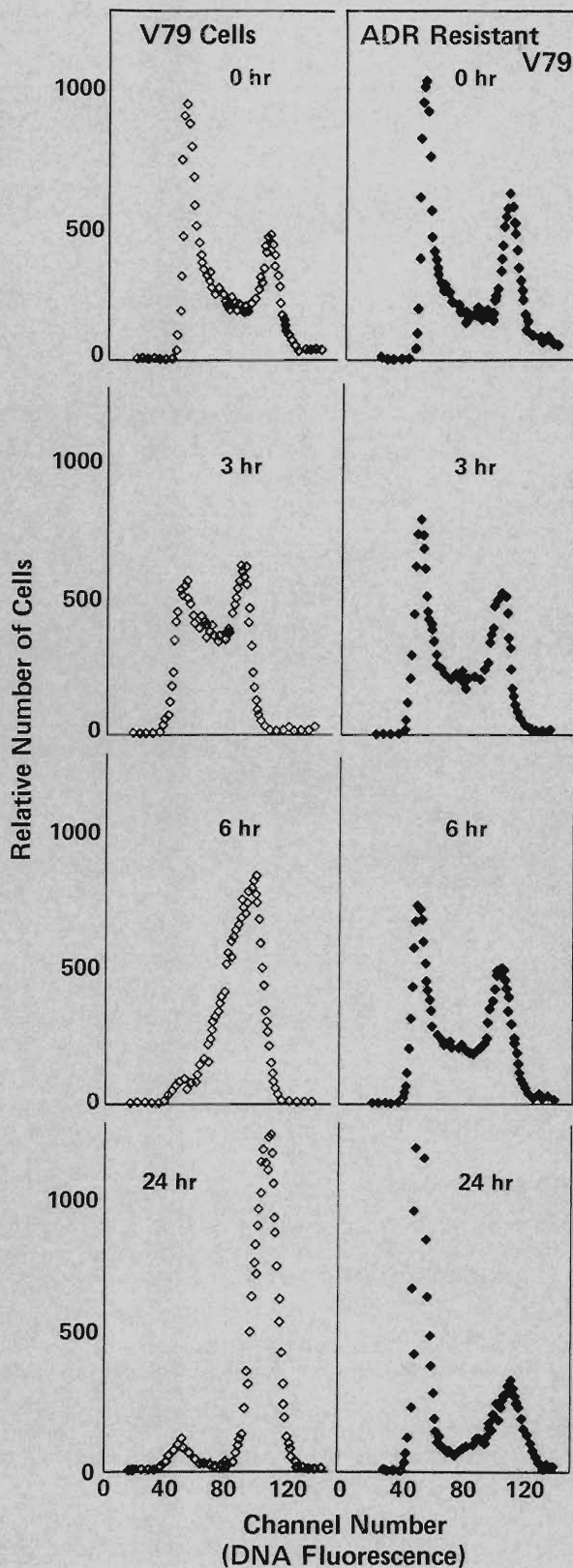


Fig. 16. Comparison of DNA distributions of normal cell line V79 and Adriamycin-resistant V79b cell line following exposures to Adriamycin (0.5 mg/ml, 1 h). (Courtesy of N. Tokita, LASL)

microscopic analysis by sorting cells from this area. Because some signals from normal cells and aggregates also appear in the area, we cannot use histogram analysis directly for automated diagnosis of cervical cancer. To increase the potential accuracy of automated diagnosis, a new independent marker correlated with cellular abnormality must be included as a third parameter.

Recently a group at the National Cancer Institute Division of Cancer Biology and Diagnosis used the LASL multi-variable cell sorter in combination with the two-color DNA/protein staining protocol. A double-blind clinical trial was conducted so that experimenters were not informed of the clinical diagnosis of the samples. That is, each sample was divided into two fractions, one screened by cytopathology and the other screened by the machine. Upon completion of the trials, the instrument results were compared with the cytopathology results, which were assumed to be correct. With this parameter set the accuracy was 80%. Accuracy depends on normal variation and on how one sets the criteria for defining an abnormal cell. Because of variations between normal cells, it may be impossible to improve the accuracy by measuring only DNA and protein. It appears that additional variables must be measured to improve definition of a cancerous cell.

Errors are expected in any diagnosis. The two important errors are the false negative rate (rate of abnormal cases misclassified as normal) and the false positive rate (rate of normal cases classified as abnormal). In the trial just described, the false negative rate was 10%, a serious error because it means that disease is missed in 10% of the incidents. The false positive rate, 27% in the trial, is less serious. In any mass screening program, the false positive error increases both work and costs because sampling must be repeated. However, a 27% false positive rate would be acceptable, if the false negative rate could be brought down to about 5%. Thus, although these results do not justify a clinical technique at present, they are encouraging.

Perhaps as we learn more about the nature of cancer, we can identify more definitive parameters that will make

automated screening more accurate than the cytopathologist. Work continues with this goal in mind.

### Monitoring Radiotherapy and Chemotherapy Effects

Flow cytometry is used extensively to study the effects of radiation and chemotherapeutic drugs on cultured cells, rodent tumors, and human tumors. Goals include an understanding of the damage in terms of cell-killing and the application of this information to radiotherapy and chemotherapy problems.

After cells are exposed to radiation or to chemotherapeutic drugs, the cell's DNA distributions and other cytometric parameters are measured at different times after exposure to complement the data obtained by other methods. For example, two cell lines of Chinese hamster origin differ considerably in their sensitivity to Adriamycin, a drug commonly used for the treatment of human cancer. When M. R. Raju (LASL) studied the DNA distributions for the two cell lines after treatment with 0.5 milligram per milliliter Adriamycin for 1 hour, he found that one cell line was more sensitive than the other (Fig. 16). Changes in DNA distributions in the normal (V79) cell line were dramatic, but the changes in the Adriamycin-resistant cell line (V79b) were small. Cell survival data obtained by colony formation indicated that cell survival of the normal (V79) cell line was 3%, but cell survival of the Adriamycin-resistant (V79b) cell line was 80%. Therefore, the magnitude of the cell cycle DNA distribution perturbations measured by flow cytometry was related to the cell killing. Since Adriamycin resistance is due to a decreased drug uptake, he measured the amount of Adriamycin per cell using flow cytometry. The measurements confirmed that the drug resistance of the V79b cell line was due to lower drug uptake. This study, together with others, in-

dicates that DNA distributions can be used to measure and predict the drug sensitivity of tumor cells rapidly. DNA distributions also may be useful for studying drug-induced perturbations of cells from a patient biopsy as a prognostic test for cancer patients.

### Search for Key Malignancy Parameters

Because it is likely that a single cell (a stem cell) is the origin of every cancer, the isolation and characterization of such a cell remains an important problem. A stem cell should be inherently different from other cells of the population in one or more characteristics. The differences probably originate in an altered expression of the cell's DNA. It is unrealistic to expect that initial changes in a cell's gene expression will be manifested in DNA changes large enough to be detected in the presence of the cell's total DNA content. In humans, the packaging of cellular DNA into chromosomes during a portion of the cell cycle (mitosis) provides a natural subdivision of cellular DNA into 46 chromosomes, which can be isolated and studied individually with flow cytometers and sorters. That chromosome changes are important in the etiology of tumors has long been recognized, but the number of chromosomes we can study using a microscope is too small for us to detect the rare chromosome changes that occur when a tumorigenic stem cell first begins to propagate.

Lawrence Livermore Laboratory and Los Alamos Scientific Laboratory in cooperation with the Max-Planck-Institut in Goettingen, West Germany, have demonstrated the capability to analyze isolated individual chromosomes by flow cytometry. Figure 17 is an example of such a high-resolution measurement in which all but one of the expected chromosome types isolated from Chinese hamster M3-1 cells were resolved. In addition, three "homologue"

pairs (1, 1; 7, 7; and 9,9) had small differences in DNA content. Chromosome changes reflecting the appearance of a new stem cell in the population might be manifested by changes in the position or area of one of the peaks.

More subtle changes in chromosome morphology can be quantitated by using a unique chromosome-imaging flow sorter. In contrast to conventional flow cytometers that measure total fluorescence without regard to its spatial distribution, our chromosome-imaging sorter maintains the optical image as the object passes through the laser beam. The chromosome image is formed in front of a mechanical slit placed at the image plane; the slit scans the chromosome as it flows through the observation region. The intensity profile of a chromosome contains valuable additional information as illustrated in Fig. 18. A unique intensity profile is recorded for each chromosome at the rate of 1,000 per second. Now much more subtle features associated with chromosome aberrations can be sought, and when chromosomes with these features are found, they can be sorted for verification and further analysis. Chromosomes with aberrations such as breaks, some types of translocation of chromosome pieces, fragments, and dicentric (two chromosomes joined at their ends) should be detectable. The search for additional parameters that will define chromosome types uniquely is continuing.

The role of chromosome changes in cancer is not well defined. In certain types of cancer, such as chronic myelocytic leukemia, a specific and fairly consistent chromosome abnormality has been identified. In advanced cancers, one observes a large variety of chromosome aberrations but not much consistency. These observations are in stark contrast with the remarkable process of DNA replication and cellular mitosis that results in exact maintenance of the amount of DNA per cell and the stability of chromosome number and

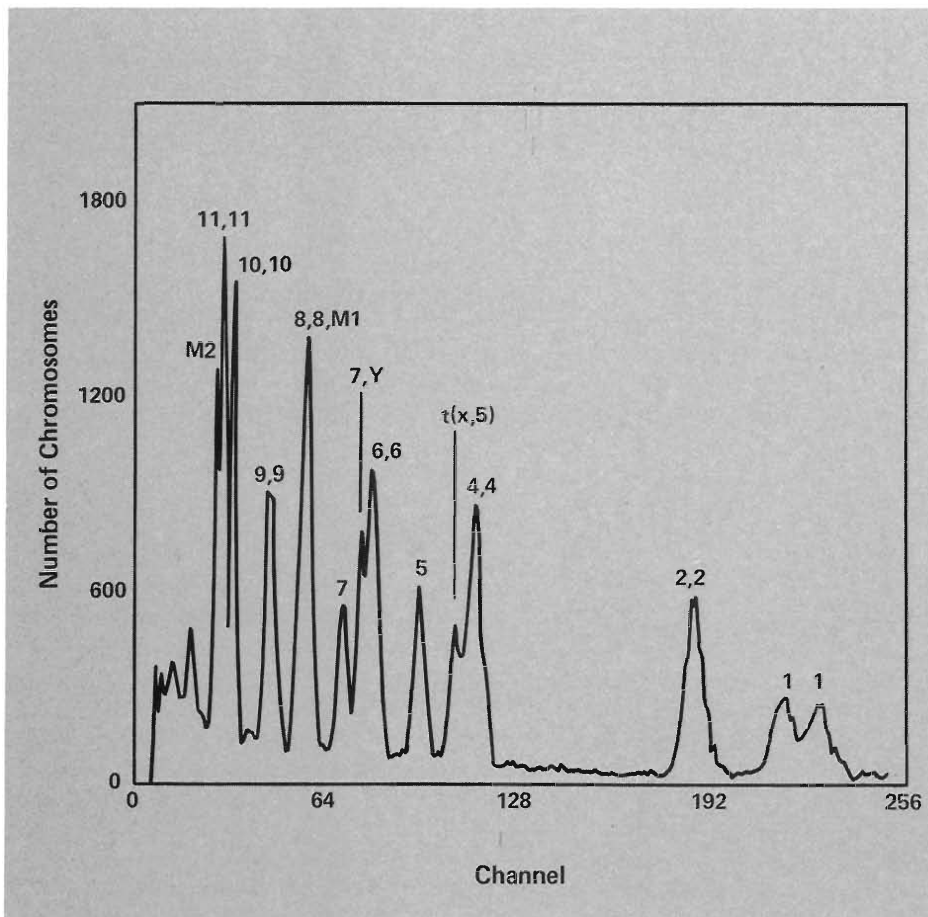


Fig. 17. The flow karyotype of chromosomes isolated from Chinese hamster M3-1 cells and stained with the bis-benzimidazole dye Hoechst 33342 (10 micromolar). The mean fluorescence intensity of a peak reflects stain content and hence DNA content for chromosomes of that type. The area of each peak is proportional to the number of chromosomes of that type. Approximately 30,000 chromosomes were analyzed. The type of chromosomes constituting each peak was verified by sorting and visual identification. The longest chromosomes (No. 1) have the most DNA, and the smallest chromosomes (M2 or markers) have the least. Each chromosome type and its corresponding peak are identified.

morphology in all normal cells. It is this very contrast that makes chromosome stability appear to be of extraordinary importance. With full implementation of the imaging flow cytometer/sorter, we will be able to characterize 100,000 or more chromosomes—a truly remarkable accomplishment that has been equated to taking that first look at the back side of the moon.

The development of flow cytometry is due to the combined efforts of many people throughout the world. Many members of the LASL Life Sciences Division and groups at Stanford University, Lawrence Livermore Laboratory, Memorial Sloan-Kettering Cancer Center, and the University of Rochester have contributed to this work. The manufacture of commercial equipment is a strong

indication that the technology has reached a fairly mature state. The instruments now can be used by a large variety of biologists throughout the world and the potential biological applications appear to be very great.

This work was supported by the Assistant Secretary for Environment of the Department of Energy and the National Cancer Institute.

### Bibliography

1. *Flow Cytometry and Cell Sorting*, M. R. Melamed, P. F. Mullaney, and M. L. Mendelsohn, Eds. (John Wiley and Sons, New York, 1979).
2. *J. Histochemistry and Cytochemistry*, 24 (1) 1976, 25 (7) 1977, 27 (1) 1979. These issues contain many articles on flow cytometry instrumentation and related biological applications.
3. *Cytometry*, a new journal (1980) published by Williams and Wilkins, specializing in analytical cell analysis.

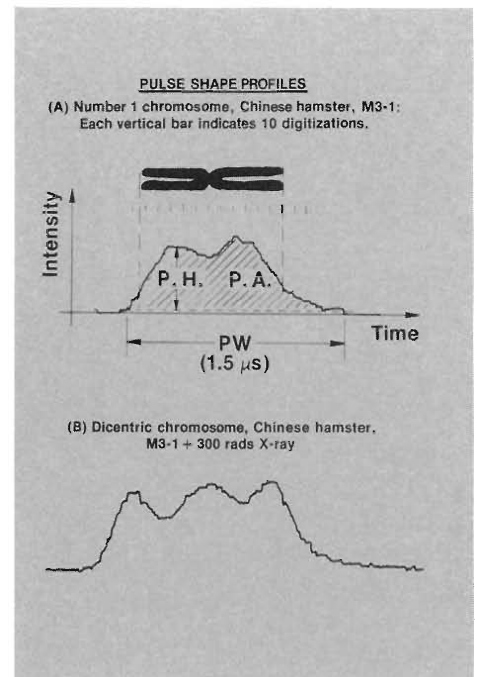


Fig. 18. Pulse shape profiles of individual chromosomes isolated from Chinese hamster M3-1 cells. (A) Pulse shape profile of a single No. 1 chromosome. The intensity of fluorescence across the length of each chromosome is analyzed by a waveform recorder, which rapidly digitizes the fluorescence intensity as a function of time. This information, recorded for each chromosome, can be used to make sorting decisions. Pulse shape profiles provide information on pulse area (P.A.) or total DNA content, pulse width (P.W.) or total chromosome length, and pulse height (P.H.) or fluorescence density per unit length. (B) Pulse shape profile of what appears to be a dicentric chromosome (two chromosomes attached end to end) isolated from the cells used in A but having received 300 rads of x ray 16 hours before chromosome isolation. The time axis is the same for both profiles.

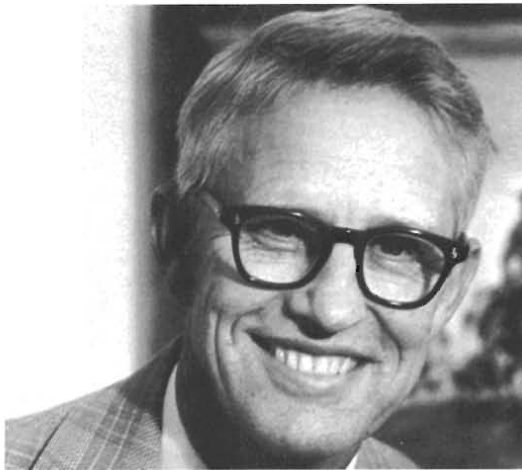
---

## THE AUTHORS

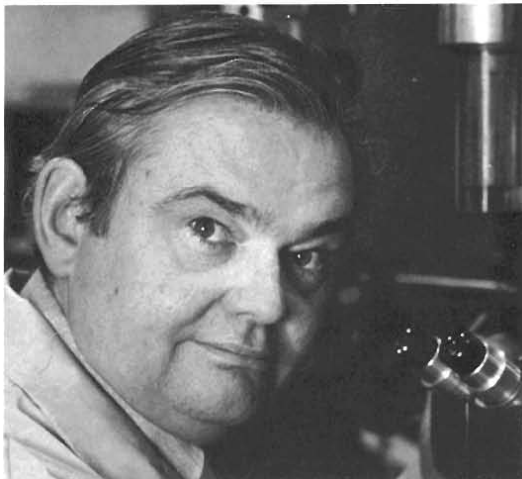
---



**L. Scott Cram**, Alternate Group Leader of the Experimental Pathology Group, came to LASL on a postdoctoral appointment in 1969 and became a Staff Member in 1971. He has been involved in research in flow cytometry and cell sorting, and biomedical applications, including DNA measurements, disease diagnosis, cellular immunology, and chromosome analysis. His current research is to determine the significance of chromosome karyotype stability in cell tumorigenesis. He is an Adjunct Associate Professor in the Pathology Department of the University of New Mexico School of Medicine. In 1978, he received a fellowship from the Max-Planck Gesellschaft and attended the Max-Planck Institut für biophysikalische chemie. Cram earned his bachelor of science degree in physics-chemistry from Emporia State University, his master of science degree in physics-mathematics from Vanderbilt University, and his Ph.D. in biophysics from Pennsylvania State University. He is on the Editorial Board of *Cell Biophysics* and serves on the Executive Council of the Society for Analytical Cytology.



**Dale M. Holm** is Alternate Group Leader of LASL's Biophysics Group, USDA Program Manager at LASL, and LASL/USDA Liaison Officer. He joined LASL in 1952, and since then, he has been involved in nuclear physics experiments; reactor design, construction, and operation; activation analysis; bomb testing; biophysics; and agricultural bioscience. He holds a patent on flow cytometry, for which he also helped to develop instrumentation. At present, his work involves application of LASL technology to many problems of the USDA, including detection of tuberculosis in cattle using flow cytometry. He earned his bachelor of science degree in physics from Lewis and Clark College, and his Ph.D. in physics from Oregon State College in 1955.



**Paul F. Mullaney**, Group Leader of LASL's Biophysics Group, received his bachelor of science degree in physics from Iona College in 1959 and his Ph.D. in physics and biophysics from the University of Delaware in 1965. He is president-elect of the Society for Analytical Cytology and serves on the editorial board of the *Journal for Analytical Cytology*. A Fellow of the Alexander von Humboldt Stiftung Foundation, Bad Godesberg-Bonn, FRG, Mullaney earned their Senior American Scientist prize in 1978. His work has resulted in patents issued for dual-parameter flow photometric apparatus and method, a potential-sensing cell analyzer, and an ellipsoidal cell flow system.

**H**heavy ions provide the newest approach to inertial confinement fusion. This scheme uses high-current beams of heavy ions instead of laser beams to heat a fusion target. The targets are similar to those used for laser fusion. The beams of heavy ions (for example, xenon or uranium) are accelerated to energies of 5-20 GeV by nuclear particle accelerators.

Recently, Donald M. Kerr, LASL's Director, formed a task force to make a complete assessment of heavy ion fusion (HIF). From our studies, we concluded that, although HIF is relatively untested, it has definite promise.

Subsequently, Duane C. Sewell, Assistant Secretary for Defense Programs in the Department of Energy (DOE), designated LASL to take a lead role in the management and technical direction of the DOE's HIF program. We will have the responsibility for directing the heavy ion efforts at Argonne National Laboratory (ANL), Lawrence Berkeley Laboratory (LBL), Lawrence Livermore Laboratory (LLL), Brookhaven National Laboratory (BNL), Sandia Laboratories (SLA), and other laboratories and universities contributing to the program.

Our studies indicate that heavy ions may have an advantage in efficiency and cost over lasers as drivers for inertial confinement fusion. The efficiency of accelerators for converting electric power into kinetic energy of heavy ions can be as high as 20-30%. In contrast, the most efficient lasers, the CO<sub>2</sub> lasers being developed at LASL, now have about 2% efficiency. A new technique for beam multiplexing may bring this up to 10%.

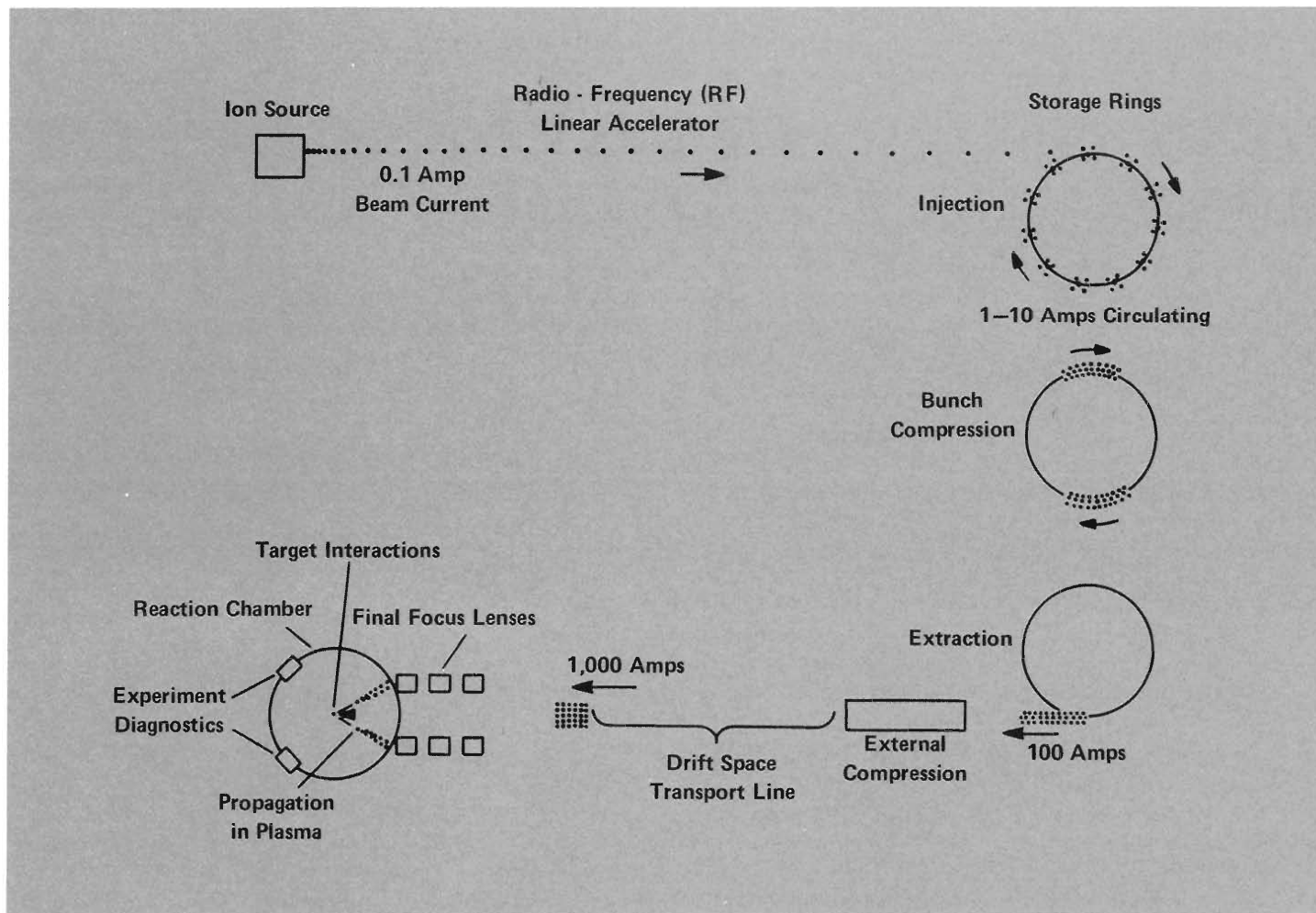
# Heavy Ion Fusion

by George Sawyer

In addition to producing driver energy efficiently, we must be able to deposit a fair fraction of this energy into the fusion target. Our estimates based on well-established theory for the electromagnetic scattering of ions in a medium suggest that heavy ion beams have a high absorption efficiency compared to glass or CO<sub>2</sub> laser beams. We must, however, add a note of caution. As yet we have no experience in stopping beams of heavy ions with (1) the high beam intensities, (2) the high energies (MeV's) per nucleon, and (3) the high temperatures in the stopping medium that will be involved in heavy ion fusion.

The interest in heavy ion fusion comes at a time of increased concern about the efficiency of laser interactions with fusion pellets. Calculations and experiments indicate that the amount of laser energy required to reach fusion break-even is much larger than it was believed to be a few years ago. At least 20 MJ/g must be deposited in the pellet in about 10 ns. Pellets must have a radius of a few millimeters determined by compromises between the pellet gain, the pellet mass, the driver energy, and the practical limits on beam focus. These requirements lead to the conclusion that the driver energy must be 1-10 MJ with a peak power of about 10 TW. Achieving such a large driver energy with lasers will be difficult and expensive. In contrast, the cost scaling with size for large

## SHORT SUBJECTS



*Inertial confinement fusion driven by heavy ions. The drawing shows the production of peak beam currents in the kiloampere range. This method of driving fusion reactions will require many storage rings operating in parallel.*

particle accelerators is favorable. Consequently, large heavy ion drivers should be considerably cheaper than lasers if, as expected, they are more efficient in producing and depositing beam energy.

However, formidable technical problems must be solved in HIF accelerators. Although the accelerators use well-established accelerator technology, they must operate at beam currents much larger than have ever been

achieved. Typical design parameters for HIF include a beam of uranium ions with a 10-GeV particle energy, a 30-kA peak beam current, a 300-TW beam power, a 3-MJ beam, and a 5-mm focal spot size.

There are two quite different approaches to heavy ion acceleration, the rf linac approach being developed at ANL and BNL and the induction linac being developed at LBL. An rf linac,

similar to the one used at the Clinton P. Anderson Meson Physics Facility (LAMPF), would accelerate a 100-mA current of heavy ions to 10 GeV (~50 MeV per nucleon). In this approach, continuous rf energy is applied to a series of drift tubes. The rf energy is phased so that a bunch of ions is accelerated across each gap between adjacent tubes. We believe that the necessary accelerator structure can be developed.

However, the 100-mA beam current from the accelerator is much too small for reactor requirements, so the beam must be accumulated in a storage ring until the current is built up to about 20 A; space-charge considerations limit the amount of current that can be stored in the ring. After the ring is filled, the circulating current is extracted and transmitted through a final beam line to the target chamber. In the final beam line, the pulse of ion current must be bunched or compressed by a factor of 20 to a length of about 10 ns before it hits the pellet. The compression is accomplished by passing the beam through a series of accelerating gaps having a voltage pulse profile that slows down the head and speeds up the tail of the beam bunch. After compression, the final beam current will be about 1000 A.

Thirty such beams must be stored to meet the typical reactor requirements of 30 kA. Multiple storage rings and final beam transport lines fed from the single main accelerator are needed to accumulate and transport this enormous current. Complex beam manipulation is required in the storage ring to stack many turns of current. In addition, large current and long storage times lead to the possibility of beam instabilities. However, the rf linac is based on proven technology and has an inherent high repetition rate required for commercial application.

The induction-linac approach being developed at LBL is based on a pulsed high-voltage technology originally developed at LLL for accelerating electrons in the Astron magnetic fusion device. A single, large pulse of ions is accelerated through a long series of gaps that are momentarily pulsed to high-voltage as the beam passes by. The ions gain only a few hundred kilovolts of energy in each gap, so thousands of gaps

are required in a structure several kilometers long. The pulse of current must be bunched, as well as accelerated, as it passes through the gaps.

The voltage profiles across all the gaps must be carefully controlled, and the timing of voltage application across successive gaps must be precise. The high-voltage pulse technology is similar to that used in magnetic fusion and laser fusion. Although the configuration is simple, because no storage ring is required, it has not yet been demonstrated for ion acceleration. Instabilities in the beam associated with the very large currents may make maintaining the required beam quality difficult. These problems are shared by the rf-linac approach because its final beam transport section is an induction linac, required because of the very high current, short-duration pulse that must be accelerated onto the target.

The real issue for HIF is whether heavy ion accelerators can be developed that are capable of accelerating large currents while maintaining very good collimation of the beam so it can be focused on a small target. Over the next 3 to 5 years, the DOE-funded program will concentrate on development of small test bed accelerators at ANL and LBL to explore the two main accelerator approaches. The test bed accelerators, each costing about \$25 million, will be used to explore critical accelerator issues. As with all inertial-fusion approaches, it is also important to determine the target-interaction physics for heavy ions. The test bed accelerators will not have enough beam energy for significant target experiments.

By fiscal year 1984, the more promising approach will be selected and a several hundred-million dollar test facility will be proposed at a site to be determined. The test facility will be

capable of doing target-physics experiments and testing accelerator performance at large beam currents. The facility, which will be larger than LAMPF, will require a site with considerable space and electrical power.

As part of its lead role in the management and technical direction of the HIF program, LASL has begun a very small effort in accelerator design, target-design physics, beam-transport studies, and reactor-system studies. Given the present funding pressures, it is likely that the entire HIF program will need to be restructured in the near future with emphasis shifted from construction of expensive facilities to the design and execution of less expensive experiments to address the crucial issues of target interaction and beam transport.

# The Search for Muon-Number Violation at

by Cy Hoffman and Minh Duong-Van **LAMPF**

**T**he ultimate goal of studies in nuclear and particle physics is to understand both the interactions between elementary particles and the structure of nuclei and particles. In the construction of a new theory, symmetry principles provide guidelines that must be followed. Noether's Theorem<sup>1</sup> tells us that whenever a physical process is invariant under a certain symmetry, there exists a corresponding conservation law. Thus, conservation of momentum follows from the invariance of physical laws under spatial translations; that is, there are no preferred positions in space. Conservation of energy follows from the invariance to time displacements; that is, there are no preferred moments in time. For a quantum mechanical system, such as a nucleus or a particle, the state of the system is characterized by certain quantum numbers corresponding to the values of conserved quantities. Some of these quantum numbers, such as electric charge and baryon number, do not change; others, such as angular momentum and strangeness, change only by discrete amounts as governed by selection rules. The discovery of a selection rule prompts us to search for the underlying symmetry principle.

In the past 25 years, we have learned that symmetries are not always univer-

sally valid. For example, the weak interaction (responsible for  $\beta$  decay, for example,  $n \rightarrow p + e^- + \bar{\nu}$ ) violates invariance with respect to both parity (the symmetry between a system and its mirror image) and charge conjugation (the symmetry between matter and anti-matter). The interaction responsible for the decay of the neutral K meson is not invariant under time reversal. In this article, we describe some origins of the law of conservation of lepton number and some efforts of physicists at LAMPF (the Los Alamos Clinton P. Anderson Meson Physics Facility) to search for violations of this law.

The muon was discovered in 1937<sup>2</sup> by exposing emulsion plates to cosmic rays. It was first thought to be Yukawa's meson,<sup>3</sup> the pi meson, which was predicted to mediate the strong interaction binding the nucleus together. However, it became clear that the muon interacted too weakly with matter to be the pi meson. The existence of the muon was neither predicted nor theoretically expected and since then, according to Feynman, "No one is quite sure what to do with the muon."<sup>4</sup>

The muon ( $\mu$ ) is an unstable particle that decays through the weak interaction into an electron ( $e$ ), a neutrino ( $\nu$ ), and an antineutrino ( $\bar{\nu}$ ):  $\mu \rightarrow e + \nu + \bar{\nu}$ . Historically, other possible decays into an electron and a photon ( $\gamma$ ) or into three electrons ( $\mu^+ \rightarrow e^+\gamma$ ,  $\mu^+ \rightarrow e^+e^+e^-$ ) and the capture process ( $\mu^- + Z \rightarrow e^- + Z$ ) were of great interest.<sup>5</sup> In the 1950s, it was believed that the muon and the electron were both leptons (light particles) with the same quantum numbers, a condition that allowed these processes to exist. The process  $\mu^+ \rightarrow e^+\gamma$  is not allowed in the first-order approximation to the four-fermion theory of weak interactions in which four fermions (spin 1/2 particles, such as the electron, muon,



and neutrino) can interact at a single point. However, the process is allowed in the second-order approximation by a more complicated virtual sequence of interactions. The branching ratio (ratio of decay rates) predicted by this model for the rare decay  $\mu \rightarrow e\gamma$  versus the known decay  $\mu \rightarrow e\nu\bar{\nu}$  was

$$\frac{\Gamma(\mu^+ \rightarrow e^+\gamma)}{\Gamma(\mu^+ \rightarrow e\nu\bar{\nu})} \sim 10^{-7},$$

where  $\Gamma$  is the rate for the process. In 1957, the experimental upper limit<sup>6</sup> for this branching ratio was  $<2 \times 10^{-5}$ ; therefore, there was no discrepancy between theory and experiment.

However, as had long been realized,<sup>7</sup> the four-fermion weak interaction increases in strength with energy, leading to infinite answers. To avoid high-energy problems, Schwinger<sup>8</sup> conjectured that an intermediate vector boson (a particle with spin 1) mediates the weak interaction. In this nonlocal interaction, the branching ratio above is calculated to be  $\sim 10^{-4}$ ,<sup>5,9</sup> in contradiction to the experimental limit. To reconcile this problem, the notion of lepton quantum number conservation was postulated.<sup>8,10</sup> In this scheme, assignments of lepton and muon numbers are made for the muon, the electron, and two kinds of neutrinos. These are shown in Table I. The conservation law then states that the sum of the lepton number and the sum of the muon number are each conserved separately. This law then forbids the unobserved processes,  $\mu^+ \rightarrow e^+\gamma$ ,  $\mu^+ \rightarrow e^+e^+e^-$ , and  $\mu^- + Z \rightarrow e^- + Z$ , but allows all observed processes including  $\mu^+ \rightarrow e^+v_e\bar{v}_\mu$ .

This scheme, however, requires the existence of distinct electron-neutrinos ( $v_e$ ) and muon-neutrinos ( $v_\mu$ ). Pontecorvo and Schwartz<sup>11</sup> discussed methods to determine if there are two kinds of neutrinos. Danby et al.<sup>12</sup> perfor-

TABLE I ASSIGNMENTS

Lepton Number Assignment

Particle	Lepton Number
$e^-, \mu^-, v_e, v_\mu$	+1
$e^+, \mu^+, \bar{v}_e, \bar{v}_\mu$	-1
All others	0

Muon Number Assignment

Particle	Muon Number
$\mu^-, v_\mu$	+1
$\mu^+, \bar{v}_\mu$	-1
All others	0

med the two-neutrino experiment in 1962 and showed that the neutrinos produced in pion decay, by the process

$$\pi^+ \rightarrow \mu^+v_\mu,$$

interacted with matter to produce muons by

$$v_\mu + n \rightarrow p + \mu^-,$$

but did not produce electrons. The neutrinos produced in beta decay ( $N \rightarrow N' + e^- + \bar{\nu}$ ) do interact with matter to produce electrons.<sup>13</sup> This observation of two distinct types of neutrinos was seen as a validation of the lepton number conservation law.

The experimental status of muon number conservation in 1964 and 1980 is shown in Table II. The prevailing attitude after the two-neutrino experiment was expressed as follows:

The results of the neutrino experiments . . . indicate that the normal weak interaction channels are closed to this decay mode [ $\mu \rightarrow e\gamma$ ]. Since it now appears unlikely that this decay is lurking just beyond present experimental resolution, any further search for the  $\mu \rightarrow e\gamma$  decay mode at this time seems futile.<sup>14</sup>

Even though lepton number conservation accounts for the failure to detect the processes in Table II, there is no fundamental reason for this conservation law to be exact. Unlike electric charge

TABLE II

STATUS OF MUON NUMBER CONSERVATION, 1964 AND 1980

Muon Number Process	1964	1980
$\frac{\Gamma(\mu \rightarrow e\gamma)}{\Gamma(\mu \rightarrow e\nu\bar{\nu})}$	$<2.2 \times 10^{-8}$	$<1.9 \times 10^{-10}$
$\frac{\Gamma(\mu \rightarrow eee)}{\Gamma(\mu \rightarrow e\nu\bar{\nu})}$	$<1.3 \times 10^{-7}$	$<1.9 \times 10^{-9}$
$\frac{\Gamma(\mu^-Z \rightarrow e^-Z)}{\Gamma(\mu^-Z \rightarrow \nu Z')}$	$<2.4 \times 10^{-7}$	$<7 \times 10^{-11}$
$\frac{\Gamma(\mu \rightarrow e\gamma\gamma)}{\Gamma(\mu \rightarrow e\nu\bar{\nu})}$	$<1.6 \times 10^{-5}$	$<5 \times 10^{-8}$

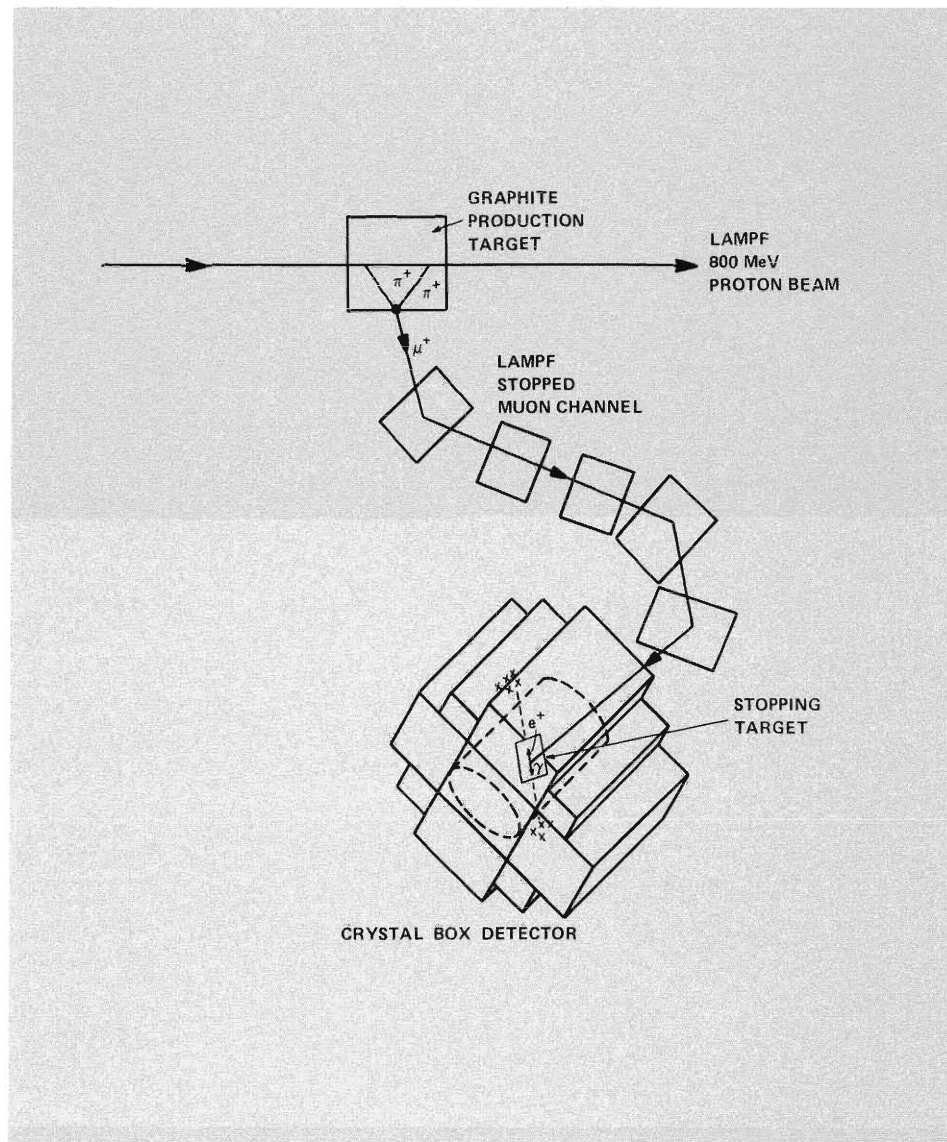
## SHORT SUBJECTS

conservation, which must be exact by virtue of the gauge invariance of the electromagnetic field whose quantum excitations are massless photons, lepton number conservation is not associated with a massless gauge field. A heuristic argument<sup>15</sup> will help explain this.

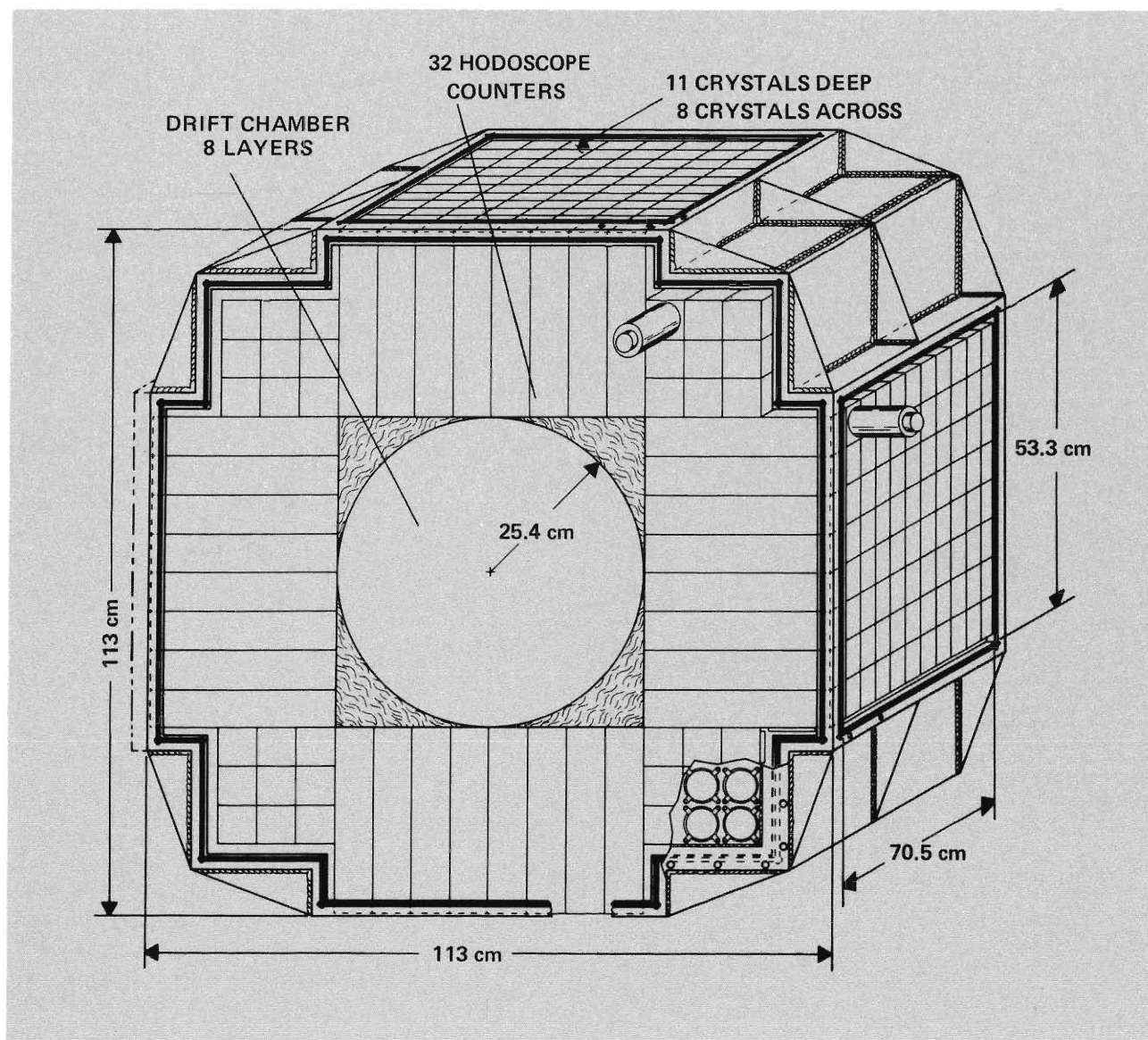
Assume that black holes exist. If a charged particle falls into a black hole, the memory of its charge is preserved by its electric field outside the black hole, so that conservation of electric charge can be verified. On the other hand, if a muon-neutrino falls into the black hole, it leaves no trace at all, so that an exact conservation law for muon number is not a measurable concept.

The most exciting development of the past decade in theoretical physics has been the successful unification of the weak and electromagnetic interactions within the framework of non-Abelian gauge theories.<sup>16</sup> (Electromagnetic interactions alone are described by the Abelian gauge theory quantum electrodynamics.) At present, there is considerable latitude in the exact composition and structure of the correct gauge model: this freedom can be reduced only by accumulating more experimental facts. The gauge models not only suggest that lepton number conservation is not exact, they also predict that muon-number violating processes may occur at rates somewhat below experimental limits. If we can measure the rates or reduce the experimental limits, we will go a long way toward pinning down the correct unified model of weak and electromagnetic interactions and place additional constraints on the many grand unified models recently proposed to unify the weak, electromagnetic, and strong interactions.

In 1977, a group of physicists from the Los Alamos Scientific Laboratory, the University of Chicago, and Stanford



*Schematic view of the experiment at LAMPF. The 800-MeV proton beam strikes a graphite production target producing pions. Some  $\pi^+$ 's slow down and stop near the surface of the target and decay into  $\mu^+$ 's. The muons are then transported by the stopped muon channel to the detection apparatus: The channel consists of dipole bending magnets and quadrupole focussing magnets. The muons come to rest in the stopping target in the center of the crystal box detector and then decay. A  $\mu^+ \rightarrow e^+\gamma$  event is shown. Positrons and electrons are detected in the drift chamber and hodoscope counters and their energy is measured in the NaI; photons do not register in either the drift chamber or the hodoscope counters but do deposit energy in the NaI. All information from the detectors is processed by an on-line computer and stored on magnetic tape for further off-line analysis.*



*The crystal box that will be used to detect rare decays of the muon.*

University<sup>17</sup> mounted an experiment at LAMPF to search for the decay  $\mu^+ \rightarrow e^+\gamma$ . The high-intensity beams of muons available at LAMPF are particularly well suited to a search for this rare decay. This experiment did not detect the decay but did push the experimental upper limit<sup>17</sup> down to

$$\frac{\Gamma(\mu^+ \rightarrow e^+\gamma)}{\Gamma(\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu)} < 1.9 \times 10^{-10},$$

about an order of magnitude more sensitive than any previous search.

Now a new collaboration from the same three institutions has embarked on an experiment to search for the muon-

number violating processes  $\mu^+ \rightarrow e^+\gamma$ ,  $\mu^+ \rightarrow e^+e^+e^-$ , and  $\mu^+ \rightarrow e^+\gamma\gamma$  with a large new experimental facility known as the *crystal box*.

As shown in the conceptual drawing of the apparatus, the basic design of the detection system calls for a large solid-angle modular sodium iodide detector,

## SHORT SUBJECTS

weighing  $\sim 2000$  kg, surrounding a thin target in which the muons stop and decay, a cylindrical drift chamber, and trigger hodoscope (plastic scintillation) counters. The approximately 400 sodium iodide modules will detect 53-MeV positrons and photons with essentially 100% efficiency, an energy resolution of  $\sim 2$  MeV (FWHM) and a timing resolution of 0.5 ns (FWHM) ( $1 \text{ ns} = 10^{-9} \text{ s}$ ). The drift chamber will record the passage of charged particles with a position resolution of  $\sim 200 \mu\text{M}$  (FWHM) in each of eight layers. Photons produced in the events will be identified by detecting energy deposited in the sodium iodide with no corresponding response from the drift chamber or hodoscope counters; electrons and positrons are detected by all of these systems.

The three processes,  $\mu^+ \rightarrow e^+\gamma$ ,  $\mu^+ \rightarrow e^+e^+e^-$ , and  $\mu^+ \rightarrow e^+\gamma\gamma$ , will be studied simultaneously with a sensitivity to branching ratios of about  $10^{-11}$ . (This represents an improvement of  $\sim 10$ , 100, and 5000, respectively, over present experimental limits.) Events will be selected by a hard-wired processor designed to use both the analog and digital information from the detector and to make a decision within 250 ns.

This speed will enable the apparatus to operate at a flux of  $5 \times 10^5 \mu^+/\text{s}$  and will provide an immediate suppression of accidental coincidences from the ordinary decays of several muons. We will begin setting up the experiment in late 1980 and will begin taking data by mid-1981.

If any of these processes is observed, it will be obvious evidence of the failure of the conservation of muon number. The strength of the failure will provide a great deal of information as to what is the correct model of the basic interactions. Should none of these processes be



*Cy Hoffman (left) and Minh Duong-Van (right) examine a prototype drift chamber, a component of the crystal box detector that will be used to search for rare decays of the muon.*

observed, the experiment will force tight constraints on many potential models and eliminate many others. If the process  $\mu^+ \rightarrow e^+\gamma$  is not observed in the crystal box, we plan to reconfigure the sodium iodide modules inside a large magnet and continue the search for muon-number violation with at least an order of magnitude greater sensitivity.

The search for muon-number violation is being pursued at LAMPF with several orders of magnitude greater sensitivity than anywhere else in the world. It must be stressed that "theory" neither predicts nor forbids this violation. The outcome of these experiments will have a great bearing on the way we view the world. Perhaps one of the few conservation laws that we believed to be exact will turn out to be violated after all.

## References

1. E. Noether, Nachrichten Akad. Wiss. Goettingen, Math Physik, Kl. Ila, Math. Phys. Chem. Abt. 1918, 235 (1918).
2. C. D. Anderson and S. H. Neddermeyer, Phys. Rev. **51**, 884 (1937) and J. C. Street and E. Stevenson, Phys. Rev. **51**, 1005 (1937).
3. H. Yukawa. Proc. Phys. Math. Soc. Japan, **17**, 48 (1935).
4. R. P. Feynman, *The Theory of Fundamental Processes*, W. A. Benjamin (1966), p 66.
5. See, for example, G. Feinberg, Phys. Rev. **110**, 1482 (1958) and the references therein.
6. S. Lokanathan and J. Steinberger, Phys. Rev. **240** (1955).
7. W. Heisenberg, Zeit. Phys. **101**, 533 (1936).
8. J. Schwinger, Ann. Phys. **2**, 407 (1957).
9. M. Gell-Mann and R. P. Feynman, Am. Phys. Soc. Meeting, 1957 (unpublished).
10. K. Nishijima, Phys. Rev. **108**, 907 (1957) and S. Bludman, Nuovo Cimento **9**, 433 (1958). The first lepton number scheme was proposed by Konopinsky and Mahmoud. Phys. Rev. **92**, 1045 (1953).
11. B. Pontecorvo, JETP **10**, 1236 (1960) and M. Schwartz, Phys. Rev. Letters **4**, 306 (1960).
12. G. Danby, J-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger, Phys. Rev. Letters **9**, 36 (1962).
13. F. Reines and C. L. Cowan, Phys. Rev. **90**, 492 (1953).
14. S. Frankel, W. Frati, J. Halpern, L. Holloway, W. Wales, and O. Chamberlain, Nuovo Cimento **27**, 894 (1963).
15. A. De Rujula, H. Georgi, and S. L. Glashow, Phys. Rev. **D12**, 147 (1975).
16. See, for example, A. De Rujula, H. Georgi, S. L. Glashow, and H. R. Quinn, Rev. Mod. Phys. **46**, 391 (1974) and E. S. Abers and B. W. Lee, Phys. Reports **9C**, 3 (1973).
17. J. D. Bowman, M. D. Cooper, M. Hamm, C. M. Hoffman, R. E. Mischke, D. E. Nagle, J. S. Sarracino, P. A. Thompson, H. L. Anderson, W. W. Kinnison, H. S. Matis, S. C. Wright, R. L. Carrington, R. A. Eichler, R. Hofstadter, E. B. Hughes, and T. McPharlin, Phys. Rev. Letters **42**, 556 (1979).

The collaborators for this experiment are H. L. Anderson, J. D. Bowman, R. Carlini, M. D. Cooper, M. Duong-Van, J. Frank, C. M. Hoffman, G. Hogan, W. W. Kinnison, R. Macek, H. Matis, R. McKee, R. E. Mischke, D. E. Nagle, V. D. Sandberg, G. H. Sanders, R. Talaga, and R. A. Williams of LASL; D. Grosnick and S. C. Wright of the University of Chicago; and R. Hofstadter, E. B. Hughes, and S. Wilson of Stanford University. Technical support is being provided by R. Bolton, C. Dalton, G. Hart, E. Iverson, G. Krausse, M. A. Johnson, J. Sandoval, J. Sena, and H. P. Von Gunten of LASL, and by J. Rolfe of Stanford University.

# Nuclear Safeguards—



# A Global Issue



*The location of nuclear power plants, operating, under construction, and planned. Of the 530 units worldwide, 341 are outside the United States; about 230 plants are in operation.*

by G. Robert Keepin

Since the dawn of the nuclear age man has experienced varying degrees of both optimism and apprehension about the use and misuse of nuclear energy. This awesome form of energy has provided both an abundant and beneficial source of energy and unprecedented capability for destruction. This basic duality is the driving force behind the unsettled, and still evolving, politics of nuclear energy.

Immediately after World War II, there was hope that placing all nuclear activities under international ownership and management would prevent the proliferation of nuclear weapons. In 1946, the Baruch plan proposed the creation of an international atomic development authority, to be entrusted with all phases of the development, use, inspection, and control of nuclear energy. The plan delineated the need for restraint in nuclear-weapon development and for international safeguards and penalties to prevent diversion of nuclear materials from civilian nuclear power programs. It also proposed that all nations forego the production and possession of nuclear weapons. Although many elements of the Baruch plan were eventually incorporated into international safeguards, in its time the plan was rejected and by 1952, three nations had produced nuclear weapons. Secrecy became the fundamental nuclear policy of the United States and other nations. By the early 1950s, many nations were seeking ways to acquire nuclear technology benefits and to develop their own nuclear energy programs. This activity had an inherent potential not only for peaceful uses but also for military applications. The situation clearly called for renewed attempts to arrive at some form of international understanding, consensus, and constraint.

President Eisenhower's 1953 proposal, the widely hailed "Atoms for Peace" program, marked a fundamental change in US nuclear policy. The program was designed to promote international cooperation in the peaceful uses of nuclear energy and, at the same time, to establish international controls to ensure that the products of this cooperation would not be diverted to military uses.

The Atoms for Peace program was adopted as part of the Atomic Energy Act of 1954. This federal legislation also authorized private ownership of nuclear materials and facilities in the United States and signalled the start of rapid development of nuclear power programs, both domestically and internationally. In 1955, the first United Nations Conference on the Peaceful Uses of Atomic Energy assembled in Geneva, Switzerland. Here, for the first time, scientists from the West and the East met to discuss the technical problems of nuclear energy.

I recall clearly that many of us in the US delegation to the Geneva Conference were filled with a sense of history, and some amazement too, at the open reporting of previously restricted information on fuel-cycle processes and plant operations. Nearly every day, after late-night meetings of the US delegation at the headquarters Hotel du Rhone, we saw new areas of cross-section and fission process data declassified and released to the public. During this historic and unprecedented conference, I could not help but remember my earlier days as a University of Chicago freshman. There, on the way to our freshman calisthenics class under the West Stands of the football stadium, we would occasionally pick up black dust on the soles of our tennis shoes as we passed a sealed-off, heavily guarded area posted with the following warning: US Government Metallurgical Project—Keep Out. As I was to learn years later, the black dust was graphite, the neutron slowing-down or “moderator” material used by Enrico Fermi and his coworkers to achieve the world’s first self-sustaining fission chain reaction on December 2, 1942. To me, the unprecedented open spirit of international cooperation that marked the first Geneva Conference was in stark contrast to the wartime secrecy that had of necessity characterized nuclear activities

just 13 years earlier in Chicago.

The International Atomic Energy Agency (IAEA), a cornerstone of the Atoms for Peace implementation, was created in 1957 to focus on the promotion and control of the peaceful uses of nuclear energy. From the standpoint of politics and economics, Eisenhower’s Atoms for Peace program was more acceptable and far more feasible to implement than the Baruch plan because it did not call for international ownership and management of sensitive nuclear activities. Instead, it proposed a system of *nationally* owned and operated nuclear programs under *international* (IAEA) safeguards inspection and control. Two years before the creation of the IAEA the United States had begun implementing the Atoms for Peace program through “bilateral agreements.” Under these agreements the United States provided other nations with nuclear reactors, enriched nuclear fuel, and technical assistance in the development of their civilian nuclear programs. In exchange, these nations accepted bilateral safeguards to ensure the peaceful use of the material and assistance. The United States administered and inspected these safeguards. Establishment of the IAEA in 1957 provided a more acceptable and effective framework for the administration of safeguards agreements than had been possible under the strictly bilateral agreements.

The IAEA’s two basic objectives are simple and direct:

**The Agency shall seek to accelerate and enlarge the contribution of atomic energy to peace, health and prosperity throughout the world. It shall ensure, so far as it is able, that assistance provided by it or at its request or under its supervision or control is not used in such a way as to further any military purpose.**

Throughout the 1960s, peaceful nuclear

energy programs flourished in many countries because supplier nations, including the United States, offered an extremely attractive, long-term source of nuclear fuel, to discourage the development of other supply sources.

Independently of this peaceful development, France and the People’s Republic of China developed and tested nuclear weapons during this period: France in 1960 and China in 1964. These events increased concerns about nuclear weapon proliferation—both the further build up within nuclear-weapon nations and the possession by new nations. In the mid-1960s, intensified efforts to reduce the risk of proliferation culminated in the 1970 Treaty on the Nonproliferation of Nuclear Weapons (NPT), drafted and signed by the United States and the Soviet Union. Nonnuclear-weapon nations that ratify this treaty give up the option to develop nuclear weapons and agree to submit all their nuclear activities to international (IAEA) inspection in exchange for the right to engage in peaceful nuclear activities with the cooperation of the nuclear-weapon nations.

One concept in international relations introduced by the NPT was unprecedented: participating nations committed themselves to international inspections within their boundaries, thereby yielding part of their national sovereignty to an international authority. During NPT negotiations, one concern of the nonnuclear-weapon nations was that applying international safeguards to their activities, and not to comparable activities in the nuclear-weapon nations, would work to their disadvantage in the competitive marketing of peaceful nuclear energy. This problem was partially resolved when the United States and the United Kingdom, by volunteering to place their peaceful nuclear facilities under IAEA safeguards, put both weapon and nonweapon nations on an equal footing. At present, 116 na-



tions, including 3 having nuclear weapons, are parties to the NPT; and 61 nonnuclear-weapon nations have concluded the required safeguards agreements now in force with the IAEA.

The bases for international safeguards on proliferation and diversion of nuclear materials are the NPT, and the IAEA safeguards and inspection system. Other instruments and activities playing a role include the 1967 Treaty for the Prohibition of Nuclear Weapons in Latin America, known as the Treaty of Tlatelolco; the 1976 Nuclear Suppliers Agreements, the so-called "London Club"; and the Nuclear Non-Proliferation Act of 1978.

In the mid-1970s, a broad re-examination of the policies and practices underlying the NPT was undertaken. The re-examination stemmed in part from India's nuclear explosion in 1974 and in part from the concern that the worldwide growth of nuclear power and the reprocessing of spent fuel would make available large quantities of plutonium. Rightly or wrongly, many persons believe that the step from available plutonium to a nuclear weapon is relatively short. This belief leads in turn to the question of whether safeguards inspection and detection systems can provide "timely warning" of plutonium diversion in one nation quickly enough for the international community of nations to take necessary diplomatic actions, including possible sanctions. This line of reasoning resulted in the conclusion that an unacceptable risk of proliferation exists even if the safeguards system could detect diversion at the moment it occurs, and led to the new US position, announced by President Carter in April of 1977. The new position has deferred breeder reactor development, the reprocessing of spent fuel, and the so-called plutonium economy until after evaluation of alternative fuel cycles by the 66-nation International Nuclear Fuel Cycle Evaluation

(INFCE) study. After 2 years of extensive effort, the final INFCE report was published in March 1980. Although there are many differences among the INFCE participating nations, a significant degree of consensus was reached on the future directions of nuclear energy. Recognizing the worldwide growth of nuclear energy, the INFCE final summary report calls for continued development of nuclear energy under strengthened nonproliferation measures, and for specific endorsement of plutonium, properly safeguarded, as an important fission energy resource for the future. It was further recognized that countries with large electrical power grids, limited uranium resources, and appropriate experience in nuclear technology will be employing the plutonium-burning fast breeder reactor; accordingly the INFCE report calls for placing the associated sensitive fuel-cycle materials under the most highly effective safeguards and nonproliferation measures.

### **LASL's Safeguards Program**

The safeguards program began at the Los Alamos Scientific Laboratory (LASL) in 1966, at a time when nuclear power programs were expanding in the United States and several other industrial countries at an unparalleled rate. After 2 years with the IAEA Headquarters Staff in Vienna, I became convinced of the coming importance—both politically and technically, of the worldwide nuclear safeguards problem. I returned to the United States in late 1965 equally convinced that LASL should launch a vigorous program to develop new nondestructive assay (NDA) techniques and instruments that would in time provide the technical basis for meeting the increasingly stringent safeguards requirements that were inevitable. Following a lengthy series of briefings, hearings, button-holing, and

budget reviews with the Atomic Energy Commission (AEC) and the Congressional Joint Committee on Atomic Energy, the nation's first safeguards research and development program was funded and launched at LASL. The program began in a small laboratory at Pajarito Site; as the program grew, this was augmented a year later by the addition of a second, larger laboratory at Ten Site. Six months after the LASL program was launched, the AEC in Washington established the Office of Safeguards and Materials Management as well as a Division of Safeguards in the AEC Regulatory Branch. The Regulatory Branch is now the Nuclear Regulatory Commission (NRC).

Typical of new projects in their early stages, the new safeguards staff at LASL was highly enthusiastic and dedicated to the challenge before us. The LASL Safeguards program got off to a head start in the safeguards field with a commanding lead that I believe has been retained ever since. With the encouragement and cooperation of Dick Baker, Chemistry-Materials (CMB) Division Leader at the time, and the patient tolerance of Bill Maraman and his Plutonium Chemistry and Metallurgy Group, a special technical liaison committee was set up in 1967 between safeguards researchers and the CMB staff. The committee identified needed applications of newly developed NDA technology to materials measurement, accountability, and safeguards problems. Such problems were not uncommon in the materials processing, fabrication, and recovery operations carried out routinely at the CMB plutonium facility. Through the years, the close liaison between safeguards researchers and the CMB staff has contributed significantly to LASL's leadership position in US and international safeguards technology and to LASL's designation as the Department of Energy's (DOE) lead laboratory in safeguards material accountability

and control research and development.

Today, the LASL safeguards program encompasses all aspects of the design, development, testing, and in-plant evaluation of new techniques, instrumentation, and integrated systems for safeguarding fissionable materials in all types of civilian and national defense nuclear facilities. These activities involve over 150 staff and support persons mainly concentrated in 7 technical groups: Safeguards Technology, International Safeguards and Training; Detection, Surveillance, Verification and Recovery; Safeguards Subsystems Development and Evaluation; Integrated Safeguards Systems and Technology Transfer; International Safeguards; Analytical Chemistry; and Computer and Telecommunications Security.

### **Safeguards Objectives—Domestic and International**

From the beginning of the US nuclear program, nuclear materials and facilities—in both civilian power and defense-related activities—have been recognized as potential targets for theft, diversion, extortion, and sabotage. Accordingly, a substantial national program of safeguards and security was established early on and has been operational ever since. The goal of the national system is to protect nuclear material in facilities and in transit from subnational threats, such as overt attack by an armed group; from diversion, theft, or other unauthorized activity by facility employees; or from a combination of these “external” and “internal” threats. An external threat—for example, an overt attack by 4 to 8 adversaries armed with automatic weapons, the hijacking of a shipment, or sabotage—is countered by physical protection measures. The more subtle internal threat—covert diversion or theft of nuclear materials—is countered by materials accountability and control

systems together with appropriate containment and surveillance measures.

In practice, an integrated system of materials accountability and control together with physical protection is structured to provide, for a given facility, a high-confidence, defense-in-depth safeguards and security system. The record speaks well for the effectiveness of US operational safeguards and security. According to a 1980 Rand Corporation document on threat analysis:

**No nuclear installations in the United States have been attacked, seized, or sabotaged in a manner that caused public risk by release of radioactive materials. No nuclear weapons have been stolen or illegally detonated. No nuclear materials have been diverted or taken by force from installations or while in transit and used for blackmail or made into bombs. No radioactive matter has been maliciously released, endangering public safety.**

The document also states that, although threats have been made to use nuclear materials, all but one proved to be hoaxes. In the one exception low-enriched uranium was removed from a facility, but was recovered within 3 days, and the thief was apprehended. A few cases of minor sabotage have occurred in the United States and occasional incidents of sabotage or attempted sabotage at nuclear facilities in other countries have been reported in the international press.

Physical protection measures include fences, alarms, the prohibition of unauthorized vehicles, random searches of packages or containers entering secured areas, written records of visitors, DOE clearance requirements, portal monitors to detect illicit movement of nuclear material, dual communications systems for protective force personnel, and response force procedures. Operational

procedures, such as the two-man rule, requiring the presence of two cleared persons for access to nuclear materials; special training in emergency procedures; and drills and operational tests of system effectiveness, further support and strengthen the physical protection measures.

NRC regulations for civilian facilities and DOE regulations for government facilities implement material accountability and control in the United States. Under these regulations, measurement and control of special nuclear materials (SNM), such as plutonium and  $^{235}\text{U}$ , are typically required to be better than 1%. The control requirement for reactor fuel fabrication plants, for example, is to within 0.5% of plant throughput. For plutonium and highly enriched uranium (enriched to 20% or more in  $^{235}\text{U}$ ), physical inventories are typically required at bimonthly intervals and/or semiannual intervals. Physical inventories are based on measurements of all material categories including difficult-to-measure scrap and waste, but excluding certain categories of sealed containers and storage vaults. In addition, control and accountability procedures and records must be independently reviewed and audited by NRC or DOE at established intervals.

The concept of “graded” safeguards, used in both domestic and international systems, provides the greatest amount of control and protection to the most sensitive nuclear materials. In the US system, nuclear materials are divided into three categories depending on how difficult it is to convert the material into weapons-usable form. Thus plutonium, highly enriched uranium, and  $^{233}\text{U}$ , in the form of metal or pure compounds, such as oxides or carbides, are designated Category I materials; these require the highest priority and most stringent safeguards. Scrap, residues, and mixtures that must be processed, transmuted, or enriched to become

usable in an explosive device are designated Category II or III materials, depending on the amount of SNM involved.

In contrast to national safeguards systems, which are designed to counter subnational adversaries, international systems are designed to verify that governments have not used nuclear activities as a source of material for clandestine nuclear weapon programs. The objectives as well as the technical requirements and methods used in the two systems are quite different in some important respects. International safeguards systems are aimed at detecting the diversion of nuclear material to unauthorized purposes and at deterring such diversion by the risk of early detection. While a national safeguards system has the authority and capability to physically protect facilities and material and to recover diverted material, the international (IAEA) system is neither intended to, nor able to, *prevent* diversion. Its main objective is to *detect* discrepancies in inventories and to *deter* diversion by providing a *timely warning* intended to *trigger* international reaction, including possible sanctions.

IAEA safeguards objectives and requirements contain two important quantitative expressions: *significant quantity* and *conversion time*. A significant quantity is the approximate amount of nuclear material, including allowance for loss, deemed necessary to construct an explosive device. Conversion time is the estimated minimum time required to produce the nuclear components of an explosive device. For materials in direct weapon-usable form, such as the metallic state, the IAEA has defined the significant quantity of plutonium and  $^{233}\text{U}$  as 8 kilograms and the significant quantity of highly enriched uranium as 25 kilograms. Designated significant quantities are of course larger for low-enriched uranium and for materials in less directly usable forms. Similarly,

IAEA has adopted estimated conversion times for different material categories. For example, conversion times for plutonium,  $^{233}\text{U}$ , or highly enriched uranium in the metallic form are taken as the order of days (7-10 days), whereas conversion times of oxides or other pure compounds of plutonium,  $^{233}\text{U}$ , or highly enriched uranium are taken as the order of weeks (1-3 weeks). The IAEA timely warning criterion requires that the *detection time*, defined as the maximum elapsed time between an indicated diversion and its detection by IAEA safeguards, be less than the estimated minimum conversion time.

In general, international safeguards criteria and requirements are not as stringent as the corresponding national safeguards and security requirements; this is due to the distinctly different objectives of national and international safeguards. In a national system, diversion of a relatively small amount of SNM, such as a threat to disperse 100 grams of plutonium, would be a matter of immediate concern. Thus, performance goals for a national system would typically include the detection of relatively small quantities of SNM in minutes or hours. Likewise, an alarm indicating unauthorized entry into a nuclear facility should bring armed guards to the scene within minutes. The IAEA, on the other hand, does not have the task of prevention or interception of such malevolent acts, or even the detection of such small target amounts of materials in such short times. Instead, the international safeguards system inspector must detect the larger IAEA significant quantities and, depending on the type of material, IAEA detection times may be days, weeks, or months rather than minutes or hours.

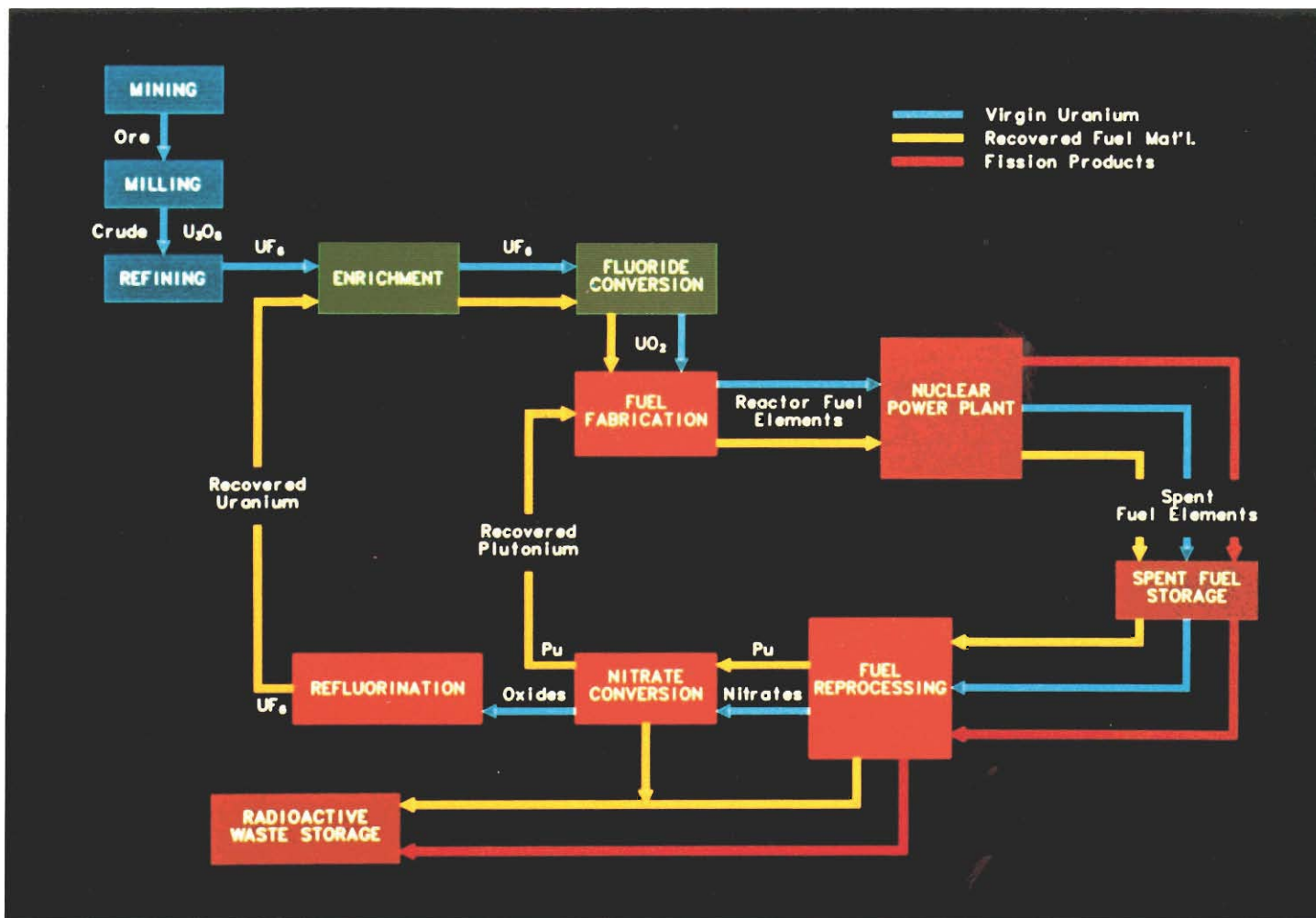
Like the record of the US system, the record of the IAEA international safeguards system is indeed reassuring. Thus far, there has been no diversion of nuclear material under IAEA

safeguards, and the likelihood of future diversions can reasonably be expected to remain small, in part because of the ongoing operation, and continuous upgrading, of the IAEA system. The IAEA currently carries out some 800 inspections annually in well over 300 facilities around the world having an aggregate of some 70 tons of plutonium, over 10,000 tons of enriched uranium, and 30,000 tons of natural uranium. If one divides the annual IAEA safeguards budget by the kilowatt hours of electricity generated annually in all nuclear power plants, the result is roughly \$0.00002 per kilowatt hour—or about 0.1% of the nominal cost of electricity.

### Safeguarding the Nuclear Power Fuel Cycle

According to a recently completed 5-year study by the US National Academy of Sciences (the so-called “CONAES report”), the only choice the United States has to meet large-scale electricity demands for the next 30 years or more is to burn coal and build and operate nuclear power plants. This study of nuclear and alternative energy systems further concluded that nuclear-generated electricity may be the nation’s only choice for the 20-year period beginning in 1990. Around that time, operation of coal-burning plants may be curtailed sharply by the future strong demand for coal as a valuable source of synthetic liquid and gas fuels, and by the threat that carbon dioxide accumulation from coal combustion could alter climatic conditions through the heat-trapping *greenhouse effect*.

The CONAES report addressed US domestic energy needs. Similar conclusions on the international level were reached independently by the 66-nation INFCE study, which addressed worldwide nuclear energy needs and fuel-cycle alternatives on a worldwide basis. The INFCE report calls for con-



**Fig. 1. Power Reactor Fuel Cycle.** Uranium ore is mined, milled, and refined, and the resulting  $U_3O_8$  is converted to  $UF_6$  for enrichment to approximately 3%  $^{235}U$ . Reactor fuel, fabricated from virgin, enriched uranium, or mixtures of uranium and plutonium, provides power in the reactor. Spent-fuel elements are stored at reactor sites or at specially designed away-from-reactor storage facilities. In a "once-through" fuel cycle, the spent fuel is stored permanently, and the remaining  $^{235}U$  and the plutonium formed as a by-product of power generation are not used. In a complete fuel cycle, the uranium and plutonium are recovered from the spent fuel and recycled to provide raw materials for new, "mixed-oxide" fuel elements. Safeguards efforts are concentrated on preventing diversion of separated plutonium between the fuel reprocessing and fuel fabrication steps.

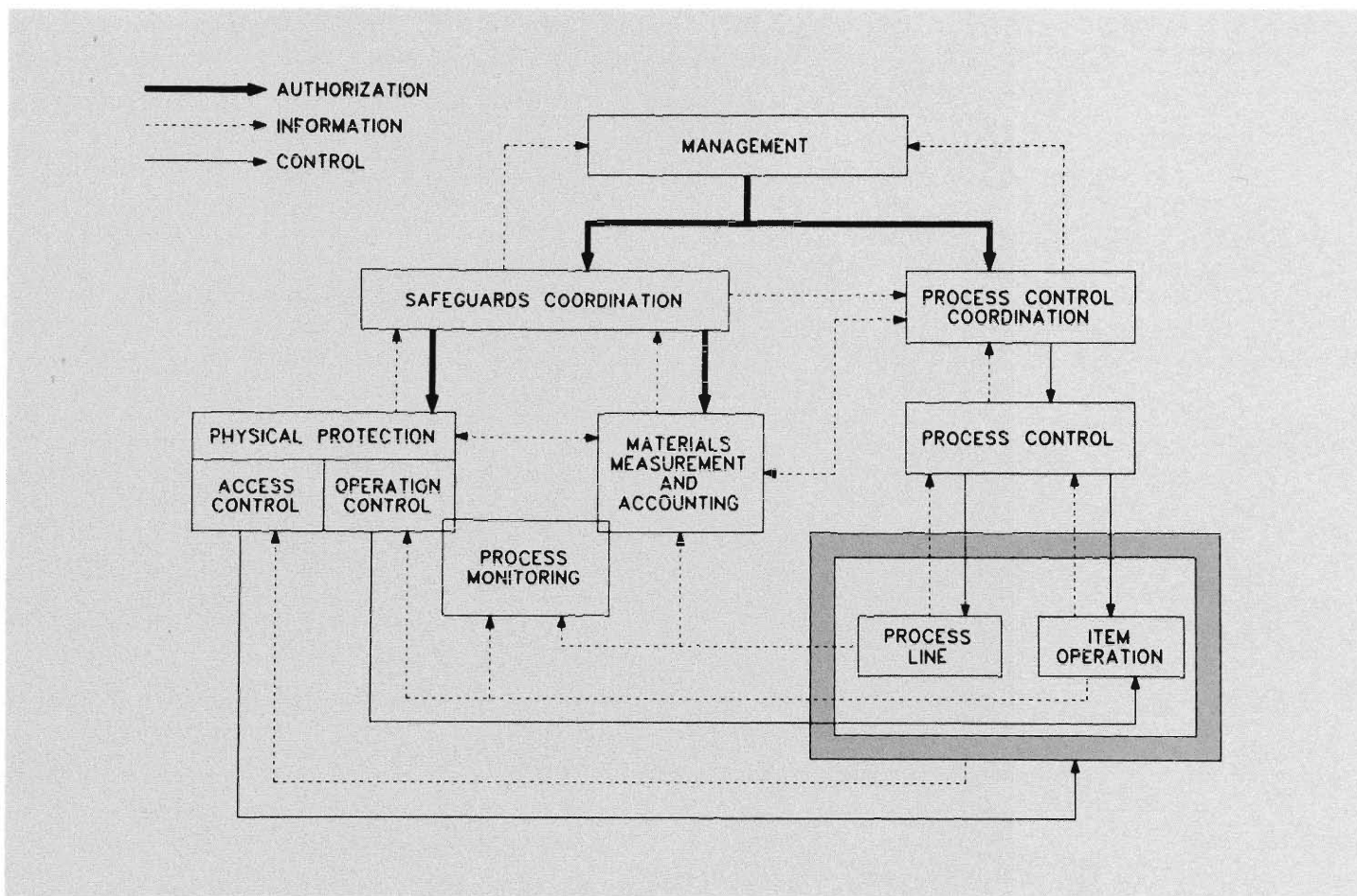
continued development of nuclear power and endorsement of plutonium-based nuclear energy systems, including commercial development of the fast breeder reactor in appropriate countries, in order to avoid a projected shortage of uranium fuel by the end of this century. The INFCE report urged that technical safeguards against proliferation be applied in "a consistent and predictable" (reasonably standardized) way that would not discourage the peaceful development of nuclear energy by creating doubts and uncertainties about the future availability of fuel supplies.

Despite numerous problems and difficulties, nuclear energy is rapidly becoming a major energy source in an

increasing number of countries. It currently supplies over 10% of all electric power generated in the United States and 20% or more of total electric power in some industrialized countries, such as Belgium, Sweden, and Switzerland. Some of the more advanced developing countries, such as India and South Korea, have significant and growing nuclear power programs, while many other developing countries are actively seeking to acquire this new source of energy. Recent projections indicate that nuclear power plants will supply nearly one-quarter of the world's electrical energy by approximately the year 2000. As IAEA Director General Sigvard Eklund noted at a recent LASL collo-

quium, the driving force behind the worldwide growth of nuclear energy is not difficult to understand when viewed against the background of economic, political, and supply-assurance problems associated with the world's shrinking supply of hydrocarbon fuels. With the growing demands for fossil fuel, the cost of oil, for example, has risen by a factor of 5 or 6 in nearly as many years.

Hand in hand with the promise of nuclear energy come some challenging, and recently much publicized, problems and concerns. The accident at Three Mile Island near Harrisburg, Pennsylvania, in 1979, focused worldwide attention on the problems of nuclear reactor safety. TMI also has had



**Fig. 2. Structure of the Safeguards System.** The functional relationships among the elements of a safeguards system and the normally required management and process control elements of a nuclear fuel cycle facility are indicated by the arrows. Process and item operations are contained within a physical protection barrier (dark outline box) that is part of the physical protection and materials control components of the safeguards system. Materials control is provided by monitoring both the process line and the item operations; the item operations also are controlled. Materials measurements and accounting data are derived from measurements of nuclear materials in process operations. Coordination of each of the components of the safeguards system provides facility safeguards status information to both management and process control coordination.

implications for nuclear energy generally, bringing increased attention to the problems of nuclear waste, weapons proliferation, and nuclear material safeguards. Full realization of nuclear power's great potential for meeting world energy needs will clearly depend on how effectively such problems are addressed, including how effectively nuclear safeguards can be implemented on both the national and the international levels.

During their *lifetime*, nuclear reactor fuel materials undergo a variety of physical and chemical processes in various plants and facilities collectively known as the nuclear fuel cycle (Fig. 1). To maintain strict accountability and control of sensitive fissionable materials throughout the nuclear fuel cycle, we

must be able to take a rapid and accurate inventory of these materials in each facility at any given time. This requirement is especially important if a diversion, theft, extortion, or blackmail threat should occur. We must be able to ascertain quantitatively what, where, and how much material is present in any facility at any time. Even more to the point, we must be able to ascertain how much material may be missing from a facility at any time.

### Unique Role of Measurement and Accounting Systems

Effective safeguards (Fig. 2) depend on a combination of three basic components: (1) physical protection, (2) materials measurement and accounting,

and (3) materials control, including process monitoring. Each component is necessary for a fully effective overall safeguards and security system, but only the materials measurement and accounting component can determine the amount and location of material in a plant at any given time. This capability for determining nuclear material inventories with adequate sensitivity and timeliness provides an overall quantitative check on the combined effectiveness of all other safeguards and security measures at a facility.

Under DOE sponsorship, LASL has developed and demonstrated new automated chemical analysis and NDA instruments that can measure the various forms of nuclear materials rapidly and accurately and thereby

provide the high degree of incisiveness required of modern materials measurement and accounting systems. In the application of analytical chemistry methods for safeguards and accountability, it is extremely important to obtain analysis samples that are truly representative of the material being measured. Reliable inventory confirmation further requires precise and accurate analyses of the amounts and isotopic compositions of fissionable materials (uranium, plutonium, and thorium) in widely diverse physical and chemical forms, including pure products, reactor fuels having complex chemical compositions, and numerous types of scrap. Multiphase scrap and materials containing highly refractory components are particularly difficult to dissolve and analyze, while characteristically heterogeneous solid-waste materials in general are simply not amenable to meaningful assay by conventional sampling and chemical analysis techniques.

Major objectives of the LASL analytical chemistry safeguards program are (1) development of fast, effective dissolution techniques and analytical methods for uranium, plutonium, and thorium determinations; (2) design and construction of automated analyzers for these determinations; (3) evaluation of mass spectrometric measurements of uranium and plutonium isotopic distribution; (4) preparation of well-characterized plutonium standard reference materials for distribution by the National Bureau of Standards and for use in DOE safeguards standards intercomparison programs; (5) preparation of plutonium and uranium reference materials for calibration of NDA instrumentation used in the dynamic materials accountability (DYMAC) system at the LASL plutonium processing facility; and (6) participation in an interlaboratory program devoted to measurement of plutonium isotope half-lives.

An example of newly developed automated chemical analysis instrumentation is LASL's automated controlled-potential plutonium analyzer, which determines low-milligram amounts of plutonium with high (0.1%) precision at an average rate of one sample per 30 minutes. The combination of high measurement precision and a specially developed high tolerance for impurity elements makes this relatively low cost analyzer directly applicable to the analysis of a wide variety of nuclear materials.

Because representative sampling of some types of scrap and particularly of heterogeneous solid waste is a particularly plaguing problem, it is not surprising that in the early days of the LASL safeguards program one of the first CMB-identified requirements was for NDA instruments to measure scrap and waste materials. The inherently rapid NDA methods also offered the capability for measuring essentially every individual contained unit of feed or product material. For example, in the assay of reactor fuels, NDA techniques made it possible to measure the total fissionable material loading of each individual reactor fuel rod and to certify, on a routine production basis, the pellet-to-pellet uniformity of uranium fuel loading. Such certification of uniform loading is an important quality control factor in avoiding "hot spots" in the fissioning fuel, and thereby also an important factor in reactor safety. Other "spin-off" benefits of modern non-destructive and destructive measurement techniques developed for safeguards include better in-plant process control, quality assurance, operational safety, and more efficient management of recycle and waste materials.

Major goals for acceptable performance of NDA instruments were set forth in the period from 1965 to 1970, concurrent with steadily increasing pressures to rigorously quantify and

reduce uncertainties in measured nuclear material inventories. Characteristic measurement times for individual items were usually under 10 minutes and desired accuracies for the various material categories were 0.2-3.0% for well-characterized, uniform feed and product materials; 2-10% for recoverable scrap materials; and 5-30% for poorly characterized nuclear waste.

Fissionable nuclide characteristics exploited for "passive" assay are the gamma-ray, neutron, and alpha-heat emissions accompanying the natural radioactive decay of the nuclides. Supplementing passive NDA techniques, "active" assay methods use external neutron sources to induce fissions in a sample; the fissions are then measured by counting fission neutrons or gamma rays. Gamma-ray and x-ray densitometry also provides rapid, accurate determination of the concentrations of uranium, plutonium, and thorium in typical solutions and solids.\* The principal neutron and photon measurement techniques and instruments currently in use or being developed for measuring fuel-cycle materials are summarized in Tables I and II. Calorimetry, a technique based on the measurement of radioactive decay heat of contained materials, also has been implemented widely for measurement of plutonium.

### Advanced Materials Accountancy and Control

In conventional safeguards practice, the accountability of nuclear materials within a facility and the detection of unauthorized removals have relied almost exclusively on discrete-item counting (as opposed to the more difficult task of measuring bulk process materials) and on material-balance accounting following periodic shutdown, cleanout, and

\*See "Nondestructive Assay for Nuclear Safeguards," in this issue.

**TABLE I**  
**GAMMA- AND X-RAY ASSAY SYSTEMS**

Instrument or Technique	Operating Principle	Application
Segmented gamma scanner <sup>a</sup>	Passive gamma-ray analysis; transmission of external source gammas used for attenuation correction	Quantitative assay of <sup>239</sup> Pu, <sup>235</sup> U in scrap and waste
MEGAS <sup>a</sup>	High-sensitivity passive x- and gamma-ray detection	Screening of low-density transuranic waste at the 10 nCi/g fiducial
Gamma-ray spectroscopy <sup>b</sup>	Passive gamma-ray analysis with Ge or NaI detectors; sometimes augmented by measurement of transmission of external source photons	Assay of U, Pu; Pu isotopic analysis; U enrichment
X-ray edge densitometry <sup>b</sup>	Photon transmission in the region of L <sub>III</sub> or K-edges	Elemental concentrations of Th, U, Pu
XRF <sup>b</sup>	X-ray fluorescence	Elemental analysis of Th, U, Pu

<sup>a</sup>Well developed for many fuel cycle applications; instruments commercially available.

<sup>b</sup>Developed for some fuel cycle applications and being evaluated for others.

physical inventory. The classical materials balance usually is drawn around the entire facility or a major portion of the process. It is formed by adding all measured receipts to the initial measured inventory and subtracting from this sum all measured removals and the final measured inventory. During routine production, material control is vested largely in administrative and process controls, augmented by secure storage for discrete items and sealed containers.

Although periodic shutdown-cleanout operations will always be important in determining the amount of bulk nuclear material holdup in process equipment, pipes, pumps, traps, and filters, the use

of this procedure alone has inherent limitations in sensitivity and timeliness. Sensitivity is limited by measurement uncertainties that might obscure the diversion of relatively large quantities of SNM in a large throughput plant. Timeliness is limited by the practical difficulties, the expense, and hence the infrequency of process shutdown, cleanout, and physical inventory; thus a loss of material could remain undiscovered until the next physical inventory.

Recently developed NDA technology, state-of-the-art conventional (destructive) measurement methods, and special implant sensors, combined with computer and data-base management technology,

provide the necessary technical basis for much more effective methods of safeguarding nuclear facilities. For example, the greater sensitivity and timeliness requirements on SNM control now being imposed by DOE and NRC can be achieved by subdividing a nuclear facility into discrete accounting envelopes, called unit processes, and drawing individual material balances around them. A unit process is chosen on the basis of process logic, the time material resides within the unit process, and the ability to perform quantitative measurements and draw a material balance. Thus, by subdividing a facility into unit processes and measuring all material flows across unit process boundaries, the

**TABLE II**  
**NEUTRON ASSAY SYSTEMS**

Instrument or Technique	Operating Principle	Application
Random driver <sup>a</sup>	Am-Li source neutrons; fast coincidence detection of prompt neutrons	Fissile assay of Pu and highly enriched U; passive assay of <sup>240</sup> Pu and <sup>242</sup> Pu
<sup>252</sup> Cf shuffler <sup>b</sup>	Cyclic irradiation with moderated <sup>252</sup> Cf neutrons; detection of fission delayed neutrons	Fissile assay of a wide range of U, Pu material categories
<sup>252</sup> Cf fuel rod scanners <sup>a</sup>	Irradiation with moderated <sup>252</sup> Cf neutrons; detection of fission prompt neutrons, delayed neutrons, or delayed gamma rays	Fissile assay of FBR, LWR fuel rods
Antimony-beryllium <sup>b</sup>	Irradiation with Sb-Be photoneutrons; integral counting of fission prompt neutrons	Fissile assay of cold and spent fuels
Thermal-neutron coincidence counter <sup>a</sup>	Time-correlated detection of spontaneous fission neutrons with polyethylene-moderated <sup>3</sup> He well counter	Assay of <sup>240</sup> Pu, <sup>242</sup> Pu, and <sup>239</sup> Pu

<sup>a</sup>Well developed for many fuel cycle applications; instruments commercially available.

<sup>b</sup>Developed for some fuel cycle applications and being evaluated for others.

location and movement of all SNM can be localized in both space and time. Materials balances drawn around unit processes are called *near-real-time* or *dynamic* materials balances to distinguish them from conventional balances drawn after a shutdown, cleanout, and physical inventory.

In a direct application of the foregoing general principles, LASL is integrating newly developed NDA technology with automated data processing, monitoring and surveillance techniques, modern data base management, and decision analysis methods into an overall

dynamic materials accountability system, called DYMAC. The system provides direct in-plant implementation of modern near-real-time accountability and control. A specific application of the DYMAC concept and associated safeguards technology at the LASL plutonium processing facility is shown in Fig. 3. To date, the DYMAC system at the LASL plutonium facility has completed over 100,000 material transactions with no significant errors or discrepancies.

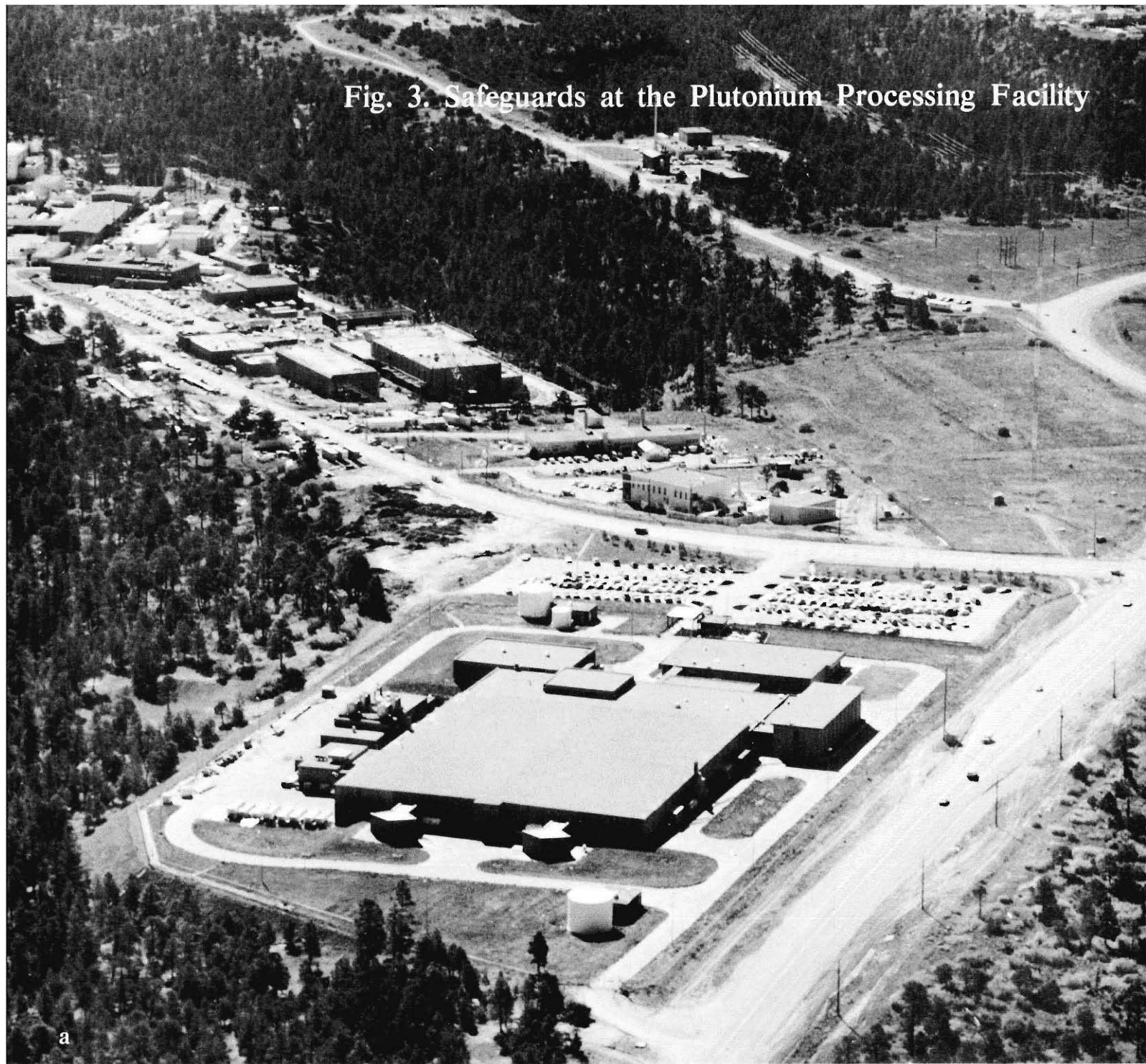
The least disruptive time to develop and evaluate new safeguards systems

designs is before a facility is built, preferably at the conceptual-design stage. Accordingly, the LASL safeguards program has developed and evaluated conceptual designs of cost-effective integrated systems for most fuel-cycle facility types. Concepts, design criteria, and recommendations from the LASL effort are being increasingly implemented both domestically and internationally.\* In addition, advanced materials account-

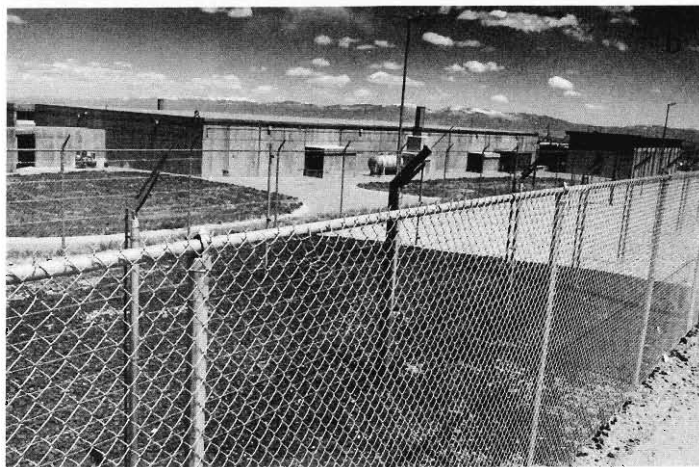
\*See "Dynamic Materials Accounting Systems," in this issue.



Fig. 3. Safeguards at the Plutonium Processing Facility



a



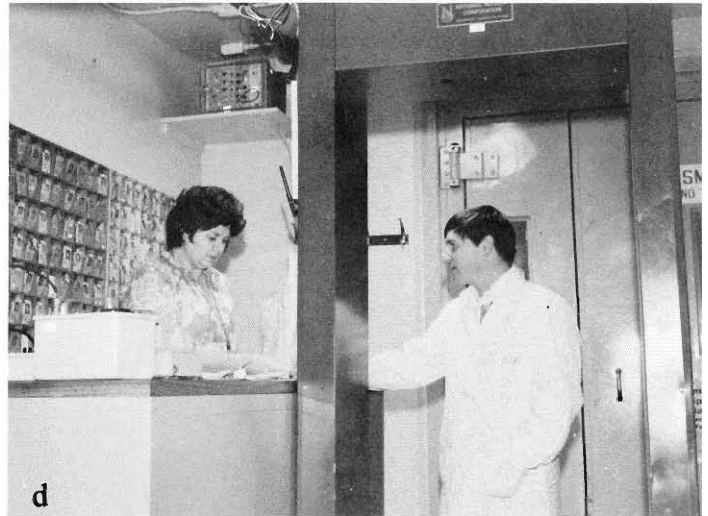
(a) The new LASL Plutonium Processing Facility is the large concrete structure in the left half of the photograph. Many proven security measures including fences, limited access, and alarms are employed to safeguard the facility. Other measures are so new they are in the demonstration stage, such as DYMAC, the dynamic materials accountability system that keeps track of the facility's inventory of nuclear material. Processing operations began here in January 1978. The facility now has over 6000 inventory items to safeguard.

(b) To study the effectiveness of various types of perimeter fences and intrusion devices, Sandia Corporation recently erected a 100-meter-long test bed outside part of the existing perimeter fence at the plutonium facility.

(c) Security guards inspect all items that persons carry into the facility grounds and search vehicles entering and leaving the grounds with portable gamma meters developed at LASL.



(d) The two main entry points into the processing building are manned stations where persons exchange badges and pass through commercially available portal monitors, which detect radiation. The performance of the portal monitor is being evaluated.



(e) Inside the facility's vault, all items are stored in containers that have individual seals. Some of the vault spaces have LASL-developed shelf monitors that can detect whether a container or part of its contents has been removed.

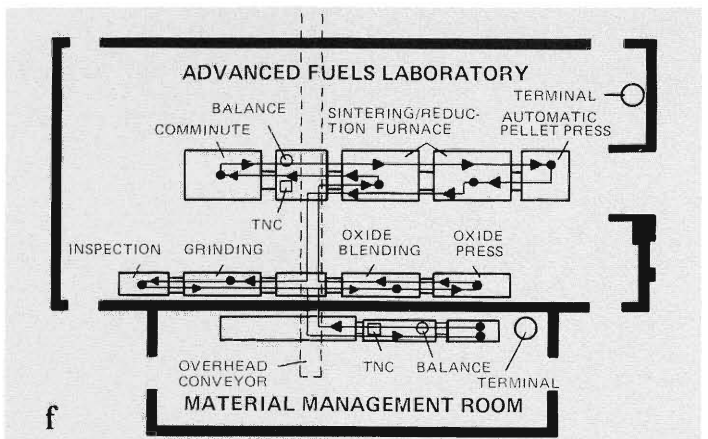
(f) The Advanced Fuels process, one of 23 processes currently operating in the plutonium facility, produces fuel pellets for the FFTF (fast-flux test facility) reactor in Richland, Washington. The process constitutes one accountability area, within which seven subareas have been defined, each corresponding to a particular step in the fabrication process. A materials balance can be drawn around the entire process and around each subarea to determine exactly how much plutonium is present and where it is located.



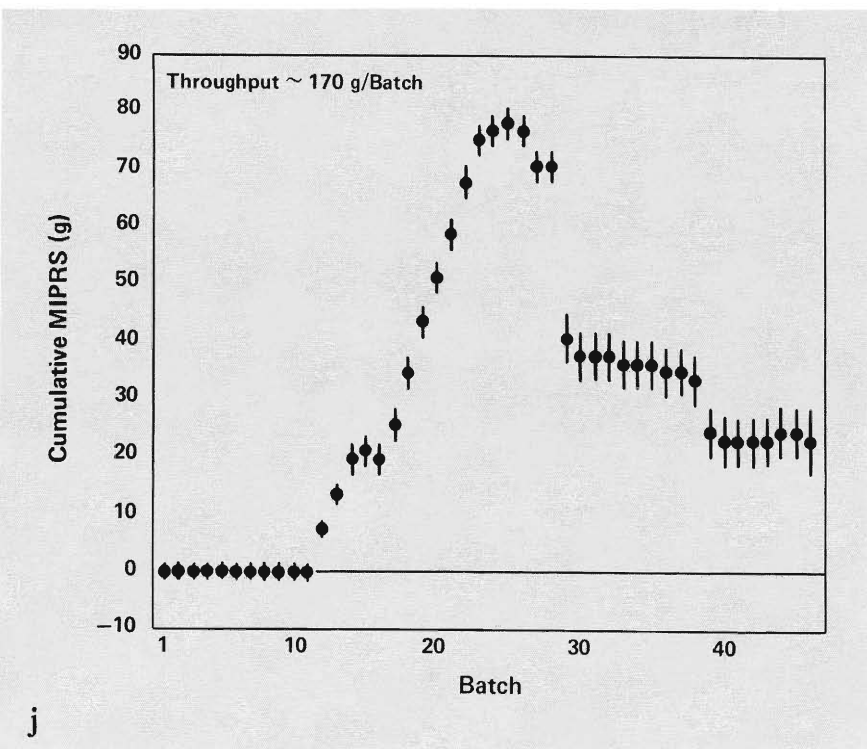
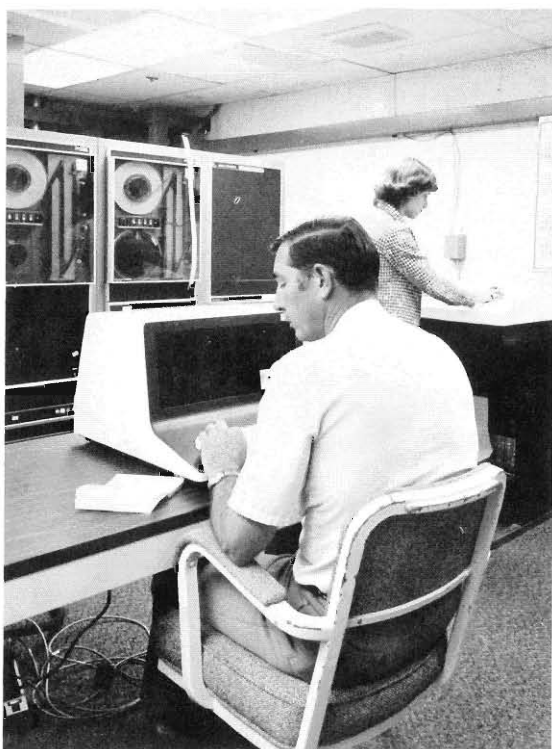
(g) A technician uses a microprocessor control unit to control operation of the thermal neutron coincidence counter on top of the glove-box line. She is measuring the amount of plutonium in a batch of finished fuel pellets for a materials balance.

(h) Following the measurement of finished fuel pellets, a technician makes a transaction to update the computer inventory with the correct amount of plutonium contained in the pellets.

(i) The central computer keeps track of the facility's inventory. It accumulates, sorts, and stores the information from individual transactions, and makes it available to any authorized requestor at a terminal. Programmers and computer operators in the DYMAC computer room keep a constant check on the system to make sure it is operating properly.



(j) A materials balance is the difference between material introduced into a unit process and the material removed from the process. Results of material balancing for the sintering/reduction furnace in the Advanced Fuels process is shown on an MIP (material-in-process) chart. Small amounts of plutonium accumulate on the boats that transport the pellets into the furnace. The chart indicates the amount of plutonium that each batch contributes to the MIP buildup on the boats. When the MIP grew to about 80 g, the supervisor conducted a cleanout to recover as much plutonium as was practical. The plutonium recovered from the boats was measured with the thermal neutron coincidence counter and sent to scrap recovery.



tability and control systems, many quite similar to DYMAC, are being tested and evaluated at a number of nuclear facilities in the US and abroad. As these in-plant test and evaluation programs are completed, the resulting technology and operational experience will be available for introduction into various types of domestic and international fuel-cycle facilities.

### Emergency Response and Recovery

Its excellent record notwithstanding, if a safeguards system should fail and nuclear materials are missing from a facility, there must clearly be a demonstrated response capability to recover materials rapidly, and to apprehend the offender. Likewise, an emergency response plan and demonstrated field-operational capability is essential in responding to nuclear emergencies, accidents, acts of terrorism, blackmail, and sabotage. LASL's special qualifications and experience in both national defense programs and safeguards technology provide a unique capability for innovative design and development of instrumentation for surveillance and search-and-recovery applications. This capability includes the design of handheld monitors for searching personnel and vehicles at facility-access areas and the development, testing, and evaluation of SNM portal monitors, vault monitoring systems, and enclosure detector arrays. It also includes passive and active NDA techniques for SNM identification and verification as applied, for example, to a variety of thorny problems that arise in safeguarding SNM movements into and out of a rigidly proscribed "perimeter" around sensitive technology areas in domestic or international fuel-cycle facilities, or in safeguarding defense-related activities and facilities.

A major component of emergency-

response capability is the NEST—Nuclear Emergency Search Team—activity. This program provides emergency response to incidents of nuclear extortion, nuclear weapon accidents, lost or stolen nuclear materials, and terrorist threats. Portable and mobile nuclear detection systems having high sensitivity and real-time data processing and analysis capability have been developed and deployed for field test, evaluation, and operational use. A related effort involves the development and field testing of instrumentation and procedures for detection, diagnosis, and disabling of improvised or otherwise unknown nuclear devices. Suffice it to say, these efforts require extensive coordination with other DOE laboratories and federal agencies, primarily the FBI and the Department of Defense, all of whom share with LASL major responsibilities in the nation's emergency response system.

### The 1980s as the Decade of Technology Transfer

If the 1970s can be regarded as the decade of modern safeguards technology development, the 1980s must be the decade of the transfer of this technology to nuclear facilities—both existing and new. As indicated in Table III, interactions between the LASL program and nuclear facilities of all types, in both the government and private industry sectors, involve the gamut of safeguards R&D activities from instrument development, calibration, test, and in-plant evaluation to the design, optimization, and performance analysis of overall facility safeguards systems. On the international level, there is growing interest, particularly among other industrialized nations, in the design, optimization, and practical in-plant implementation of integrated safeguards systems incorporating state-of-the-art materials



**TABLE III**

**LASL INTERACTION WITH US FUEL CYCLE FACILITIES**

Fuel Cycle Facilities	MC&A <sup>a</sup> Systems Studies	Location for Test and Evaluation of NDA Instrument	Process Info to LASL	LASL Instrument Specs to Vendor or Facility	Training	Consultation
<b>Fuel Fabrication</b>						
Westinghouse Corporation	●	●	●	●	●	●
General Atomic Company		●	●	●	●	●
General Electric Corporation		●	●	●	●	●
Nuclear Fuel Services			●	●	●	●
Babcock & Wilcox				●	●	●
<b>Spent-Fuel Reprocessing</b>						
Allied-General Nuclear Services	●		●	●	●	●
INEL, Idaho Falls		●	●		●	●
Oak Ridge National Laboratory	●		●	●		●
Savannah River Plant	●	●	●		●	●
<b>Nitrate-to-Oxide Conversion</b>						
Allied-General Nuclear Services			●	●	●	●
Savannah River Plant	●	●	●		●	●
General Electric Company	●		●		●	●
Rocky Flats Plant			●	●		
<b>Waste Handling and Solidification</b>						
Allied-General Nuclear Services			●		●	●
SROO-SRL	●		●		●	●
INEL, Idaho Falls			●	●	●	●
Rockwell, Hanford			●	●	●	●
<b>Uranium Enrichment</b>						
Union Carbide	●	●	●	●	●	●
Goodyear Atomic		●	●	●	●	●
TRW	●		●			●
LLL, Livermore	●		●			●
<b>Critical Facilities</b>						
Argonne National Laboratory	●	●	●	●	●	●
<b>Nuclear Instrumentation Vendors</b>						
				●	●	●

<sup>a</sup>Materials control and accountability.

measurement and accountability technology, materials control, and physical security including effective use of containment and surveillance techniques.

A major component of effective technology transfer is education and training in the use of modern NDA instrumentation and information-handling and analysis systems. The entire area of safeguards professional training has received marked impetus from the Three Mile Island nuclear reactor accident and the resultant three main "lessons learned": the need for (1) better professional training of reactor operators, (2) better measurement instrumentation, and (3) better emergency response. One notable example of the effective transfer of modern safeguards technology to plant operators and safeguards inspectors alike is DOE's ongoing Safeguards Technology Training Program conducted by LASL through four separate course offerings per year:

1. **Fundamentals of Nondestructive Assay of Fissionable Material Using Portable Instrumentation.**
2. **In-Plant Nondestructive Assay Instrumentation (to be succeeded in 1981 by a course on advanced instrumentation based on neutron detection methods).**
3. **Gamma-Ray Spectroscopy for Nuclear Materials Accountability.**
4. **Advanced Systems for Nuclear Materials Accounting.**

These training courses attract well over 100 participants annually. Participants from the United States represent both the government and private sectors and those from the IAEA inspectorate represent a large number of countries around the world.

Technology transfer and assistance to the IAEA encompasses not only development, test, and evaluation of instruments, but also personnel training (of highest priority to IAEA), technical con-

sultation, and direct assistance to the IAEA safeguards staff by visiting consultants and resident experts on loan from member states. Two examples are US participation in the IAEA International Working Group on Reprocessing Plant Safeguards and in the IAEA Advisory Group on Fuel Element Fabrication. Both groups are concerned with the application of IAEA safeguards to the advanced large-scale fuel-cycle facilities that are foreseen for the future. Four LASL safeguards staffers are currently assigned to the IAEA Department of Safeguards at Agency headquarters in Vienna.

A new component in the safeguards technology transfer program at LASL is the International Training Course on Nuclear Materials Accountability sponsored by DOE in cooperation with IAEA. This course, authorized by the US Nuclear Non-Proliferation Act, was conducted May 27-June 6, 1980, at Bishop's Lodge near Santa Fe, New Mexico. The course provided to foreign governmental and institutional managers the basic knowledge needed to develop national safeguards regulations and requirements for their individual countries, and to plan toward implementation of domestic safeguards systems that will serve national needs as well as those of the IAEA International Safeguards System of inspection and verification. Lecturers for the course were experts drawn from the IAEA, United States, Canada, Czechoslovakia, Germany, and Japan. Delegates from over 25 countries participated in the course. A similar DOE/IAEA-sponsored course on the physical protection of nuclear materials is conducted by Sandia Laboratories each fall.

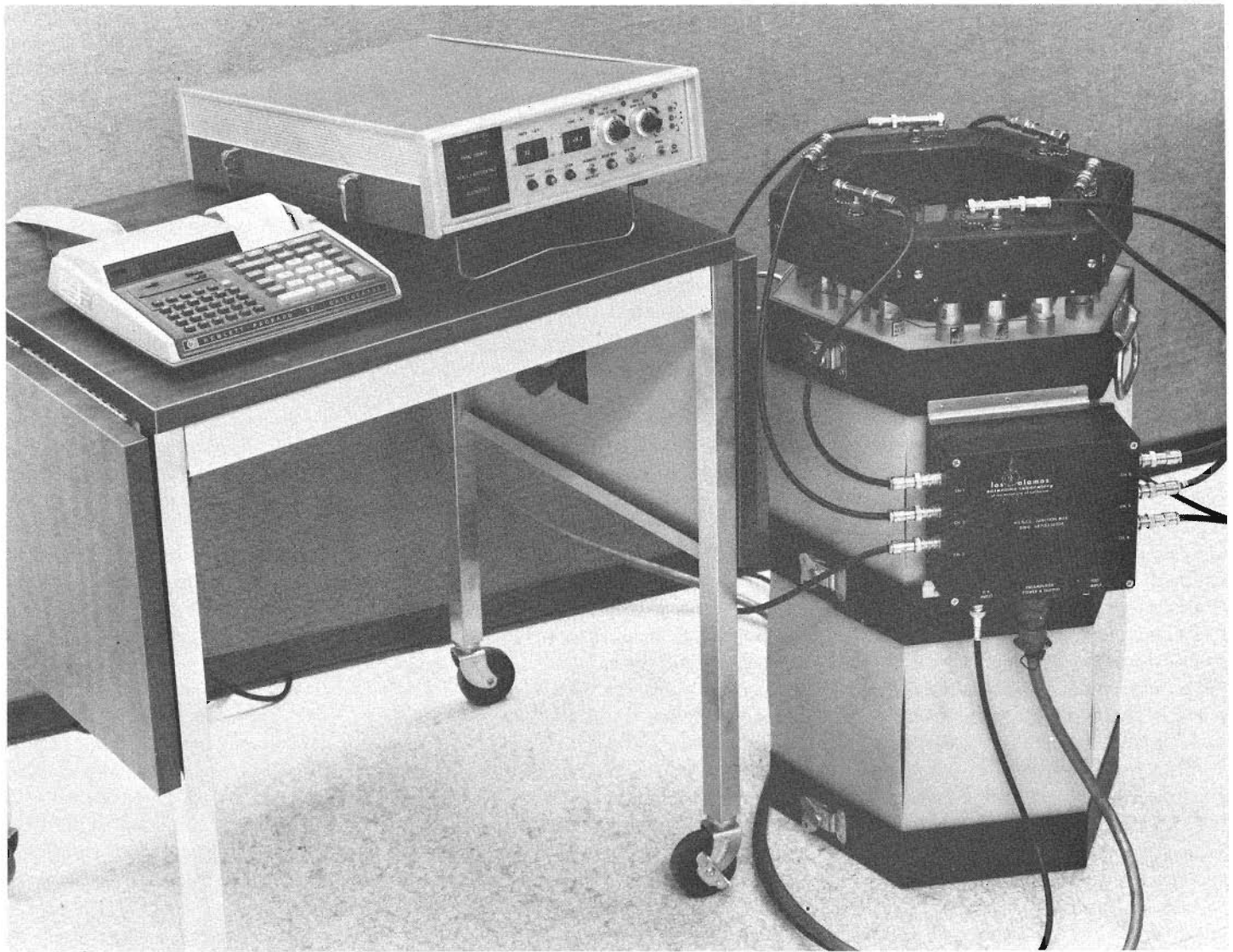
### **Emerging Impact and Role of International Safeguards**

Recent expansion of the US safeguards program in areas of technical

support for the IAEA and cooperative agreements with other countries reflect the growing importance of international safeguards. IAEA needs can be grouped into two major categories: (1) present requirements for portable measurement instrumentation, inspection and verification capability in direct field inspection applications (for example, the HLNCC instrument shown in Fig. 4) and (2) future requirements for methods, instruments, and techniques to be developed for independent verification of different types of advanced in-plant material accountability and control systems, such as DYMAC.

A major international effort is the TASTEX program, in which the United States, Japan, and IAEA are participating jointly in the development, test, and evaluation of advanced instrumentation and safeguards techniques at the Tokai spent-fuel reprocessing plant in Tokai Mura, Japan. In this program, a K-edge densitometer is used for nondestructive assay of plutonium nitrate product solution. The densitometer, which measures *elemental* (total plutonium) concentrations in solutions, provides a valuable complement to gamma-ray spectrometry, which measures plutonium *isotopic* composition. Successful in-plant experience with this type of new NDA instrumentation is expected to lead to the deployment of a wide range of automated NDA instruments at nuclear processing facilities. This should, in turn, provide a sound technical basis for future implementation of near-real-time material measurement and accountability systems in various types of plants and facilities throughout the nuclear fuel cycle.

As regards the outlook for the future, it is significant that this first year of the 1980s will see a number of important developments in international safeguards and nonproliferation. In March, INFCE endorsed stringently safeguarded plutonium-based nuclear energy systems



*Fig. 4. The high-level neutron coincidence counter (HLNCC) detects neutrons from the spontaneous fission of  $^{240}\text{Pu}$  using  $^3\text{He}$  proportional counters in a polyethylene moderator. A shift register coincidence technique is used to distinguish fission neutrons from background. The instrument is portable for use by IAEA inspectors. The electronics to operate the detectors and analyze the coincidence data are contained in the package on the table, next to a programmable calculator that is interfaced to the shift register unit.*

for the future, including the judicious deployment of plutonium breeder reactors (again under strict safeguards and controls) as the only means of avoiding future shortages of uranium fuel. Today the total plutonium inventory of irradiated civilian reactor fuels is easily the order of 100 metric tons and is increasing at a rate of 25-30 tons per year. Although breeder reactors eventually will reduce this inventory, concerns about such potentially large stockpiles of plutonium—in whatever form—have given rise to several international studies

and evaluations, involving both technical improvements and institutional arrangements, designed to place sensitive materials and fuel-cycle facilities under multinational or international control.

Proposed institutional arrangements include (1) regional fuel-cycle centers, in which large fuel reprocessing and fabrication plants would be co-located to provide economy of size and operational efficiency and to minimize vulnerability to theft and diversion; (2) an international fuel authority responsible for providing fuel service and allocating fuel resources; (3) establishment of international plutonium storage centers under IAEA control (foreseen in the Agency's Statute, Article XII, A.5); and (4) the concept of regional nuclear waste repositories, fuel reprocessing plants, and enrichment facilities under international or multinational authority. Working out the details of any such international or multinational arrangements would be a monumental task indeed, and could only be done by the potential participants themselves. With such proposals, some of them strikingly similar to the international ownership/custody/management concepts in the original Baruch plan, we have, in some sense, come almost full circle in the evolution of international safeguards.

Also in this pivotal year, 1980, two important international safeguards agreements are pending ratification by the US Senate. The first is the US-Australian Agreement on the Peaceful Uses of Nuclear Energy, the first renegotiated safeguards agreement under the new, more stringent safeguards provisions of the NNPA. The second is the US-IAEA Agreement for the Application of IAEA Safeguards in the United States, pursuant to the US 1967 offer to implement IAEA safeguards in all US facilities except those having direct national security significance. A similar voluntary agreement, already in force with the United Kingdom, is enabling the

IAEA to gain valuable experience in the inspection of a fast-breeder plant and related reprocessing facility. President Carter recently asked the US Senate to take up the US-IAEA Agreement this spring so that ratification can be completed before the (potentially contentious) NPT 5-year review conference of the 116 NPT signatory nations at Geneva in August of this year. The US-IAEA Agreement, an act of good faith on the part of the United States, may help to alleviate a certain hardening of position by some countries against the NPT, which some nations view as an unequal treaty that discriminates in favor of the nuclear-weapons states and thereby against all others.

To make the NPT as equitable and acceptable as possible, the IAEA is working hard to upgrade and standardize the applications of NPT "full-scope" safeguards. Measurement and surveillance techniques used by IAEA inspectors are being improved continually both by the IAEA staff and through technical support programs of the United States and other IAEA member nations. Also, through IAEA field-inspection experience, better methods of inspection, inventory verification, reporting, and assessment are being evolved constantly to maximize inspection efficiency and effectiveness while minimizing intrusion into plant operations and production. Implementation of the US offer to place its peaceful facilities under IAEA safeguards should do much to facilitate further improvement of the IAEA system.

Another key aspect of NPT acceptability and workability is the assurance of an available supply of nuclear fuel—at present, uranium. Irrevocable fuel supply assurances are essential to the fundamental *quid pro quo* of the NPT agreement and should be extended promptly to nations that meet their non-proliferation undertakings. Uncertainties and doubts about supply assurances in

recent years have had serious repercussions throughout the world nuclear community. An oft-quoted international safeguards slogan succinctly states the basic *quid pro quo* of the NPT Treaty: "Irrevocable safeguards for irrevocable supply."

As many have pointed out (especially to safeguards technologists!), there is no question that safeguards and non-proliferation issues are first and foremost a political problem. However, it is also clear that safeguards technology development, coupled with "real world" operational experience, is indispensable in (1) providing the technical understanding and input essential to prudent planning and decision making, even at the highest political levels, and (2) providing the demonstrated technical means to implement the hardware and systems called for in those plans and decisions. Within severe budget limitations, the IAEA is making every effort to anticipate and prepare for the sophisticated fuel cycles of the future and the commensurately sophisticated technical capabilities that will be needed to carry out its essential inspection and verification functions effectively.

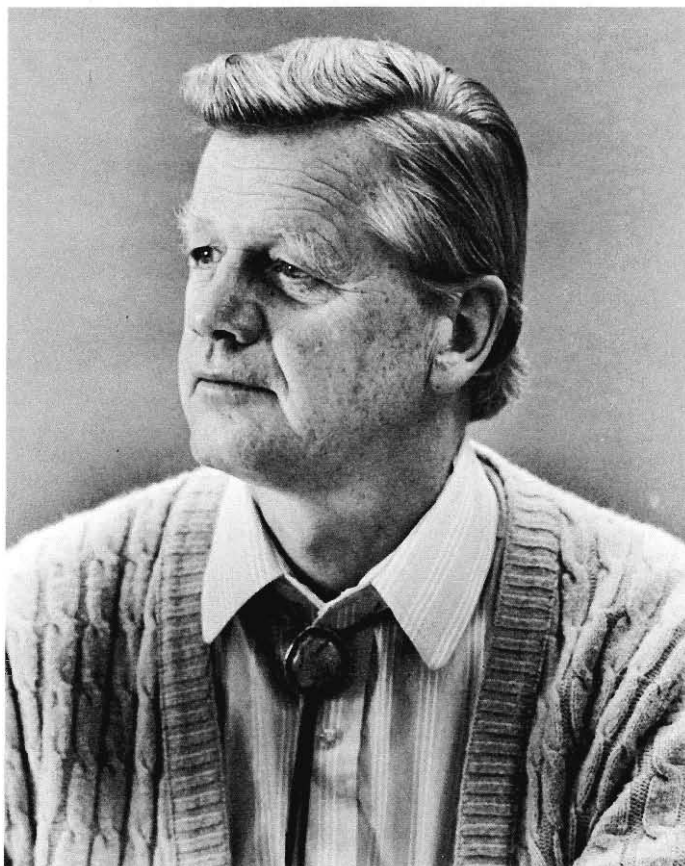
In concluding, I can do no better than to cite a poignant and timely question posed in a recent National Academy of Sciences report:

**Which represents the greater threat to peace? The dangers of proliferation associated with the replacement of fossil resources by nuclear energy, or the exacerbation of international competition for fossil fuels that could occur in the absence of an adequate worldwide nuclear-power program.**

Many hope, as I do, that this first year of the new decade will prove to be a milestone of significant progress toward worldwide implementation of effective, workable, and acceptable nuclear safeguards as an indispensable, vital contribution to safe, and safeguarded, nuclear energy for the benefit of mankind.



## THE AUTHOR



**G. Robert Keepin** joined the LASL staff (Critical Assemblies Group) in 1952 after being an Atomic Energy Commission Postdoctoral Fellow at the University of California, Berkeley, and a Consultant to Argonne National Laboratory and to LASL. He was a US Delegate to the First United Nations Atoms for Peace Conference in Geneva in 1955, and IAEA Technical Advisor to the Third Geneva Conference in 1964.

From 1963 to 1965 Keepin was with the Headquarters Staff of the International Atomic Energy Agency in Vienna. Following his return to the United States in 1965, he established the Nuclear Safeguards research and development program at LASL. In 1973 he received a Special Award for Nuclear Materials Safeguards Technology from the American Nuclear Society for his early recognition of the need for NDA instrumentation, his demonstration of practical passive and active assay methods, and his leadership in implementing these techniques and gaining wide acceptance for their use. He is now Program Manager at LASL for Nuclear Safeguards affairs.

Keepin is a Fellow of the American Physical Society and the American Nuclear Society and is National Chairman of the Institute of Nuclear Materials Management. He is widely published in the fields of nuclear and fission physics, reactor kinetics and control, and nuclear safeguards technology, and is an internationally recognized authority in the field of nuclear safeguards and nondestructive assay technology.

## BIBLIOGRAPHY

1. Committee on Nuclear and Alternative Energy Systems of the National Academy of Sciences, *Energy in Transition*, (W. H. Freeman, San Francisco, 1980).
2. Bertrand Goldschmidt, *Le Complexe Atomique*, (Fayard, Paris, 1980), now being translated into English.
3. "International Nuclear Fuel Cycle Evaluation," International Atomic Energy Agency, Vienna, February 1980.
4. G. Robert Keepin, "Safeguards Implementation in the Nuclear Fuel Cycle," *J. Instit. Nucl. Mat. Mgmt.* **VII**, 3 (1978).
5. USDOE-IAEA International Training Course on Nuclear Materials Safeguards (Lecture Text and Video Library) held in Santa Fe, New Mexico, May 27-June 6, 1980.

*Do these residues sent from Rockwell-Hanford at Richland, Washington contain salvagable plutonium? Members of LASL's Plutonium Facility determine the kinds and amounts of special nuclear materials in such heterogeneous waste with nondestructive assay techniques (the segmented gamma scanner and the thermal neutron coincidence counter). Measurements of this sample indicated that the contents were not worth recovering.*

# *Nondestructive Assay for Nuclear Safeguards*

by Roddy B. Walton and Howard O. Menlove

Sophisticated nondestructive assay techniques have solved many difficult measurement problems in the nuclear fuel cycle, but their potential for keeping track of sensitive nuclear materials on a nearly continual basis is just beginning to be exploited in nuclear facilities.

**R**adioactive nuclear materials reveal their presence by the radiation they produce spontaneously or otherwise. These characteristic signatures form the basis for nondestructive assay (NDA) of sensitive nuclear materials and for modern safeguards measurement technology. NDA techniques are used now by the nuclear industry, defense facilities, and safeguards inspectors to make rapid, accurate measurements of sensitive nuclear materials in diverse forms and compositions and thus to close many gaps in inventory measurements. However, this technology has not drastically changed the overall materials accounting practices of most nuclear facilities. Instead, the nuclear industry has placed heavy emphasis on physical protection to safeguard nuclear materials against overt threats of diversion. Further, a major part of the National Safeguards Program is devoted to increasing these

physical protection measures.

The Los Alamos Scientific Laboratory (LASL) program, on the other hand, has pioneered the development of NDA technology needed not only for inventory closure but also for near-real-time measurement and accountability systems. NDA techniques can be used to assure that no sensitive nuclear materials (SNM) are being lost or diverted from their defined flows and containments. Rapid measurements of SNM in feed materials, process lines, finished products, scrap and waste, and holdup in the plant are all possible with NDA techniques. When put together in an integrated materials control and accountability system, they can deter protracted diversion of SNM by a knowledgeable insider by detecting the amount and location of losses in a timely fashion. Such systems can also determine the validity of threats that significant amounts of nuclear material have

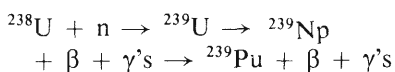
been diverted for unauthorized purposes.

At present several NDA-based near-real-time materials accounting systems are in existence and we are helping to plan such systems for future high throughput fuel production and spent-fuel reprocessing facilities. We also continue to work closely with members of the nuclear industry and safeguards inspectors to solve individual measurement problems associated with accurate inventory measurements, quality assurance of nuclear fuel production, and adequate safeguards inspection procedures. For many existing facilities, accounting systems that combine conventional chemical analysis with NDA techniques can provide adequate safeguards, but advanced facilities will require upgraded approaches.

Uranium and plutonium are the principal raw materials of fission weapons and hence the most sensitive nuclear materials to safeguard and to measure.



Moreover, because their fissile content or isotopic composition determines their strategic value as well as their value for nuclear reactor applications, it is important to measure isotopic compositions as well as absolute amounts of these elements. For example, natural uranium contains 0.7% of the fissile isotope  $^{235}\text{U}$ , with the remainder being the fertile isotope  $^{238}\text{U}$ . Both fissile and fertile isotopes will fission when bombarded with neutrons, but fissile isotopes are the important ingredient in nuclear fuels because they fission with high probability when bombarded by low-energy neutrons, whereas fertile isotopes will not fission unless they are bombarded by high-energy neutrons. Plutonium-239 is the principal fissile isotope in all plutonium fuels. This isotope is produced in nuclear reactors through the reaction



Uranium-238 is called fertile because it produces the fissile isotope  $^{239}\text{Pu}$  by neutron capture. Other fertile isotopes are  $^{240}\text{Pu}$ , which produces fissile  $^{241}\text{Pu}$  by neutron capture, and  $^{232}\text{Th}$ , which produces  $^{233}\text{U}$ . Thus safeguards measurements are concerned with determining the presence of a large number of fissile and fertile isotopes of uranium, plutonium, and thorium.

In this article we discuss nuclear NDA methods based on detecting either the natural radioactivity of these materials (passive assay) or the radiations produced when these materials are bombarded by external sources of neutrons and gamma rays (active assay). The methods are rapid, usually requiring only a few minutes to complete a measurement. They are nondestructive in the sense that the materials can be measured without removing them from their containers and that the measurement is based on observation of the

decay or transmutation of a negligible number of nuclei relative to one mole. Moreover, they interrogate the entire bulk of the material rather than just a sample and are therefore capable of accurately measuring heterogeneous materials such as scrap and wastes from processing.

One of the earliest, most critical needs for NDA methods was for the measurement of nuclear process scrap and waste. Chemical analysis was not reliable because of the problem and expense of obtaining representative samples from these characteristically heterogeneous materials. Some large inventory discrepancies (for example, the problem encountered at the NUMEC facility in Apollo, Pennsylvania, in 1965) have been difficult to resolve because of deficiencies in such measurements. Good measurements of scrap also were needed to assess the material's value for recycle, and NDA methods for the screening and assay of low-level waste were needed to dispose of the materials safely and economically.

NDA techniques were also required to extend the ability of safeguards inspectors to check facility operator compliance with safeguards requirements—in particular, to check that measured material inventories matched declared “book” inventories within the limits of measurement uncertainties. To accomplish this objective, inspectors carry out detailed auditing of records and procedures; perform on-the-spot inventories of containers of nuclear materials, which involve weighing and verifying seals; collect random samples for chemical analysis; and examine analytical laboratory procedures. With portable and in-plant NDA instrumentation, together with certified standards, an inspector can independently verify the types and amounts of materials in a facility, including materials held up in process lines.

International Atomic Energy Agency

(IAEA) inspectors present special challenges for measurement instrumentation because they work under difficult political and physical constraints. The terms of the safeguards agreements between the country being inspected and the IAEA define the political constraints. For example, their terms limit materials accessibility and inspection time. Further, the inspector must work in a foreign environment without the normal supporting services, such as calibration standards, that are available for domestic safeguards. Equipment therefore must be lightweight, reliable, rugged, and easy to calibrate.

The LASL Safeguards Program has studied and continues to study these and other challenging measurement problems. In most cases, the problems have been identified by domestic and IAEA safeguards inspectors, Department of Energy (DOE) field offices, DOE facility operators, DOE safeguards systems analysts, and designers of integrated safeguards systems for new fuel cycle facilities, as well as the operators of LASL facilities, principally CMB Division staff. Because we want to gain widespread acceptance of our new technology, we consider test and evaluation of a fully engineered prototype instrument in the operating environment of a host nuclear facility to be the most important phase of the development of a measurement method. We also prepare comprehensive design and performance documentation to facilitate the work of instrumentation vendors and potential users.

In the remainder of this article we describe (a) techniques for nondestructive assay of fissionable material including some applications of major methods, (b) the status of NDA technology development and its implementation in the fuel cycle and DOE facilities, and (c) some future challenges in this field of measurement technology. Many laboratories over the world have



**Fig. 1. Examples of containers used for nuclear materials in fuel cycle facilities. The large containers are used for low-concentration scrap and waste, the intermediate-size containers (1- and 4-L) hold high-purity process fuel materials or high-concentration scrap, and the small vial is a typical container for samples withdrawn from process lines for chemical analysis.**

contributed significantly to the development of this technology, but rather than present a comprehensive review, we will focus on LASL-developed instrumentation and methods.

### Characteristics of a Nondestructive Assay System

No single NDA instrument or method will suffice for the assay of the diverse forms and containments of nuclear materials in the fuel cycle. Some typical samples are shown in Fig. 1. Typical "feed" materials for input to the fuel fabrication process are  $UF_6$  in metal cylinders with capacities for up to  $\sim 10^4$  kg, depending on enrichment; plutonium nitrate solution in 10-L plastic bottles (120 cm long); and uniform fuel blends in few-liter cans. Reactor fuel materials include pellets, plates, rods, and bundles. Recoverable scrap, such as defective product material and calcined process residues, is stored in few-liter containers. Slightly contaminated wastes, such as paper, rags, and rubber gloves, are often stored in 120- and 220-L drums.

Clearly then, the capability to "view" bulk materials quantitatively is essential for most applications; hence, for both active and passive assay the interrogating beams and radiation signatures of the nuclear materials being measured must be highly penetrating. Neutrons and gamma rays are obvious choices from the domain of low-energy nuclear physics, as opposed to less penetrating alpha and beta rays. Additionally, calorimetry, which is based on the measurement of heat generated by radioactivity, can be used to measure plutonium with high precision, provided the isotopic composition of the plutonium is known from independent mass spectrometry or gamma-ray spectrometry measurements. (Uranium does not generate enough heat to be measured with calorimetry.)

The characteristics that are common to all practical NDA systems are:

1. an assay principle based on the detection of suitably copious fundamental signatures of the isotopes or elements to be measured, that is, neutrons or

gamma rays from natural radioactivity or from interrogation with an external source;

2. a detection system optimized for the signature selected and the types of samples to be measured;

3. electronic and mechanical systems for data acquisition and assay controls; and

4. a consistent means for obtaining the desired mass of the isotope or element from observed counting data.

Detectors for NDA instruments are essentially solids or gases that ionize as they interact with incoming neutrons or gamma rays. The amount of ionization is proportional to the energy deposited in the detector by each interaction. The generated charge forms an electronic pulse that is then amplified and processed and either counted or analyzed by electronic and microprocessor units. Gamma-ray detectors used in NDA instruments are solid crystals of NaI and germanium, standard tools of the gamma-ray spectroscopist. In NaI detectors, the ionization in the crystal produces light (scintillation), which is converted to an electronic pulse by a photomultiplier tube. In germanium and other solid-state detectors, the charge produced is converted directly into an electronic pulse. The neutron detectors most commonly used for NDA instruments are gas proportional counters, filled with  $^3He$ ,  $BF_3$ , or  $^4He$ ; however, plastic or liquid scintillators also can be used.

The conversion of counting-rate data to mass determinations or isotopic abundances of SNM involves the most subtle and difficult problems. Among these are calibration of the instruments and corrections for absorption and scattering of the signature radiation by the sample being measured. The sample includes not only the particular isotope or element of interest, but also "matrix" materials in which the material of in-

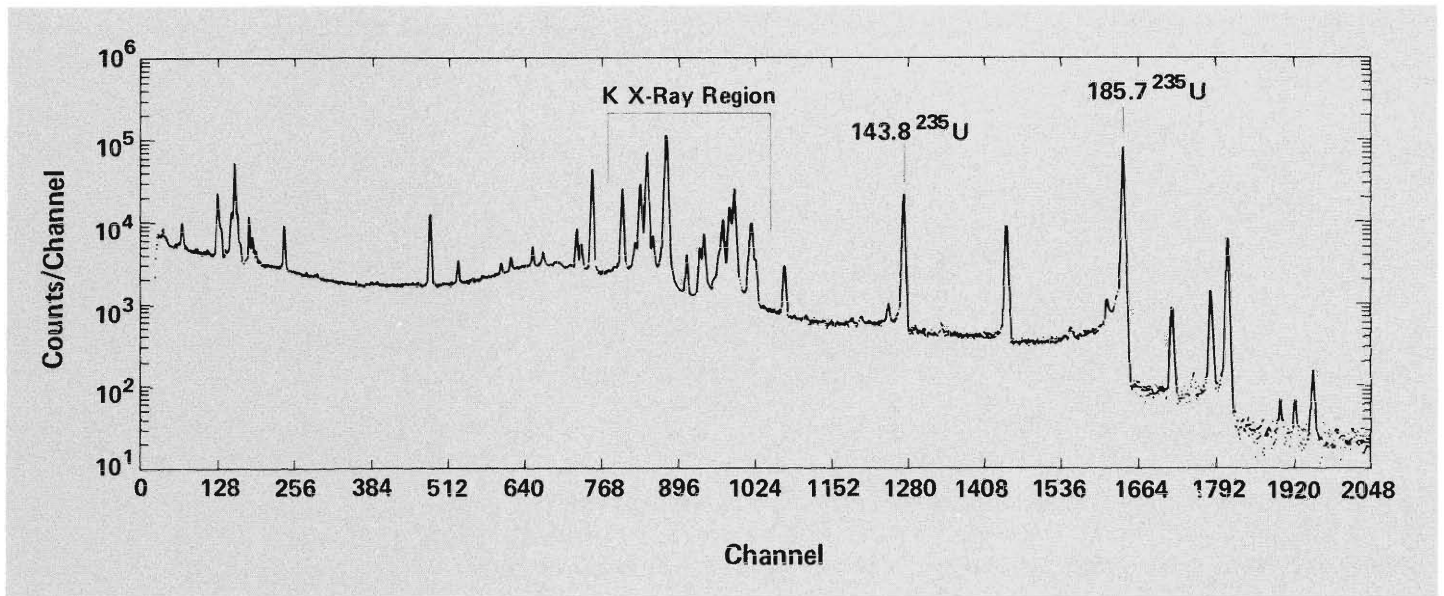


Fig. 2. Pulse-height distribution of gamma rays from high-enriched (93%  $^{235}\text{U}$ ) uranium, measured with a high-resolution germanium detector. Gamma-ray energies (in keV) are noted on some of the peaks. The prominent peak at 186 keV is used for passive assay of  $^{235}\text{U}$ .

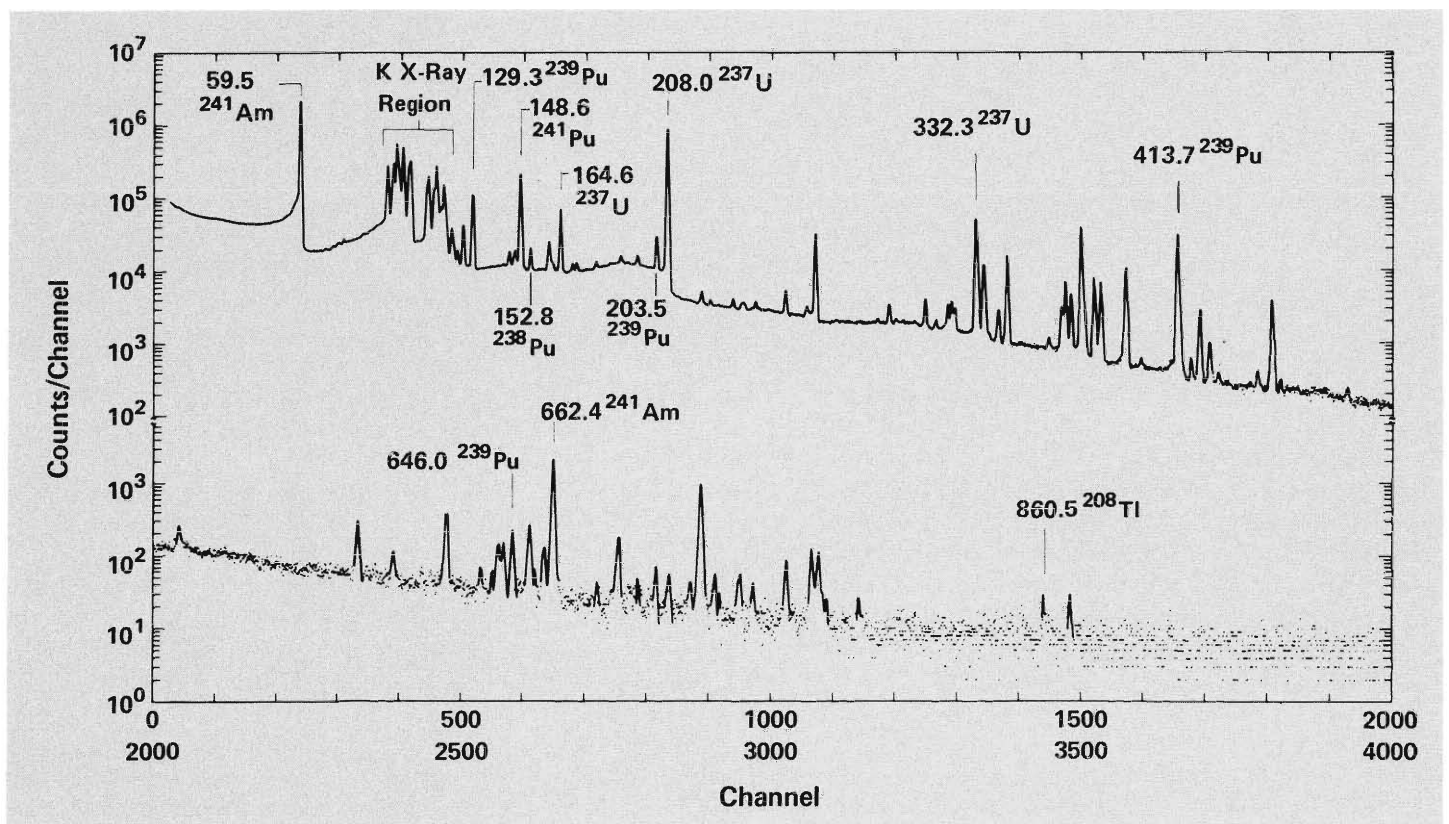


Fig. 3. Pulse-height distribution of reactor-grade plutonium (15%  $^{240}\text{Pu}$ ), measured with a high-resolution gamma-ray detector. Gamma-ray energies (in keV) are noted on some of the peaks.

terest is embedded. In an active method, one also has to account for the absorption and scattering of the interrogating beam of photons or neutrons as it penetrates the sample. If the chemical and isotopic compositions of samples are well known, as with identical cans of PuO<sub>2</sub> powder, then the correction problem may be circumvented simply by calibrating the measurements with a set of standards that closely cover the range of unknowns. However, we prefer a method whose response is as independent of sample matrix materials as possible, to reduce the number of calibration standards. It is even more important to reduce matrix dependence in measurements of poorly characterized scrap and waste materials. We use the sophisticated radiation transport codes developed over the years for defense projects and reactor design to guide the design of an assay system.

### General Methods for Passive Gamma-Ray Assay

Each isotope of SNM is radioactive and decays at a known rate through alpha or beta decay (for example,  $^{235}\text{U} \rightarrow ^{231}\text{Th} + \alpha + \gamma\text{'s}$  and  $^{239}\text{Pu} \rightarrow ^{235}\text{U} + \alpha + \gamma\text{'s}$ ). These spontaneous decays produce a characteristic spectrum of gamma rays, specific in both quantity and energy. Thus a direct measurement of the amount of SNM in a sample can be obtained by counting the gamma rays it emits at specific energies.

The pulse-height spectra of gamma rays from high-enriched (93%  $^{235}\text{U}$ ) uranium and from reactor-grade (15%  $^{240}\text{Pu}$ ) plutonium are shown in Figs. 2 and 3, respectively. The isotopic abundances of the reactor-grade plutonium are:  $^{238}\text{Pu} = 0.15\%$ ,  $^{239}\text{Pu} = 81.6\%$ ,  $^{240}\text{Pu} = 15.2\%$ ,  $^{241}\text{Pu} = 2.4\%$ ,  $^{242}\text{Pu} = 0.66\%$ , and  $^{241}\text{Am} = 0.8\%$ . These data were taken with a high-resolution germanium gamma-ray detector and a multichannel pulse-height analyzer. The

plutonium spectrum is very complex, with contributions from several hundred gamma rays from  $^{238}\text{Pu}$ ;  $^{239}\text{Pu}$ ;  $^{240}\text{Pu}$ ;  $^{241}\text{Pu}$ , and the daughters or products of  $^{241}\text{Pu}$  decay;  $^{237}\text{U}$ ; and  $^{241}\text{Am}$ . (In Fig. 3, a few of the peaks are labelled with their energies and isotopic origins.) Furthermore, the gamma-ray spectrum of plutonium changes with time because of the relatively short half-lives of  $^{241}\text{Pu}$  and  $^{237}\text{U}$ .

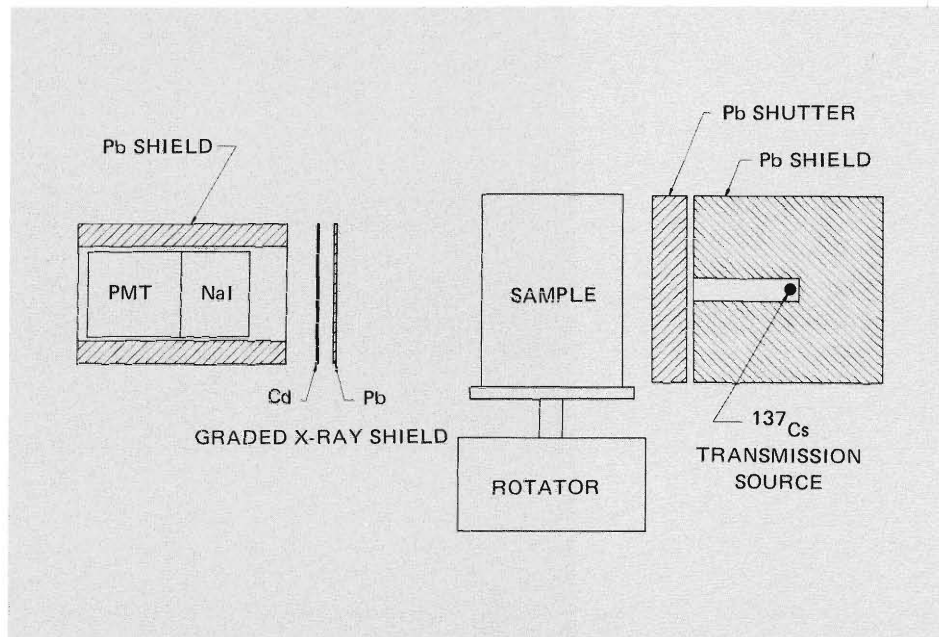
An assay is based on determining the areas under one or more peaks in the spectrum, usually calculated with a minicomputer that is an integral part of a multichannel-analyzer, data-acquisition system. We use either NaI or germanium gamma-ray detectors, but we prefer germanium because its higher (30X) resolution permits cleaner separation and more accurate analysis of the gamma-ray spectrum peaks. On the other hand, for field measurements, such as the assessment of material holdup in process equipment, portable assay units comprising a NaI detector and single-channel electronic analyzers set to bracket specific peak and background regions of the gamma-ray spectrum are satisfactory.

For measurement of uranium, we use the prominent 186-keV gamma ray from  $^{235}\text{U}$ ;  $4 \times 10^4$  gamma rays are emitted per gram  $^{235}\text{U}$  per second. Similarly we use the 414-keV gamma ray from  $^{239}\text{Pu}$ , which has a comparable intensity, to measure plutonium. To convert counting rates of these isotopes to total uranium or plutonium, their isotopic abundances must be known or measured independently. In all applications except spent-fuel recovery, uranium isotopes other than  $^{235}\text{U}$  and  $^{238}\text{U}$  are present in such small amounts that they can be neglected. Uranium-238 can be measured using as the signature the 1.001-MeV gamma ray from the decay of its daughter  $^{234\text{m}}\text{Pa}$ . However, the measurement must be made at least 3 months after chemical purification of

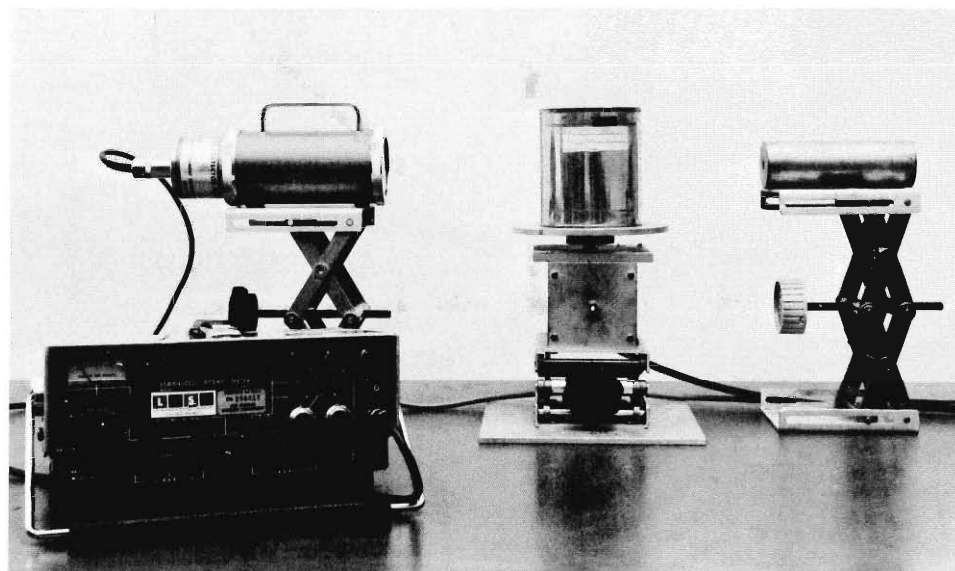
uranium so that the daughters are in equilibrium. The isotopic abundances of plutonium are much more complex, but still they can be measured by performing a detailed analysis of gamma-ray pulse-height spectra (described below).

The simplest gamma-ray arrangement, a portable system that has found extensive application in plant surveys, is shown in Figs. 4 and 5. The assay system consists of a NaI detector housed in a lead collimator, pulse-processing electronics (contained in the small box in Fig. 5), a sample turntable, and an external gamma-ray source for determining the gamma-ray attenuation of the sample. An assay is performed by collecting counting-rate data with the unknown sample, measuring the transmission of an external beam of gamma rays through the sample, and applying a correction for the sample self-attenuation of the SNM gamma rays derived from the transmission measurement. We calibrate the system by performing measurements of known standards in the same manner.

The central problem in the NDA of bulk samples is the correction for sample self-attenuation. That is, the emitted gamma rays are scattered and absorbed within the sample itself. Attenuation is large and difficult to anticipate because the gamma rays have low energies, typically 100 to 400 keV; the samples frequently contain high-Z elements that absorb strongly; and the chemical composition of the sample is often unknown. The attenuation of gamma rays in a sample is given by  $e^{-\mu\rho X}$  where  $\rho$  ( $\text{g}/\text{cm}^3$ ) is the density of the medium,  $\mu$  ( $\text{cm}^2/\text{g}$ ) is the mass attenuation coefficient, and  $X$  ( $\text{cm}$ ) is the thickness of the sample. The gamma-ray attenuation also can be written as  $e^{-\mu_t X}$ , where  $\mu_t$  is the linear absorption coefficient. Table I gives the mean free path,  $1/\mu_t$ , of the  $^{235}\text{U}$  185.7-keV and the  $^{239}\text{Pu}$  413.7-keV gamma rays for several materials. Note that high-Z materials cause a dramatic



**Fig. 4.** A passive gamma-ray assay measurement setup. The sample is placed on a rotating table, and the detector (at the left) counts the gamma radiation from the sample. Correction for the attenuation of the SNM gamma rays by the sample is determined by measurement of transmitted gamma rays through the sample from an external radioactive source (in this case,  $^{137}\text{Cs}$ ) shown at the right.

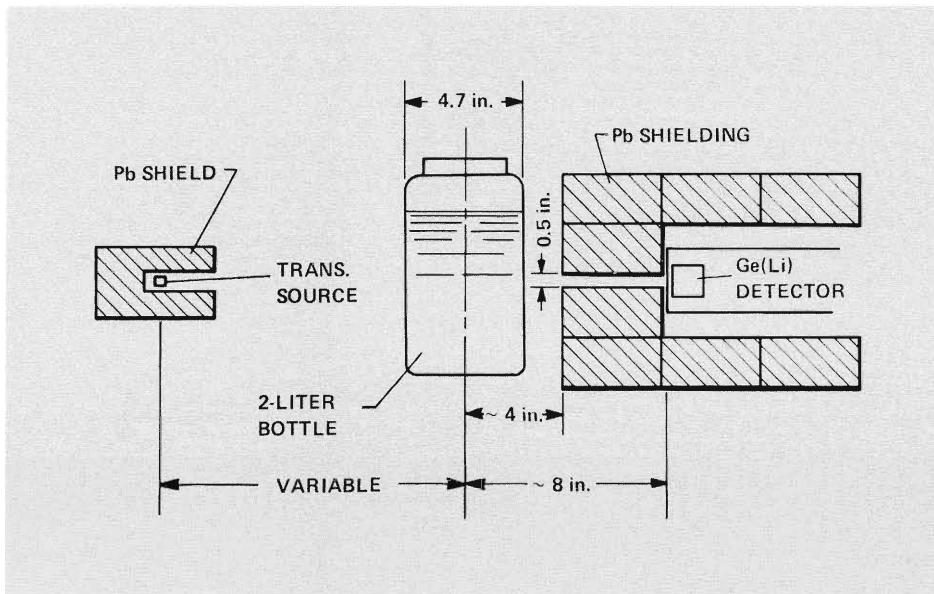


**Fig. 5.** A simple passive gamma-ray assay setup. At the left is the gamma-ray detector, a NaI crystal attached to a photomultiplier tube. The sample is mounted on a rotating table, and the transmission source is positioned to the right of the sample. (The arrangement is the same as that shown in Fig. 4.) The pulse processing electronics is an Eberline SAM-2 portable instrument located below the detector.



**TABLE I**  
**GAMMA-RAY MEAN FREE PATH**

Medium	186 keV ( $^{235}\text{U}$ ) (cm)	414 keV ( $^{239}\text{Pu}$ ) (cm)
H <sub>2</sub> O	7.1	9.7
Al	3.0	4.1
Fe	0.84	1.4
Pb	0.071	0.44
U	0.033	0.21



**Fig. 6.** The segmented gamma-ray scanning system used for passive gamma-ray measurement of heterogeneous uranium and plutonium scrap and waste. The sample is rotated and assayed by segments as it is translated vertically past the collimated germanium detector shown on the right. Transmissions of gamma rays from the source on the left are used to derive self-attenuation corrections for each assay segment.

decrease in the gamma-ray mean free path, especially for the lower energy gamma ray. Even small particles ( $\sim 0.02$  mm) of metallic uranium or plutonium are highly self-absorbing.

To determine the self-attenuation correction, most gamma-ray assays employ a separate transmission measurement of the linear absorption coefficient  $\mu_t$  of the sample. The external gamma-ray source must supply gamma rays with energies close to the energy of the signature gamma ray of interest. Once we know  $\mu_t$ , the dimensions of the sample, and its distance from the detector, we can calculate the self-absorption correction provided the mixture of the material to be assayed and the matrix

are reasonably uniform and the particles of assay material are small enough to ignore self-attenuation within individual emitting particles. Few closed forms exist for the correction factors, but often semiempirical analytical forms are sufficiently accurate. J. L. Parker and T. D. Reilly have developed analytical forms for self-attenuation corrections for most practical passive gamma-ray assay problems.

### Segmented Gamma Scanner

The simple procedures described above are inadequate to measure containers of scrap and waste because they cannot take into account the vertical

variations in SNM and matrix densities characteristic of these containers. Radial inhomogeneities are usually less pronounced, and their effects are substantially reduced by sample rotation. We developed the Segmented Gamma Scanner (SGS), which is both an instrument and a procedure, to improve the assay accuracy for a wide range of scrap and waste. It was also the first fully automated NDA instrument for Safeguards and is now used widely in fuel processing facilities.

The basic principle of the SGS is to divide the sample into thin horizontal segments and assay each segment independently using the transmission-corrected passive assay technique described above. After all the segments have been measured, the results are summed to give the total assay for the container. The method is shown in Fig. 6. A germanium detector views a segment of the container through an open slit, or collimator, in a lead shield. The transmission source is on the opposite side of the container, in line with the detector. For  $^{235}\text{U}$  assay, the transmission source is  $^{169}\text{Yb}$ . It emits 177.2-keV and 198.0-keV gamma rays, which closely bracket the 185.7-keV assay gamma ray of  $^{235}\text{U}$ . The  $^{75}\text{Se}$  400.6-keV gamma rays serve for transmission measurements in  $^{239}\text{Pu}$  assays based on detection of 414-keV gamma rays.

Computer control of SGS instruments allows automatic data acquisition, analysis, and management of all hardware. An operator places a container on the sample table, and the SGS does the rest. The SGS begins the assay sequence by positioning the sample table so that the top of the sample is just below the detector axis. Automatic controls rotate the sample continually and elevate it in discrete steps until all segments are assayed. Vertical profiles of individual segment assays and transmissions are available as data output, as well as the total SNM in the container

(Fig. 7). Typically a complete SGS assay requires 3 to 5 minutes, and accuracies lie in the 1-5% range for liter-size samples.

SGS instruments are now commercially available and are used in most major fuel cycle facilities to measure scrap and waste. Figure 8 shows an SGS instrument for 220-L and other large containers of low-level waste. We also have designed smaller units to assay containers less than 20 L.

### The Enrichment Meter Principle

The end products of uranium fuel production are high-concentration, homogeneous forms of enriched uranium, such as metal, uranium oxide in powder and ceramic fuel forms, and rich uranium scraps. We have developed a very simple method to measure the  $^{235}\text{U}$  enrichment in these materials, which is widely used in nuclear facilities not only to safeguard nuclear fuels but also to assure their quality. Called the *enrichment meter* technique, it uses to advantage the severe attenuation of the  $^{235}\text{U}$  186-keV gamma ray by high-Z materials. For homogeneous samples whose thickness is essentially infinite relative to the mean free path for this gamma ray, the intensity of the 186-keV gamma ray emitted will be independent of the sample thickness because only the gamma rays emitted near the surface will reach the detector. With reference to the sample-detector geometry in Fig. 9, performing the integration over the sample thickness indicated in the figure shows that for a metallic uranium sample, the counting rate (CR) of 186-keV gamma rays is given by the following expression.

$$CR = (\epsilon IA^{235}\text{U}/\mu_1) \propto (^{235}\text{U}/\mu_1),$$

where

- $\epsilon$  = detection efficiency,
- $I$  = intensity of 186-keV gamma ray ( $\gamma/\text{s per g } ^{235}\text{U}$ ),
- $\underline{^{235}\text{U}}$  =  $^{235}\text{U}$  mass fraction (enrichment),

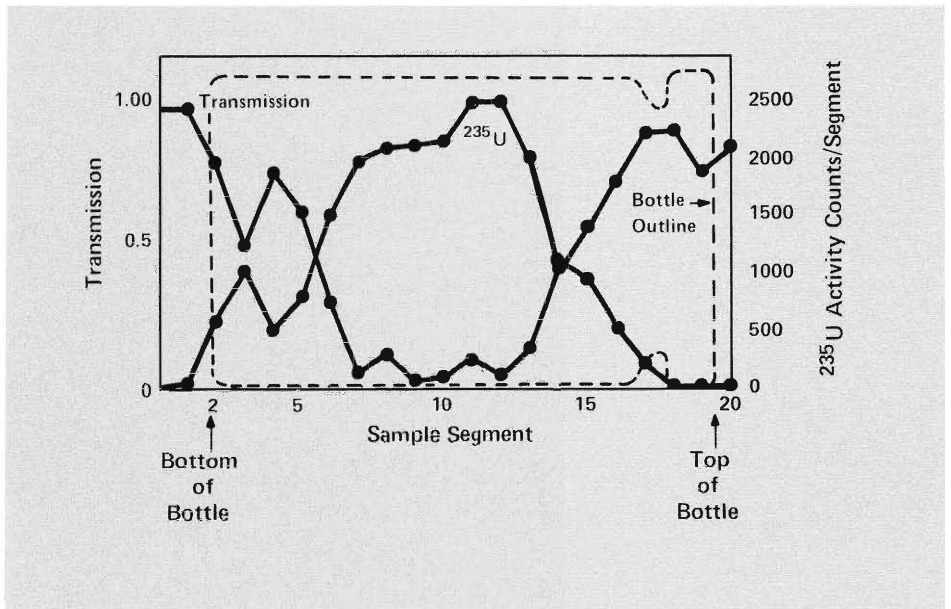


Fig. 7. Sample of data produced by segmented gamma scanner. The transmission of external rays (left-hand ordinate) and  $^{235}\text{U}$  186-keV gamma-ray activity (right-hand ordinate) are shown as a function of vertical segment of a 2-L bottle of high-enriched uranium. Note the profile of the bottle (lying on its side). The mirror effect in the transmission and activity data indicates that the sample attenuation is caused principally by its uranium content.

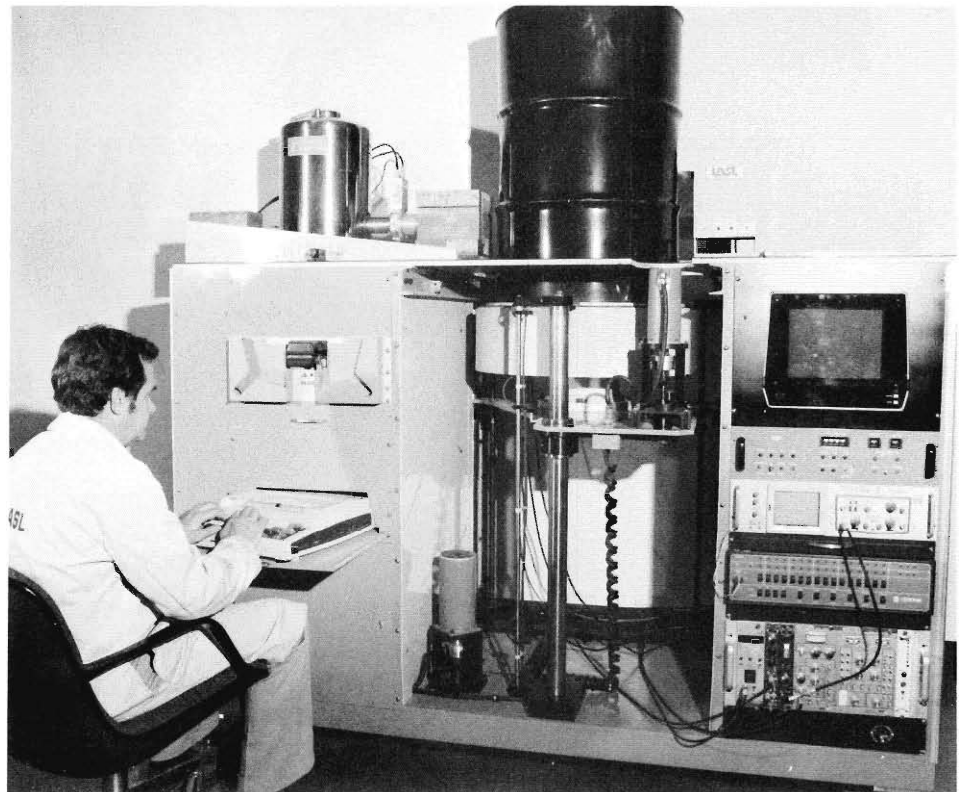


Fig. 8. An automated segmented gamma scanner for assay of uranium and plutonium in large containers. The germanium detector is located on the left of the 55-gallon drum sample, and the transmission source is on the right.

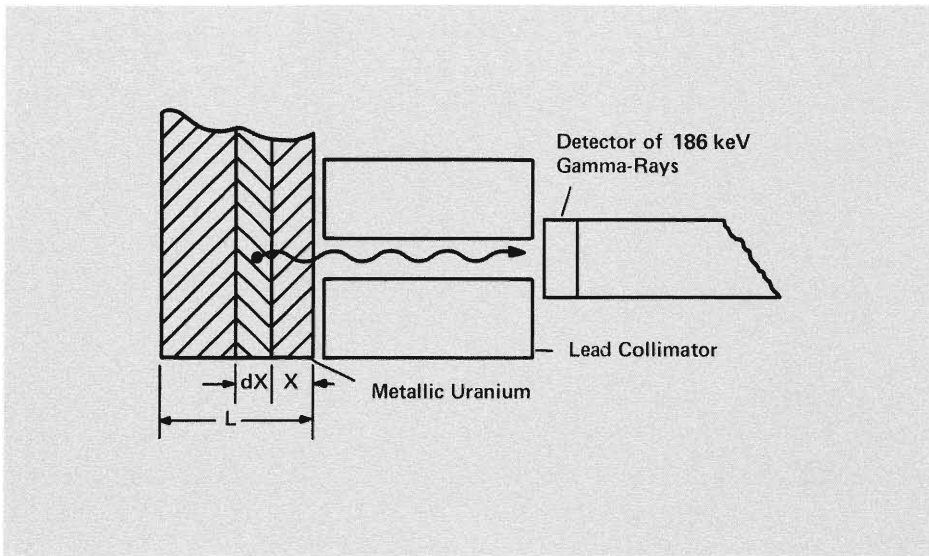


Fig. 9. The measurement geometry used for the enrichment meter principle. A measurement of 186-keV  $^{235}\text{U}$  gamma rays from a thick high-concentration uranium sample can yield the  $^{235}\text{U}$  enrichment directly.

$A$  = area of sample ( $\text{cm}^2$ ) defined by collimator, and  
 $\mu_1$  = mass attenuation coefficient of uranium for 186-keV gammas ( $\text{cm}^2/\text{g}$ ).

The counting rate is directly proportional to the  $^{235}\text{U}$  mass fraction, or enrichment, and does not depend on sample density or other factors as long as a fixed detector and collimation geometry is maintained for calibration and measurements.

For uranium compounds such as  $\text{UO}_2$  or  $\text{UF}_6$ , the expression for CR is

$$\text{CR} \propto ({}^{235}\text{U}/\mu_1) F,$$

where

$$F = [1 + (\mu_2\rho_2/\mu_1\rho_1)]^{-1},$$



Fig. 10. Field measurement of  $^{235}\text{U}$  enrichment of  $\text{UF}_6$  in a 2-1/2-ton shipping container. A portable NaI detector is used to measure 186-keV  $^{235}\text{U}$  gamma rays. A portable ultrasonic gauge is used to determine the cylinder wall thickness for the attenuation correction.

$\mu_2$  = mass attenuation coefficient of the low-Z element, and  
 $\rho_1, \rho_2$  = the density (g/cm<sup>3</sup>) of uranium and the low-Z element, respectively.

For matrix materials with  $Z < 30$ ,  $\mu_2/\mu_1 \leq 0.11$  and, in many practical cases,  $F$  is nearly equal to unity, being 0.988 for UO<sub>2</sub> and 0.930 for residues containing only 50 wt% uranium. Note also that  $F$  depends on concentration ratios instead of absolute quantities.

The enrichment meter principle may be used with a NaI or germanium detector; the latter gives more accurate results. Figure 10 shows a field measurement of the enrichment of UF<sub>6</sub> in a standard 2½-ton cylinder used to ship this material to light water reactor (LWR) fuel fabricators.

### Plutonium Isotopic Analysis

The isotopic abundances of plutonium are needed for safeguards and accounting, particularly to distinguish weapons-grade from reactor-grade material and to assure the quality of product fuel. Furthermore, all NDA methods developed so far for the quantitative assay of plutonium in bulk materials depend on a prior knowledge of the isotopic abundances of the isotopes whose signatures are to be used for the measurement technique, that is, <sup>239</sup>Pu for gamma-ray assay based on the 414-keV gamma ray, the fissile isotopes <sup>239</sup>Pu and <sup>241</sup>Pu for active neutron methods, and all plutonium isotopes and <sup>241</sup>Am for calorimetry. In many instances, particularly for scrap and waste materials and recycle streams, the isotopic abundances of plutonium are not specified reliably.

To meet the need for a rapid non-destructive isotopic analysis, we are developing and adapting gamma-ray spectrometry. The methods involve analysis of pulse-height spectra of

plutonium gamma rays measured with high-resolution germanium detectors. Peak areas are obtained by using either a simple channel summation procedure with a straight-line background subtraction or sophisticated peak-fitting algorithms. As evident from Fig. 3, the plutonium isotopes <sup>238</sup>Pu, <sup>239</sup>Pu, and <sup>241</sup>Pu, as well as the <sup>241</sup>Pu daughters <sup>241</sup>Am and <sup>237</sup>U, emit gamma rays that are useful for isotopic analysis. Unfortunately, the intensities of the gamma rays from <sup>240</sup>Pu are weak and often obscured by gamma rays from other isotopes. Plutonium-242 does not have a useful gamma ray, but its isotopic abundance is small for low-burnup plutonium. For high-burnup fuels such as LWR spent fuels, it can be estimated from correlations of other plutonium isotopes or inferred from results of an independent measurement of elemental plutonium concentration combined with gamma-ray spectrometry data for the other isotopes.

Samples of controlled and constant form and known chemical composition are the easiest for isotopic analysis. For example, if we prepare "thin" samples of low-concentration solutions in precision vials, we can obtain quantitative abundances of individual plutonium isotopes by comparing gamma-ray peak areas from standard solutions and applying self-attenuation corrections.

We have also developed a relatively simple method for the isotopic analysis of plutonium in samples of arbitrary shape and composition. The procedure requires no detailed peak fitting; thus it is fast, and computer cost and speed requirements are minimal. This method can be programmed into portable multiple-channel analyzers for use in field inspections. We measure areas under the peaks of closely spaced gamma-ray lines in the region from 120 to 414 keV to determine the following isotopic ratios.

<sup>238</sup> Pu/ <sup>241</sup> Pu	152.7 keV/148.6 keV
<sup>239</sup> Pu/ <sup>241</sup> Pu	203.5 keV/208.0 keV or 345 keV/332.3 keV
<sup>240</sup> Pu/ <sup>241</sup> Pu	160.3 keV/164.6 keV
<sup>241</sup> Pu/ <sup>241</sup> Pu	332.3 keV/335.4 keV
<sup>242</sup> Am/ <sup>239</sup> Pu	125.0 keV/129.3 keV

In some energy regions, the measured lines are not clean peaks; that is, they must be corrected for contributions from neighboring lines of other isotopes. For example, the measured <sup>240</sup>Pu peak at 160.3 keV includes contributions from the <sup>241</sup>Pu line at 160.0 keV and from the <sup>239</sup>Pu line at 160.2 keV. By measuring the clean peak at 164.6 keV from the <sup>241</sup>Pu daughter <sup>237</sup>U, using the known intensity branching ratios for the 160.0- and 164.6-keV lines, and applying small corrections for changes in relative efficiencies, we can determine the 160.0-keV contribution and subtract it from the measured peak; similarly the <sup>239</sup>Pu peak at 161.5 keV is used to obtain the <sup>239</sup>Pu contribution to the 160-keV complex. The corrected peak is then divided by the <sup>241</sup>Pu (<sup>237</sup>U daughter) peak at 164.6 keV to give the <sup>240</sup>Pu/<sup>241</sup>Pu isotopic ratio.

For each sample, the relative detection efficiency is derived empirically from the areas of isolated, clean peaks of <sup>239</sup>Pu and <sup>241</sup>Pu in different regions of the spectrum, adjusted for their respective gamma-branching intensities. The isotopic ratios are then combined to yield the actual plutonium mass fractions using the constraint that their sum is unity. Results can be obtained by this method in 1 hour and, in some cases, with accuracies comparable to those obtained from mass spectrometry.

### Active Gamma-Ray Assay

We are adapting two active gamma-ray assay methods for measuring fissionable materials in plant environments: x-ray edge densitometry and x-ray fluorescence. These methods com-

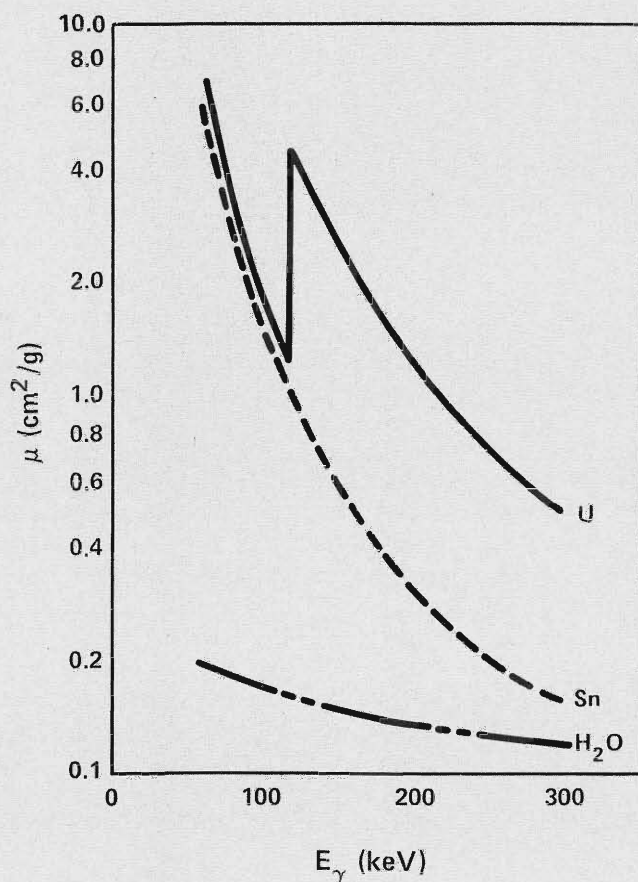


Fig. 11. Gamma-ray mass attenuation coefficient versus energy. The sharp K-edge discontinuity in the coefficient of uranium is contrasted with the monotonically varying coefficients of lower-Z materials.

TABLE II

ABSORPTION EDGES AND  $\Delta\mu$  VALUES FOR URANIUM AND PLUTONIUM

Absorption Edge	Element	Edge Energy (keV)	$\Delta\mu$ (cm <sup>2</sup> /g)
L <sub>III</sub>	U	17.17	56.6
	Pu	18.05	51.9
K	U	115.6	3.65
	Pu	121.76	3.39

plement passive gamma-ray assays in that they measure total elemental content rather than specific isotopes. The sample is bombarded with an external beam of photons or particles at energies tailored to induce specific atomic interactions characteristic of the element, such as removal of an electron from an inner atomic shell. Traditionally, analytical chemistry laboratories have implemented x-ray edge densitometry and x-ray fluorescence with x-ray generators and high-resolution wavelength dispersive techniques. Our adaptation uses high-resolution gamma-ray detectors to measure uranium and plutonium concentrations in solutions. The measurements can be made either in the process line itself (in-line) or by removing a sample and making the measurement alongside the process line (at-line). These techniques can be very useful in reprocessing plants to assay low-concentration solutions containing both uranium and plutonium.

We will limit this discussion to x-ray edge densitometry. The technique utilizes the fact that gamma rays with energies just above the threshold for removing an electron from the K or L<sub>III</sub> atomic shells of uranium and plutonium are absorbed with much greater probability than lower energy gamma rays. These thresholds create jumps in the mass attenuation coefficient, known as K and L<sub>III</sub> edges. By bombarding a sample with gamma rays just above and below the K or L<sub>III</sub> edges, and measuring the difference in transmissions, we can measure the amount of plutonium or uranium in the sample.

The K edge for uranium is shown in Fig. 11, together with the smoothly varying mass attenuation coefficients of lower-Z elements in the same energy region. Table II presents the positions of the energies of the K and L<sub>III</sub> edges of uranium and plutonium and the magnitudes of the corresponding mass attenuation discontinuities,  $\Delta\mu$ .

The basic experimental arrangement for densitometry is shown in Fig. 12. Narrow-beam collimators eliminate the detection of multiple-scattered gamma rays and thereby minimize matrix effects. The transmission source is either an x-ray generator or appropriate radioactive isotopes. X-ray generators have the advantages of variable intensity and a continuous bremsstrahlung spectrum that can be tailored to cover the absorption edge, but radioactive sources are simpler and less expensive. High-resolution germanium detectors normally are used for K-edge measurements, whereas Se(Li) and germanium detectors must be used for L-edge measurements.

The transmissions  $T_1$  and  $T_2$  just below and above the absorption edge are related to the total concentration  $\rho$  of the element of interest by the relation

$$T_i = T_{m_i} e^{-\mu_i \rho X},$$

where

$\mu_i = \mu_1$  or  $\mu_2$ , the mass attenuation coefficient of the element below and above its edges, respectively;

$T_{m_i}$  = corresponding transmissions of the matrix (all other materials); and

$X$  = sample thickness.

If the energies of the transmission gamma rays very closely bracket the edge, then  $T_{m_1} \cong T_{m_2}$ , and the concentration of the subject element is given by

$$\rho = \frac{-\ln T_2 / T_1}{\Delta \mu X}; \Delta \mu = \mu_2 - \mu_1.$$

Corrections for effects of matrix materials may be applied when the transmission gamma rays are not so closely spaced.

X-ray edge densitometry measurements for multiple simultaneous elemental determinations are shown in Fig. 13.

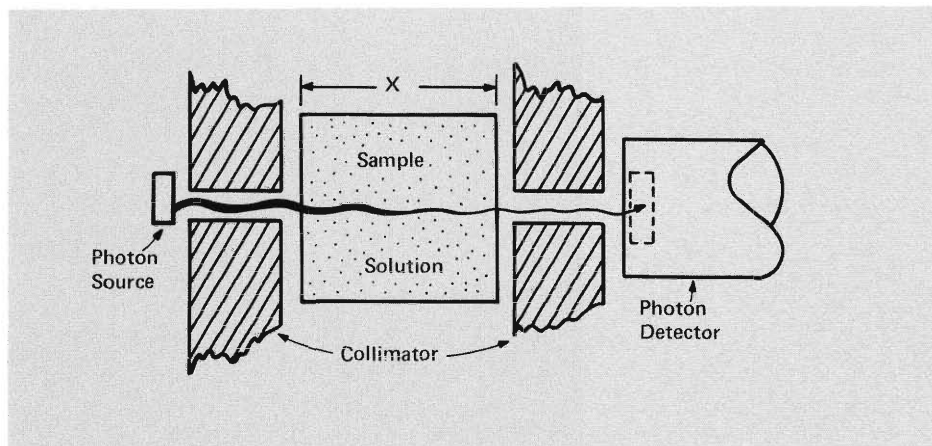


Fig. 12. Typical experimental setup for x-ray absorption-edge measurement of uranium and plutonium solution concentrations. The gamma-ray source is collimated, and the detector measures the intensity of gamma rays transmitted through the sample (thickness  $X$ ) at energies above and below the characteristic x-ray absorption edges of the elements of interest.

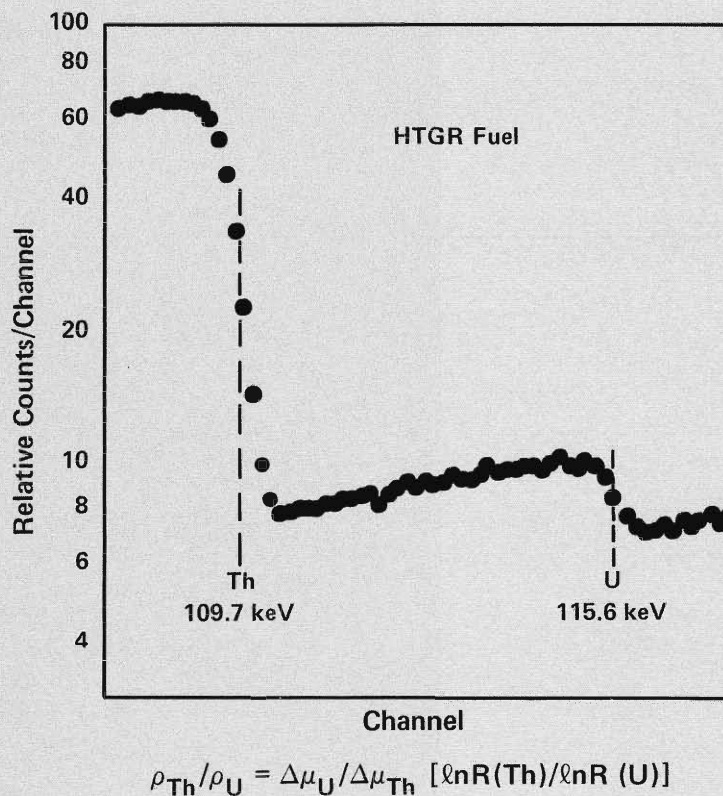
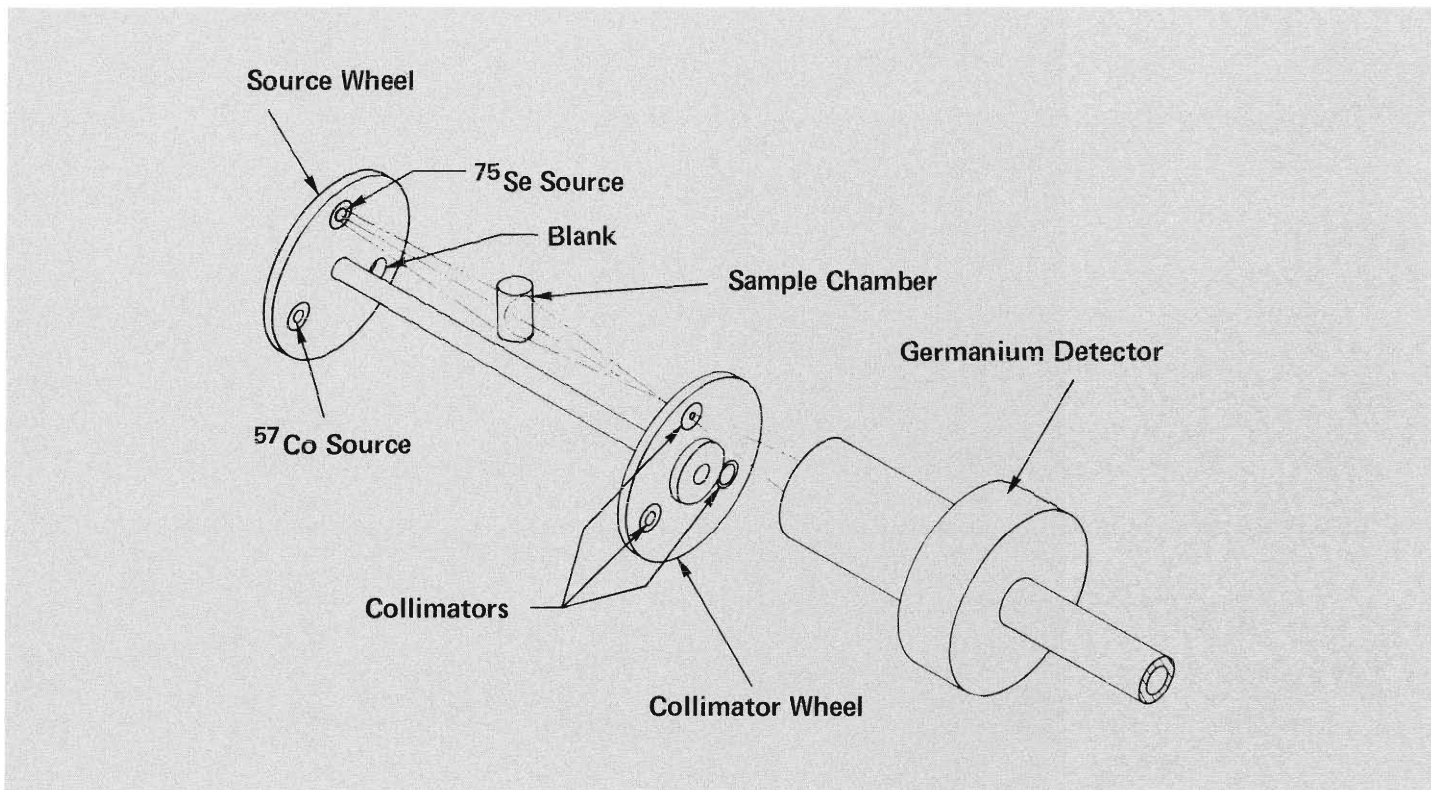


Fig. 13. Spectrum of x rays transmitted through a sample of high-temperature gas-cooled reactor fuel. The K edge of thorium at 109.7 keV and uranium at 115.6 keV are shown. Elemental concentrations are derived from the observed discontinuities at those edges.



**Fig. 14. Conceptual view of measurement configuration of the Tokai-Mura densitometer. The plutonium solution sample is shown between the source ( $^{75}\text{Se}$ , in this case) and the collimator in the detector line of sight. Sources and collimators are mounted on separate wheels driven synchronously.**

The x-ray spectrum is transmitted through a sample of high temperature gas-cooled reactor (HTGR) fuel, which contains thorium and uranium in a graphite matrix. The discontinuities produced by the K-absorption edges of both uranium and thorium are clearly evident, allowing accurate determination of the respective densities of thorium and uranium, in this case in proportion 4:1.

To resolve low-concentration solutions of uranium and plutonium, we must use the  $L_{III}$  edge with its larger  $\Delta\mu$  values rather than the K edge. The  $L_{III}$  edge is useful for solution concentrations ranging from 2 to 100 g/L. However, an x-ray generator is required for the transmission source because no adequate radioactive sources are available. We have developed an x-ray generator-based absorption-edge densitometer for the simultaneous measurement of uranium and plutonium in solutions con-

taining both elements. Some future reprocessing and nitrate-to-oxide conversion plants will coprocess uranium and plutonium in such solutions. We have conducted a preliminary test of this instrument, operating in-line, at the Savannah River Laboratory coprocessing demonstration facility, as well as at LASL. The instrument performed well for mixed solutions with plutonium concentrations between 2 and 10 g/L and uranium concentrations 4 to 10 times higher and is thus a candidate for use in advanced, near-real-time accounting systems for fuel reprocessing and conversion facilities.

For K-edge measurements of higher concentration plutonium solutions, we can use two discrete gamma-ray transmission sources, the 121.1- and 122.2-keV gamma rays from  $^{75}\text{Se}$  and  $^{57}\text{Co}$ , respectively. These gamma-ray energies very closely bracket the 121.8-keV electron-binding energy of the K

shell of plutonium.

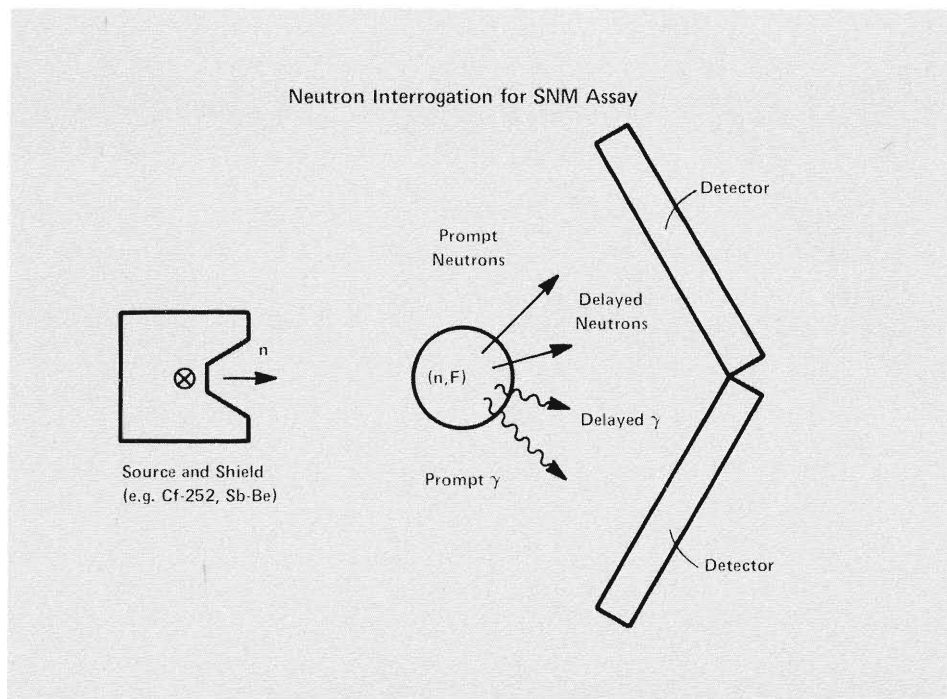
We have recently developed such a densitometer to measure 150-350 g/L plutonium product solutions of the spent-fuel reprocessing in Tokai-mura, Japan. The instrument was installed and calibrated in this facility in the latter part of 1979 and is being evaluated for cooperative use by the facility operator and the IAEA for rapid off-line plutonium analysis. The Tokai densitometer rotates transmission sources in and out of position (Fig. 14) to measure the x-ray edge for plutonium concentration. This instrument also measures isotopic abundances by passive gamma-ray assay. Results obtained with the Tokai densitometer indicate that it will measure in times less than 1 hour plutonium concentrations to an accuracy of  $\pm 0.3\%$  (1 standard deviation) or better, and isotopic abundances in the accuracy range 0.3-1.5%, depending on the isotope and its abundance.

## Passive Neutron Assay

Passive neutron assays detect the neutrons emitted when an isotope undergoes spontaneous fission. This is perhaps the easiest assay to perform. It is useful for measuring bulk materials because neutrons penetrate much farther than gamma rays. In practice,  $^{240}\text{Pu}$  and  $^{238}\text{U}$  are the only isotopes having sufficient spontaneous fission rates and isotopic abundances for passive neutron assay.

The assay is complicated by the fact that alpha particles from radioactive decays of uranium and plutonium interact with light elements in the matrix materials to produce additional neutrons. For example, if  $\text{PuF}_4$  is present in the sample, neutrons will be produced through the reaction  $^{19}\text{F} + \alpha \rightarrow ^{22}\text{Na} + \text{n}$ . However, each  $(\alpha, \text{n})$  reaction produces only one neutron in contrast to a spontaneous fission that produces a group of two or three. Consequently, we can differentiate spontaneous fission neutrons from the others by neutron coincidence counting. We have developed special time-correlation counting circuitry, called the shift register, to count coincidence neutrons and thus the spontaneous fission rate.

Neutron coincidence counting of  $^{240}\text{Pu}$  spontaneous fission is the simplest method for the assay of plutonium provided the plutonium isotopic composition is known ahead of time. For high-burnup plutonium reactor fuels, a coincidence count may also contain a significant contribution from  $^{242}\text{Pu}$  spontaneous fission. Since the spontaneous fission rate for  $^{240}\text{Pu}$  is low, approximately 460 fissions/s per g  $^{240}\text{Pu}$ , and the  $^{240}\text{Pu}$  isotopic abundance is typically in the range 8-20%, a high-efficiency detection system is essential for rapid, quantitative assay. For uranium fuels,  $^{238}\text{U}$  can be assayed by counting its spontaneous fissions, but the specific rate is much lower (approximately  $7 \times$



*Fig. 15. The basic components of an active neutron assay system showing the interrogation source, the sample, and the detectors.*

$10^{-3}$  fissions/s per g  $^{238}\text{U}$ ) and kilogram quantities are required. We have developed cylindrical detectors (with a central sample cavity) that have efficiencies as large as 20% for coincidence detection of fission events. These detectors are made of  $^3\text{He}$  or  $\text{BF}_3$  proportional counters and polyethylene. The sensitivity of the highest efficiency units is about 1 mg  $^{240}\text{Pu}$ . Organic scintillation detectors have also been used for coincidence detection of spontaneous fission.

## Active Neutron Interrogation

When the gamma-ray attenuation of materials is too high for passive assay, and the spontaneous fission rate is too low for neutron coincidence counting, we can interrogate the sample with highly penetrating beams of neutrons. The bombarding neutrons induce fissions in the sample and various gamma-ray and neutron signatures of the neutron-fission reactions are detected. During the past few years, active neutron interrogation has been used extensively to measure uranium and plutonium fuel rods, unirradiated fuel assemblies, and high-enriched uranium scrap. Either accelerator or radioisotopic neutrons could be used. Although accelerator neutron sources

offer large yields and source modulation flexibility (that is, pulsing), they usually require a substantial amount of initial investment and technical support. On the other hand, radioactive neutron sources are reliable, simple, and less costly and are well suited for in-plant measurements and quality control of reactor fuel components as well as for some categories of scrap-material and process-stream measurements.

The basic components of an active neutron assay system are the neutron source, the sample, and the detectors to count the induced activities (Fig. 15). The neutron source is normally surrounded by materials to moderate or slow down the neutrons and tailor their energy for interrogation. These materials also are needed for personnel shielding. The fissions induced by the irradiation emit prompt and delayed neutrons as well as prompt and delayed gamma rays. One or more of these emitted signatures can be used for the assay depending on complexity of the measurement problem and the detection method.

Most fissionable material contains a mixture of both fissile and fertile isotopes. The fertile isotopes (for example,  $^{238}\text{U}$  and  $^{232}\text{Th}$ ) have a threshold energy (about 1 MeV) for induced fis-



**TABLE III**  
**A NEUTRON IRRADIATION TECHNIQUE**

Moderator Configuration <sup>a</sup> (Thickness in cm)	Total Leakage (%)	Median Energy (MeV)	Fraction of Neutrons With E < <sup>238</sup> U Threshold	Fission Ratios <sup>b</sup> ( <sup>235</sup> U/ <sup>238</sup> U)	
			(%)	Measured	Calculated
Pb: 7:5 (Cd)	146	1.76	39	003.19	003.10
W: 7:5 (Cd)	135	0.47	78	005.68	006.53
W/C: 7:5/7:5 (Cd)	122	0.19	79	015.3	013.3
W/C/CH <sub>2</sub> : 7:5/7:5/2:5 (no Cd)	109	0.002	80	600.	788.

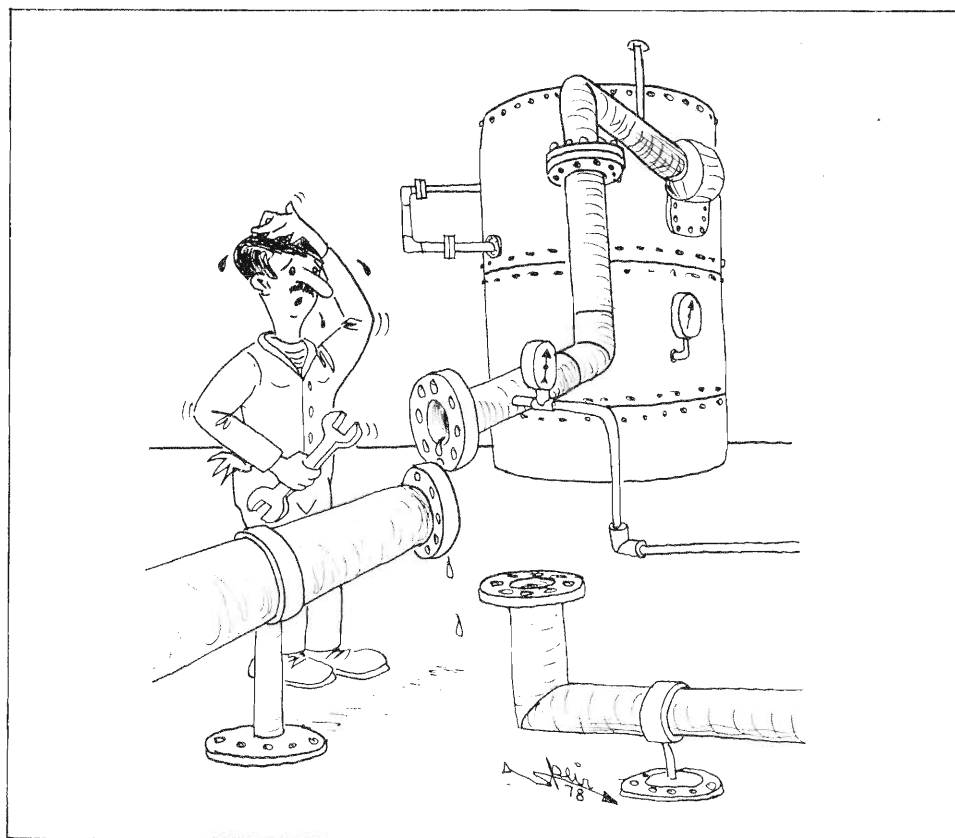
<sup>a</sup>The moderators and detectors were covered with 0.76-mm-thick cadmium where noted. DTF-IV calculations were used for the moderators with no cadmium; all others used the Monte Carlo code.

<sup>b</sup>Fission ratios are for equal weights of each fissionable isotope.

sion reactions, whereas fissile isotopes (for example, <sup>235</sup>U, <sup>233</sup>U, and <sup>239</sup>Pu) fission for all neutron energies. Active neutron interrogation techniques use subthreshold and superthreshold irradiation to separate the fissile and fertile components in a sample. The D-T, 14-MeV neutron generators are a convenient source of these interrogation neutrons, but the neutron energy is so high that both fissile and fertile isotopes undergo fission. To separate these components, we have developed neutron tailoring techniques to give both superthreshold and subthreshold irradiations from the same 14-MeV source. We have used Monte Carlo and DTF-IV neutron transport calculations to obtain maximum neutron leakage and desired neutron energy characteristics from various moderator assemblies. For example, for a 14-MeV neutron source, we have designed moderating assemblies that consist of concentric spherical shells of different materials including tungsten, graphite, and polyethylene. The large cross sections for the (n,2n) and (n,n') reactions (incoming neutrons, outgoing neutrons) in tungsten reduce the energy of the 14-MeV source neutrons to a median energy of approximately 500 keV, and elastic scattering of neutrons in the graphite further reduces the average neutron energy to the low-keV region. Table III lists typical moderator configurations, where the notation

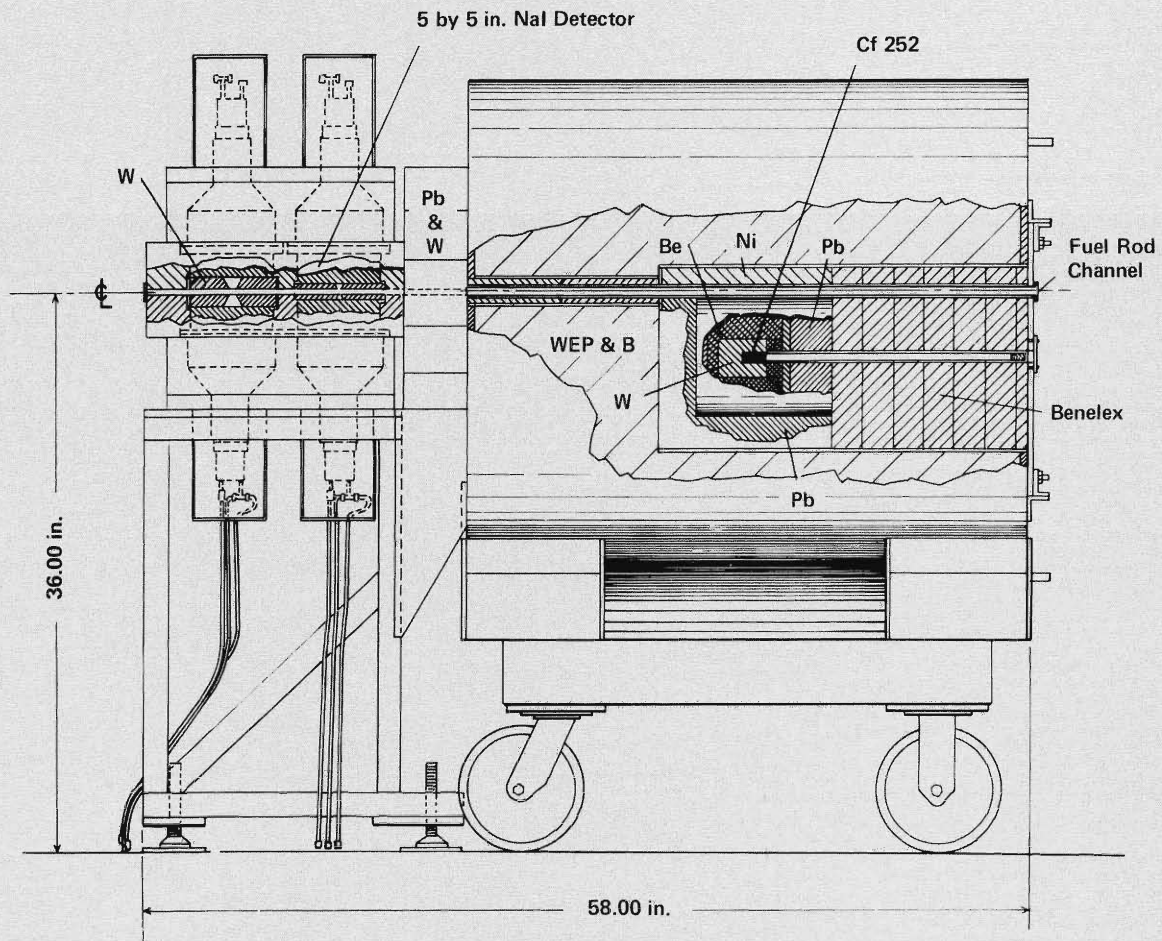
W/C/CH<sub>2</sub>: 7.5/7.5/2.5 indicates a 7.5-cm-thick core of tungsten surrounded by a 7.5-cm-thick shell of graphite surrounded by a 2.5-cm-thick shell of polyethylene. The total neutron leakage from the moderators exceeds 100% because of the large (n,2n) cross section in the heavy isotopes. Table III shows that the median neutron energy can be

lowered from 14 MeV to a few keV using moderators with radii less than 18 cm. The tailoring assemblies given in Table III result in <sup>235</sup>U/<sup>238</sup>U fission ratios from 3-600, a range that is very adequate for the separation of the fissile and fertile components in assay applications.



**TABLE IV**  
**RADIOACTIVE SOURCES FOR ACTIVE ASSAY APPLICATIONS**

Neutron Source	Prompt n and $\gamma$		Delayed n and $\gamma$
	Integral	Coincidence	Integral
Sb-Be; 25 keV	n		
Ra-Be; 200 keV	n		
$^{238}\text{Pu}$ -Li; 400 keV	n		
Am-Li; 400 keV		n and $\gamma$	
$^{252}\text{Cf}$ ; moderated		n and $\gamma$	n and $\gamma$
$^{252}\text{Cf}$ ; thermalized	n	n and $\gamma$	n and $\gamma$



*Fig. 16. Schematic of the  $^{252}\text{Cf}$  fast-neutron assay system for the fast breeder reactor fuel rods. The delayed gamma rays induced by the fast-neutron irradiation are subsequently counted with the two NaI detectors, which also measure the passive gamma rays to determine pellet-to-pellet uniformity.*

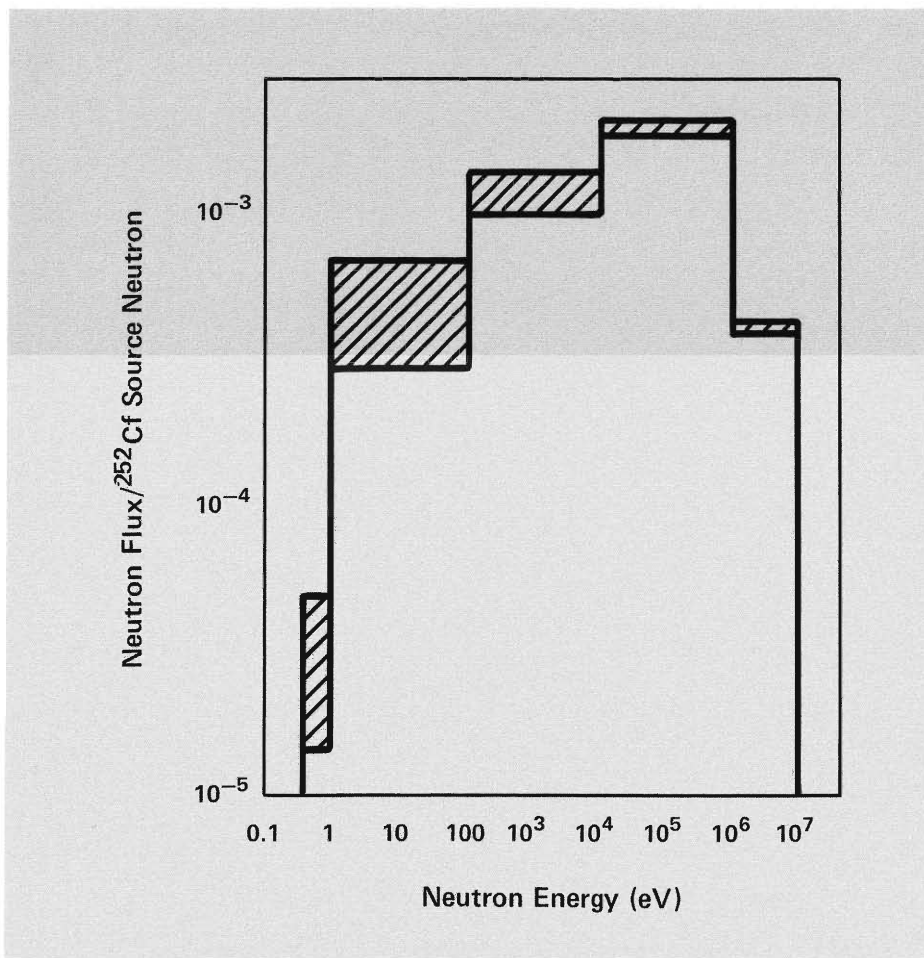


Fig. 17. Calculated neutron spectrum from a  $^{252}\text{Cf}$  source for the system shown in Fig. 16 with (cross-hatched area) and without the cylindrical nickel fast-flux trap.

For most safeguards applications employing active neutron assay, radioactive sources have been preferred because of their simplicity and reliability.

To apply radioactive sources to a given measurement problem, three primary variables should be optimized: (1) the type of neutron source, (2) the moderator and shield for the source, and (3) the detector to count the induced neutrons and/or gamma rays. Radioactive neutron sources under investigation at LASL include  $^{124}\text{Sb}-\text{Be}(\gamma, n)$ ,  $^{88}\text{Y}-\text{Be}(\gamma, n)$ ,  $^{226}\text{Ra}-\text{Be}(\gamma, n)$ , and  $\text{Am}-\text{Li}(\gamma, n)$  (which have neutron energies primarily below the fission thresholds of  $^{238}\text{U}$  and  $^{232}\text{Th}$ ), and  $^{252}\text{Cf}$  (which emits higher energy neutrons).

Table IV gives a summary of radioactive sources and corresponding detection methods that have been used by LASL for assay applications. These applications range from portable field instrumentation used by inspectors to large in-plant installations used for waste and spent-fuel assay.

### Fuel Rod Scanners

A high-performance instrument that incorporates several of the assay signatures previously described is the fuel rod scanner for the Fast Flux Test Facility (FFTF) at Hanford, Washington, a low-power reactor used to test designs for the Breeder Reactor Program. The design of the scanner required extensive use of Monte Carlo computer calculations to determine an optimum combination of materials, geometry, and detectors.

This hybrid assay system uses active neutron interrogation to determine total fissile content and passive gamma-ray spectroscopy to measure selected plutonium isotopes. The system measures the plutonium fissile content in a fuel rod to better than 0.5% accuracy and verifies that the plutonium content in individual fuel pellets within each rod is uniform from pellet to pellet. The fuel rods contain 70-80% natural uranium and 20-30% plutonium. The rods have

an active length of 914 mm and a diameter of about 5 mm. Many thousands of these rods are used in the reactor core, and tight accountability of the plutonium is required for nuclear material safeguards.

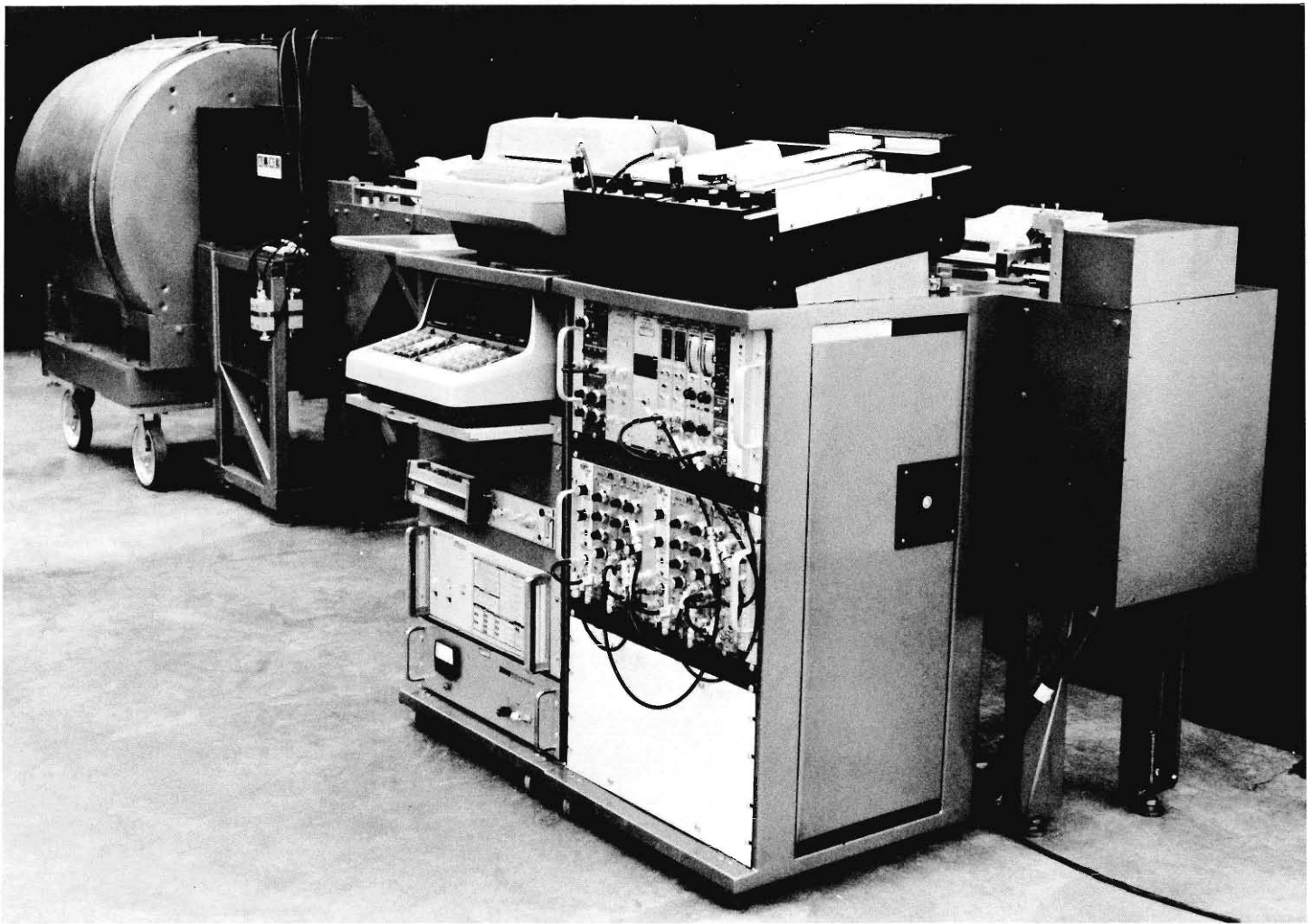
A schematic of the system (Fig. 16) shows the 600- $\mu\text{g}$   $^{252}\text{Cf}$  neutron source used for the interrogation, the moderator, the radiation shield, and the NaI detectors used to count the gamma rays. The neutron moderator is designed to give a fast-neutron irradiation (the irradiation channel is lined with cadmium) and a high fissile/fertile fission ratio (400/1). The fissile/fertile fission ratio corresponds to

$$\frac{\int \sigma_{239}^f(E) \phi(E) dE}{\int \sigma_{238}^f(E) \phi(E) dE}, \quad E > 0.4 \text{ eV},$$

where  $\sigma_{239}^f$  and  $\sigma_{238}^f$  are the  $^{239}\text{Pu}$  and  $^{238}\text{U}$  fission cross sections, respectively, and  $\phi(E)$  is the energy-dependent neutron flux in the irradiation channel.

Neutron transport calculations using both  $S_n$  and Monte Carlo techniques were used in the design of the  $^{252}\text{Cf}$  source moderator. The calculations led to the design of the cylindrical moderator assembly shown in Fig. 16 with a tungsten core surrounded by a shell of beryllium followed by nickel and lead for fast-neutron reflection. The calculated neutron spectrum from a  $^{252}\text{Cf}$  source for this system is shown in Fig. 17 with (cross-hatched area) and without the cylindrical nickel fast-flux trap. The presence of the nickel reflector increased the fission rate by approximately 70% over the lead reflector. This moderator resulted in a  $^{239}\text{Pu}/^{238}\text{U}$  fission ratio of approximately 400/1 for irradiation neutrons above the cadmium cutoff energy (0.4 eV). It is necessary to keep this ratio high to avoid a significant contribution from  $^{238}\text{U}$  fission, which is not of safeguards interest.

During operation, up to 30 fuel rods are placed in the loading magazine of the



**Fig. 18.** Fast-neutron  $^{252}\text{Cf}$  assay system for fast breeder reactor fuel rods. The system includes a  $600\text{-}\mu\text{g } ^{252}\text{Cf}$  source and shield, two 5- by 5-in. NaI detectors to count the delayed gamma rays, automated fuel-rod handling, and a data-reduction system.

fuel-rod translator. The automated translator picks up the rod to be assayed and moves it through the NaI crystals to take a background count of the un-irradiated rod and then into the  $^{252}\text{Cf}$  assembly for irradiation with neutrons. The direction of travel is then reversed and the rod is withdrawn through the NaI detectors, which measure the delayed fission gamma rays for the total fissile assay and the passive gamma rays for the pellet-to-pellet uniformity. The two 5-in.-long by 5-in.-diam NaI detectors count the high-energy delayed gamma rays ( $>1200$  keV) to determine the total fissile measurement because of their higher penetration through the rod. Simultaneously, the same NaI detectors determine pellet-to-pellet uniformity by measuring the much more intense lower energy passive gamma rays from the fuel. Each of the two NaI detectors examines a different low-energy window

on a pellet-to-pellet basis, using a tungsten sleeve with a narrow collimation slit. One detector counts 60-keV gamma rays, which are mainly from  $^{241}\text{Am}$ , and thus serves as a batch monitor because the  $^{241}\text{Am}$  content of the fuel is time dependent. The second detector counts primarily plutonium gamma rays in the range of 100 to 500 keV. This second window provides a qualitative measure of the fissile plutonium in the rod. After the irradiation and scan of the fuel column, the rod is unloaded in the tray directly below the loading magazine and the cycle is repeated for the next rod.

The complete assay system, which includes automated fuel rod handling and computer-based data reduction, is shown in Fig. 18. This system was installed in 1973 at the Hanford Engineering Development Laboratory (HEDL) Plutonium Facility and has been used to

measure the plutonium fissile content (better than 0.5% accuracy) and uniformity of over 70 000 FFTF fuel pins, about 40 000 of which are now being used in the reactor.

### The $^{252}\text{Cf}$ -Based Assay System for the FAST Facility

We have designed a neutron interrogation assay system to measure waste solids from reprocessing and spent-fuel packages at the Fluorinel and Storage (FAST) Facility, a new addition to the Idaho Chemical Processing Plant used to reprocess spent fuels from the national defense program. Measurements of the canisters of solid waste will be performed for  $^{235}\text{U}$  process control and accountability of the waste solids. The spent-fuel packages will be assayed to provide assurance that the  $^{235}\text{U}$  content is below 10.5 kg, the criticality limit

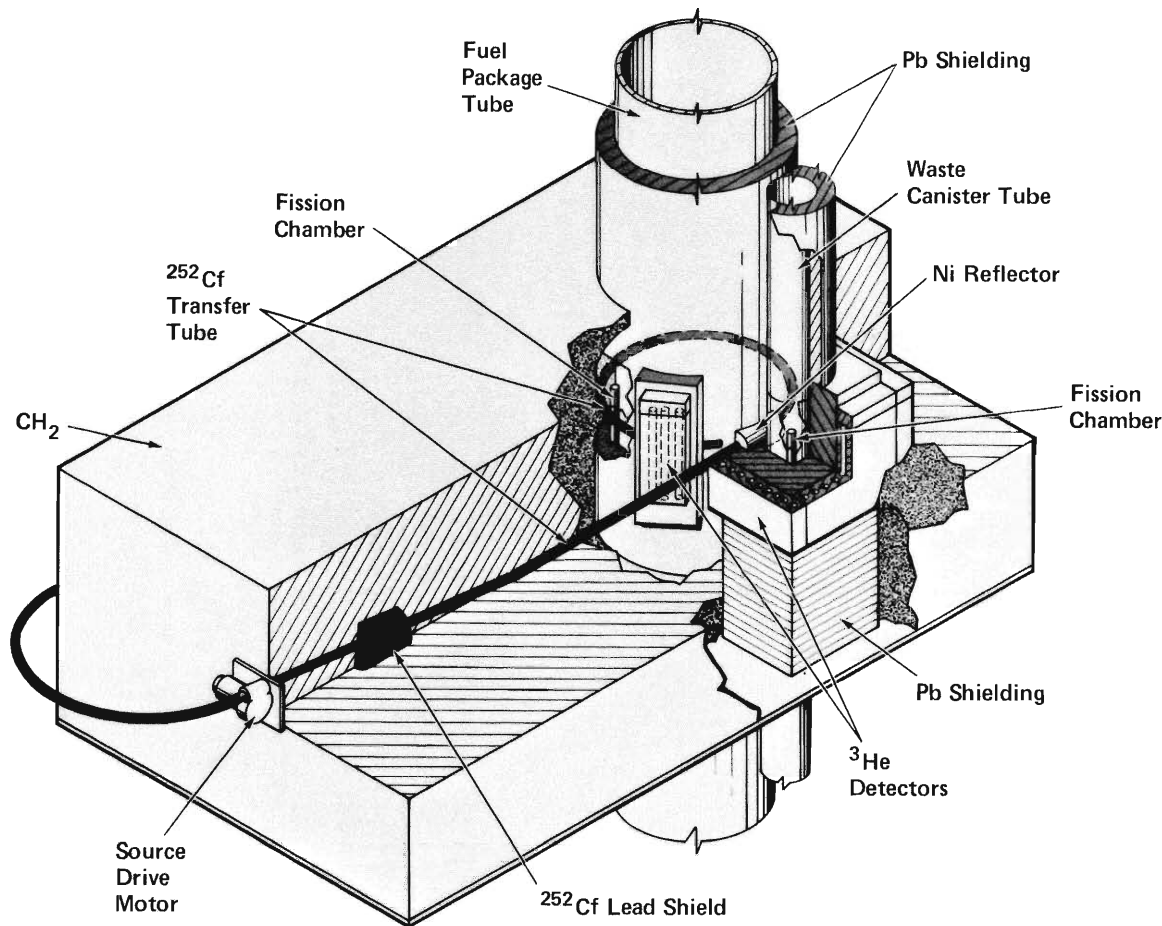


Fig. 19. The fast-neutron assay system for the Idaho Chemical Reprocessing FAST facility showing  $^{252}\text{Cf}$  source transfer system,  $^3\text{He}$  detectors, and the sample tubes for irradiated fuel elements and waste containers.

The assay system is based on the principle of the *shuffler*, which is a modulated  $^{252}\text{Cf}$  source assay system that repetitively transfers the neutron source from the interrogation position to a shielded storage position while the delayed neutrons are counted. The assay system shown in Fig. 19 includes a source shield tank, a decoupling  $\text{CH}_2$  shield, and an irradiation area. The assay sample is irradiated in a chamber surrounded by  $^3\text{He}$  neutron detectors that are used to count delayed neutrons produced following each irradiation. We have evaluated the shuffler system for the assay of fuel rods, inventory samples, scrap and waste, uranium ore, irradiated fuel, and plutonium mixed-oxide fuel.

At the FAST Facility, the fissile con-

tents vary from 0 to 400 g in the waste and from about 8 to 12 kg in the spent-fuel elements. Measurements must be obtained in the presence of large neutron ( $1.2 \times 10^7$  n/s) and gamma (50 000 R/h) backgrounds. The system employs fast-neutron interrogation of the sample using a large (5-mg)  $^{252}\text{Cf}$  source to override the neutron backgrounds. Preliminary Monte Carlo calculations indicate that measurement precisions of better than  $\pm 5\%$  ( $2\text{-}\sigma$ ) are obtainable for waste loadings in excess of 200 g  $^{235}\text{U}$ .

We are designing the assay instrument as an integral part of the FAST Facility with close cooperation from Exxon and Ralph M. Parsons Company (Architectural Engineer). The facility will house the assay instrument in a separate cubicle connected by through

tubes to the fluorinel dissolution cell. Waste canisters and spent-fuel packages will be lowered by crane into the assay instrument through the tubes that penetrate the cell floor and extend below the assay system. The assay system will be remotely controlled and operated by a dedicated mini-computer system. 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 Active Well Coincidence Counter (AWCC) for Portable Inspection Applications

An important safeguards area that has very different requirements from in-

*Fig. 20. The Active Well Coincidence Counter including detector body, cart, and portable electronics for automated data collection and analyses.*



plant instrumentation is field inspection using portable instrumentation. Equipment for this application must be lightweight, rugged, and simple to operate. A leading example of this instrumentation is the High-Level Neutron Coincidence Counter (HLNCC), which measures the effective  $^{240}\text{Pu}$  content in bulk plutonium samples by coincidence counting the spontaneous fission neutrons from the plutonium.

To extend the same type of equipment to uranium assay applications, we use active neutron interrogation to induce the fission reactions because the spontaneous fission rate in  $^{235}\text{U}$  is too small to be useful in the passive mode.

The basic principle of the AWCC is fast-neutron interrogation using a random neutron source (for example, AmLi) and counting the induced fission reactions using coincidence techniques to suppress the signal from the random interrogation source. The AWCC uses 42  $^3\text{He}$  detectors to count the induced fission neutrons. It also has two neutron sources of similar strength ( $\sim 5 \times 10^4$  n/s), one in the lid and one in the bottom plug. The use of two neutron sources produces a rather uniform vertical response. Also, the  $\text{CH}_2$  plugs act as neutron shields to reduce the background neutrons counted in the  $^3\text{He}$  tubes from the AmLi sources. This shielding technique improves the induced signal-to-interrogation neutron background ratio by a factor of 10.

The AWCC has been designed to take advantage of the portable electronics package that was developed for the HLNCC. The electronics unit is interfaced directly to the HP-97 programmable calculator shown in Fig. 20. A microprocessor in the unit reads out the run time, total counts, real plus accidental counts, and accidental counts to the HP-97. The HP-97 reduces the data using the software package selected by the operator.

The unit is useful for measuring bulk

$\text{UO}_2$  samples, high-enrichment uranium metals, LWR fuel pellets, and  $^{233}\text{U}$ -Th fuel materials, which have very high gamma-ray backgrounds. By removing the AmLi source, the unit can measure  $^{238}\text{U}$  and plutonium in the passive neutron coincidence mode. We recently have supplied the first AWCC system to the IAEA for field test and evaluation.

### Status of Development and Implementation

Acceptable performance criteria for NDA techniques for safeguards, established well over a decade ago, include measurement times in the range from a few seconds to 1 hour and the following accuracies for the various material categories: 0.2-3.0% for uniform feed and product materials, 2-10% for recoverable scrap, and 5-30% for waste. We have achieved or exceeded these goals for many of the materials associated with the following fuel processes: uranium enrichment, uranium and plutonium fuel fabrication, and uranium and plutonium scrap recovery. The development of NDA methods and instruments has been greatest for measurements of scrap and waste, holdup, high-purity raw materials, and finished fuels—the material categories that were first identified as priority problems for NDA development.

Portable instruments and a number of stand-alone instruments for measuring containers of SNM were the first to be developed and tested in operating facilities. Some of these instruments then were adapted for on-line applications so that items or samples of plutonium material could be measured without being removed from the glove box containment.

Current research is focused on extending the range of applicability (to other compositions and concentrations) and improving accuracies of existing technologies, adapting techniques for in-



*Fig. 21. In-line  $UF_6$  enrichment monitor installation at the extended-range product-withdrawal station of the DOE/Goodyear Atomic gaseous diffusion plants. Passive gamma-ray and neutron instruments continuously measure  $^{235}U$  and  $^{234}U$  enrichments of liquid  $UF_6$  to provide criticality safety and process controls.*

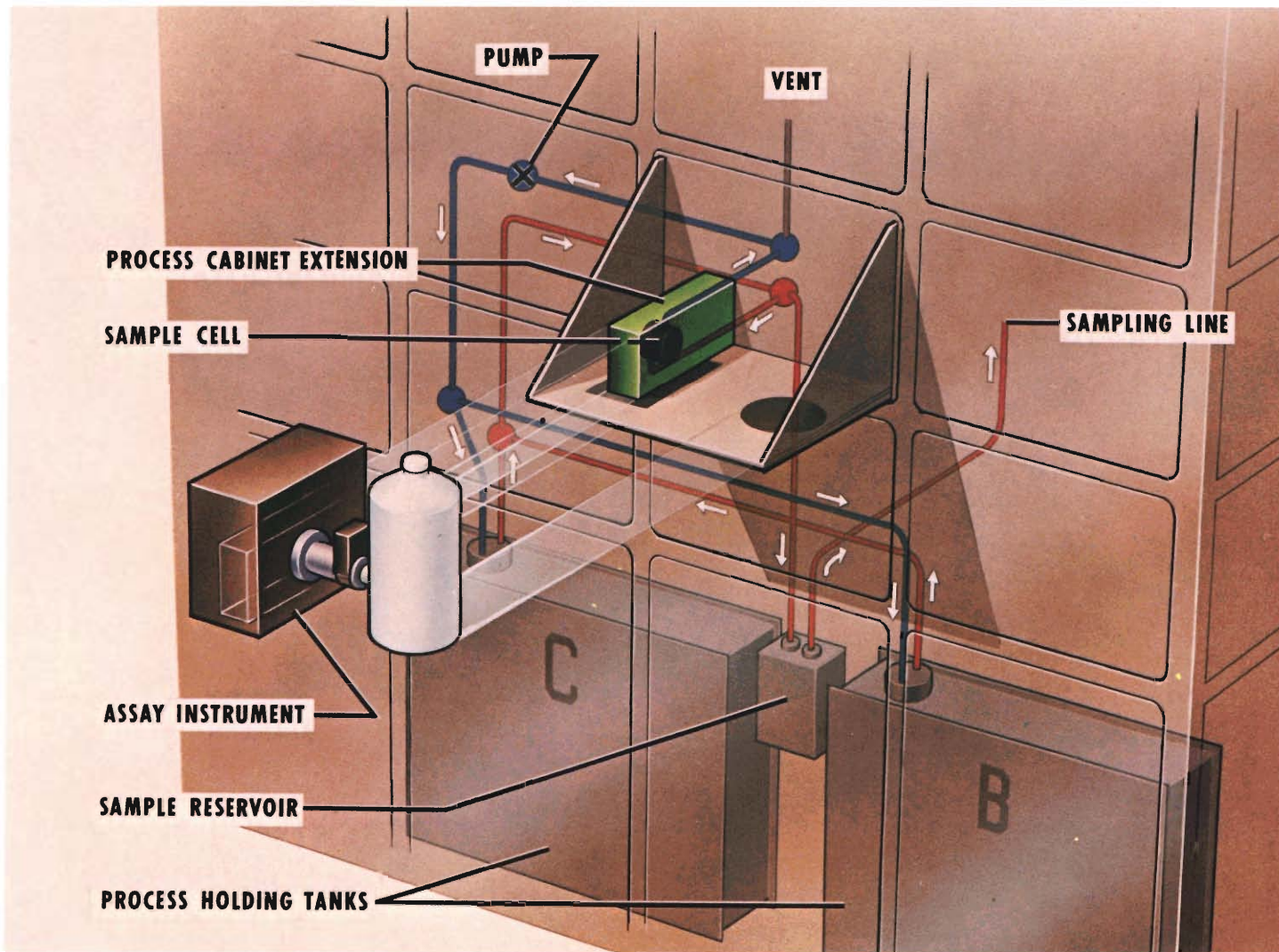
line measurement and monitoring of bulk process materials, developing specialized instrumentation for inspectors and for containment and surveillance, developing methods to measure spent fuels and "hot" reprocessing materials, and finding techniques to measure advanced fuel materials such as mixed uranium-plutonium solids and solutions.

Ultimately, the best gauge of the success of these developments is the level of acceptance and implementation by the nuclear industry, regulatory agencies such as the US Nuclear Regulatory

Commission, and for the international arena, the IAEA. In most operating nuclear facilities that process low-enriched uranium for use in commercial power reactors and high-enriched uranium and plutonium for defense programs and research reactors, the implementation of the NDA methods described here to complement or replace analytical chemistry measurements for safeguards and process control has been

extremely gratifying. An example is the in-line  $UF_6$  enrichment monitor at the DOE/Goodyear Atomic gaseous diffusion plant (Fig. 21), which serves principally for process control and criticality safety. Major exceptions are spent-fuel reprocessing plants, where the slowdown or postponement of reprocessing has in turn delayed the testing and implementation of promising measurement methods for the fission-product-contaminated





**Fig. 22.** *The in-line absorption-edge densitometer for the Savannah River Reprocessing Plant. Plutonium product solution is pumped from holding tanks through a measurement cell, which is located inside an extension of the process line containment.*

uranium and plutonium materials processed in these plants. A project is now under way to test an in-line K-edge densitometer (Fig. 22) for measuring plutonium concentration in the product solution line of the Savannah River spent-fuel reprocessing plant, which is operated for DOE national defense programs.

A typical nuclear facility now has an NDA counting room in which scrap and

waste and samples from the process lines are measured for materials accounting, process control, and waste management. Portable NaI and germanium gamma-ray detectors and neutron counters are used routinely to measure holdup for inventories or investigate process anomalies. Fuel rod scanners are used in all light water reactor fuel fabrication facilities in the United States and most of those in

Belgium, France, Germany, Italy, and Japan to quality-assure all fuel rods. Some facilities, such as the plutonium plants operated at LASL and Rocky Flats for defense programs and the General Electric/Wilmington light water reactor fuel fabrication plant, have installed near-real-time material accounting systems that use NDA instruments for measurement or verification of material movement throughout the plant.

We are equally gratified by the acceptance of NDA technology by regulatory agencies, both as a part of safeguards systems for plants and for independent

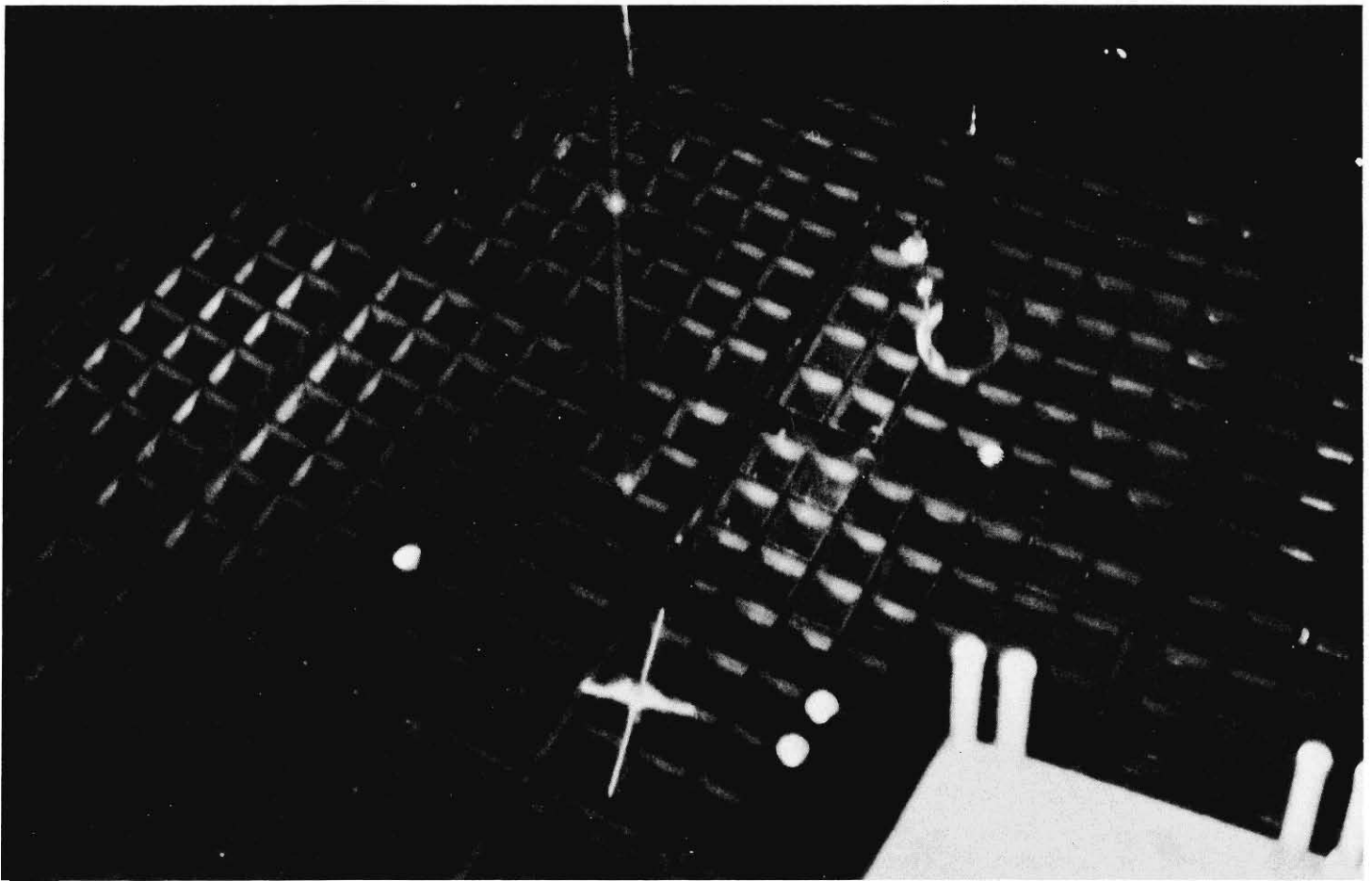
**TABLE V**  
**INSTRUMENTATION DEVELOPED BY LASL FOR IAEA APPLICATIONS**

Instrument	Principle	Applications
SAM-II	Passive gamma-ray enrichment meter	UF <sub>6</sub> cylinders, UO <sub>2</sub> powder, fuel pellets and rods
Segmented gamma scanner	Transmission-corrected gamma-ray measurements	UO <sub>2</sub> , PuO <sub>2</sub> scrap and waste
HLNCC	Passive neutron coincidence counting	Pu metal, PuO <sub>2</sub> powder, MOX fuel
AWCC	Active neutron interrogation and coincidence counting	U metal, UO <sub>2</sub> powder, U-Al alloy
Coincidence collar	Active neutron interrogation	PWR, BWR, and HWR fuel assemblies
K-edge densitometer	Active gamma-ray transmission	Pu and U solutions reprocessing plants
Reactor power monitor	Passive neutron	LWR reactor power monitor
Spent-fuel verification instruments	Cerenkov glow, passive gamma-ray, passive neutron	LWR and HWR spent fuel assemblies in storage pools

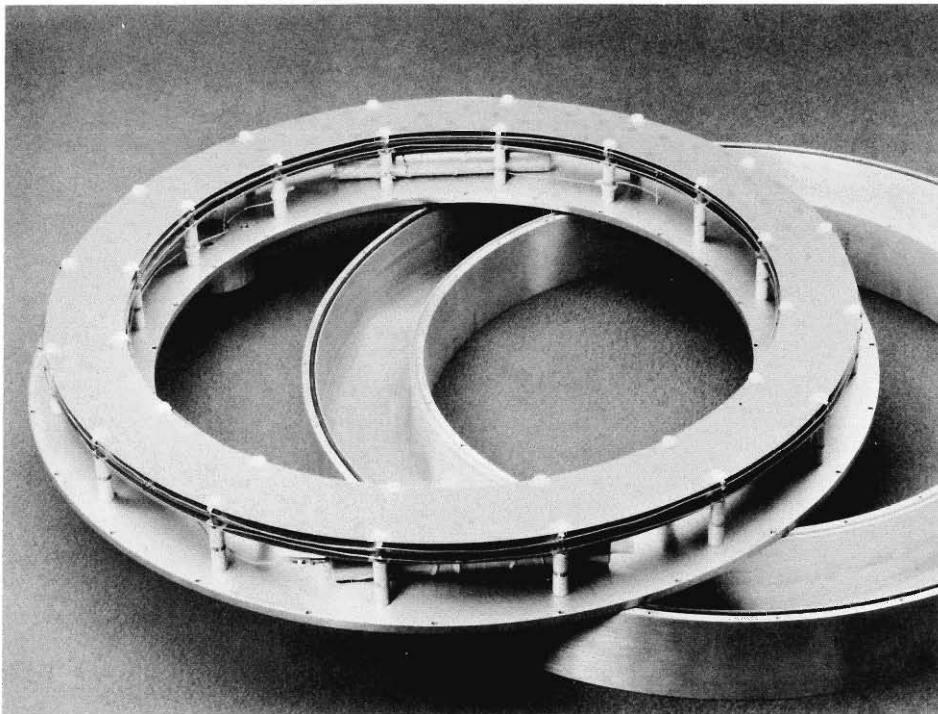
inspector verifications. Through the US program of technical assistance to the IAEA, we have developed a number of key NDA instruments and applications for international inspection of nuclear plants, and have provided inspectors

with training, manuals, calibration, and implementation assistance. Table V lists instrumentation we have developed for IAEA applications. One technique helps to verify relative burnup and cooling time of spent reactor fuel assemblies to

infer the resulting plutonium inventory (Figs. 23 and 24). E. Dowdy has pioneered a convenient method using the Cerenkov glow for confirming that the spent-fuel assemblies stored in an underwater array as shown in Fig. 23 are



*Fig. 23. A light water reactor fuel element storage array under about 10 m of water in a spent-fuel storage pool. One fuel element is raised about one-half its length through the ring ion and fission chamber (see Fig. 24) to measure the gamma-ray and neutron emissions.*



*Fig. 24. Ring detector containing ion and fission chambers to measure intense gamma-ray and neutron emissions from spent-fuel assemblies to help verify burnup and cooling time in the fuel.*

highly radioactive. Inspectors can turn off the overhead lights and measure the intensity of the Cerenkov glow, which is proportional to the activity level in the spent-fuel assembly. This approach is rapid and requires no instrumentation to be put in the contaminated water of the pool.

### Challenges of the Future

While great progress has been made in filling the voids in safeguards measurement technology with NDA methods, the implementation of new technology has been confined largely to retrofitting existing facilities and conventional safeguards procedures. We have introduced high technology for measurements and accounting into plants that were designed both for manual operations and to rely on administrative controls and chemical assay for safeguards accounting. Implementation of NDA-based vehicle and personnel monitors has enhanced the containment and sur-

veillance elements of safeguards for existing facilities.

Heretofore, the development of instruments has been focused on the need to measure uranium and plutonium generated in fabrication processes that keep these elements isolated from one another and are free of fission product contamination.

Current studies and conceptual process and facility designs involve both reprocessing and coprocessed uranium and plutonium fuels. These material forms and compositions pose a whole new set of measurement problems, which we have only begun to address. In the international arena, the fact that fuel reprocessing and development of the breeder are proceeding now makes solving the associated measurement and accounting problems an urgent need. Some of the materials for which NDA measurement techniques should be developed, tested, and evaluated are: leached spent-fuel hulls and other "hot" scrap and waste, mixed uranium and plutonium solutions with the various levels of fission product contamination characteristic of reprocessing, "cold" mixed plutonium and uranium solutions and solids characteristic of co-conversion (from nitrate to oxide), and final breeder and (perhaps plutonium recycle) fuels. Furthermore, until commercial (LWR) spent fuels are reprocessed, they must be stored at the reactor sites or at special away-from-reactor (AFR) facilities. The spent fuels should be verified and measured quantitatively, if possible, to establish their economic value and to control batch make-up for future reprocessing, as well as for safeguards.

Concerns are increasing about safeguards for uranium enrichment, the front end of the fuel cycle, as advanced isotope separation processes based on centrifuges, lasers, or plasma devices are implemented. Because of their inherently large isotope separation factors, fewer separation stages are required to produce high-enriched uranium than are needed for classical gaseous diffusion separators. Hence the advanced isotope separation techniques could be more easily used for covert production of bomb materials by the facility operator. New safeguards measurement problems associated with these advanced methods must be addressed.

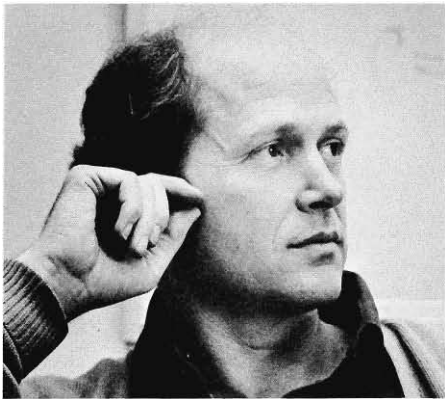
Process lines in future facilities will use advanced processing and control technology, including remote operation and maintenance, to minimize personnel radiation exposure, assure material containment, and handle high throughput efficiently. Almost certainly more, not less, measurement instrumentation will be needed in process lines. The instruments must function very reliably, even in such high-radiation environments as reprocessing canyons, because of the restricted access. Individual measurement stations must be designed to meet the well-ordered requirements of cost-effective integrated, automated systems of materials accounting and control. At present, LASL and the Hanford Engineering Development Laboratory are planning NDA instrumentation to be incorporated at DOE's new Fuel Materials Examination Facility at Richland in an automated mixed uranium and plutonium oxide fuel fabrication line. With the DOE/ORNL

Consolidated Fuel Reprocessing Program, LASL also is working on the conceptual safeguards design for an advanced fuel-reprocessing plant.

Implementation of advanced automated materials accounting and control systems may constitute an enigma for safeguards inspectors. The question is how an inspector can independently verify that a large complex integrated system has been operated as declared. Providing system integrity with sufficient transparency for verification will undoubtedly bring changes in the design of individual instruments and their measurement controls. Inspectors will need improved portable instruments to confirm the results of the large in-plant systems.

Safeguards technology, like waste management, was once viewed by many scientists and planners as a supportive, ancillary factor in the development of nuclear power for the generation of electricity. These factors, together with reactor safety, have now become dominant in the determination of the future of nuclear energy. Thus levels of activity in future safeguards instrumentation development just described will depend to a large extent on the acceptance and use of nuclear energy, which in turn will be strongly influenced by public perception of the safety of reactors, the capabilities for safeguarding these fuels for peaceful uses, and safe management of the associated nuclear waste materials. In this chicken-and-egg cycle, we hope that the significant advances in safeguards technology, as well as in reactor safety and waste management, will be given due consideration.

## THE AUTHORS



**Howard O. Menlove**, Group Leader of the International Safeguards Group, has worked with the Nuclear Safeguards Program at LASL for 12 years. He has been active in research and development of advanced techniques for nondestructive assay of fissionable materials. At present, his work is in inspector instrumentation development and implementation, nondestructive assay standards and calibration, spent-fuel verification techniques, training, and technology transfer. Before joining LASL, he had considerable experience in neutron and fission physics and in gamma-ray spectroscopy. After earning his Ph.D. in nuclear engineering at Stanford University, he spent a year at the Kernforschungszentrum in Karlsruhe, FRG, supported by a Fulbright Award.



**Roddy B. Walton** earned his bachelor of science degree in physics at Texas A&M University, and his Ph.D. in nuclear physics at the University of Wisconsin in 1957. From 1959 to 1967, he was with the General Atomic Company, where he was involved in electron linac instrumentation and experiments in neutron thermalization, neutron capture cross sections, delayed gamma rays from fission, and photonuclear reactions. In 1967, he joined LASL, and has been working in development and implementation of methods for the measurement of fissionable materials for the nuclear safeguards program. He has been instrumental in the initiation of new projects, including the Mobile Nondestructive Assay Laboratory and its operation in the field, DYMAC (dynamic materials accounting), and systems analysis. He is a Fellow of the American Nuclear Society.



*Fig. 1. The Allied-General Nuclear Services spent-fuel reprocessing plant in South Carolina.*

# Dynamic Materials Accounting Systems

by Darryl B. Smith, Dante Stirpe, and James P. Shipley

The problem of maintaining strict control over nuclear material will be made more difficult by the nuclear power demands of the future, which will require large facilities—enrichment plants and reprocessing plants, for example—that process great quantities of high-quality fissile materials. The scale of these operations has forced a reassessment not only of facility design, construction, and process operation, but also of safeguards methods to prevent unauthorized use of the nuclear materials contained in the facilities. A comprehensive domestic safeguards system combines the functions of materials accounting and physical protection.

The Los Alamos Scientific Laboratory has been designated the Department of Energy's lead laboratory for the design and evaluation of materials accounting systems for nuclear facilities of the future. In this article, we examine these systems and the techniques for their design. Nuclear materials accounting systems must keep track of large quantities of materials as they

***Systems analysis suggests that near-real-time materials accounting systems designed for future large-throughput nuclear facilities can meet high performance standards.***

move through the various processing stages and must keep track of them so well that the absence of even small amounts can be detected. The uncertainties inherent in any measurement process and the difficulties of measuring in high-radiation fields behind heavy shielding complicate this task.

We also illustrate the potential benefits of these systems by describing the development and expected performance of a materials accounting system we have designed for the Allied-General Nuclear Services (AGNS) spent-fuel reprocessing plant at Barnwell, South Carolina (Fig. 1). This plant was designed to process large amounts of irradiated fuel from power reactors. The accounting system was designed after the plant was built and with simulated data because the plant is not yet operating.

The potential of system performance is based on projected measurement capabilities of instruments, some of which are still under development. These projections cannot be tested without access to an operating facility. However, our preliminary evaluations suggest that we can design dynamic materials accounting systems for large bulk-processing facilities that meet detection standards close to those recommended by the IAEA.

### **The Basis for Materials Accounting**

The ultimate aim of nuclear safeguards is to be able to state with confidence, "No significant amount of nuclear material has been diverted." The philosophy underlying the development of materials accounting systems is that the truth of the statement can and should be verified. Thus, materials accounting systems are designed to account for or keep track of the amounts and locations of sensitive nuclear materials by periodic measurements. Materials balances are drawn about suitable areas of the facility according to the equation

$$\begin{aligned} \text{Materials balance} &= \text{initial inventory} \\ &+ \text{transfers in} \\ &- \text{transfers out} \\ &- \text{final inventory} \end{aligned}$$

defined over a reasonable time interval. In principle, if all nuclear material in each term of the equation has been measured, the materials balance should be zero in the absence of diversion. In practice, however, it is never zero because of the uncertainties inherent in all measurement processes. The measurement uncertainties produce a corresponding uncertainty in the materials balance, so statistical techni-

ques are used to decide whether a balance indicates diversion of material.

At present, materials balances are drawn around an entire plant or a major portion of it after the facility has been shut down and cleaned out to inventory the material present. Although such accounting methods are essential to safeguards control of nuclear materials, they have inherent limitations in sensitivity and timeliness. The sensitivity is limited by measurement uncertainties that may conceal losses of significant quantities of nuclear material in large plants. The timeliness is limited by the frequency of physical inventories; that is, the practical limits on how often a facility can be shut down for inventory and still remain productive.

Both sensitivity and timeliness can be improved by implementation of *dynamic materials accounting*. This approach combines conventional chemical analysis, weighing, and volume measurements with the on-line measurement capability of NDA (nondestructive assay) instrumentation to provide rapid and accurate assessment of the locations and amounts of nuclear material in a facility. Materials balances are drawn without shutting down the plant: in-process inventories are measured, or otherwise estimated, while the process is operating.

To implement the approach, the facility is partitioned into several discrete accounting areas. Each accounting area contains one or more chemical or physical processes and is chosen on the basis of process logic and the ability to draw a materials balance, rather than on geography, custodianship, or regulatory requirements. By measuring all material flows in each area separately, quantities of material much smaller than the total plant inventory can be controlled on a timely basis and any discrepancies can be localized to the portion of the process contained in the accounting area.

Control by dynamic materials accounting is rigorous. It forces a potential divertor to steal nuclear material in quantities small enough to be masked by measurement uncertainties. Thus, to obtain a significant quantity of material, the divertor must commit many thefts and run the concomitant high risk of detection by the accounting system, surveillance instruments, and physical protection system.

### Designing a Materials Accounting System

The performance, or diversion detection sensitivity, of a materials accounting system depends on the details of the measurement system, which in turn depend on the details of the process. Because these details vary from one plant to another, the Los Alamos safeguards systems studies focus on specific designs of existing or planned nuclear facilities.

The first step in the development of a facility's accounting system is to determine the flows of nuclear materials through the facility from design data and operator experience. Then, the facility is partitioned into logical accounting areas, and an appropriate measurement system is postulated for each area. Wherever

possible, the designer incorporates the measurement processes already in the plant design into the measurement system and augments them with any additional measurements necessary to draw a materials balance.\* The final step is to examine the expected performance of the accounting system design.

To develop preliminary designs of materials accounting systems, we model and simulate the in-plant processes and measurement systems by computer because no large fuel-cycle plants are yet in operation. Detailed dynamic models of material flows are based on actual process design data. They include bulk flow rates, concentrations of nuclear materials, holdup of materials in the process line, and the variability of all these quantities. Design concepts for the accounting systems are evolved by identifying key measurement points and appropriate measurement techniques, comparing possible materials accounting strategies, developing and testing appropriate data-analysis algorithms, and quantitatively evaluating the proposed system's capability to detect losses. The use of modeling and simulation allows us to study the effects of process and measurement variations over long operating periods and for various operating modes in a short time.

Computer codes simulate the operation of each model process using standard Monte Carlo techniques. Input data include initial values for all process variables and values of statistical parameters that describe each independent process variable. These data are best estimates obtained from process designers and operators. Each accounting area is modeled separately. When a process event occurs in a particular area, the values of the flows and in-process inventories associated with that part of the process are computed

\*See "Nondestructive Assay for Nuclear Safeguards."

and stored in a data matrix. These data are available for further processing and as input to computer codes that simulate accounting measurements and materials balances.

The flow and inventory quantities from a simulated process model are converted to measured values by applying simulated measurements. Each measurement type is modeled separately; measurement errors are assumed to be normally distributed (Gaussian), and provisions are made for both additive (absolute) and multiplicative (relative) errors. Significant measurement correlations are included explicitly. In most cases the measurement models are derived from the performance of similar instrumentation that has been used and characterized in laboratory and field applications involving similar materials. Simulated measurements are combined to form dynamic materials balances under various accounting strategies.

### Data Analysis

We combine the most promising measurement and accounting strategies with statistical techniques in comparative studies of loss-detection sensitivities. One of the major functions of the materials accounting system is to indicate loss, or possible diversion. Diversion may occur in two basic patterns: abrupt diversion (the single theft of a relatively large amount of nuclear material) and protracted diversion (repeated thefts of nuclear material on a scale too small to be detected in a single materials balance because of measurement uncertainties). Protracted diversion usually is the most difficult to detect.

The use of dynamic materials accounting enhances the ability to detect both diversion patterns, but it results in the rapid accumulation of relatively large quantities of materials accounting data. For example, if an area's materials



balance is closed once each 8-hour shift, after 1 month the safeguards operator will have a sequence of 84 materials balances and estimates of their associated uncertainties. Analysis of a single materials balance may be sufficient to detect a large abrupt theft of material, but the entire sequence of data contains the information necessary to detect small protracted diversions. Because small diversions may be masked by measurement uncertainties, they often are difficult to detect, and the operator must use one or more statistical tests of the accounting data to decide whether diversion has taken place.

### Decision Analysis

We have developed or adapted a variety of statistical tools for the analysis of materials accounting data that become available sequentially in time. These tools and their implementation are known collectively as *decision analysis*.

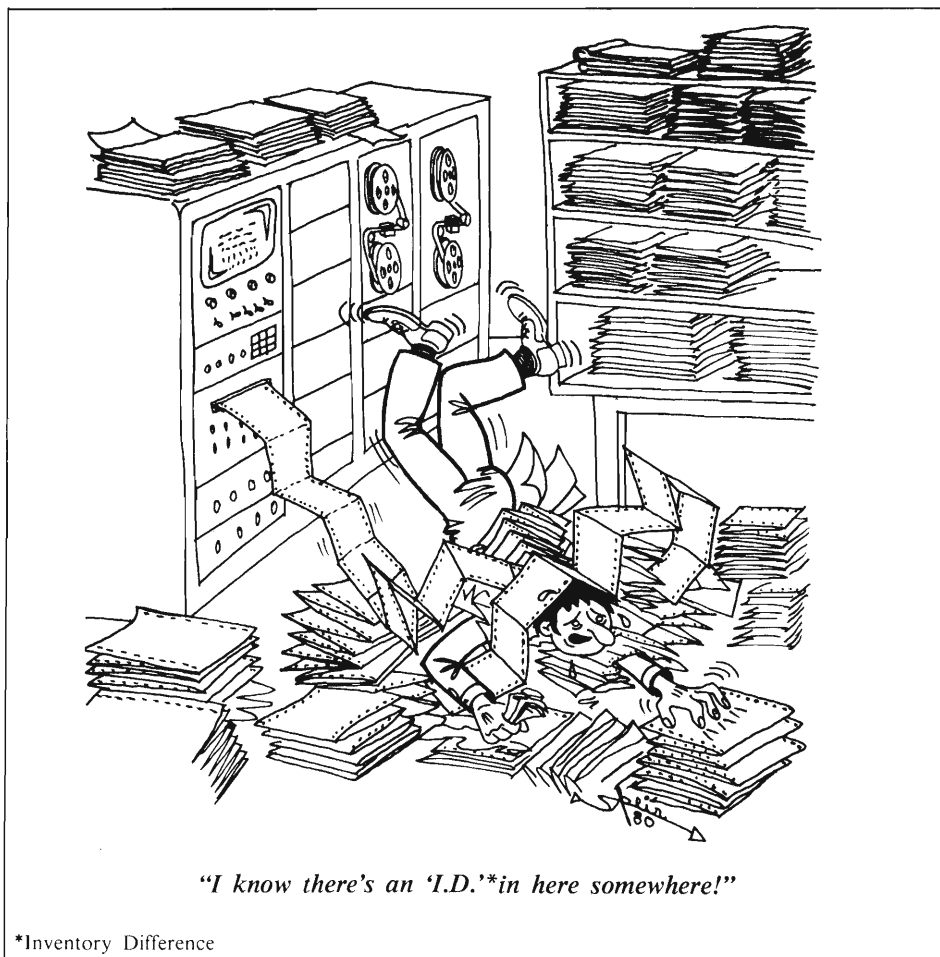
A simple, specific example shows what is involved in decision analysis. Suppose we have a sequence of 10 materials balances (47, 2, -109, 76, 2, 40, 62, -20, 34, 18 g)\* and an estimate of each balance's standard deviation. The standard deviation  $\sigma$  is a measure of

the uncertainty in a materials balance calculated from individual measurement errors. To analyze these data we must select an appropriate *statistical test*, construct a *test statistic* from the materials balance data, and establish one or more *test thresholds*. Then we can compare the value of the test statistic to the threshold(s) and draw a conclusion as to whether material has been diverted.

In our example, we use the *cusum test*, a statistical test that uses the cumulative summation (cusum) of the materials balances as the test statistic. The cusum is used often because it provides an estimate of the total amount of material diverted during an accounting period. Other test statistics include a single materials balance or a weighted average of the materials balances. Our cusum test will have a single test threshold determined by the false-alarm probability—that is, the probability of concluding (because of measurement uncertainties) that nuclear material has been diverted when, in fact, no diversion has occurred. The false-alarm probability (FAP) is a measure of the significance of the test results. The FAP value used in setting up the test usually depends on the false-alarm rate that can be tolerated in the plant. The rate often depends on the consequences (shutting down the plant, perhaps) of incorrectly concluding that diversion has taken place.

Our cusum test is illustrated in Fig. 2. In the absence of diversion, we would expect the value of our test statistic—the cusum—to be zero. However, because of measurement uncertainties, the cusum value we get from our accounting data will almost never be zero. The curve in Fig. 2 represents the probability distribution of getting various cusum values when no diversion has occurred. The curve is centered at zero—the expected cusum value. The total area under the curve is 1 because the probability is 100% that the cusum has *some* value.

\*The values are the result of a Monte Carlo simulation: they were obtained from a sequence of 10 normally distributed random numbers having mean zero and standard deviation 1 by using the relationship  $MB = RN \times \sigma + D$ , where MB is the materials balance, RN is a random number, and D is the diversion.



The width of the curve is determined by the uncertainty (measured by the standard deviation  $\sigma$ ) in the cusum, which can be computed from the uncertainties in the individual materials balances. In our example, the standard deviation of the cusum is 100 g. The area under the curve to the left of any cusum value represents the probability that—in the absence of diversion—the cusum will have this value or less.

Now we must set our test threshold. We assume that a 5% false-alarm rate is acceptable. In this case, the test threshold (labeled  $Z$  in Fig. 2) is set at 165 g. The 5% of the area under the curve lying to the right of the threshold represents the false-alarm probability, labeled FAP.

We are finally ready to test our materials accounting data for evidence of diversion. The materials balances in our example sum to 152 g. Because this cusum value is less than our test threshold, we conclude that there has been no diversion.

Had our cusum value been greater than 165 g, we would have concluded that material had been diverted, but we recognize that there is a significant chance that this would be an incorrect conclusion. If there were no other considerations, the false-alarm probability could be reduced to any arbitrarily small number by increasing the value of the test threshold—but then what would happen to our ability to detect the diversion of a significant amount of material? The relationship between the test threshold and the detection probability is illustrated in Fig. 3.

Suppose 250 g of material have been diverted. Our probability curve is now centered at 250 g because this is the expected cusum value under the hypothesis that this amount of material has been diverted. The width of the curve has not changed because our cusum still has the same associated uncertainty, and the test threshold has not moved. The area un-

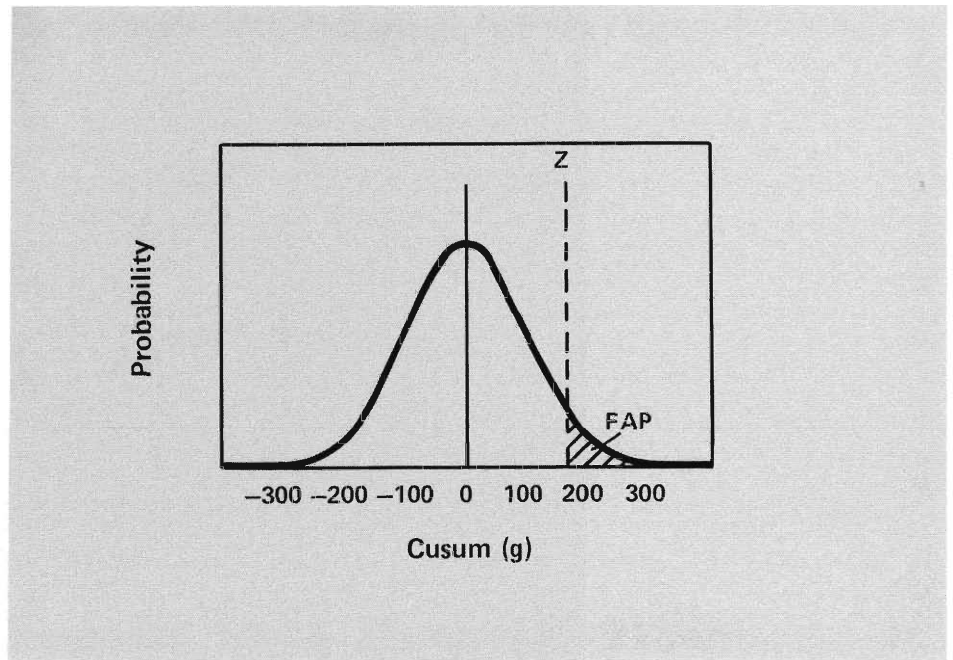


Fig. 2. A cusum test having a single test threshold,  $Z = 165$  g. The value of  $Z$  is determined by the false-alarm probability (FAP) desired. The FAP is the area under the curve to the right of  $Z$ , and in the example, this area is 0.05.

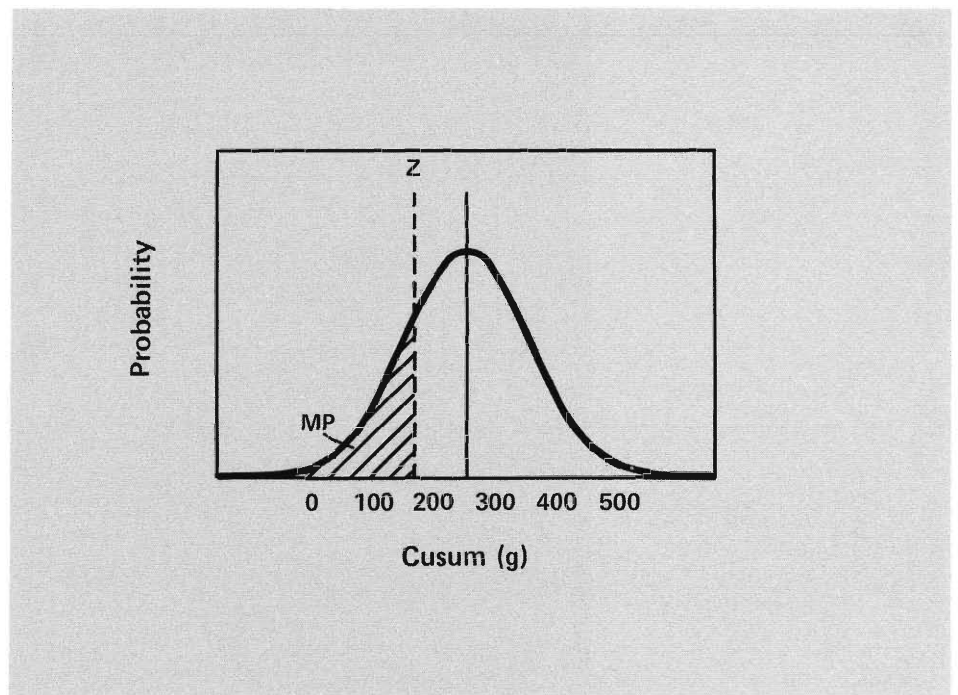


Fig. 3. A cusum test with a total diversion of 250 g and a test threshold kept at  $Z = 165$  g. The area under the curve to the right of  $Z$  is the detection probability, which, in this illustration, is 0.80. The miss probability (MP) is  $1 -$  the detection probability.

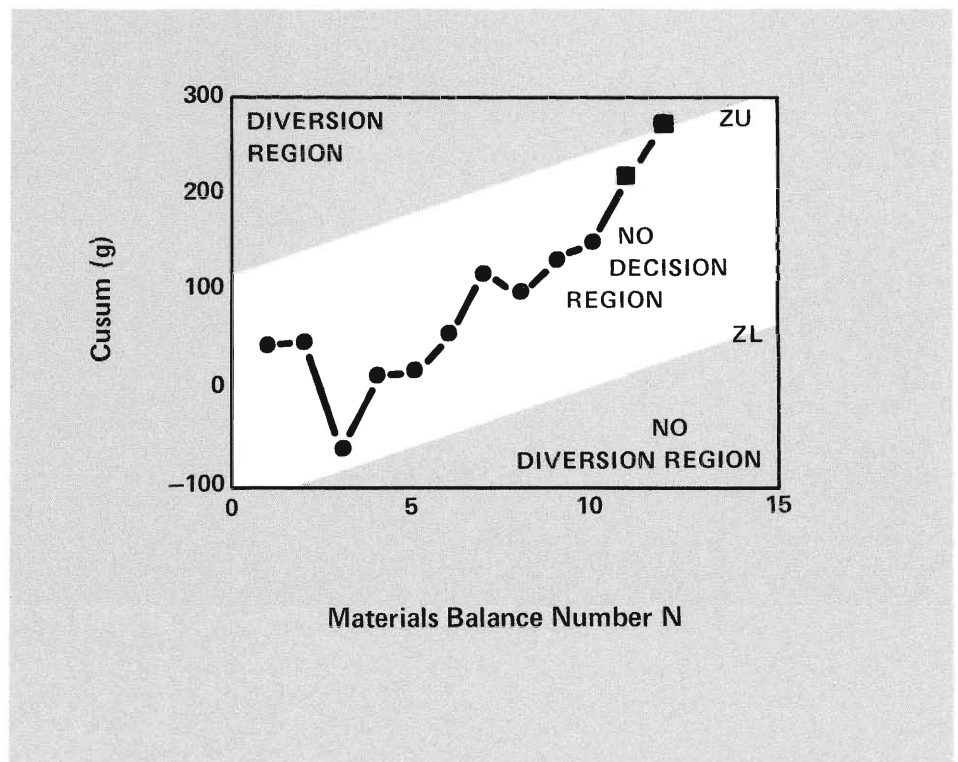
der the curve to the right of the threshold represents the probability of detecting the diversion of 250 g of material. (In Fig. 3 this area is 0.8.) The shaded area on the left of the test boundary is the miss probability (MP). It is the probability of concluding that there has been no diversion when, in fact, 250 g have been diverted; it is equal to  $1 -$  the detection probability.

In our example, the cusum is 152 g. This value is smaller than the test boundary, and we have already concluded that there was no evidence of diversion in our materials accounting data. However, if material has been diverted (as illustrated in Fig. 3), we have failed to detect this fact. With our test boundary set for a 5% false-alarm probability, we have only an 80% probability of detecting the diversion of 250 g of material. As the amount of material diverted increases, so does our ability to detect the diversion.

In our example, we considered a single cusum test of 10 materials balances and found that we could choose our test threshold based on an acceptable false-alarm rate. In practice, the 10 balances in our cusum would have been accumulated over a period of time: 10 weeks, if a materials balance were drawn at the end of each week. Perhaps we would like to test each materials balance as it becomes available or test the current cusum as each new materials balance is added. However, if we test each cusum, and if the false-alarm and miss probabilities are fixed for each test, the overall false-alarm and miss probabilities become unacceptably large after several such tests.

## Sequential Tests

Another kind of test, the sequential probability ratio test (SPRT) is particularly suited for analyzing data that become available sequentially. The SPRT allows us to guarantee that



*Fig. 4. An example of the Sequential Probability Ratio Test with diversion detected at the 12th materials balance. The test thresholds are established to detect a diversion of 25 g of material during each materials balance period with a detection probability of 95% and a false-alarm probability of 5%. Because the test allows for a no-decision region, an incorrect decision was not made.*

neither the false-alarm probability nor the miss probability will exceed desired values, no matter how long the sequence. The cusum remains an appropriate test statistic.

A sequential test has an upper and a lower threshold. Thus, at any time the test result may be that no diversion has occurred, that diversion has occurred, or that no decision can be made until more data are available. Both test thresholds depend not only on the false-alarm probability but also on the desired detection probability, the average rate of diversion, and the number of materials balances in the cusum.

A typical SPRT is illustrated by Fig. 4 for our example sequence of materials balance data. In this test, as each new materials balance is drawn, it is added to the previous cusum to obtain a new cusum value; the value is plotted against the materials balance number. The upper and lower test thresholds are the two parallel lines labeled ZU and ZL, respectively, which divide the cusum chart into three regions indicating diversion, no decision, and no diversion. If the current

cusum value falls above ZU, we conclude that diversion has taken place. If it falls below ZL, there is no evidence of diversion. If it lies between ZU and ZL, we wait for the next materials balance to be drawn. The thresholds have a positive slope because, if a pattern of protracted diversion is present, the total amount of material diverted increases as the number of materials balances in the cusum increases.

The thresholds in Fig. 4 were set for 5% false-alarm probabilities and 5% miss probabilities and for an average 25-g rate of diversion. The settings mean that we would like to detect, with at least 95% probability, the removal of 25 g of material during each balance period and that we can tolerate a false-alarm rate no greater than 5%. The circular symbols correspond to the 10 cusums computed from our example data sequence. The 10th cusum ( $N = 10$ ) lies in the region between the test thresholds so, at the time the 10th materials balance was drawn, we were unable to make a decision. Earlier, we saw that the single cusum test applied after this materials

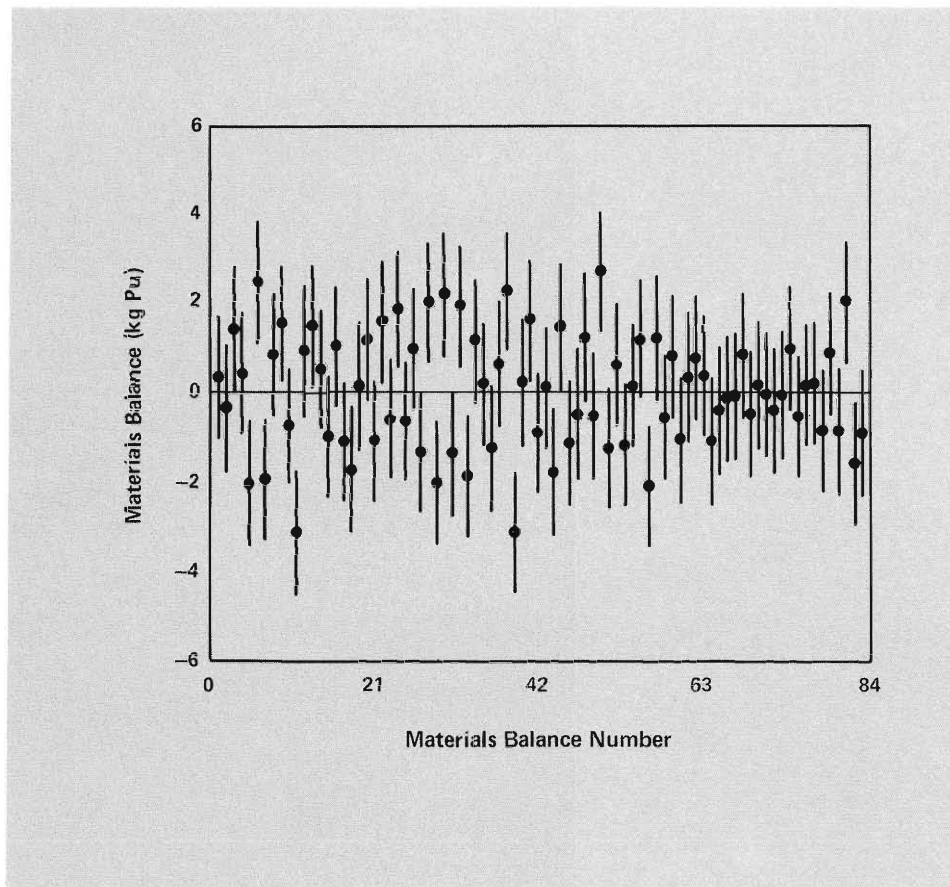
balance resulted in the conclusion that no diversion had occurred. Indeed, a similar test applied to each of the nine previous cusums would have resulted in the same conclusion. Such conclusions are incorrect for the simulated process from which our example sequence of materials balances was taken: 25 g had been diverted from each materials balance. On the other hand, the SPRT still has not permitted a decision after the 10th materials balance, and thus no incorrect decision has been made. However, after two additional materials balances are drawn, the cusum exceeds the upper test threshold, resulting in the (correct) conclusion that material has been diverted. The current ( $N = 12$ ) cusum value of 271 g provides an estimate of the total amount diverted: the true quantity was 300 g.

### Test Statistics

A variety of test statistics can be formed from the materials accounting data and tested sequentially for indications of diversion. Each statistic is based on a different assumption concerning the state of prior knowledge of the measurement errors and of the diversion strategy. Three of the most useful test statistics are the Shewhart, cusum, and uniform diversion statistics.

The Shewhart chart (Fig. 5) is the oldest graphical-display tool to be used widely by industry for process control. In the standard form, measured data are plotted sequentially on a chart where  $2\text{-}\sigma$  and  $3\text{-}\sigma$  levels are indicated. In safeguards applications, the Shewhart chart is a sequential plot of the materials balance data with  $1\text{-}\sigma$  error bars. This chart is most sensitive to large, abrupt shifts in the materials balance data.

The cusum statistic is computed after each materials balance period. It is the sum of all materials balances since the beginning of the accounting interval. Cusum charts are sequentially plotted



*Fig. 5. The Shewhart chart is a graph of sequential materials balance values and their respective materials balance numbers. For each materials balance number, the short, horizontal line gives the materials balance value, and the vertical lines above and below represent the  $\pm 1\text{-}\sigma$  deviations from this value. This chart is rather insensitive to protracted, low-level material diversions, but is sensitive to large, abrupt diversions.*

cusum values that are used to indicate small shifts in the materials balance data (Fig. 6). The cusum variance ( $\sigma_c^2$ ) is a complex combination of the variances of individual materials balances, because these balances usually are not independent. Correlation between materials balances has two principal sources. The first source is the correlation between measurement results obtained by using a common instrument calibration. The magnitudes of the associated covariance terms depend on the magnitude of the calibration error and the frequency of each instrument recalibration; omission

of these terms can cause gross underestimation of the cusum variance. The second source is the occurrence, with opposite signs, of each measured value of in-process inventory in two adjacent materials balances. As a result, only the first and last measurements of in-process inventory appear in the cusum, and only the corresponding variances appear in the cusum variance.

The Kalman filter is a statistical technique applied widely to communications and control systems for signal processing. It is a powerful tool for extracting weak signals embedded in noise.

Fig. 6. The cusum chart is a graph of the sums of all materials balances drawn from the beginning of an accounting period versus the number of materials balances in the cusum. In this chart, the short, horizontal lines give these cusum values, and the vertical lines represent  $\pm 1\sigma$  deviations from these values. Because the chart is relatively sensitive to small shifts in materials balance data, it is useful for the detection of protracted, low-level diversion.

It has been applied recently to safeguards because dynamic materials accounting systems rapidly generate large quantities of data that may contain weak signals, caused by repeated, small diversions, embedded in the noise produced by measurement errors.

The uniform diversion test (UDT) is designed to detect a small, constant diversion during each materials balance period. Estimates of the average diversion and the inventory at each time are obtained using the Kalman filter. A chart of the UDT is shown in Fig. 7.

The cusum and the UDT are complementary in several respects. The cusum estimates the *total* amount of missing nuclear material at each time step, and its standard deviation is the  $1\text{-}\sigma$  error in the estimate of the total. The UDT, on the other hand, estimates the *average* amount of nuclear material missing from each materials balance, and its standard deviation estimate is the  $1\text{-}\sigma$  error in the estimate of the average. Thus, both the cusum and the UDT search for a persistent, positive shift of the materials balance data—the cusum by estimating the total and the UDT by estimating the average.

### Alarm-Sequence Charts

The decision tests examine all possible sequences of the available materials balance data because, in practice, the time at which a sequence of diversions begins is never known beforehand. Furthermore, to ensure uniform application and interpretation, each test is performed at several levels of significance (false-alarm probability). Thus, it is useful to have a graphic display that indicates the alarm-causing sequences, specifying each by its length, time of occurrence, and significance. One such tool is the alarm-sequence chart, which has proven useful in summarizing the results of the various tests and in identifying trends in the materials accounting data.

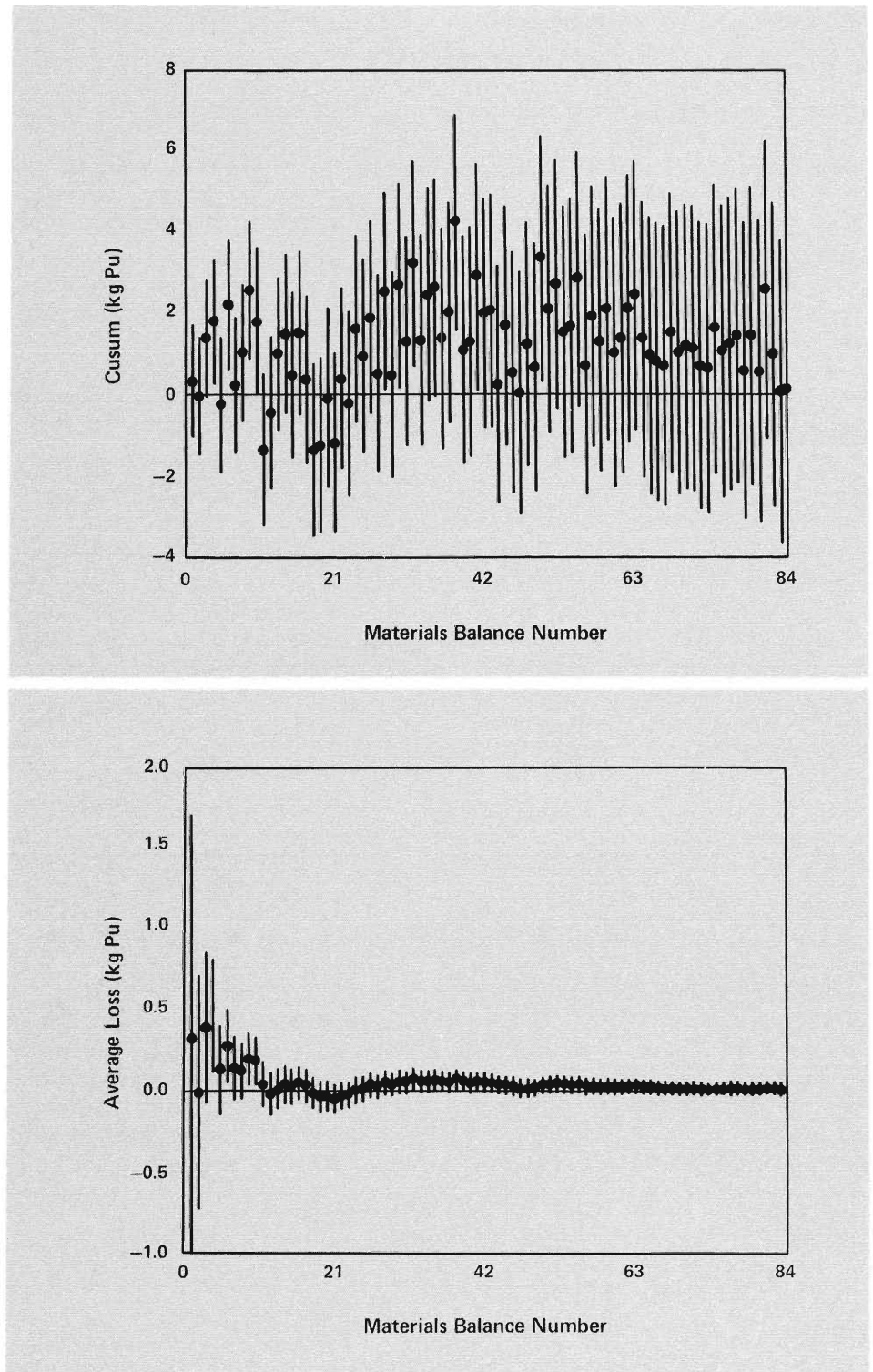


Fig. 7. In this chart of the uniform diversion test, each point on the graph is obtained by a linear combination of previous materials balances. The combination is constructed to provide an estimate of the average amount of material missing per materials balance. Like the cusum test, the UDT searches for a persistent, positive shift in the materials balance data.

An alarm-sequence chart is shown in Fig. 8.

To generate the alarm-sequence chart, each sequence in which the test statistic exceeds the upper boundary ZU and causes an alarm is assigned both a descriptor that classifies the alarm according to its significance (false-alarm probability) and a pair of integers ( $r_1, r_2$ ) that are, respectively, the indexes of the final and initial materials balances in the sequence. The alarm-sequence chart is a point plot of  $r_2$  vs  $r_1$  for each sequence that caused an alarm, with the significance range of each point indicated by the plotting symbol. One possible correspondence of plotting symbol to significance is given in Table I. The symbol T denotes sequences of such low significance (high false-alarm probability) that it would be fruitless to examine their extensions; the position of the symbol T on the chart indicates the termination point.

TABLE I

Alarm Classification for the Alarm-Sequence Chart

Classification (Plotting Symbol)	False-Alarm Probability
A	$10^{-2}$ to $5 \times 10^{-3}$
B	$5 \times 10^{-3}$ to $10^{-3}$
C	$10^{-3}$ to $5 \times 10^{-4}$
D	$5 \times 10^{-4}$ to $10^{-4}$
E	$10^{-4}$ to $10^{-5}$
F	$\leq 10^{-5}$
T	$> 0.5$

It is always true that  $r_1 \gg r_2$ , so that all symbols lie to the right of the line  $r_2 = r_1$  through the origin. Persistent data trends (repeated diversions) cause long alarm sequences ( $r_1 \gg r_2$ ), and the associated symbols on the alarm chart extend far to the right of the line  $r_2 = r_1$ .

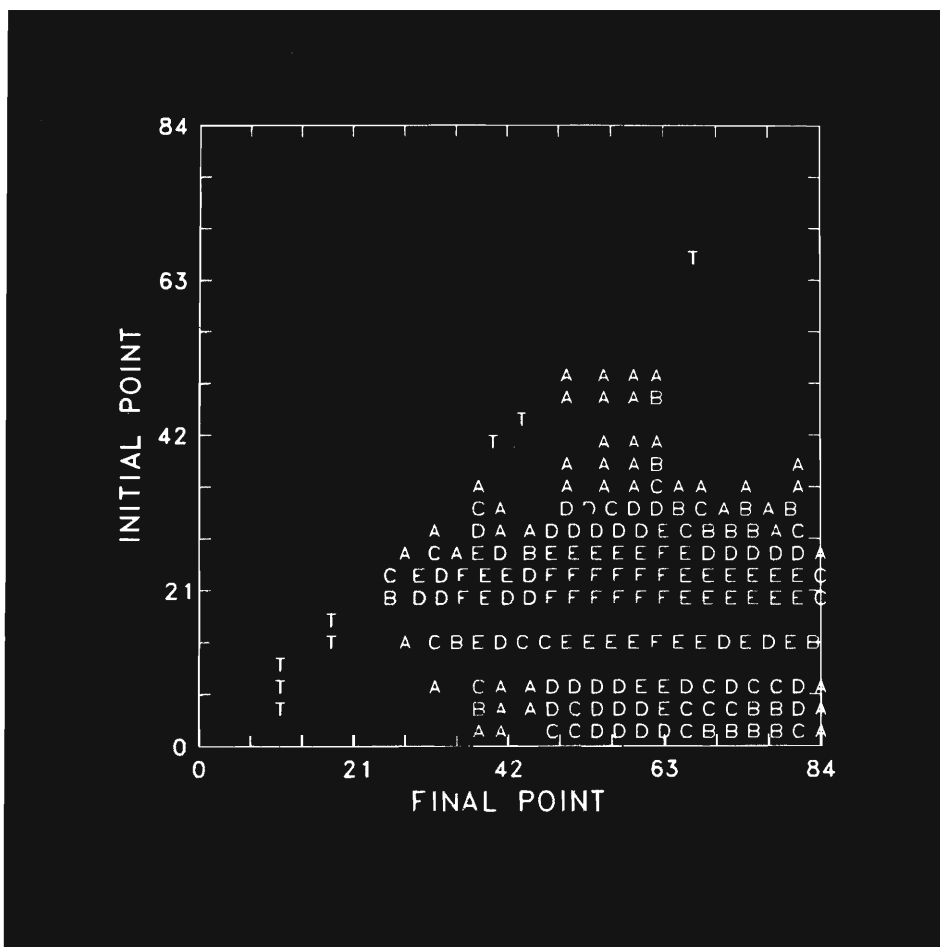


Fig. 8. An alarm-sequence chart. The false-alarm probability associated with each letter is given in Table I. To illustrate how this chart is used, consider a sequence of materials balance data beginning at balance number 21, and suppose that one of the tests gives an alarm with a false-alarm probability of  $2 \times 10^{-4}$  at balance number 30. On the alarm-sequence chart for that test, the letter D would appear at the point (30,21). Because this false-alarm probability is so small, the probability of material diversion commencing with balance number 30 is large.

Systems Performance Evaluation

An analysis of a system's performance in detecting losses of nuclear material is essential to the design of nuclear materials accounting systems. Performance measures must include the concepts of loss-detection sensitivity and loss-detection time. Because materials accounting is statistical, loss-detection sensitivity is described in terms of the probability of detecting some amount of loss while accepting the probability of some false alarms. Loss-detection time is the time required by the accounting system to reach a specified level of loss-detection sensitivity. The loss scenario is not specified in performance measures; whether the loss is abrupt or protracted, the total loss is the measure of performance. The loss-detection time refers

only to the accounting system's internal response time.

The performance of any accounting system can be described by a function

$$P [L, N, \alpha],$$

where P is the accounting system's probability of loss detection, L is the total loss over a period of N balances, and  $\alpha$  is the false-alarm probability. A convenient way of displaying system performance is a three-dimensional graph of the surface P vs L and N for a specified value of  $\alpha$ , called a *performance surface*. A single point (N, L, P) of such a surface is plotted in Fig. 9. The entire surface portrays the expected performance of an accounting system as a function of the three performance measures, loss, time, and detection probability.

Because systems performance may depend on the details of a particular diversion strategy and, therefore, on the statistical techniques used, overall performance is difficult to quantify. Fortunately, the cusum statistic does not depend on how the material was lost, but responds only to the total loss  $L$  during any time interval  $N$ . Moreover, even though the cusum test is seldom the best test for any particular scenario, it detects any loss relatively well. Consequently, it is always among the tests applied to the accounting data, and it provides a conservative, scenario-independent measure of systems performance. The performance of more powerful tests for specific loss scenarios, such as the UDT, should be compared with the cusum test performance to ensure that the cusum approximation does not generate undue pessimism.

### Measurement Error Models

Because detection sensitivity is limited by measurement errors, we must have measurement models and error estimates for various types of instrumentation to evaluate the performance of a materials accounting system. The simple measurement model given below applies when error standard deviations are expressed on a relative basis and is appropriate for measurement situations in which the associated error tends to be proportional to the quantity being measured.

$$m = M(1 + \varepsilon + \eta), \quad (1)$$

where  $m$  is the measured value of a true quantity  $M$ .

The measurement errors have been grouped in two categories, instrument imprecision  $\varepsilon$  and calibration  $\eta$ ; both are regarded as observations on random variables. The instrument imprecision  $\varepsilon$  represents the deviation of the measured value from the true quantity caused by the scatter or dispersion in a set of in-

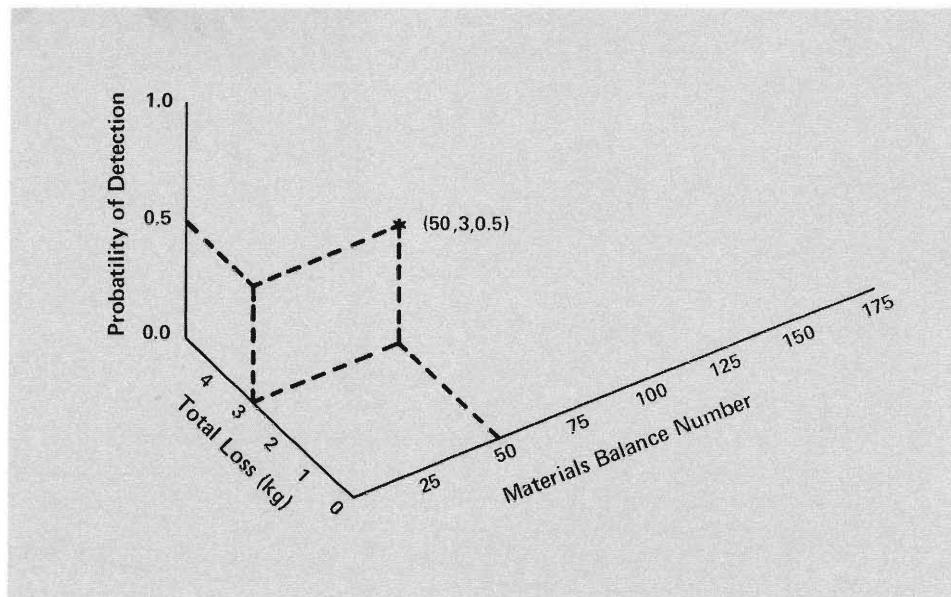


Fig. 9. The three-dimensional space of performance surfaces. Note that the Total Loss axis increases up and to the left. This graph indicates that the probability of detecting a loss of 3 kg of material at balance number 50 is 0.5.

dividual measurement results (for example, the uncertainty caused by counting statistics in NDA measurements). The calibration error  $\eta$  represents the errors that persist, unchanged, throughout a limited set of measurements as a result of the uncertainty in converting raw measurement results into the quantity of interest (for example, in converting counts to plutonium mass for NDA measurements). Calibration errors are the more difficult to estimate because they include uncertainties in standards, calibration parameters, instrument environment, and measurement control procedures.

The error random variables ( $\varepsilon$  and  $\eta$ ) have means of zero and variances of  $\sigma_\varepsilon^2$  and  $\sigma_\eta^2$ , respectively. The variance  $\sigma_m^2$  of the measured value  $m$  is given by

$$\sigma_m^2 = M^2(\sigma_\varepsilon^2 + \sigma_\eta^2). \quad (2)$$

To simulate a series of measurements from a given instrument or measurement process, a new value of  $\varepsilon$  is sampled from the appropriate  $\varepsilon$ -error distribution for each measurement, whereas a new value of  $\eta$  is sampled from the appropriate  $\eta$ -error distribution only when the instrument is recalibrated. All measurements from the same instrument having the same  $\eta$  error (calibration) are correlated. These correlations become important if an instrument cannot be

recalibrated frequently, and they may dominate the materials balance error. The covariance (a measure of the correlation) between the  $i^{\text{th}}$  and  $j^{\text{th}}$  measurements is given by

$$\sigma_{ij} = M_i M_j \sigma_\eta^2. \quad (3)$$

### An Ideal Process

A simple example illustrates dynamic materials accounting concepts and principles. Figure 10 represents an ideal process having a daily throughput of 50 kg of nuclear material consisting of twenty-five 2-kg batches. The in-process inventory is 25 kg, and the residual holdup is 5 kg after shutdown and cleanout, which occur once each month. The entire process is contained in a single materials balance area (Fig. 10a); the storage areas for feed and product are located in separate accounting areas (not shown).

Figures 10b and 10c show two possible divisions of the process into accounting areas for dynamic accounting purposes. In Fig. 10b, the process is divided into a *series* of five smaller accounting areas. To use this division, we would measure transfers of nuclear materials between adjacent accounting areas and the in-process inventory in each area. In Fig. 10c, the process is divided into five *parallel* areas

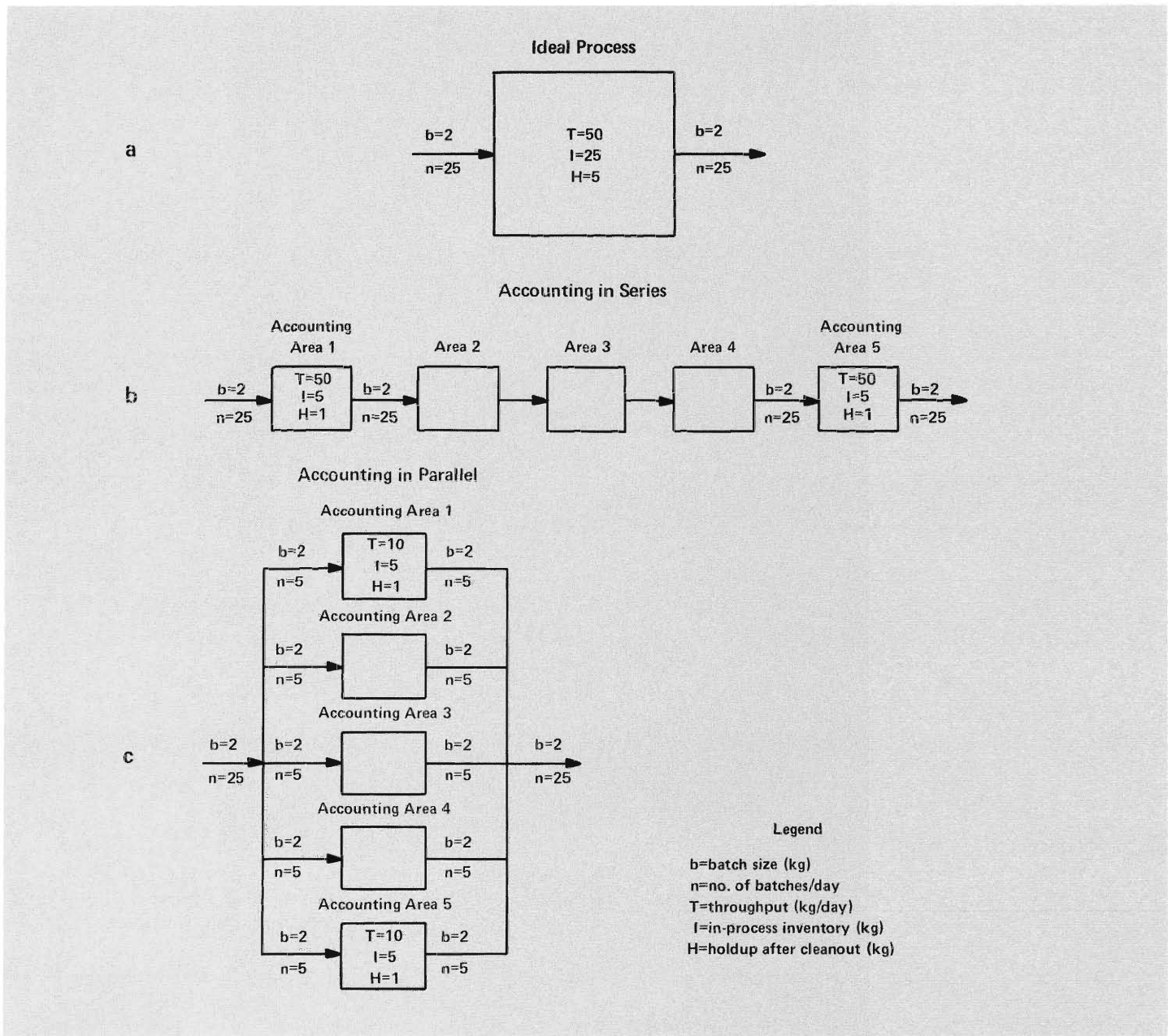


Fig. 10. Materials balance accounting areas for an ideal process. In a, the entire process is contained in a single materials balance area. In b the process is divided into five accounting areas in series; one measurement suffices to determine the transfer out of one area and into the next. The entire daily throughput passes through each area. In c, the five accounting areas are in parallel, that is, the throughput per day in each area is one-fifth of the total throughput; the transfers into and out of each area are measured independently.

corresponding to five separate process lines in parallel. In this case, we would measure the input, output, and inventory of each area separately. In practice, the division of the process depends on its configuration.

We can calculate measurement errors in dynamic materials balances applied to the ideal process by using the measurement model described in Eqs. (1)-(3). For an accounting period during which  $N$  batches are processed, the dynamic

materials balance  $MB_N$  for one accounting area is given by

$$MB_N = \Delta I_N + T_N, \quad (4)$$

where  $\Delta I_N$  is the net change in the inventory and  $T_N$  is the net transfer of nuclear material (inputs minus outputs) across the accounting area. If there were no measurement errors,  $MB_N$  would be exactly zero and, if the process were operated at steady state,  $\Delta I_N$  and  $T_N$

also would be zero.

Measurement errors produce an uncertainty in  $MB_N$  having a variance  $\sigma_{MB}^2$  (assuming no correlation between transfer and inventory measurements) given by

$$\sigma_{MB}^2 = \sigma_{\Delta I}^2 + \sigma_T^2. \quad (5)$$

Understanding the behavior of the inventory-change and net-transfer variances,  $\sigma_{\Delta I}^2$  and  $\sigma_T^2$ , is basic to the ef-



fective design of a materials measurement and accounting system.

If the initial and final inventories,  $I_0$  and  $I_N$ , are measured during the same calibration period (that is, if they have the same  $\eta$  error), the variance  $\sigma_{\Delta I}^2$  of the net inventory change  $\Delta I$  is given by

$$\sigma_{\Delta I}^2 = (I_0^2 + I_N^2) \sigma_{\epsilon I}^2 + (I_0 - I_N)^2 \sigma_{\eta I}^2, \quad (6)$$

where  $\sigma_{\epsilon I}^2$  and  $\sigma_{\eta I}^2$  are the  $\epsilon$ - and  $\eta$ -error variances of the inventory measurements. Note that if  $I_0$  and  $I_N$  are equal,  $\sigma_{\Delta I}^2$  has the minimum value

$$\sigma_{\Delta I}^2 = 2I_0^2 \sigma_{\epsilon I}^2. \quad (7)$$

For a large class of process equipment, efficiency and economy dictate that the in-process inventory be held nearly constant during normal operation. Near-steady-state operation benefits materials accounting by reducing the materials balance uncertainty because the condition  $I_0 \cong I_N$  implies that the dependence of  $\sigma_{MB}$  on  $\sigma_{\eta I}$  is weak [Eq. (6)]. Hence, a well-known value for  $\sigma_{\eta I}$  is not required. This result is important because calibration of in-process inventory measurements may be difficult, especially for process equipment located in high-radiation fields behind heavy shielding. The ideal process is assumed to satisfy the steady-state condition so that Eq. (7) holds. The inventory measurement error ( $\sigma_{\epsilon I} = 10\%$  in this example) limits the dynamic accounting sensitivity over short accounting periods.

The variance  $\sigma_T^2$  of the net material transfer  $T$  is given by

$$\sigma_T^2 = 2Nb^2 (\sigma_{\epsilon b}^2 + \sigma_{\eta b}^2) + 2N(N-1)b^2 \sigma_{\eta b}^2, \quad (8)$$

where  $b$  is the input and output batch size, and  $\sigma_{\epsilon b}^2$  and  $\sigma_{\eta b}^2$  are the  $\epsilon$ - and  $\eta$ -error variances of the batch transfer measurements. For simplicity of presen-

tation, the error variances of input and output batch measurements have been set equal in value (hence the factor of 2), but the two measurements are statistically independent; that is, they are not correlated.

The first term in Eq. (8) occurs whenever  $N$  input and  $N$  output batches are measured during the accounting period and is present even if the transfer measurements are uncorrelated. The second term accounts for pair-wise correlations among the transfer measurements [Eq. (3)]. The transfer measurements are correlated primarily because the instruments are not recalibrated during the accounting period. Note that the number of pair-wise correlations increases approximately as  $N^2$ ; if  $N$  is sufficiently large, correlations make the dominant contribution to  $\sigma_T^2$ .

The effect of measurement correlations can be reduced by recalibrating the transfer-measuring instruments. If the instruments are calibrated  $K$  times during the accounting period, and if  $n_k$  is the number of batches processed between the  $k^{\text{th}}$  and  $(k+1)^{\text{th}}$  calibrations, then  $\sigma_T^2$  is given by

$$\sigma_T^2 = 2Nb^2(\sigma_{\epsilon b}^2 + \sigma_{\eta b}^2) + 2b^2\sigma_{\eta b}^2 \sum_{k=1}^K n_k(n_k - 1), \quad (9)$$

where

$$N = \sum_{k=1}^K n_k. \quad (10)$$

The number of correlation terms in this case increases approximately as  $\sum n_k^2$  rather than as  $N^2$ .

The effect on  $\sigma_T$  of daily versus monthly recalibration of transfer-measuring instruments is shown in Fig. 11. The relative standard deviation (RSD), which is  $\sigma_T$  divided by the throughput  $Nb$ , is plotted as a function of the number  $N$  of processed batches. Values of  $\sigma_{\epsilon b}$  and  $\sigma_{\eta b}$  have been taken to

be 2% and 0.5%, respectively. The net-transfer RSD varies as  $|(\sigma_{\epsilon b}^2 + \sigma_{\eta b}^2)/N|^{1/2}$  for a small  $N$  and as  $(\sigma_{\eta b}^2/K)^{1/2}$  for a large  $N$ , that is, when the transfer correlations are dominant.

Correlations between transfer measurements limit the sensitivity of dynamic materials balances over relatively long accounting periods. Therefore, the parameters  $\sigma_{\eta b}$  and  $K$  are especially important. The value of  $\sigma_{\eta b}$  depends primarily on the measurement control procedures and on the quality of available calibration standards, whereas the value of  $K$  depends on how often the transfer-measuring instruments are recalibrated. Adequate measurement controls must include well-characterized standards for the transfer measurements. Further, provision must be made for sufficiently frequent recalibration of the transfer-measuring instruments.

Table II contains kilogram values of the standard deviation  $\sigma_{MB}$  of dynamic materials balances calculated for the ideal process. Results are given for four accounting periods: one batch, 1 day, 1 week, and 1 month (30 days), and for two transfer calibration periods, 1 day and 1 month. The inventory-change and net-transfer components of  $\sigma_{MB}$  are given separately. Calculated values are shown for one accounting area in a

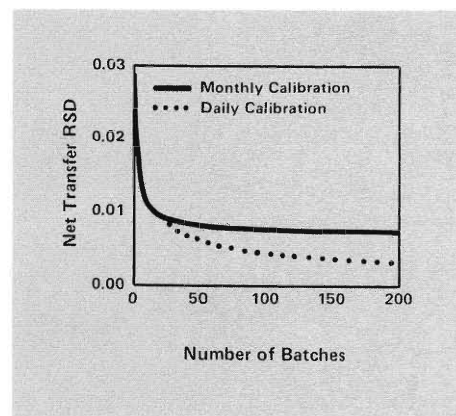


Fig. 11. Effect of calibration on transfer measurement errors.

TABLE II

## Dynamic Materials Accounting in an Ideal Process

Accounting Period	Standard Deviation (kg)					
	Monthly Recalibration			Daily Recalibration		
	Series	Parallel	Total	Series	Parallel	Total
1 Batch						
Inventory change	0.71	0.71	1.58	0.71	0.71	1.58
Net transfer	0.06	0.06	0.06	0.06	0.06	0.06
Materials balance	0.71	0.71	1.58	0.71	0.71	1.58
1 Day						
Inventory change	0.71	0.71	1.58	0.71	0.71	1.58
Net transfer	0.45	0.14	0.45	0.45	0.14	0.45
Materials balance	0.84	0.72	1.64	0.84	0.72	1.64
1 Week						
Inventory change	0.71	0.71	1.58	0.71	0.71	1.58
Net transfer	2.59	0.60	2.59	1.20	0.38	1.20
Materials balance	2.68	0.93	3.03	1.39	0.80	1.98
1 Month						
Inventory change	0.14	0.14	0.32	0.14	0.14	0.32
Net transfer	10.72	2.23	10.72	2.48	0.79	2.48
Materials balance	10.72	2.24	10.72	2.48	0.81	2.50

series arrangement, one accounting area in a parallel arrangement, and the total process (see Fig. 10). Note that the data for the total process are a synthesis of the data from the smaller accounting areas. In practical application the capability of combining the same dynamic accounting data in different ways to form materials balances for various accounting envelopes provides obvious safeguards advantages that can be exploited by the materials accounting system software.

The data in Table II support the following conclusions. For relatively short accounting periods, the materials balance standard deviation ( $\sigma_{MB}$ ) is determined primarily by the size of the inventory (I) and the inventory instrument-precision RSD ( $\sigma_{e1}$ ). For longer accounting periods,  $\sigma_{MB}$  is determined by the sizes of the transfers (b), the transfer calibration-error RSD ( $\sigma_{nb}$ ), and the number (K) of transfer-instrument recalibrations.

Reduction of in-process inventory and accessibility of process equipment for inventory measurements are important design considerations. Since the use of parallel process lines reduces throughput and inventory in each accounting area for the same total plant throughput, it often can markedly improve materials accounting sensitivity. This practice, however, requires independent instrumentation for each accounting area. Large-capacity tanks present special accounting problems, and strict surveillance (process monitoring) measures, in addition to materials accounting measures, should be considered for them. Processing relatively small batches and operating the process near steady state generally enhance the capability of materials accounting.

Materials measurements require rapid in-line or at-line assay techniques that provide precise inventory measurements and accurate transfer measurements, and provision for frequent recalibration

of the transfer-measuring instruments. The period between physical inventories should be coupled to the buildup of transfer-measurement correlations; that is, after the materials-balance error standard deviation for the accounting area becomes unacceptably large, a physical inventory is necessary to "restart" the dynamic accounting system.

### Application to a Reprocessing Plant

We have applied the principles of a dynamic materials accounting system to a real plant, the fuel reprocessing facility built by Allied-General Nuclear Services (AGNS) at Barnwell, South Carolina. Since this plant is not yet operating, process and materials balance data are simulated for analysis. AGNS is designed to receive and process irradiated power-reactor fuel containing  $^{235}\text{U}$  and plutonium. The plant capacity, which is 50 kg plutonium/day, is typical of plants that will be required in the 1990s to support a mature nuclear industry. The AGNS plant uses the Purex recovery process, a process in large-scale use for over 25 years and still used, with minor variations, by most of the reprocessing plants now operating or planned throughout the world. The products are concentrated uranyl nitrate and plutonium nitrate solutions.

Spent-fuel assemblies arrive at the plant by rail or truck and remain in a fuel-storage pool while awaiting processing. The fuel elements are chopped into small pieces, and the fuel is dissolved with a concentrated nitric acid solution. Following dissolution, a paraffin hydrocarbon solvent is used to separate most of the fission products from the plutonium and uranium. The solvent stream containing the plutonium and uranium then enters a partitioning step, where the bulk of the uranium is separated from the plutonium. The uranium stream is further decontaminated, concentrated, and passed

through silica-gel beds to remove traces of zirconium and niobium. The plutonium stream is also further purified, concentrated, and stored to await conversion to plutonium oxide. The wastes from the processes are treated in either liquid- or solid-waste processing systems. Off-gases are treated before being vented to the atmosphere.

Nuclear materials in a fuel-reprocessing facility are present in several different physical and chemical forms and also at different levels of radioactivity. Therefore, the accessibility and desirability of nuclear material for diversion will vary throughout the plant.

We can illustrate this point by the following example. Say we wish to determine the amount and form of material required to divert 1 kg of plutonium from various parts of the process. In the chop and leach portion of the plant, where the fuel enters the recovery process through dissolution in nitric acid, a divertor would need about 330 L of solution to obtain 1 kg of plutonium. Furthermore, because of the fission product content of this solution, the divertor would receive an immediate lethal dose of radiation without the use of very heavy shielding. This portion of the process, then, would be a poor choice for diversion of nuclear material.

If we proceed farther along the process to the chemical separations portion, we find the diversion of plutonium somewhat more attractive. Here the uranium and plutonium are separated from each other, and the fission products are partially removed from solution. The radiation level of the solution in this area is still high, but not immediately lethal. To obtain 1 kg of plutonium, about 200 L of solution must be drained from the storage and sampling tank. This amount is less than that required from the chop and leach portion of the process, but it is still considerable.

Still farther along in the process stream, after chemical concentration of

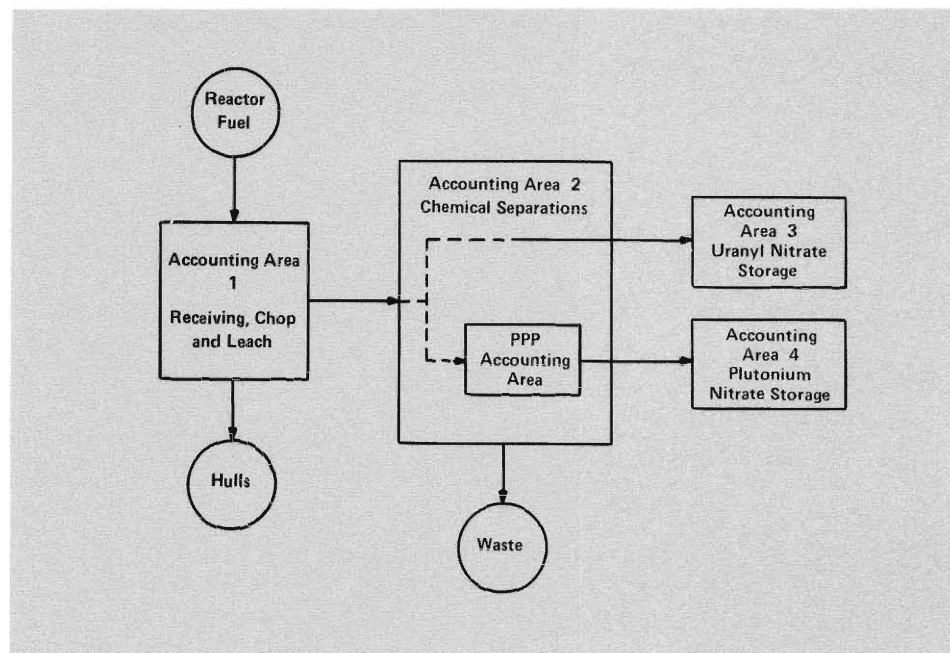


Fig. 12. Accounting areas in the AGNS facility. The plutonium purification process (PPP) accounting area has a total plutonium inventory of about 40 kg and a plutonium throughput of 50 kg/day.

the plutonium nitrate product, only about 4 L of solution would be required to obtain 1 kg of plutonium, and the radioactivity level is so low that no special shielding precautions would be necessary. This portion of the process is especially attractive to a divertor.

The example shows that a *graded* materials accounting system is both useful and economical in developing safeguards for a reprocessing facility. Where the accessibility and attractiveness of nuclear material are low, a safeguards system need not be so stringent. However, the plutonium product near the end of the recovery process is of paramount importance, and rigorous materials control and accounting must be maintained in this area.

Dividing the reference process into several materials accountability areas should be advantageous for materials accounting. For the AGNS facility, we have separated the process into the four accounting areas shown in Fig. 12. In Area 1, fuel is received for storage in pools and, as demanded by the process flow, is removed for chopping and dissolution. The concentrations of nuclear material at the downstream end of this area are about 300 g uranium/L and 3 g plutonium/L. Material of this concentration is transferred to Area 2, where the

plutonium and uranium nitrates are separated, and fission products are removed from solution. At the downstream end of Area 2, product batches of uranium are concentrated to 375 g uranium/L and transferred to Area 3 for storage, while product batches of plutonium are concentrated to 250 g plutonium/L and transferred for storage into Area 4.

For the purposes of this discussion, we will concentrate on the chemical separation process in Area 2 and will restrict our attention to the plutonium purification process (PPP) within that area (Fig. 13 and Tables III and IV).

### Materials Accounting for Plutonium Purification Process

Flow and concentration measurements at the 1BP tank (input) and 3P concentrator (output) isolate the PPP as an accounting area. In addition, acid recycles (2AW, 3AW, 3PD) and organic recycle (2BW, 3BW) must be monitored for flow and concentration, and the total in-process inventory must be estimated. Table V gives the required measurement points and some possible measurement methods and associated uncertainties.

The relative precision of dynamic volume measurements is estimated to be 3% (1  $\sigma$ ) for the 1BP tank, threefold

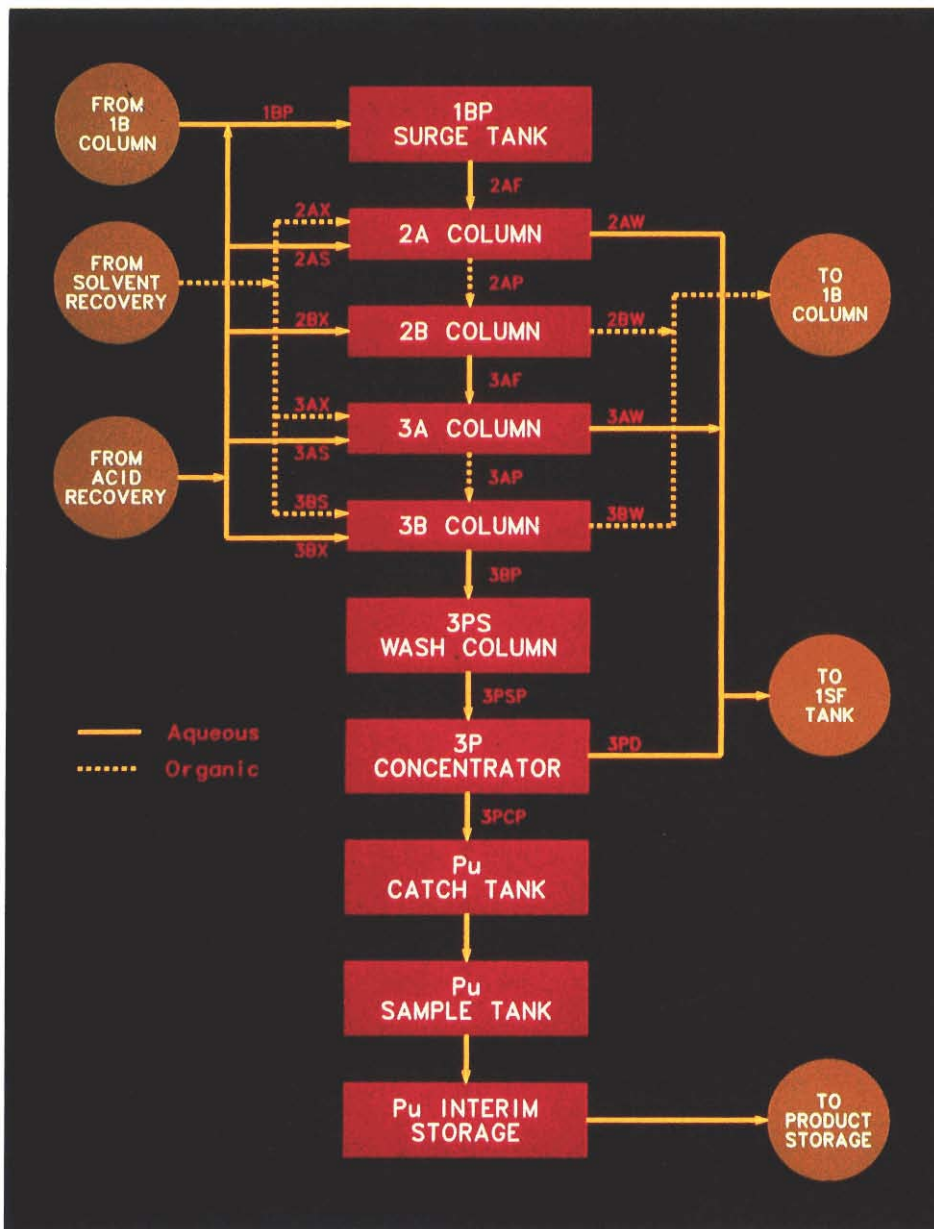


Fig. 13. The plutonium purification process (PPP) accounting area. Tables III, IV, and V describe the plutonium concentrations and flow rates, the in-process inventories, and the materials accounting measurements, respectively, that were used to design and evaluate the performance of a materials accounting system for this area.

more than for a conventional physical-inventory measurement because liquid flows into and out of the tank continuously during processing. Dynamic estimates of plutonium concentration in the 1BP and 3P concentrator tanks can be obtained from direct, in-line measurements (by absorption-edge densitometry, for example) or from combinations of adjacent accountability and process-control measurements.

Pulsed columns 2A and B and 3A and B are used to purify the plutonium. In the AGNS design, the columns are fully instrumented for process control, so that

measurements of plutonium concentration and inventory are possible. Relative precision for column-inventory measurements is estimated to be in the range of 5-20% ( $1\sigma$ ). The 20% limit appears to be conservative in terms of discussion with industry and DOE personnel. A precision of 10% should be practicable with the use of current process-control instrumentation. Improvements toward the 5% figure (or better) will require additional research and development to identify optimum combinations of additional on-line instrumentation and improved models of column behavior.

TABLE III  
Concentrations and Flow Rates  
in the PPP

Stream	Flow (L/h)	Plutonium Concentration (g/L)
1BP	400	5
3PCP	8	250
2AW	500	trace
3AW	215	0.1
3PD	32	trace
2BW	150	trace
3BW	105	trace

TABLE IV  
In-Process Inventories in Tanks  
and Vessels of the PPP<sup>a</sup>

Identification <sup>b</sup>	Volume (L)	Plutonium Inventory (kg)
1BP tank	1500	7.4
2A column	700	4.6
2B column	500	2.8
3A column	600	5.4
3B column	440	4.8
3PS wash column	20	1.2
3P concentrator	60	15

<sup>a</sup>These values are not flow sheet values of any existing reprocessing facility but represent typical values within reasonable ranges of a workable flow sheet.

<sup>b</sup>See Fig. 13.

Waste and recycle streams from the columns and the concentrator in the PPP are monitored by a combination of flow meters and NDA alpha detectors for plutonium concentration. The alpha monitors are already used for process control in the AGNS design and require only modest upgrading (primarily calibration and sensitivity studies) to be

used for accountability. Flow measurements in the waste and recycle streams can be simple and relatively crude (5-10%) because the amount of plutonium is small.

Because the PPP processes nuclear material semicontinuously, materials balances could be drawn as often as once per hour. However, our studies have shown that an 8-hour balance period gives a reasonable diversion detection sensitivity and matches normal process operating conditions. The following results are based on drawing materials balances every 8 hours.

### Performance Evaluation

Simulated results of diversion detection for 1 month of process operation in the PPP accounting area are given in Figs. 14-16. The figures show results obtained with the Shewhart, cusum, and UDT decision analysis tests. Each figure also shows plots of the test statistic and the corresponding alarm chart for the case of no diversion (upper) and for the case of diversion (lower). In the plots of the test statistics, the horizontal marks indicate the values of the statistics, and the vertical lines are 1- $\sigma$  error bars about those estimates. In each case a strategy of low-level uniform diversion is simulated during the 21st to 63rd materials balances. The Shewhart test is so insensitive to this pattern of diversion that no alarms appear on its alarm chart, while alarms appear on the charts almost immediately after diversion begins for both the cusum and UDT tests.

In the course of evaluation, many such sets of charts are examined so that the random effects of measurement errors and normal process variability can be assessed; that is, we perform a Monte Carlo study to estimate the sensitivity to diversion. However, in applying decision analysis to data from a facility operating under actual condi-

TABLE V

Materials Accounting Measurements for the PPP

Measurement Point	Measurement Type	Instrument Precision ( $1\sigma$ , %)	Calibration Error ( $1\sigma$ , %)
1BP, 3PCP streams	Flow meter	1	0.5
	Absorption-edge densitometry	1	0.3
1BP surge tank	Volume	3	a
	Absorption-edge densitometry	3	a
2A, 2B, 3A, 3B columns	See text	5-20	a
2AW, 2BW, 3AW, 3BW, 3PD streams	Flow meter	5	1
3PS column	See text	5-20	a
3P concentrator	Volume (constant)	a	a
	Absorption-edge densitometer	1.5	a

<sup>a</sup>Not important.

TABLE VI

Plutonium Purification Process Diversion Detection Sensitivity<sup>a</sup>

Accounting Period	Number of Materials Balances	Total at Detection (kg Pu)	
		Case 1 <sup>b</sup>	Case 2 <sup>c</sup>
.8 h	1.	4.2	2.6
.1 day	3.	4.4	2.9
.1 wk	21.	9.7	5.3
.2 wk	42.	17.8	7.1
.1 month	84.	34.8	9.7

<sup>a</sup>Detection probability = 0.5, false-alarm probability = 0.001.

<sup>b</sup>No recalibrations within the accounting period, and 10% estimates of column inventories.

<sup>c</sup>Two-day recalibrations of input/output concentration and flow measuring instruments, and 5% estimates of column inventories.

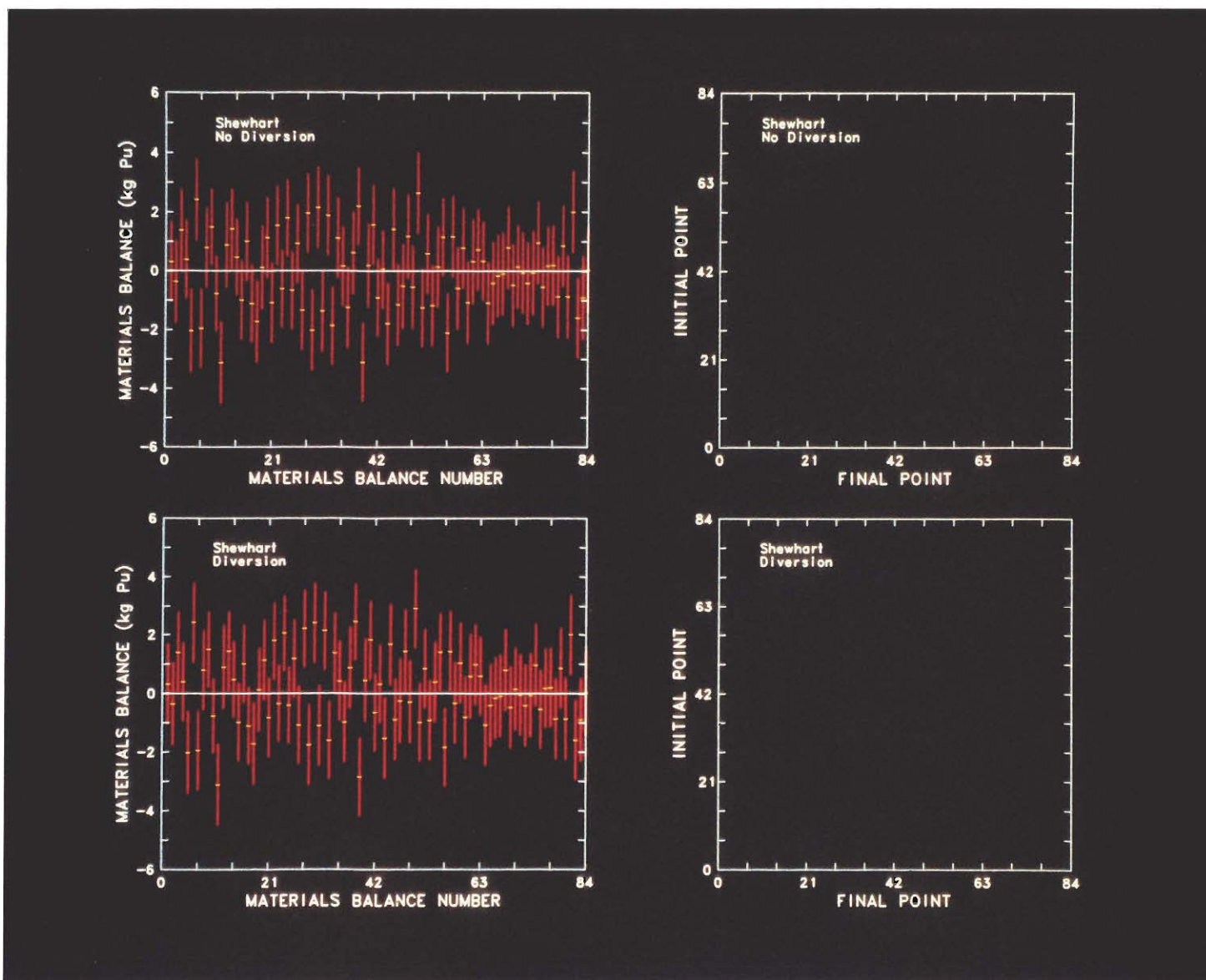


Fig. 14. No-diversion and diversion Shewhart charts with associated alarm charts. The diversion strategy is a low-level, uniform diversion of material between balances 21-63. That diversion is occurring is not obvious from the diversion Shewhart chart and no alarms show on the alarm chart.

tions, only one set of data will be available for making decisions, rather than the multiple data streams generated from a simulation. In particular, direct comparison of charts with and without diversion, as shown here, will be impossible. The decision-maker will have to extrapolate from historical information and from careful process and measurement analysis to determine whether diversion has occurred.

The results of the evaluation for two measurement strategies are given in Table VI. The diversion detection sensitivity for 1 week and less is limited by the uncertainties in the in-process inventory, which is both large ( $\approx 40$  kg of plutonium) and difficult to measure. For

longer times, the sensitivity is limited by the systematic errors in the transfer measurements.

The short-term sensitivity to diversion could be improved by modifying equipment at the codecontamination-partitioning step. In the Purex process, plutonium and uranium are coextracted from the dissolver solution and then selectively extracted in what are called solvent-extractor contactors. In the reference facility, a series of pulsed-column contactors are used for the uranium-plutonium partitioning. These contactors have a relatively large plutonium inventory ( $\approx 25$  kg), which not only varies under normal operating conditions but also is not amenable to ac-

curate measurement. Replacing the pulsed-column contactors with centrifugal contactors would decrease the plutonium inventory by an order of magnitude and improve the short-term (inventory-dominated) diversion sensitivity.

However, at about 1 month, the diversion sensitivity becomes throughput dominated; that is, errors in measuring the plutonium throughput determine the detection sensitivity. Even the best-case 1-month sensitivity (9.7 kg) seems rather large. However, the throughput of this facility is also large (1400 kg plutonium/month), so the sensitivity is 0.7% of throughput, which is really rather good. For this facility, though, im-

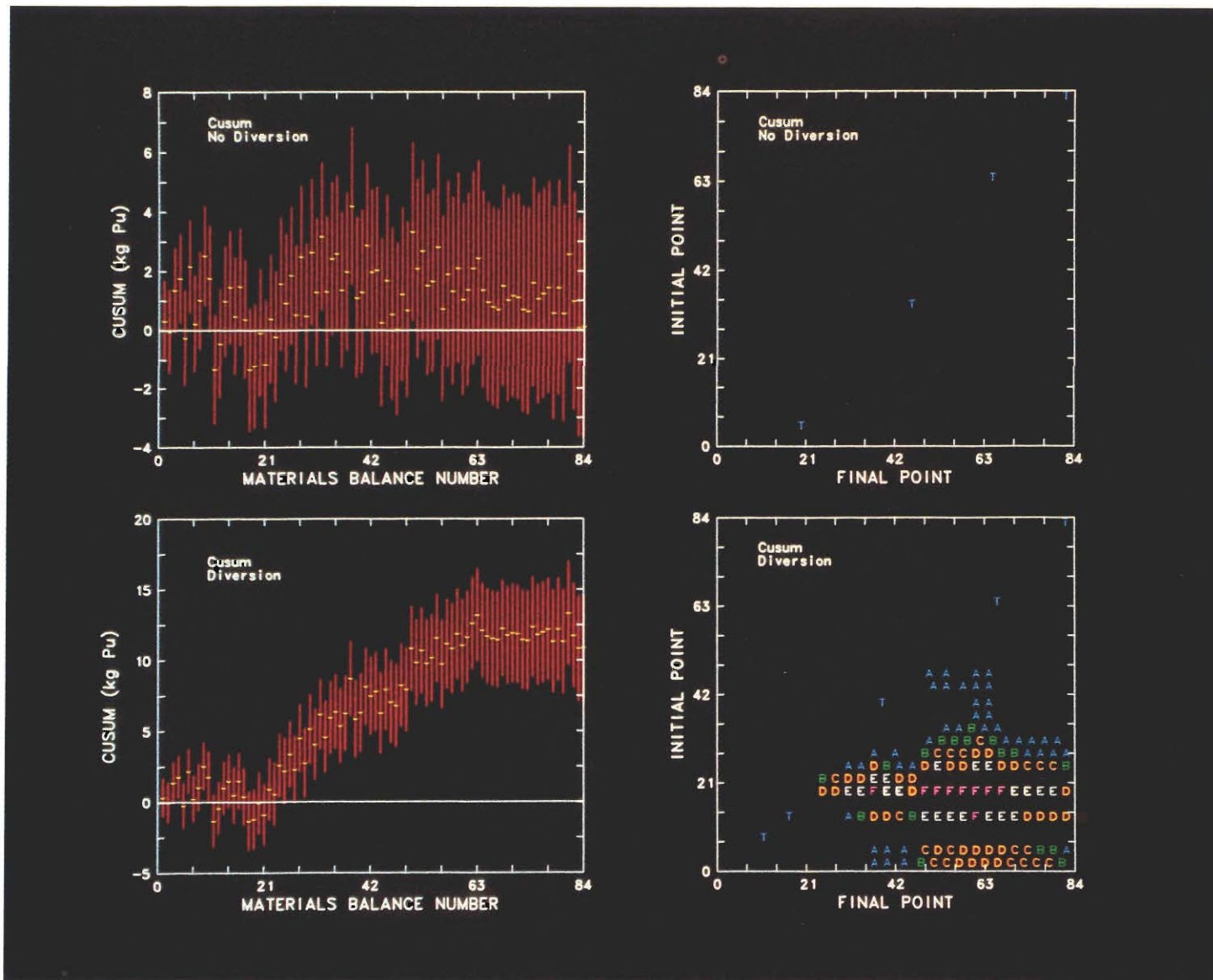


Fig. 15. No-diversion and diversion cusum charts with associated alarm charts. The diversion strategy is as described for Fig. 14. It is obvious from the diversion cusum chart that material is being diverted at about balance number 23. The cusum increases, indicating a continued diversion of material, until about balance number 63. Subsequently, the cusum maintains a roughly steady, high ( $\approx 12$ -kg) value, indicating the total loss of a fixed quantity of material. To confirm these observations, the associated alarm charts begin to show alarms having small values of false-alarm probability ( $\approx 10^{-3}$ ) at initial and final balance numbers of about 21 and 23. Balance numbers higher than 63 have high ( $\approx 0.5$ ) false-alarm probability, which indicates that material is probably not being diverted.

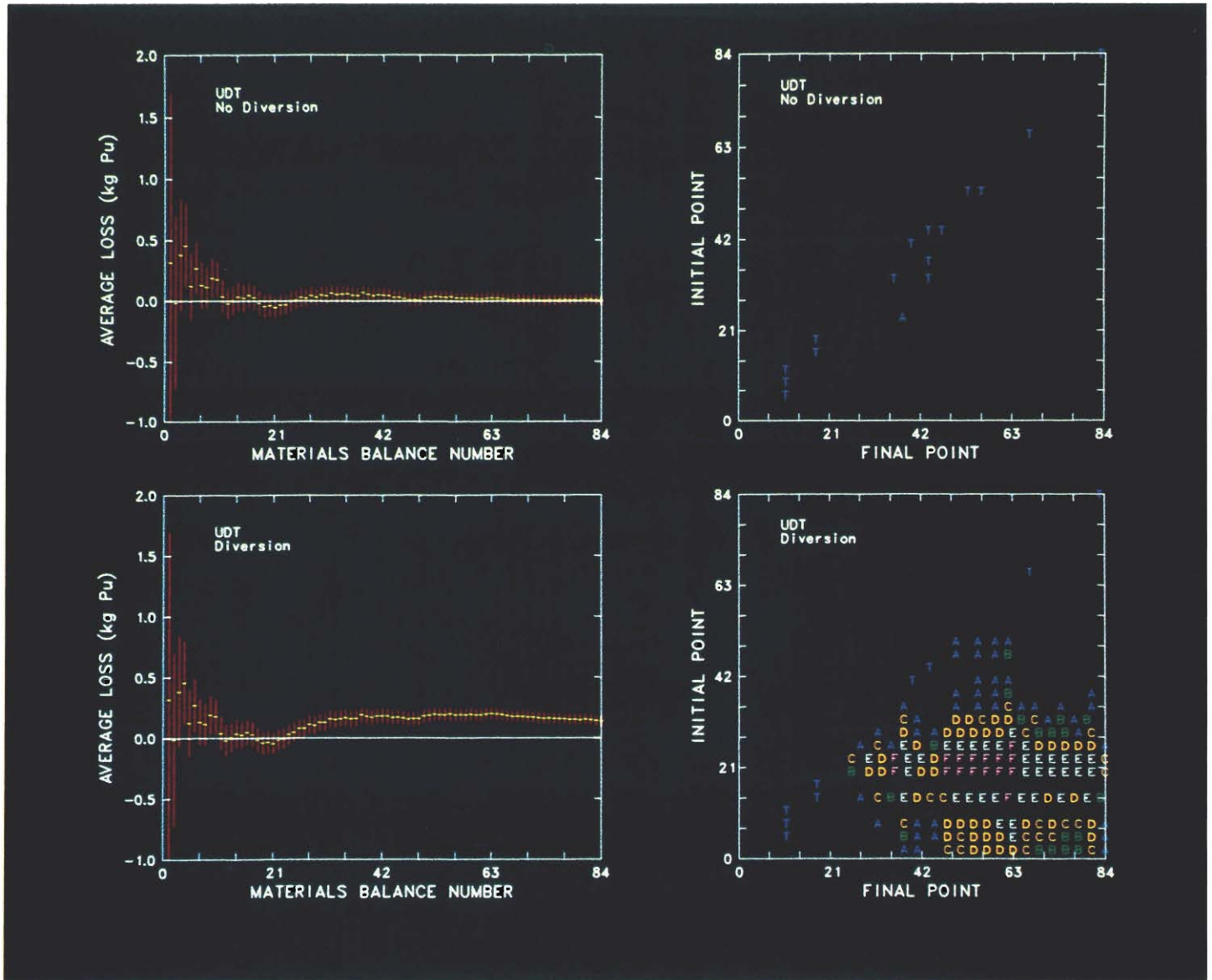


Fig. 16. No-diversion and diversion UDT charts with alarm charts. Again, the diversion strategy is as described for Fig. 14. The UDT diversion chart shows diversion commencing at about balance number 23, and the average material loss does not begin to decline until after number 63, when diversion has ceased. The alarm chart confirms these observations by the appearance of alarms at about balance numbers (21,23) and the absence of alarms in the vicinity of numbers (63,63).



provement in the long-term diversion sensitivity can be obtained only by better measurements of the throughput and better control of the correlated errors (such as calibration errors) in the throughput measurements.

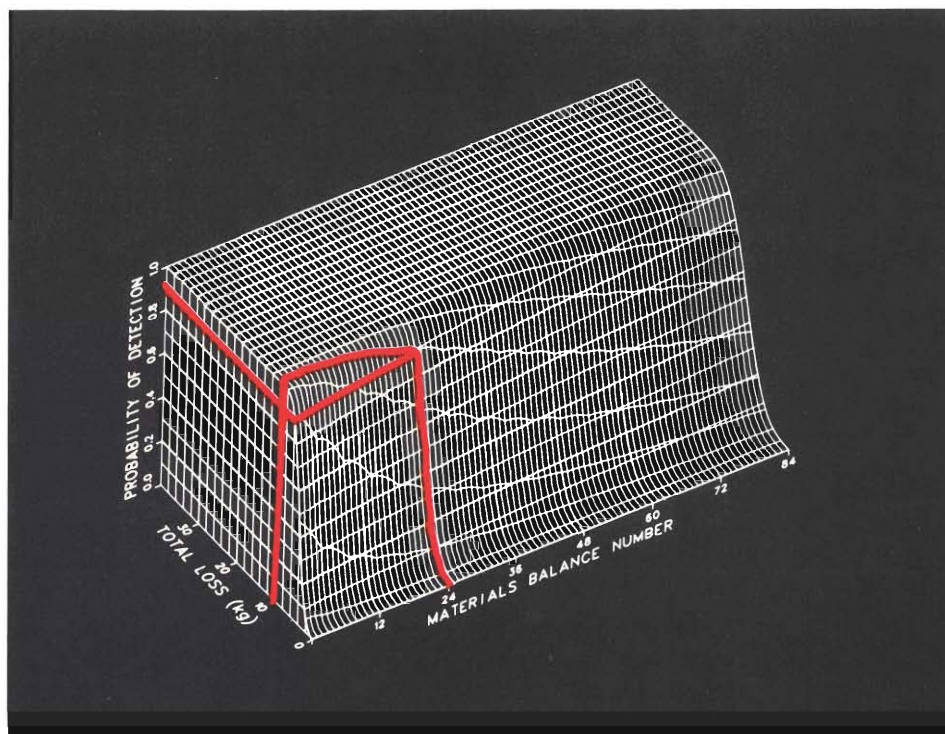
Figures 17a and 17b show examples of cusum performance surfaces from the simulated materials accounting data used to generate Table VI. Results for Case 1 (the worst case) are shown in Fig. 17a, and results for Case 2 (the best case) are shown in Fig. 17b. The figures illustrate the use of cusum performance surfaces in the design and evaluation of materials accounting systems. The improvement in sensitivity obtained by periodically recalibrating feed and product measuring devices is obvious when the figures are compared.

## Discussion

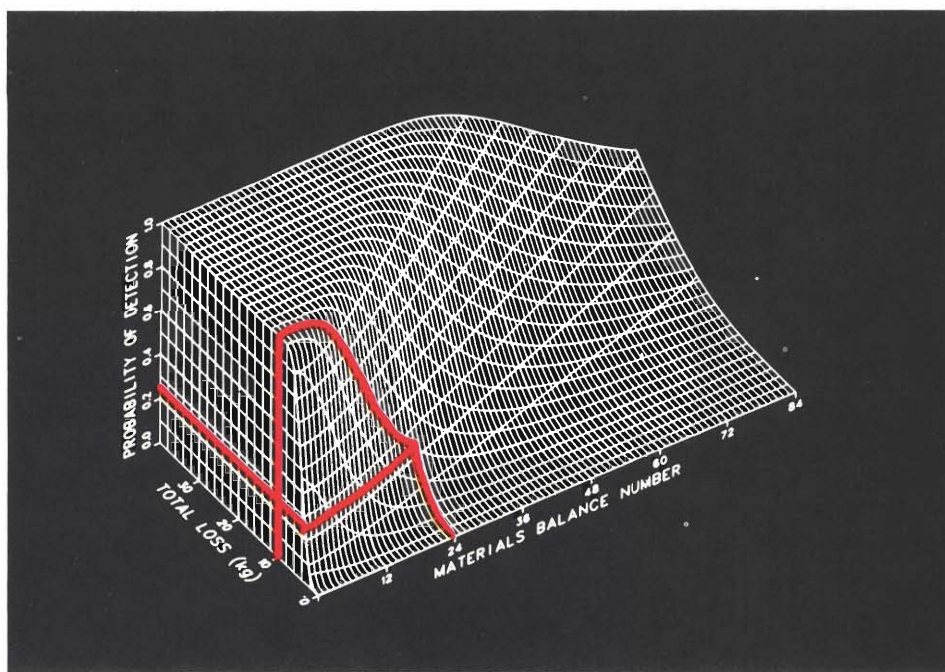
Until recently, almost no consideration was given to nuclear safeguards accounting requirements during the design of fuel-cycle facilities, the AGNS plant included. Instead the safeguards system designers were presented with either an existing facility or a relatively complete and fixed plant design. While the results of systems studies might introduce additional measurement instruments or bring about minor changes in operating equipment, they usually did not have any input to the choice of the process to be used in the facility or its mode of operation.

Increased recognition of the importance of nuclear safeguards and the need to integrate materials accounting into the process is bringing about a change. Safeguards designers are being consulted early in the design stages of fuel-cycle facilities. The resulting close cooperation between safeguards experts and process and facility designers should identify design alternatives that are both beneficial to safeguards and benevolent to the process.

The kind of materials accounting systems discussed above can provide better information on the locations and amounts of nuclear material than is



a.



b.

*Fig. 17. Cusum performance surfaces for two accounting cases. In the worst case (b), the loss of 10 kg of material can be detected at the 24th materials balance with a probability of 0.25. For the best case (a), the loss of the same quantity of material at the 24th materials balance can be detected with a probability of 0.90. The importance of good measuring instruments and a good measurement program is clear.*

currently provided by conventional methods. Such systems are beginning to be implemented at several facilities in the United States, including the new Los Alamos Scientific Laboratory Plutonium Facility, but much development work remains to be done. The process of com-

bing measuring instruments, data handling and analysis, and performance evaluation methodology into coherent effective safeguards systems is still in its infancy. The extension of these systems to international safeguards is just beginning.

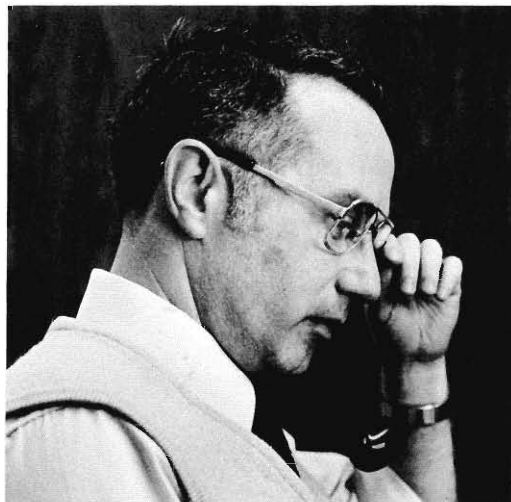
## BIBLIOGRAPHY

- IAEA Safeguards Technical Manual, Part A, "Safeguards Objectives and Criteria Requirements," International Atomic Energy Agency technical document IAEA-174 (Vienna, 1976).
- IAEA Safeguards Technical Manual, Part F, "Statistical Concepts and Techniques," International Atomic Energy Agency technical document IAEA-174 (Vienna, 1977).
- R. H. Augustson, "Dynamic Materials Control Development and Demonstration Program," *Nucl. Mater. Manage.* VII(3), 305-318 (1978).
- D. D. Cobb and J. P. Shipley, "Performance Analysis of Nuclear Materials Accounting Systems," *Nucl. Mater. Manage.* VIII(2), 81-92 (Summer 1979).
- H. A. Dayem, D. D. Cobb, R. J. Dietz, E. A. Hakkila, J. P. Shipley, and D. B. Smith, "Dynamic Materials Accounting in the Back End of the LWR Fuel Cycle," *Nucl. Technol.* 43, 222-243 (1979).
- E. A. Hakkila, D. D. Cobb, H. A. Dayem, R. J. Dietz, E. A. Kern, E. P. Schelonka, J. P. Shipley, D. B. Smith, R. H. Augustson, and J. W. Barnes, "Coordinated Safeguards for Materials Management in a Fuel Reprocessing Plant, Vol. I," Los Alamos Scientific Laboratory report LA-6881 (September 1977).
- E. A. Hakkila, J. W. Barnes, T. R. Canada, D. D. Cobb, S. T. Hsue, D. G. Langner, J. L. Parker, J. P. Shipley, D. B. Smith, R. H. Augustson, and J. W. Barnes, "Coordinated Safeguards for Materials Management in a Fuel Reprocessing Plant, Vol. II," Los Alamos Scientific Laboratory report LA-6881 (September 1977).
- E. A. Hakkila, D. D. Cobb, H. A. Dayem, R. J. Dietz, E. A. Kern, J. T. Markin, J. P. Shipley, J. W. Barnes, and L. A. Scheinman, "Materials Management in an Internationally Safeguarded Fuels Reprocessing Plant, Vol. I," Los Alamos Scientific Laboratory report LA-8042 (April 1980).
- E. A. Hakkila, D. D. Cobb, H. A. Dayem, R. J. Dietz, E. A. Kern, J. T. Markin, J. P. Shipley, J. W. Barnes, and L. A. Scheinman, "Materials Management in an Internationally Safeguarded Fuels Reprocessing Plant, Vol. II," Los Alamos Scientific Laboratory report LA-8042 (April 1980).
- E. A. Hakkila, A. L. Baker, D. D. Cobb, H. A. Dayem, R. J. Dietz, J. E. Foley, R. G. Gutmacher, J. T. Markin, H. O. Menlove, C. A. Ostenak, J. R. Phillips, T. D. Reilly, J. P. Shipley, D. B. Smith, C. C. Thomas, Jr., W. J. Whitty, J. W. Barnes, J. E. Bennett, A. L. Beyerlein, L. E. Burkhart, D. R. Engler, and A. F. Cermak, "Materials Management in an Internationally Safeguarded Fuels Reprocessing Plant, Vol. III," Los Alamos Scientific Laboratory report LA-8042 (April 1980).
- J. L. Jaech, "Statistical Methods in Nuclear Material Control," TID-26298, Technical Information Center, Oak Ridge, Tennessee (1973).
- G. R. Keepin and W. J. Maraman, "Nondestructive Assay Technology and In-Plant Dynamic Materials Control—DYMAC," in *Safeguarding Nuclear Materials*, Proc. Symp., Vienna, 1975 (International Atomic Energy Agency, Vienna, 1976), IAEA-SM-201/32, pp. 305-320.
- T. I. McSweeney, J. W. Johnston, R. A. Schneider, and D. P. Granquist, "Improved Material Accounting for Plutonium Processing Facilities and a 235-U HTGR Fuel Fabrication Facility," Battelle-Pacific Northwest Laboratories report BNWL-2098 (October 1975).
- T. D. Reilly and M. L. Evans, "Measurement Reliability for Nuclear Material Assay," Los Alamos Scientific Laboratory report LA-6574 (January 1977).
- J. P. Shipley, "Decision Analysis and Nuclear Safeguards," in *Nuclear Safeguards Analysis—Nondestructive and Analytical Chemical Techniques*, E. A. Hakkila, Ed. (Am. Chem. Soc., Washington, DC, 1978).
- J. P. Shipley, "Efficient Analysis of Materials Accounting Data," *Nucl. Mater. Manage.* VII(3), 355-366 (1978).

### Acknowledgment

The authors are indebted to their colleagues in Los Alamos Scientific Laboratory Group Q-4 and especially to H. A. Dayem, D. D. Cobb, and E. A. Hakkila for helpful discussions and for material from which portions of this article were adapted.

## THE AUTHORS



**Darryl B. Smith** earned his bachelor of arts degree in physics-mathematics from Millikin University in 1958 and his Ph.D. in physics from the University of New Mexico in 1968. In 1958, he joined LASL's Physics Division, where he compiled and edited a volume of charged-particle cross section data, investigated triton-induced nuclear reactions in thin gaseous targets, and used inelastic neutron scattering to study the lattice dynamics of crystal structures. In 1967, he joined the nuclear safeguards research program and has carried out neutron and gamma-ray transport calculations and developed calibration and error-analysis techniques for NDA instrumentation. He is now working on design and evaluation of nuclear materials measurement and accounting systems for nuclear fuel cycle facilities. He has served as an advisor to the IAEA on the qualification and calibration of NDA instrumentation. He is chairman of the Institute of Nuclear Materials Management subcommittee INMM-9, which is responsible for developing ANSI consensus standards on NDA methodology and was chairman, 1974-1975, of the Trinity Section of the American Nuclear Society.

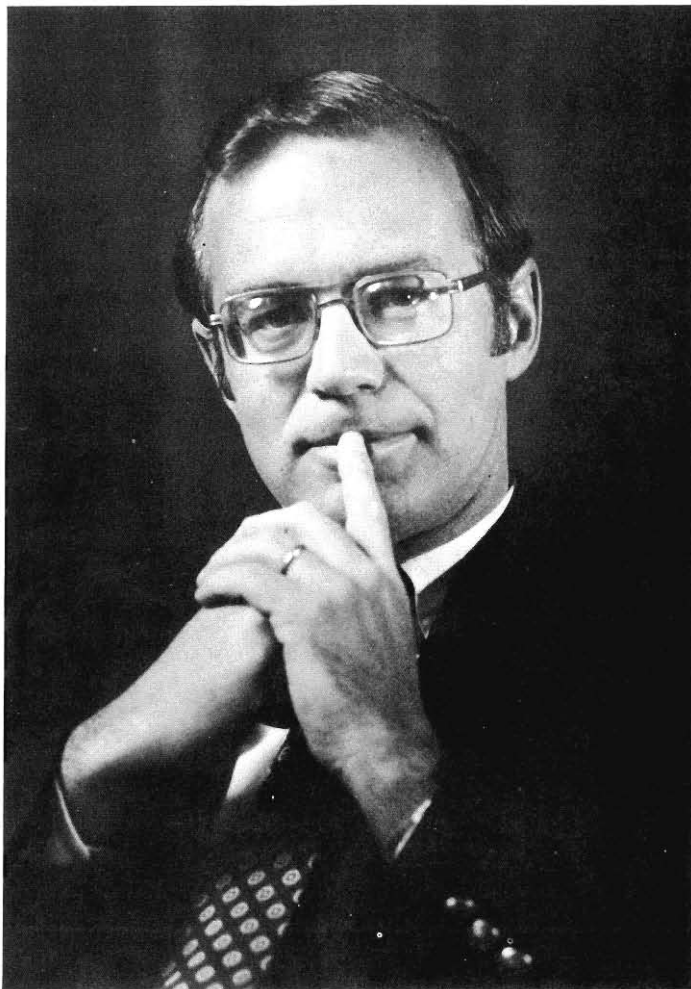


**James P. Shipley**, Group Leader of LASL's Safeguards Systems Group, received his bachelor of science degree in electrical engineering from New Mexico State University in 1966 and his Ph.D. in electrical engineering and computer science from the University of New Mexico in 1973. His background is in electronics, automatic control systems, filtering theory, and applied mathematics for systems analysis. Since joining LASL in 1966, he has been involved in research, development, and design of control systems, including those for solar heating and cooling of buildings. He joined the nuclear safeguards program in 1976, and was instrumental in setting up the Systems Control Group. He contributed Chapter 4, "Decision Analysis for Nuclear Safeguards," to *Nuclear Safeguards Analysis: Nondestructive and Analytical Chemical Techniques* (American Chemical Society, Washington, DC, 1978) and has served on the editorial board of the *Review of Scientific Instruments*.



**Dante Stirpe** earned his bachelor of science degree in physics from Union College in New York, his master of science degree at Iowa State University, and his Ph.D. from the University of Missouri. Before coming to Los Alamos in 1962, he worked on electron resonance at General Dynamics/Electronics at Rochester, New York. At LASL, he has done work in shock hydrodynamics and timing and firing devices at the Nevada Test Site, and technology assessments. He is now involved with materials accounting at isotope enrichment facilities.

# Don Kerr on Nuclear Safeguards



**F**or nuclear energy to remain an indispensable part of the United States energy supply, three major problems must be overcome both technically and institutionally: assured safety, acceptable waste disposal, and effective safeguards. The Los Alamos Scientific Laboratory is making significant contributions to all three, but improved nuclear safeguards may be the most pressing requirement today.

Since the beginning of the atomic era, the thrust of United States nuclear policy has been twofold: to build a strong national defense by developing nuclear weapons and naval propulsion systems and to support the private sector's development of a safe and efficient energy resource. As these efforts have expanded, so have the risks incurred by our accumulation of weapons-usable nuclear materials. The proliferation of nuclear weapons states has become a major national security problem. Similarly, the potential for diversion of nuclear materials from the nuclear fuel cycle by a subnational group has added to the growing arsenal of mass terrorism techniques.

Recent guerrilla actions to capture and hold diplomatic hostages in Tehran and Bogota call attention to the need for all major governments, regardless of ideology, to join in common action to protect institutions and communications from the spread of terrorism. In the nuclear age, the threat of terror may be greater than the threat of war between the superpowers. As the world runs short of fossil fuels or the cost to acquire them becomes too high, our reliance on nuclear energy is bound to increase. Among the industrialized countries, France, Japan, West Germany, and the United Kingdom have chosen nuclear energy; Brazil, Argentina, and other developing countries are likely to make the same choice. In the future, unless the tyranny of political terrorists is brought under control, a few may be able to seize, not an embassy, but a nuclear power plant or some other nuclear facility and hold hostage not ambassadors, but entire communities and even nations. Far less difficult actions could produce the same tragic results. The capture of a few kilograms of fissile material and the threat to detonate an improvised nuclear explosive, or the seizure of a few barrels of radioactive wastes and the threat to disperse them in rivers and harbors near large cities could render whole populations defenseless. The challenge to protect nuclear materials from illegal possession is enormous and urgent. It can be met only by a combination of technological and institutional developments.

International and national safeguards and security measures to limit the risks of the nuclear era have evolved over several decades. They must continue to evolve. In fact, recent studies (the International Nuclear Fuel Cycle Evaluation and the Non-Proliferation Alternative System Assessment Program) affirm the need for continuing improvements.

The goal of these studies was to determine whether we could avoid nuclear proliferation problems associated with current LWR and LMFBR fuel cycles by developing alternative fuel cycles that do not produce nuclear materials in forms suitable for use in a nuclear weapon. The studies concluded that proliferation-proof, diversion-proof fuel cycles are unattainable and that the urgent need of some countries for nuclear power will result in the continued growth of the uranium-based nuclear power industry. In fact, some countries see a need not only for nuclear power facilities, but for their own reprocessing plants and centrifuge enrichment plants in order to assure an adequate, economical supply of nuclear fuel for power production. From a safeguards point of view, reprocessing plants that recover weapons-usable uranium and plutonium from burnt or spent reactor fuel and centrifuge plants that enrich uranium are the most vulnerable elements in the nuclear fuel cycle. Not only do they produce weapons-grade uranium and plutonium in large quantities, but the safeguards technologies necessary to account for the materials on a timely basis are in the development stages prior to test and evaluation. The 1980 General Accounting Office report on problems of nuclear fuel reprocessing concludes that "new technology is needed if the United States is to further its own goals of preventing the spread of nuclear weapons and influence other countries to adopt strengthened safeguards at reprocessing facilities."

As the Department of Energy's lead laboratory for research and development in special nuclear material control and accountability, LASL has made major contributions to these and other safeguards technology problems. We have developed nondestructive assay instrumentation for timely measurement of sensitive nuclear materials in all stages of processing. We have also designed near-real-time material control and accountability systems based on the use of these measurement techniques and are now demonstrating such a system at our Plutonium Processing Facility. Further we have developed the systems methodology necessary to design similar systems for new facilities and to evaluate their detection sensitivities. These integrated systems are necessary to deter and detect diversion of sensitive nuclear material by a knowledgeable insider. With regard to materials accountability for reprocessing, the appropriate measurement instruments are under development at LASL. Based on projected measurement capabilities our design analyses suggest that adequate safeguards accountability systems can be implemented in reprocessing facilities. However, the final assurance can come only after instrument development is complete and measurement systems are tested and evaluated at an operating plant. Whether or not commercial reprocessing is carried out in this country, I believe that in order to fulfill our commitment to international safeguards, we

must continue to develop, test, and evaluate the materials control and accountability technology needed to safeguard all types of nuclear facilities including reprocessing plants.

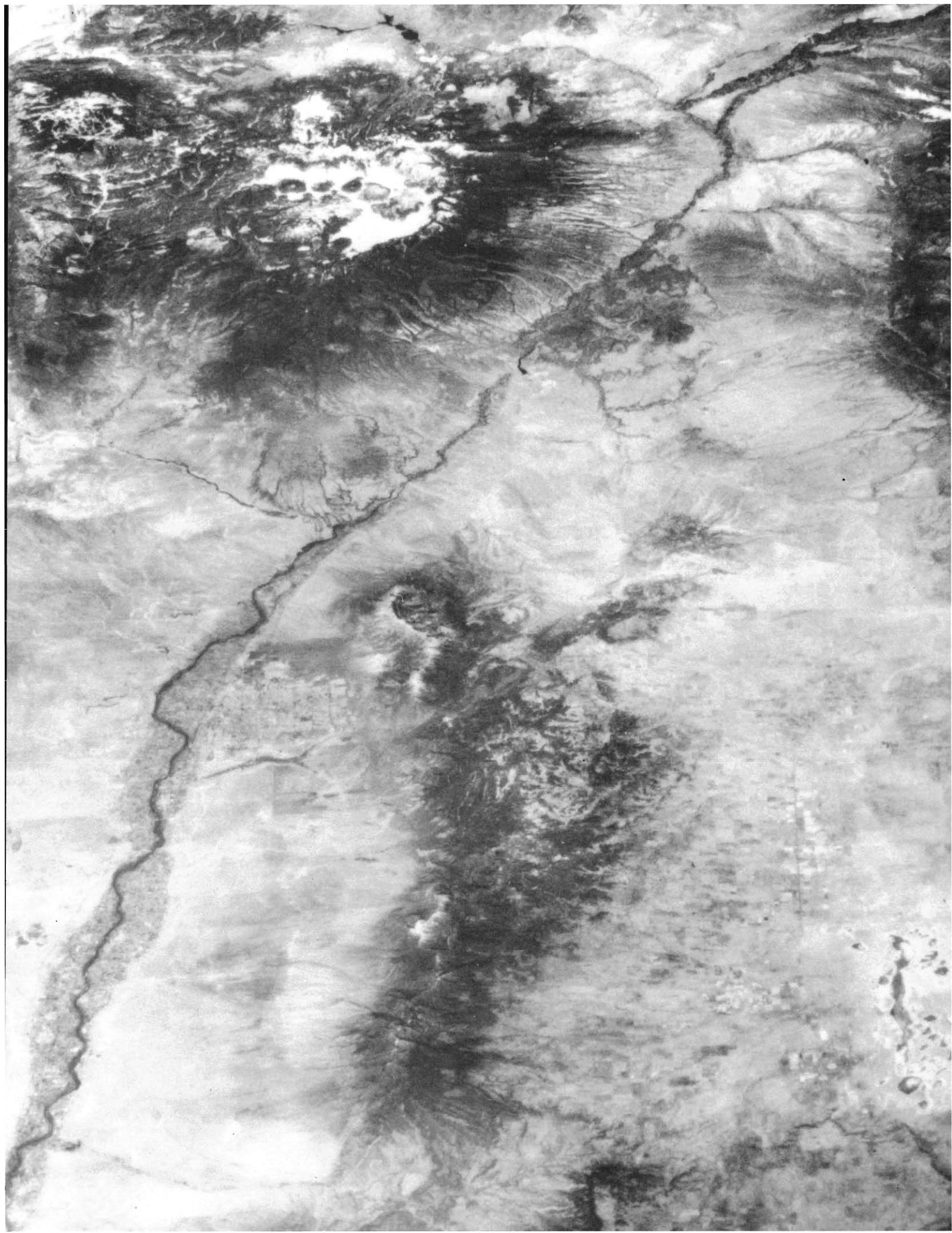
Los Alamos also has the principal responsibility for transferring this developing technology to industry, the Nuclear Regulatory Commission, International Atomic Energy Agency, and other countries. In this role, the Laboratory sponsors an extensive program of special training courses and participates directly in the design of new facilities and the addition of new safeguards to existing facilities.

To meet national goals for the prompt recovery of lost or stolen materials, Los Alamos also plays a lead role in defining and organizing DOE's nuclear emergency search team (NEST). The Laboratory's unique capabilities in weapon-related fields are applied to development and operational deployment of search, hazard-assessment, and render-safe equipment and procedures that may be needed to deal with improvised nuclear devices and nuclear terrorism. Nondestructive assay techniques contribute significantly to these NEST capabilities. The weapons development program continues to integrate protective measures into the design of nuclear weapons to deny malefactors the ability to achieve a nuclear yield without resorting to extraordinary measures.

For the future, Los Alamos is developing cost-effective methods to integrate material control and accountability with physical protection. The Laboratory is also taking initiatives to

- develop the NEST technology and organization for extension to nuclear sabotage and reactor accidents.
- develop special international safeguards technology for application to gas-centrifuge plants, advanced isotope separation plants, the fast-breeder fuel cycle, and high-throughput, spent-fuel reprocessing plants.
- develop safeguards systems for away-from-reactor spent-fuel storage.

However, technology only augments the institutional controls on nuclear materials and weapons technology. Institutional developments must proceed in parallel with technology if we are to achieve the level of deterrence we require. To that end, three proposals are under consideration: an international plutonium storage supported by about 85% of the nations participating in the INFCE, international or multinational nuclear fuel cycle facilities, and nuclear power parks that close the fuel cycle within their borders. These institutional initiatives promise to address some of the vulnerabilities of the current nuclear economy. Advanced technology must be developed to make them even more effective in order that nuclear energy may be retained as an alternative for the future.



*Skylab 4 photo with view north up the Rio Grande rift from altitude of 432 km over central New Mexico. The rift crosses a wide zone with a strong northeast trending grain. It passes through the Jemez lineament, one of the most active volcanic zones in the U.S. for the past 10 million years. Note the Jemez caldera top left.*

# Plate Tectonics— Where the Action Is

by Bob Riecker

**D**uring your lifetime, North America and Europe will separate by the average height of a person; the Pacific Ocean will shrink by the average width of a single-family home. Next year, Los Angeles will move closer to San Francisco by the length of your little finger. Last year, some 100 cubic kilometers of new earth crust were born at crests of ocean ridges. In the last 76 million years, the north magnetic pole has reversed its polarity at least 171 times, and in the last 100 million years, an area equal to the size of the Pacific Ocean basin has disappeared beneath surrounding continents.

Plate tectonics is the unifying concept in geosciences. The Earth's surface is divided into a mosaic of possibly 10 to 12 large (and many smaller) rigid, moving plates, varying from 10 to 100 kilometers thick and often many thousands of kilometers wide. The plates that move apart on our globe must collide somewhere, and when they do, earthquakes and volcanism remind us of turmoil. In an ocean chain over 30,000 kilometers long, new earth crust forms to add rock to continents. At other places, mountains rise because of plate collisions.

We ride northwestward on our namesake, the North American plate,

which extends from the Pacific coast to the middle Atlantic Ocean. We grind against the Pacific plate along the San Andreas Fault in California. Plate motion insures that, in 50 million years, Los Angeles will become a western suburb of San Francisco. That trip is not easy geologically, as major earthquakes attest periodically.

The gurus of plate tectonics predicted at the paradigm's inception, about 15 years ago, that the new theory would require rewriting of geology textbooks. In fact, that's happened. But we are just now beginning to understand how dynamic and changing the Earth's surface really is.

Problems remain. The simple model of 100-kilometer-thick, rigid plates, or rafts, floating about on "greasy skids" of the Earth's upper mantle looks naive. Seismic research based on study of propagation of elastic waves through the Earth suggests that continents have much deeper roots. Also, we don't really know yet what makes the plates go! Certainly the Earth's natural heat engine drives the rafts, but how, and over what vertical dimensions? The Earth loses heat, generated mainly by radioactive decay, at a rate about five times more than the rate at which man uses energy ( $2 \times 10^{12}$  calories/second). How long

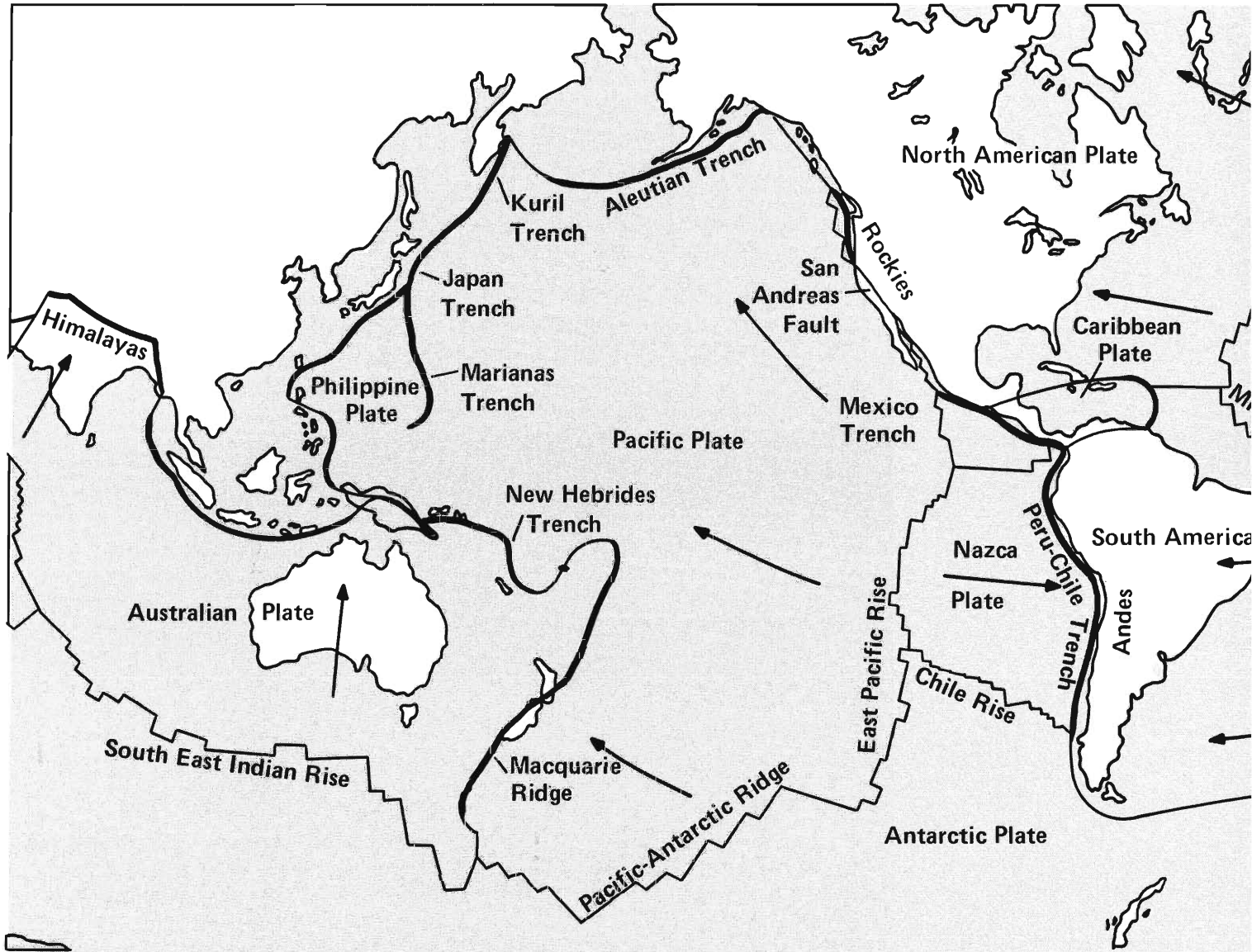
have plates moved over the Earth's surface since the planet's birth? As the Earth's heat engine cools, are plate motions slowing, and is mountain building ceasing? Often earthquakes and sometimes volcanism appear in the interior of plates, not only along their edges as predicted by the theory.



Normally the hottest spots are where the action is. That's where plates originate or where they collide or are consumed. At ocean spreading centers, new crust forms and enlarges the plates.

One such hot zone centers along the Rio Grande rift. A rift is a long, narrow, usually down-faulted valley in the Earth's crust. The Rio Grande rift extends from Leadville, Colorado, through New Mexico, and across the border into northern Mexico. Along this 800-kilometer rift, significant volcanism, earthquake activity, uplift, and separation have occurred during the last 30 million years.

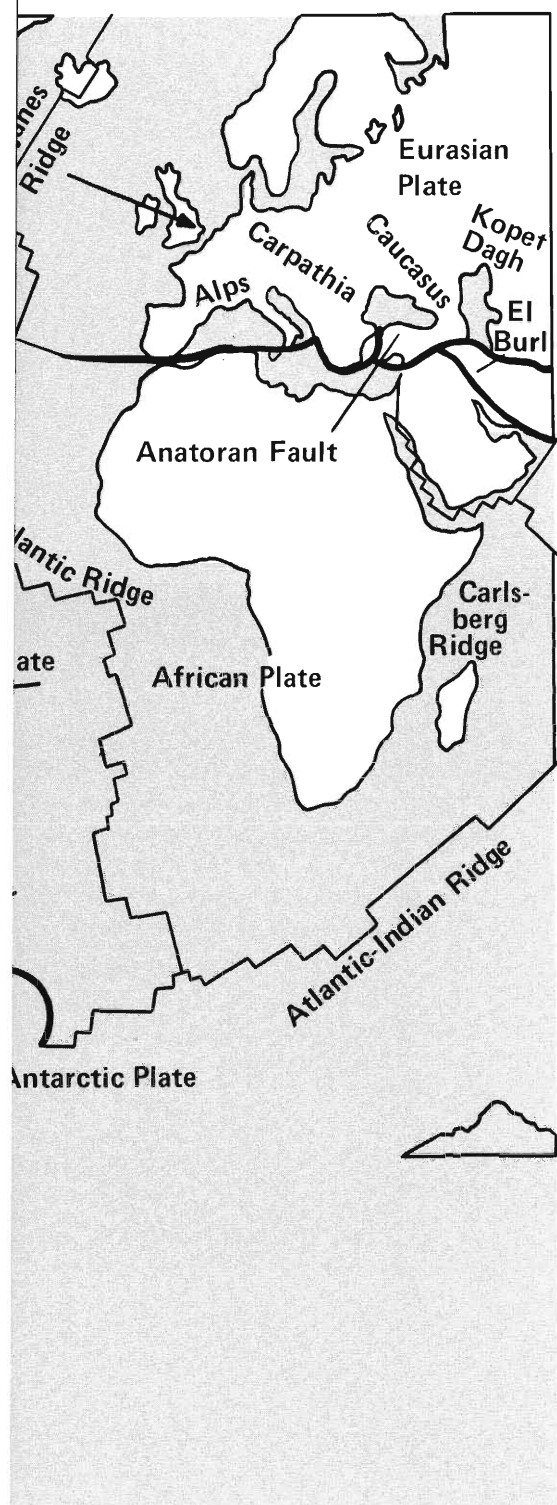
For the preceding hundreds of millions of years, the entire zone along which the Rocky Mountains stand had been an active mountain-building region. Then 30 million years ago, regional separation initiated rifting. The rift opened along a series of northward and northeastward flaws or lineaments, which reach down through the overlying

SCIENCE IDEAS



 Subduction Zone  
 Direction of Plate Motion





sedimentary cover. The series of deep basins that eventually formed in the rift now contain thousands of meters of sediment and volcanic fill.

The deeply cutting flaws that cross the rift are zones of weakness that probably extend deep into the Earth's mantle. They leak vast quantities of molten rock into and through the crust. The Jemez lineament, which extends across the Jemez caldera northeast, was especially active volcanically during the last 10 million years as indicated by large volumes of young volcanic rocks blanketing the surface. Many of these volcanics, silicic in composition, erupted during violent explosions much larger than modern man has seen. Layers of fine ash from some of the eruptions reached as far east as Iowa and Illinois before falling from the atmosphere. The giant scale of these events can be imagined as one rides past the Jemez caldera in the center of the Jemez mountains. This grassy pastureland, more than 15 kilometers in diameter, is the remnant of a large volcano.

The scale of the Rio Grande rift is also quite large. Separation across the rift is as wide as 10 kilometers and uplift exceeds many thousand meters. Does this mean that the North American plate is being torn asunder in the southwestern United States? Some geoscientists believe so, and in their jargon, the rift represents an incipient spreading center where new earth crust forms and shoves aside older rock.

Major crustal stretching occurred not only along the Rio Grande rift but over the entire southwestern portion of the United States. We see evidence of the stretching in the patterns of faulting and the appearance of volcanic rocks in the Basin and Range provinces in Arizona and Nevada to the west and in the High Lava Plains of Oregon and California to the north. The exact cause of the

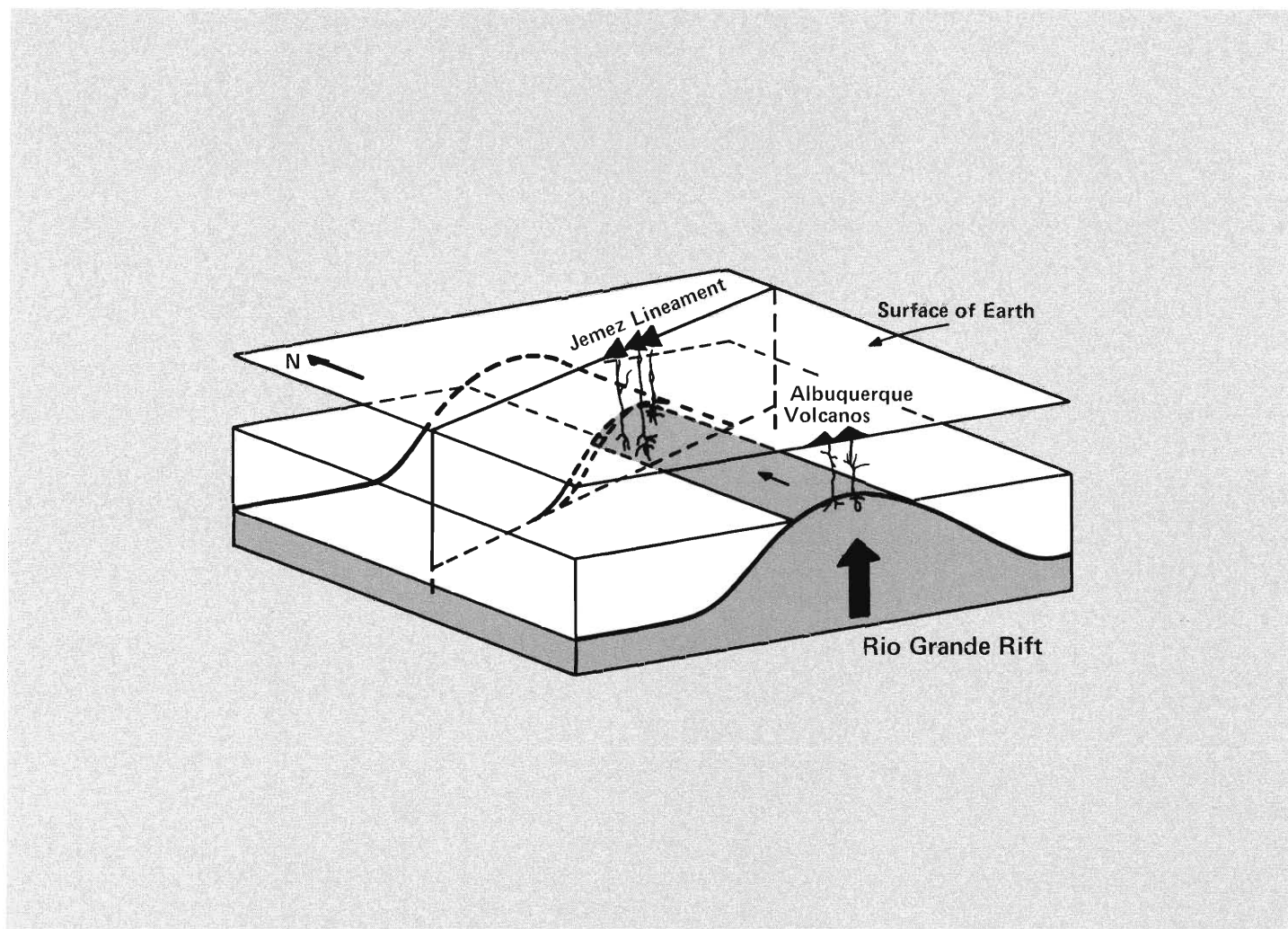
stretching remains obscure and controversial. It must involve interactions during the last fraction of geologic time among the North American and Pacific plates and the Farallon plate, which lay between them, but since has been consumed as the southwestern part of our raft overrode it. These interactions also produced and now drive the San Andreas fault splitting California.

There is no question that episodic rifting continues today in New Mexico, as evidenced by the abundant fault scarps, earthquakes, high heat flow, modern elevation changes, recent volcanism, and the geophysical evidence of an anomalous thin crust beneath the rift.

The earliest written record of earthquake activity in New Mexico is of an earthquake swarm described by an army surgeon camping near Socorro in late 1849. Most of the recent earthquakes have occurred in a 150-kilometer section between Belen and Socorro, New Mexico. During the last 20 years, seismic recording suggests a very low level of activity. The level is surprising in view of major crustal movements in the past and the presence of modern fault scarps along the edges of the rift. The rift now appears quiet, but seismically pregnant.

Detailed seismic studies continue in New Mexico with measurements performed by LASL, the New Mexico Institute of Mining and Technology, and the Albuquerque Seismological Laboratory of the U.S. Geological Survey. Each group installed seismometers for different reasons, but all cooperate closely in trying to learn more about seismic activity in New Mexico. One motivation for LASL's seismic research has been to study natural earthquake activity near the hot dry rock geothermal energy experiment at Fenton Hill, so that man-made and natural events can be distinguished. Also construction of new

## SCIENCE IDEAS



*The central Rio Grande rift. The partly molten rock (dark grey) reaches within 35 km of the Earth's surface below the Albuquerque volcanoes.*

facilities in Los Alamos, such as the half-mile-long accelerator, required knowledge of seismic risk. New Mexico is still earthquake country even though current activity is low.

Joint research shows that microseismic activity is concentrated in two areas, one between Belen and Socorro and another 15 kilometers west of Espanola. Near Socorro, seismic analysis suggests the presence of a large,

complicated-shaped magma body (molten rock) at a crustal depth of about 19 kilometers, leaking into shallower bodies. Most of the Socorro activity consists of small-magnitude earthquake swarms apparently related to movement of molten rock at depth. The swarms suggest an extension of the upper crust produced by the intrusion of a layer of magma at mid-crustal depths. Unusually high heat flows occur over the inferred

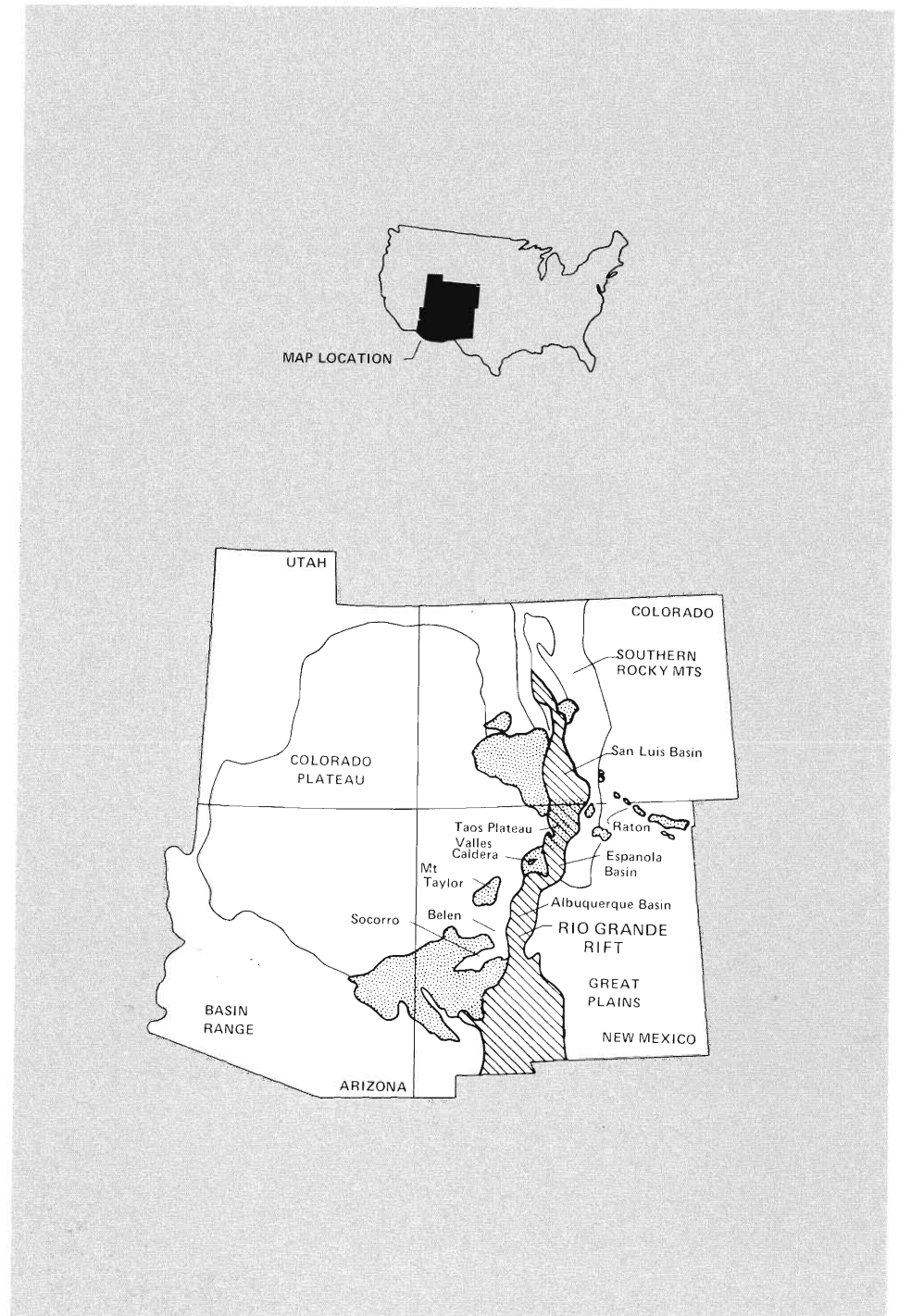
magmas. Recent seismic reflection research by the Consortium for Continental Reflection Profiling penetrated to depths of at least 35 kilometers. Results confirm the existence and shape of the molten rock mass.

West of Espanola, seismic activity also suggests a possible underlying magma body. Activity consists of small-magnitude earthquakes that recur at intervals of several months. Seismometers

record a few swarm-like sequences. Interestingly, LASL researchers recently noted a major relative subsidence northwest of Espanola associated spatially with the earthquake belt. Geodetic releveling analysis reveals a maximum subsidence of about 5 centimeters between surveys performed in 1934 and 1939. Researchers suggest that the sinking and the associated earthquakes originate in deflation of a shallow magma body. Continued studies will help to identify the processes responsible for regional extension in the Southwest.

Plate movements not only create earthquakes and produce volcanism; they also determine where mineral deposits form. In addition, most geothermal energy resources originate as a result of plate interaction. The LASL hot dry rock experiment on Fenton Hill is successful partly because of its location on the edge of the Rio Grande rift. Plate travels determine whether diamonds form in Arkansas or gold appears in California. Many ore bodies, for example, occur at present or past boundaries of crustal plates.

Geology is a young discipline with a rapidly increasing pace of surprising discoveries. Only astronomy shares with it the complexities of scale and time. In 20 A.D. Seneca wrote, "It is useful to be assured that the heavings of the Earth are not the work of angry deities. These phenomena have causes all their own." Two millennia later we are just beginning to understand what those causes are. Moreover, we are just beginning to exploit our understanding through improved exploration for mineral and energy resources. Few theories have made such an impact on their disciplines, and few disciplines, stimulated by new theories, have had such an impact on energy technologies as has plate tectonics on the geosciences.



*The Rio Grande Rift. The crosshatched areas indicate the Rift itself and the stippled areas indicate surface volcanics.*

"I feel that he is to be classed among the outstanding scientists of the twentieth century and at the top in the field of inorganic crystal structures."  
Linus Pauling



# PLUTONIUM

$SH = S(KiS + Ki)$   
 $S = HS$



**5-f Contraction**



**1906**  
**1979**

MEMORIAL TO

# Professor (Fredrik) William H. Zachariasen

Born February 5, 1906, Langesund, Norway  
Died December 24, 1979, Santa Fe, New Mexico

by Robert A. Penneman

Prof. Zachariasen, one of the outstanding scientists of the 20th Century, is gone. "Willie" leaves to his countless friends the warm remembrance of his gracious manner, his great good humor, and his generous friendship. His rich and varied scientific contributions engaged him to the last. He was one of the giants in x-ray crystallography and was involved in every major advance in that field. Linus Pauling wrote of him:

*I have known Professor Zachariasen for nearly fifty years. His principal field of work has been the determination of the structure of crystals of inorganic substances by use of the x-ray diffraction technique. This is a field in which I also have done a large amount of work, and I believe that I am in a position to form a sound opinion about his ability and his contributions. It is my opinion that he has been and is the world leader in this field.*

*I feel that he is to be classed among the outstanding scientists of the twentieth century, and at the top in the field of inorganic crystal structures.*

## The Formative Years

Zachariasen began his career, unknowingly, with his youthful exploration of the islands in Langesundfjord near his home, islands rich in well-crystallized rare earth minerals. He was to return to those islands later as a student at the University of Oslo under the great geochemist Prof. Goldschmidt, who had purchased one of the islands for his studies. "Willie" rowed him out on occasion, a short journey for the son of a sea captain. The trips to the islands combined work with pleasure. Mrs. Zachariasen supplied a photograph of a picnic on one of the islands; in it Prof. Goldschmidt is seen talking to guest Albert Einstein. The vivid memory of those days never left Zachariasen; some dozens of years later he correctly identified (on sight) one of the crystal specimens from the

Langesundfjord islands—a specimen that had been mislabeled by a commercial firm.

Zachariasen spent his student days (1923-1928) in the midst of that exciting period when Goldschmidt and his collaborators worked out general laws governing distribution of chemical elements in minerals; they were the first to apply x-ray diffraction to the study of geochemistry. During those years Zachariasen read hundreds of x-ray films, an activity that he continued throughout his life. Where the results were of particular significance, as in the recent controversy involving the  $\alpha'$  and  $\alpha''$  phases of cerium, he would read each film several times.

Zachariasen published his first paper at age 19, after presenting it before the Norwegian Academy of Science the year before. It was on x-ray diffraction studies of oxides. With its publication in 1925 he began a period of contributions to the scientific literature, most of them singly authored, which would span 55 years. At age 22, he received his Ph.D., the youngest person ever to receive it in Norway. In 1928-29, he was a Postdoctoral Fellow at the laboratory of Sir Lawrence Bragg, where he began his study of silicate structures, a study that would later culminate in the first real understanding of the structure of glass. He returned to the University of Oslo, but within a year accepted a call from Nobel Laureate Arthur Compton to join the faculty of Physics at the University of Chicago. He was 24.

Before leaving Norway in 1930, he married Ragni Durban-Hansen, the striking granddaughter of the pioneer Norwegian geochemist, W. C. Brøgger, who discovered and first described the extensive mineral deposits of the Langesund area. Willie and his bride spent their honeymoon on the ship that brought them to the United States to fulfill his commitment to the University of Chicago. There he spent the next 44 years. The Zachariasens had two children, Fredrik Zachariasen (at the California Institute of Technology, Pasadena) and Ellen Z. Erickson of Santa Monica.

# Zachariasen

The University of Chicago



Many honors came to him during the 44 years at the University of Chicago. He advanced to full Professor, then to Department Head (during the critical post-war years when the department was rebuilt), and finally to Dean of the Physical Sciences. He was a kindly administrator and solicited advice, but always maintained a policy that one had to “select the few from the many.” He was a member of the Norwegian Academy of Science, the U.S. National Academy of Science, the American Academy of Arts and Sciences, the American Physical Society, the American Crystallographic Association, and the Executive Committee of the International Union of Crystallography. He was presented with the honorary degree of Doctor of Science by the New York Polytechnic Institute.

The 1930s were years of financial hardship. Travel funds were nonexistent and Zach would sit up on a night train to New York, give his paper at a scientific meeting the next day, and return that night to save the cost of a hotel room. During that same period x-ray work was shut down for six months for lack of a \$75 tube.

But those were also years of high scientific productivity. Following his work with silicates and other oxy-anions, Zachariasen began to think about how the glass structures were built. In 1932 he published his landmark paper on the structure of glass. Referring to this paper in 1961, Charles H. Green wrote: “The present day understanding of glass rests heavily on a singly lucid paper, only 12 pages long, written in 1932 by W. H. Zachariasen.” Zachariasen continued through 1941 to study complex oxy-anions and to develop his work on the diffuse scattering of x rays caused by thermal motion.

In 1941 he became an American citizen. He was then in his mid-30s and already a world figure, having published 80 experimental and theoretical papers, including major papers on diffuse scattering, oxide structures, and the structure of glass. A significant change was soon to occur following the onset of World War II.

## The War Years

In parallel with his academic career, Zachariasen began in 1943 to immerse himself in a then-secret activity, doing all the x-ray identification work for a new project on the Chicago campus. As part of a major wartime effort, scientists had gathered at the Metallurgical Laboratory to work with new elements that had not yet been seen, new elements whose chemistry was largely a mystery.

Recall the situation—Enrico Fermi had just demonstrated the existence of the chain reaction at Chicago in December 1942, using uranium.\* Plans were being rushed for the pilot plant at Oak Ridge, Tennessee, and the production reactors at Hanford, Washington. This meant that the new element, plutonium, would be made in large quantity using neutrons from a nuclear reactor. Before this, plutonium could be made only in microgram quantities by tedious cyclotron irradiation.



*Zachariasen (top). Prof. Goldschmidt and Albert Einstein at a picnic on the islands of Langesundfjord (bottom).*

\*During the Fall of 1942 I was in Massachusetts as part of the Chicago team aiding in the production of sufficiently pure uranium metal.

There still remained the formidable chemical separations problem. How could pounds of plutonium be isolated from tons of uranium containing radioactive fission products? Chemical processes had to be devised that would work efficiently on a large scale and in high-radiation fields that required remote handling. Ultimately, plutonium metal had to be made. What would be its properties? Totally unexpected were the great complexities soon to be encountered with the plutonium metal phases, their unusual number (six), and the confusion caused by the ease of transfer among the four plutonium aqueous valence states.

Seaborg, the co-discoverer with Kennedy, Segrè, and Wahl of the fissionable isotope  $^{239}\text{Pu}$ , was responsible for directing the efforts of some 60 chemists attempting to elucidate its chemical properties and to develop reliable plutonium separation processes. This was a heavy responsibility and work was conducted with great urgency (July 4th was just another work day). Zachariasen's x-ray analyses provided the information essential to the understanding of plutonium chemistry, deciphering single-handedly the composition of countless samples that were prepared by the chemists. His work was indispensable in replacing mystery with fact and guesses with quantitative structure identification.

Zachariasen was uniquely equipped for this challenge, given his prodigious skill in elucidating structures from powder diffraction data, plus his long familiarity with the rare earths and their 4f-contraction [the decrease in size with increasing atomic number at constant oxidation state (valence)]. Zachariasen's application of crystallography to the elucidation of the nature and chemistry of the transuranic elements, their compounds and metals, is probably his most celebrated work, covering 37 years and several dozen papers.

In the introduction to a book, which did not (and now cannot) get beyond that preliminary stage, Zachariasen wrote:

*In 1943 Pu metal and some compounds were prepared in microgram amounts; but the ultra-microchemical studies of plutonium presented great difficulties of interpretation. Because of the small amounts of material, it was difficult to prepare single phase samples and to establish chemical identities.*

*In the autumn of 1943 I demonstrated that satisfactory x-ray diffraction patterns of Pu-preparations on the 10 microgram scale could be obtained and that the interpretation of the x-ray pattern often could provide positive identification of the phase or phases present in the preparation. Thus, in the period from November 1943 to June 1944 a number of plutonium compounds were identified and their crystal structures determined in full or in part.*

A fledgling chemist who was impressed by Zachariasen, and who later became an important crystallographer in his own right and Dean of Chemistry at Berkeley, David Templeton wrote:

*Of those early days, I know from personal observation that Zachariasen's work strongly influenced the development of the separation processes for the Hanford plant. As a young chemist, quite unblemished with any understanding of*

*crystallography, I heard him report to the Seaborg team the identity of this or that new compound as they tried to identify oxidation states. Zachariasen's identifications were almost the only reliable analyses available. He solved the structures of hundreds of substances... .*

## The 5f Series: Thorides vs Actinides

In "History of the Met. Lab, Section C-1," Seaborg recounts that on June 21, 1944, a sample of what was thought to be  $\text{NpO}_2$  was sent to Zachariasen. By 11:00 A.M. on June 22, 1944, his x-ray analysis had confirmed the existence of  $\text{NpO}_2$ , and Zachariasen had written his memo discussing the thoride series. It states in part:

*The radius of  $\text{Np}^{+4}$  is thus  $0.015\text{\AA}$  larger than that of  $\text{Pu}^{+4}$ ,  $0.016\text{\AA}$  smaller than that of  $\text{U}^{+4}$ , and nearly identical with that of  $\text{Ce}^{+4}$ .*

*I believe that a new set of "rare earth" elements has made its appearance. I believe that the persistent valence is four, so that thorium is to be regarded as the prototype just as lanthanum is the prototype of the regular rare earth elements.*

*W. H. Zachariasen*

There had been earlier qualitative observations\* supporting formation of a new inner transition series of elements; in particular, the narrow absorption features in the plutonium spectra suggested it. However, Zachariasen's quantitative data showing the progressive 5f contraction provided the key confirmation.

On July 14, Seaborg dictated a memo containing the sentence: "...I suggest that the elements heavier than actinium be placed in the Periodic Table as an 'Actinide Series'." The actinide name prevailed for the 5f series, but Zachariasen wrote, "The name actinide is not acceptable because thorium is never actinium-like..." He called them the 5f-series and would point out that not until the elements 95 and 96 were the metals rare-earth-like, and that the dioxide structure persisted from  $\text{ThO}_2$  to  $\text{CfO}_2$ , elements 90 through 98.

Twenty years earlier, Niels Bohr had predicted the occurrence of trans-uranium elements as a 5f series, with the series beginning at element 95. We now know conclusively that in the metals of this series, localized 5f electrons first appear with element 95 (americium) and in neighboring trivalent curium, element 96, the 5f shell is just half filled.

In an interview recounting those days and events pertaining to the history of Zachariasen's association with the University of Chicago, we find the following quotation with its significant ending:

*... usually we found that the compounds which were present in the sample were not what they had intended to make. We had a very exciting time struggling with all these patterns over the various plutonium compounds, identifying what the chemists had made and, hence, getting information about the chemistry of plutonium that was essential... . I remember working like hell on New Year's Day and all holidays; often I worked late for many, many hours to get the work done. I had a wonderful time... .*



## The Book

In 1945, Zachariasen published his classic and tightly written book, *The Theory of X-Ray Diffraction in Crystals*. Professor Pepinsky, himself a noted crystallographer, wrote of it: "The physical chapters serve as a basis for most of the developments in scattering theory since their publication; and many a contemporary paper is no more than a direct expansion of one or another paragraph." In 1948-49, Zachariasen published 26 papers, an heroic effort.

## Direct Methods

In 1952, Prof. Zachariasen pioneered the use of "Direct Methods," for determining crystal structures. The extraction of detailed structural information from x-ray measurements requires the knowledge of two fundamental classes of information: the magnitudes and the phases of the structure factors. The values of the x-ray intensities are obtained directly, and are equal to the squared magnitude of the structure factors. For a complete solution of the structure, knowledge of the phases of the structure factors are necessary. Before 1952 the inability of the x-ray experiment to provide this latter informa-

tion was referred to as the *phase problem*. This problem severely limited the application of the potentially powerful x-ray technique. The structure of solids containing small numbers of heavy atoms could be deduced more or less directly from the intensity pattern, but other substances were impossible to decipher except for the rare occasion when a model could be devised intuitively. The more powerful methods for deriving phase information depend on the statistical distribution of the intensities. This methodology is referred to as the *direct method* in crystallography.

In 1952 Zachariasen devised an algorithm for applications of this new method and successfully used it to determine the structure of boric acid. Although direct method techniques have been refined considerably in the ensuing years, Zachariasen's work remains the cornerstone of the modern methodology.

## The Extinction Problem

In a series of brilliant papers published over the years 1963-1967, he described a practical solution to the problem of secondary extinction, one of the major unsolved problems in x-ray crystallography. Secondary extinction arises for reflections of such large intensity that an appreciable amount of the inci-



dent radiation is reflected by the first planes encountered by the beam. Since deeper planes are thus presented with less incident radiation, their reflections are reduced in intensity. This effect is strongly dependent on the types and degree of imperfections of the scattering crystal. The stronger the scattered intensity the more serious is the extinction. Scattered rays closest to the incident beam are generally most intense and it is specifically these low-angle reflections that are important in the accurate determination of valence electron distributions and the positions of light atoms such as hydrogen. In severe cases, secondary extinction problems can lead to erroneous results even in routine crystal structure determinations.

Some exceptionally accurate measurements on quartz caused Zachariasen to recognize a fundamental error made by diffractionist Darwin more than 50 years before in deriving the equations for the extinction effect. By the end of 1966, Zachariasen had circulated a preprint of a paper in which a practical solution to the extinction problem was described. It is a landmark of modern diffraction theory, ameliorating many of the problems that had held back precise structure determination.

## Los Alamos Associations

In the early 1950s Zach began to consult at Los Alamos during the summers. During that period he accomplished a scientific *tour de force*. He succeeded in solving the complicated structures of plutonium metal, where others had labored fruitlessly even with computer assistance. In their book, *The Metal Plutonium*, Coffinberry and Miner state:

*It is highly improbable that any scientist other than Zachariasen could have solved the three structures as complex as those of alpha, beta, and gamma plutonium from powder patterns alone.*

He was simply unique in his ability to derive quantitative information from a powder diffraction pattern of a complex mixture. Often after what seemed like a glance, he would suggest cell size and composition. He had an uncanny knack for suggesting the correct atom positions. Only in recent years would he use a pocket calculator. He memorized trigonometric and log values to save time, and combined this with an incredible store of crystallographic information.

I was exposed to this ability and his extraordinary recall in 1953. We were then working with americium compounds. I showed him a film of one and he remarked, "Oh yes, I remember a similar pattern from an unknown plutonium compound, about six years ago." Our knowledge of the composition of the americium compound in combination with his x-ray data provided compositions and structures of both compounds. That was the beginning of a collaboration that continued, often including Larry Asprey, on the structures of the transplutonium metals. Zachariasen also continued his collaboration with Finley Ellinger and Fred Schonfeld. In such collaborations he never took offense in the rare instances when he was wrong and was able to offer guidance in such a tactful way that you were never offended.

With retirement from the University of Chicago and his

move to Santa Fe in the 1970s, there were dinner parties given by Willie and Mossa (his pet name for her) in their Santa Fe home and gracious entertaining of friends from Santa Fe and Los Alamos. On the Hill there were parties involving the Agnews, Argos, Bradburys, Cowans, Evanses, Halls, Hoerlins, Hoyts, Kings, Marks, Matthiases, Metropolitises, Richardsons, Rosens, Spences, Steins, Suydams, Tucks, Turkeviches, and Ulams. There were also poker sessions with a precise dichotomy: the *Chamber Music Society* allowed poker only in the purest form, while the *Symphony Society* permitted a wider variety of play.

For several years a Thursday ritual was established: Zach would come to Los Alamos to consult, first with Finley Ellinger and Fred Schonfeld in the morning; then after their lunch he would come to visit Larry Asprey, Bob Ryan, and me. In mid afternoon he would then go to the labs of Al Giorgi and Gene Szklarz, where he was invaluable in identifying the complicated mixtures involved in their superconductivity work. He went yearly to Bernd T. Matthias's lab in La Jolla for identification of minor but important components in the compounds that resulted from their search for high-temperature superconductors. That particular collaboration produced a series of important papers.

Recently he returned to the theme of 5f element bond lengths and bond strengths, providing the definitive tabulation of their values and simple equations to reproduce them.

## A World Figure

On the world scene, in 1975 I had occasion to introduce Prof. Zachariasen at the Baden-Baden International Conference on Plutonium and The Other Actinides. I mentioned that he was then in his 50th year of publication, publishing his first paper before most in the audience were born. He received a standing ovation. In his self-effacing way he dismissed the tribute by saying, "I was born young." This year is his 55th year of contributing to the literature of science, and I shall always be proud of sharing in his final publication.

To an extent he was an anachronism, a scientific giant out of the times when science was funded from personal resources. For years, he paid his own way to meetings and was proud that he did not seek grants. Others took full advantage of the ready support for post-war science to build personal scientific empires. Zach did not; he felt it took the scientist-administrator too far from the science of the matter.

One measure of scientific impact is the extent that others make use of your work. It is scientific courtesy to acknowledge that debt by citing the former work. The system breaks down, because work of great importance often becomes just part of the lore and the more recent users are cited, not the original. This is true of Zach's work. None-the-less, he was cited 3600 times in the period 1965-1975. This is an average of once for each day of that 10 years. It is clear that his name will remain bright in the literature of science.

To close, it is appropriate to repeat the quotation from Zachariasen's discussion of his arduous work during the war years, "I had a wonderful time."

Willie, so did those who know you!

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.