Engineering Research, Development and Technology

FY 96

Acknowledgments

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.,
Springfield, VA 22161

Engineering Research, Development and Technology

FY 96

Introduction

Spiros Dimolitsas, Associate Director for Engineering

1. Computational Electronics and Electromagnetics

2. Computational Mechanics

3. Microtechnology

4. Manufacturing Technology

Overview
Kenneth L. Blaedel, Thrust Area Leader

5. Materials Science and Engineering

Overview
Donald R. Lesuer, Thrust Area Leader

6. Power Conversion Technologies

7. Nondestructive Evaluation

8. Information Engineering

Author Index

Spiros Dimolitsas, Associate Director for Engineering

The mission of the Engineering Research, Development, and Technology Program at Lawrence Livermore National Laboratory (LLNL) is to develop the knowledge base, process technologies, specialized equipment, tools and facilities to support current and future LLNL programs. Engineering's efforts are guided by a strategy that results in dual benefit: first, in support of Department of Energy missions, such as national security through nuclear deterrence; and second, in enhancing the nation's economic competitiveness through our collaboration with U.S. industry in pursuit of the most cost-effective engineering solutions to LLNL programs.

To accomplish this mission, the Engineering Research, Development, and Technology Program has two important goals: (1) identify key technologies relevant to LLNL programs where we can establish unique competencies, and (2) conduct high-quality research and development to enhance our capabilities and establish ourselves as the world leaders in these technologies. To focus Engineering's efforts, technology *thrust areas* are identified and technical leaders are selected for each area. The thrust areas are comprised of integrated engineering activities, staffed by personnel from the nine electronics and mechanical engineering divisions, and from other LLNL organizations.

The thrust area leaders are accountable to me for the quality, programmatic relevance and progress of their activities, but have sufficient latitude to manage resources allocated to them.

Thrust area leaders are also expected to establish strong links with LLNL programs and industry; be knowledgeable in the state of the art outside of LLNL; use university contacts to supplement and complement their efforts; and use outside and inside experts to critically review the quality and direction of their work.

For the past several years, the Engineering Research, Development, and Technology Program has played a significant role in leveraging LLNL technology to enhance U.S. competitiveness, and in allowing us to access the best of industrial technologies for the benefit of LLNL programmatic missions. As such, it has served as the foundation for many of the Cooperative Research and Development Agreements now in place between LLNL programs, universities, and U.S. industries.

This annual report, organized by thrust area, describes Engineering's activities for fiscal year 1996. The report provides timely summaries of objectives, methods, and key results from eight thrust areas: Computational Electronics and Electromagnetics; Computational Mechanics; Microtechnology; Manufacturing Technology; Materials Science and Engineering; Power Conversion Technologies; Nondestructive Evaluation; and Information Engineering.

Readers desiring more information are encouraged to contact the individual thrust area leaders or authors.

Introduction

Clifford C. Shang, Thrust Area Leader

The Computational Electronics and Electromagnetics thrust area at Lawrence Livermore National Laboratory serves as the focal point for engineering R&D activities for developing computer-based design, analysis, and tools for theory.

Key representative applications include design of particle accelerator cells and beamline components; engineering analysis and design of high-power components, photonics and optoelectronics circuit design; EMI susceptibility analysis; and antenna synthesis.

The FY-96 technology-base effort focused code development on (1) accelerator design codes; (2) 3-D massively parallel, object-oriented time-domain EM codes; (3) material models; (4) coupling and application of engineering tools for analysis and design of high-power components; (5) 3-D spectral-domain CEM tools; and (6) enhancement of laser drilling codes.

Joint efforts with the Power Conversion Technologies thrust area include development of antenna systems for compact, high-performance radar, in addition to novel, compact Marx generators.

Computational Electronics and Electromagnetics

1

1. Computational Electronics and Electromagnetics

Three-Dimensional Electromagnetic Modeling of Wakefields in Accelerator Components

Brian R. Poole and Wang C. Ng
Defense Sciences Engineering Division
Electronics Engineering

George J. Caporaso
Inertial Confinement Fusion
Laser Programs

This report describes continuing work on the development of computational wakefield tools for the design of accelerator components for the advanced radiographic machine (ARM). New massively-parallel 3-D time-domain electromagnetic codes now under development using conforming unstructured meshes with advanced radiation boundary conditions allow a substantial increase in the geometric fidelity of the structures being modeled, as well as a more suitable truncation of the problem space. A variety of 3-D structures are tested with an existing cubical cell finite-difference time-domain (FDTD) code and the wake impedances are compared with simple analytic models for the structures. These results provide a set of benchmarks for testing the new time-domain codes and provide insight into the wakefield physics of the structures. Structures under consideration include a stripline beam position monitor and a circular aperture in a circular waveguide. We have achieved excellent agreement for the monopole and dipole impedances for the models below the cut-off frequency of the beamline for a stiff electron beam moving parallel to the beamline axis. We are extending the wakefield calculations to include beams that have more complex trajectories within the structure, which will allow more realistic physics to be included in calculations for ARM.

Introduction

In FY-95 and FY-96, we began development of tools to analyze the interaction of relativistic charged-particle beams with accelerator components. This work is primarily motivated by the need for advanced modeling techniques for problems relevant to the design of electron beam accelerators for ARM. The ARM requires multiple electron beams to be directed from various lines of sight onto a target for the generation of x rays for use in pulsed radiography applications. A cost-effective design requires that the electron beams be generated from a single linear induction accelerator.

To accomplish the generation of multiple electron beams, several beam kickers and beamline septums are used. A beam kicker is a structure designed to spatially separate an electron beam pulse into two separate electron beam pulses. The kicker structure and septum region are shown schematically in **Fig. 1**. The kicker, which consists of two pairs of transmission lines enclosed in a cylindrical beamline, is designed to steer the relativistic electron beam (REB) into the correct trajectory for transport into a magnetic field that bends the beam into the downstream beamlines.

One pair of transmission lines is grounded at each end by connecting them to the beam pipe. The other pair is connected to external transmission lines that can be driven by an external modulator at one end of the structure. The other ends of the driven transmission lines are connected to matched loads. By alternating the polarity of each applied electromagnetic pulse on the driven transmission lines, alternate electron beam pulses can be directed into each of the two downstream beam transport systems.

After the electron beam is separated into two beam pulses, each beam pulse is directed into the appropriate beamline by the applied magnetic field in the septum region. The magnetic field has opposite field directions on each side of the plane separating

the downstream beam transport systems to bend each beam into the appropriate system.

Understanding the wake properties of these structures is crucial to the design of accelerator components for the ARM. For high current electron beams, the ability to steer the beam through the kicker and septum regions is strongly dependent on the wake properties of these structures. If the kicker possesses a strong transverse dipole impedance, a high current electron beam can induce large voltages on the steering transmission lines.

Since these voltages may be comparable to the externally applied voltage, the ability to steer the beam may be impaired, and improved dynamic control of the external pulser voltage will be required to compensate for the beam induced steering. Similarly, the presence of a large aperture in the septum region may lead to deleterious effects on beam quality.

As discussed in a previous report,[1] it is not practical with our current cubical cell FDTD codes to model the full scale kicker and septum regions, due to the large mesh sizes required for the model and the long run times required for low frequency resolution. Also, because of the lack of advanced radiation boundary conditions, boundary conditions cannot be applied to the oblique surfaces associated with the exit beamlines of the septum region. However, confidence in analysis techniques can be achieved by modeling smaller scale devices similar in design to these structures, such as the stripline beam position monitor and the waveguide aperture.

Wakefield Analysis

A crucial area in the design of linear induction accelerators and beam transport systems is the understanding of the wakefields generated by a charge bunch as it is transported through the system. As this charge bunch passes a perturbation in an accelerator structure, an electromagnetic wakefield is generated which can re-interact with the charged particle beam. This process can lead to degradation of the beam, or worse, can lead to beam breakup instabilities with the resultant loss of beam quality at the target.

Consider a beam transport system consisting of a cylindrical drift tube with some perturbation in its structure followed again by a cylindrical drift tube, as shown in **Fig. 2**. The perturbation is not necessarily cylindrically symmetric and may consist of features such as changes in the cross-section, apertures in the wall, or different material properties. An axial current of the form defined in **Eq. 1** is used to represent the current due to a moving source charge Q in the beam pipe.

$$J_z = I_0 \delta(r - r_0)\delta(\phi - \phi_0)\delta(vt - z) \qquad (1)$$

Here r_0 and ϕ_0 represent the transverse location of the source charge in the accelerating structure, z is the axial location of the source at time t, I_0 is the strength of the current source, and $\delta(vt - z)$ represents the location of the source charge, Q. If we were modeling a charge bunch, the delta function $\delta(vt - z)$ would be replaced by a function $f(vt - z)$ representing the shape of the charge bunch. Typically a narrow Gaussian pulse is used as the source function in our modeling to approximate a delta function while preserving a "clean" turn on of the sources.

In this analysis we are interested in the response of a test charge that is trailing the source charge by a distance $s = vt - z$. The electromagnetic fields in the structure are considered to be the sum of multipole modes as shown in **Eq. 2**:
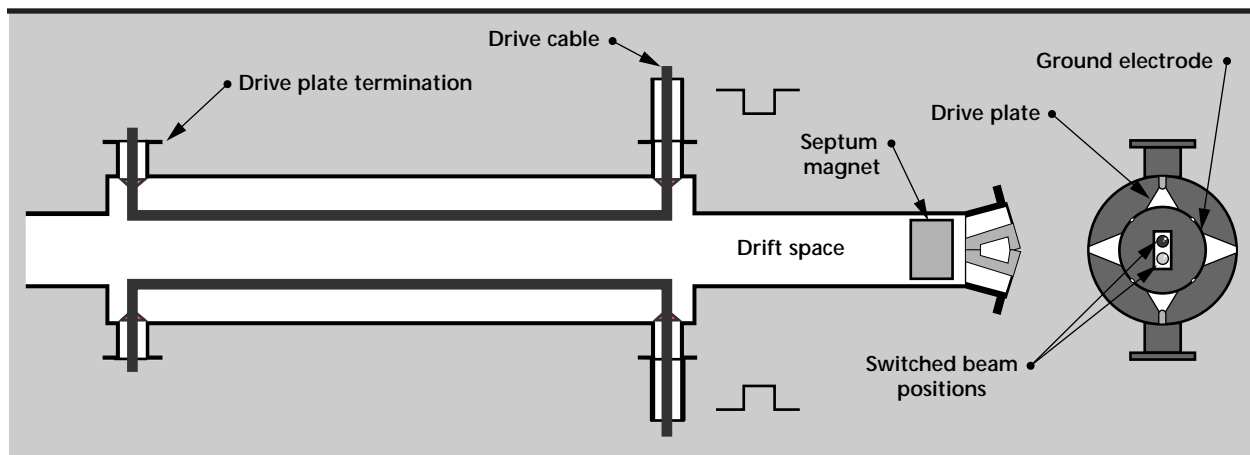
Figure 1. Schematic representation of the kicker. The kicker spatially separates the electron beam and the septum magnet bends the individual beams into the appropriate beamline.

$$\psi(r,z,t) = \sum_{m=0}^{\infty} \left[\psi_m^p(r,z,t)\cos(m\phi) \right. \tag{2}$$
$$\left. + \psi_m^q(r,z,t)\sin(m\phi) \right]$$

where ψ represents any electromagnetic field component, and ψ_m^p and ψ_m^q represent the various multipole coefficients for the specific field component. Since the force on a test charge q due to the electromagnetic field generated by the source point charge Q is given by

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{3}$$

it is convenient to define the concept of an impulse wake potential as

$$\mathbf{W}(s) = \int_{-\infty}^{\infty} (\mathbf{E} + \mathbf{v} \times \mathbf{B})\Big|_{t=(s+z)/v} dz. \tag{4}$$

The impulse wake potential for each multipole mode can be determined by applying **Eq. 4** to each field component as defined in **Eq. 2**. A wake impedance function can be defined by taking the Fourier transform over the variable $s = vt - z$ of the wake potential for the longitudinal and transverse components of the force for each multipole mode. Since the strength of the multipole component of the source current is proportional to r_0^m, this factor is included in the definition of the vector wake impedance as

$$\mathbf{Z}_m(\omega) = \frac{1}{r_0^m I_0} \int_{-\infty}^{\infty} \mathbf{W}_m(s) e^{-i\omega s/v} ds \tag{5}$$

and represents a figure of merit defining the strength of the interaction of a charged-particle beam with an electromagnetic structure in both the longitudinal and transverse directions. It is desirable to keep the wake impedance function as small as possible to prevent undesirable coupling between the beam and the structure.

Progress

Modeling

TSAR is the primary 3-D electromagnetic field code used for analysis of these structures. TSAR uses a 3-D grid composed of cubical cells for the FDTD analysis, with each cell being assigned a given material property. The cell size is typically chosen to resolve the shortest wavelength expected in the problem and, in addition, must be chosen small enough to accurately model the smallest dimension in the model. To satisfy stability requirements for the FDTD field solver, the Courant stability condition for the time step must be satisfied. This condition is defined in **Eq. 6**

$$dt < \frac{1}{2c} dx \tag{6}$$

where dx is the length of the cell edge.

Beam Position Monitor

The stripline beam position monitor is discussed in some detail by Ng.[2] The beam position monitor consists of two pairs of strip transmission lines,
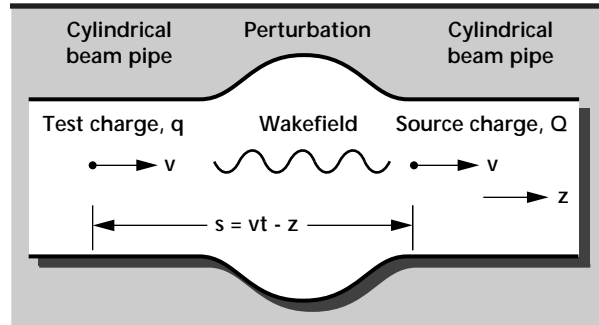
Figure 2. Geometry of a wakefield calculation. The source of the electromagnetic wakefield is the charge, Q. The spectral characteristics of the longitudinal and transverse forces are determined on a test charge, q subjected to the wakefield.
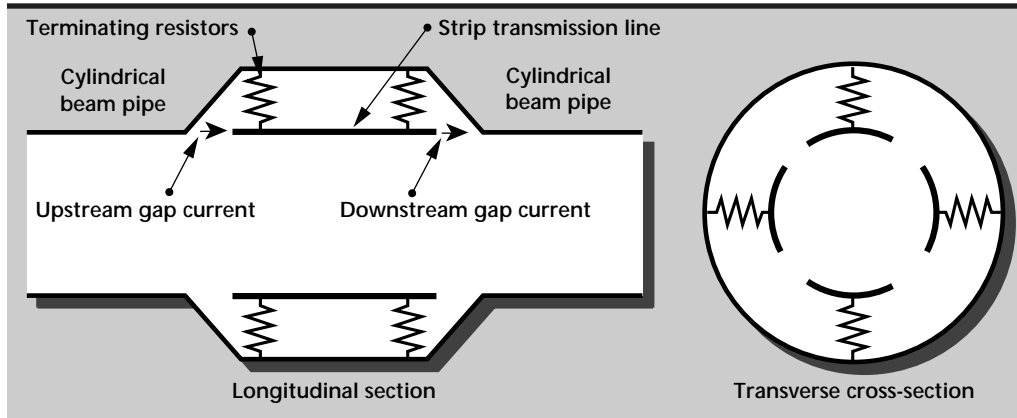
Figure 3. Schematic representation of the stripline beam position monitor.

shown schematically in **Fig. 3**. The 25-Ω terminating resistors were chosen based on a frequency domain calculation of the TEM impedance of the structure. For modeling purposes, the terminating resistors represent external coaxial transmission line feeds to the striplines. Even to represent the gross geometric characteristics of the structure using cubical cells requires a cell size of 0.5 mm, making the computational size of this problem 8×10^6 cells and requires a time step of .83 ps, as dictated by the Courant stability condition.

To accurately represent the details of the coaxial feeds would have required a much smaller cell size to have adequate geometric fidelity in the feed region. Therefore, it was decided to use simple wire resistors to model this region to minimize the computer memory requirements. The length of the striplines was 10 cm, which indicates a quarter-wavelength resonance of 750 MHz, allowing us to

sufficiently resolve the low frequency behavior using 16000 time steps. Approximately 25 cm of beam transport line was used on each side of the beam position monitor to allow evanescent waves to decay and to minimize the effects of the MUR radiation boundary conditions applied to the axial boundaries of the problem space.

The structure is excited with a current distribution of the form defined in **Eq. 1**. Analytic expressions for the longitudinal monopole ($m = 0$) and transverse dipole ($m = 1$) impedances for a matched transmission line model of the beam position monitor are[2]

$$Z_\parallel = Z_s \left(\frac{\phi_s}{2\pi} \right)^2 \tag{7}$$

$$\times \left[\sin^2 \left(\frac{\omega L}{c} \right) + j \sin \left(\frac{\omega L}{c} \right) \cos \left(\frac{\omega L}{c} \right) \right]$$
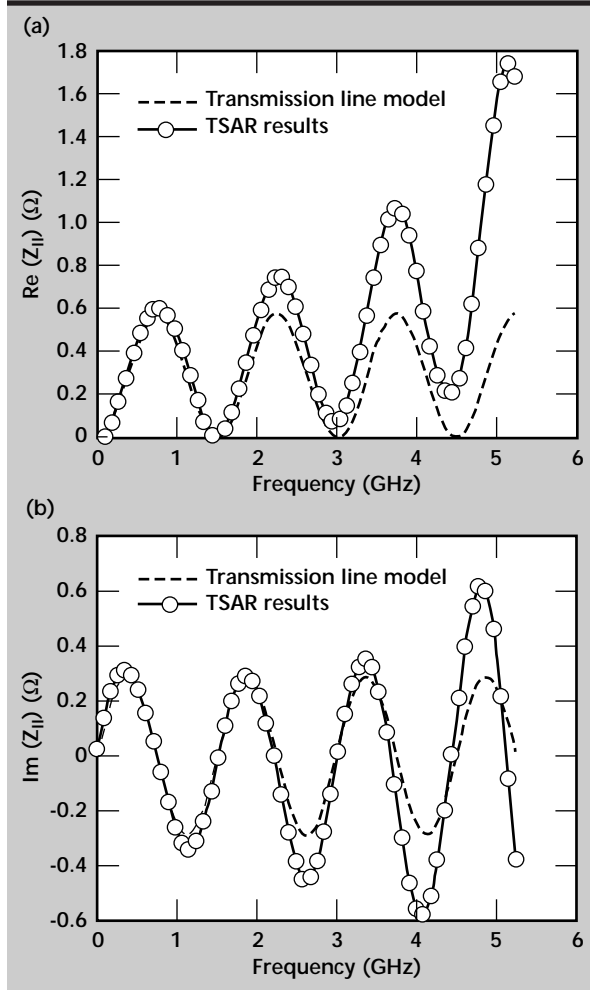
Figure 4. *Code results and analytic results for the real (a) and imaginary (b) components of the longitudinal impedance of the stripline beam position monitor.*
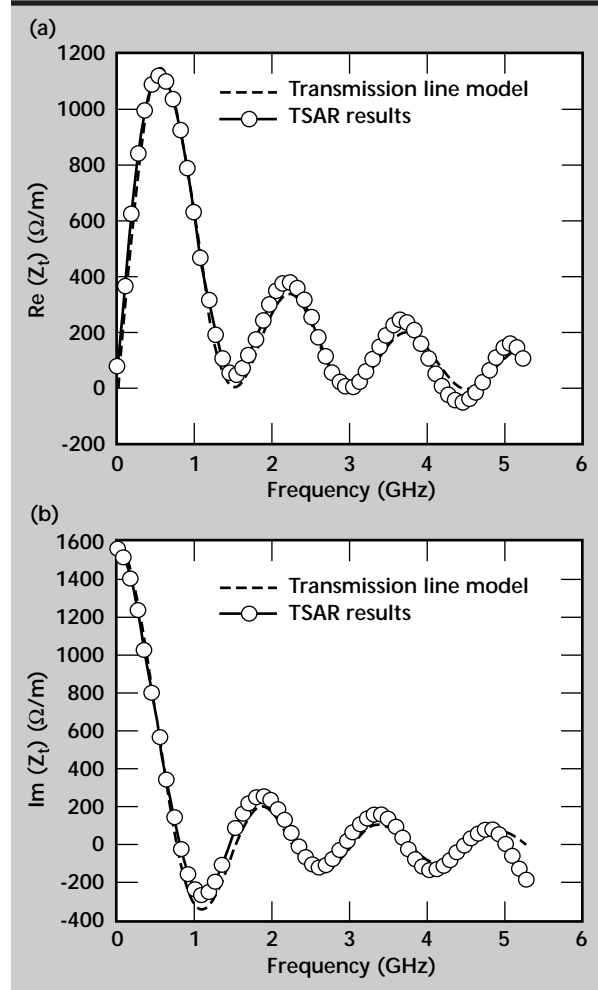
Figure 5. *Code results and analytic results for the real (a) and imaginary (b) components of the transverse dipole impedance of the stripline beam position monitor.*

$$Z_\perp = \frac{8cZ_s}{\pi^2 R_s^2}\left(\frac{1}{\omega}\right)\sin^2\left(\frac{\phi_s}{2}\right)$$

$$\times\left[\sin^2\left(\frac{\omega L}{c}\right) + j\sin\left(\frac{\omega L}{c}\right)\cos\left(\frac{\omega L}{c}\right)\right]. \qquad (8)$$

Figures 4 and **5** show comparisons of the analytic models with results obtained numerically through the TSAR modeling code for the longitudinal and transverse impedances, respectively. There is excellent low frequency agreement. The deviation at higher frequencies may be attributed to the high frequency impedance of the wire resistors which are not accurately represented by a cell size of 0.5 mm.

Circular Aperture in Circular Beamline

The impedance of coupling apertures in beam transport lines has been treated in the literature by Gluckstern[3] and Kurennoy.[4] In this section we consider a 3-cm diameter cylindrical beamline with either an 8-mm or a 3-cm circular waveguide connected at 90° to the beamline, as shown in **Fig. 6**. The 3 ports of the problem space are terminated with MUR radiation boundary conditions. A cell size of 0.5 mm is used to model both aperture problems. The length of the beamline is about 30 cm and the coupling waveguide is 3 cm long, resulting in a mesh of about 1.2×10^7 cells. An analytic expression for the low frequency transverse dipole impedance is given by **Eq. 9**.

$$Z_\perp = j\frac{2a^3 Z_0}{3\pi^2 R_0^4}f(w) \qquad (9)$$

where Z_0 is the impedance of free space and $f(w)$ is a correction factor based on the length of the coupling guide and has a value of 0.562 for $w/a > 2$, as defined in Gluckstern.[3] In our model, $w = 3$ cm. For the 8-mm coupling guide **Eq. 9** gives an inductive transverse impedance of 18.1 Ω/m. The TSAR model predicts an inductive transverse impedance of 18.3 Ω/m for frequencies up to 6 GHz.

Figure 7 shows the TSAR results for the 3-cm large aperture case for the transverse impedance along with the low frequency asymptote for the reactive component of the impedance. The real part arises from the fact that the secondary coupling waveguide is only 3 cm long in the model and does not let low frequency fields evanesce sufficiently before reaching the MUR boundary condition on the coupling guide. This accounts for the resonance near the TE_{11} cutoff at 5.9 GHz.

Future Work

The calculations and modeling of wake impedances in this report are based on simple models of complex structures. The structures that need to be modeled are typically two orders of magnitude larger in volume, requiring more efficient volume meshing to effectively model the structures with a high degree of geometric fidelity. In addition, improvements to the radiation boundary conditions need to be implemented to model structures with more complicated truncations to the problem space, such as the oblique surfaces on the exit beamlines of the kicker septum region.

The new DSI-TIGER code under development includes such features as conforming grids, hybrid structured/unstructured grids, massively parallel capability, and new radiation boundary conditions. Conforming grids eliminate stair-casing errors associated with cubical cell codes and allow more realistic models to be developed for complicated structures. However, conforming unstructured non-orthogonal grids require typically 30 times the memory overhead compared to structured orthogonal grids. Hybrid grids can overcome this limitation by using conforming grids only near appropriate surfaces and structured grids elsewhere. New advanced radiation boundary conditions can also reduce the computational effort by allowing a closer truncation of the problem space.

New physics needs to be incorporated in existing and new codes to put prescribed electron beam trajectories in the accelerator problems. For example, in the kicker the electron beam does not follow a simple straight line trajectory parallel to the beamline axis. The interaction of the beam with the externally applied fields produces a trajectory in the structure as governed by the dynamical equations of motion in the externally prescribed electromagnetic field. Extensions to
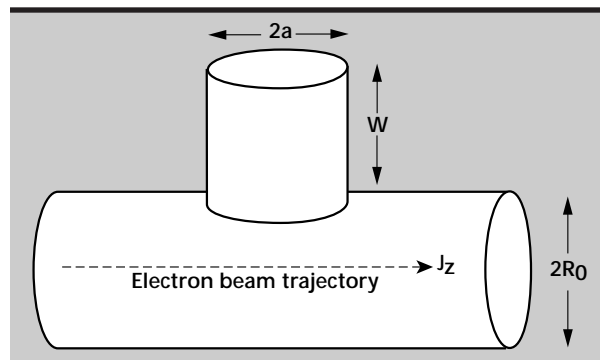
Figure 6. Schematic representation of the waveguide junction aperture problem.

Figure 7. Transverse dipole impedance for the 3-cm large aperture waveguide junction. The resonance near 5.9 GHz is due to the TE_{11} cutoff. The real part of the impedance results from the fact that the short length of the coupling guide does not allow evanescent waves to sufficiently decay before reaching the MUR radiation boundary condition.

$$\text{Im } Z_t (0) = \frac{2a^3 f (w)}{3\pi^2 b^4} Z_0$$

- - - - - Small aperture theory (imag. part)
—○— TSAR results (imag. part)
—□— TSAR results (real part)
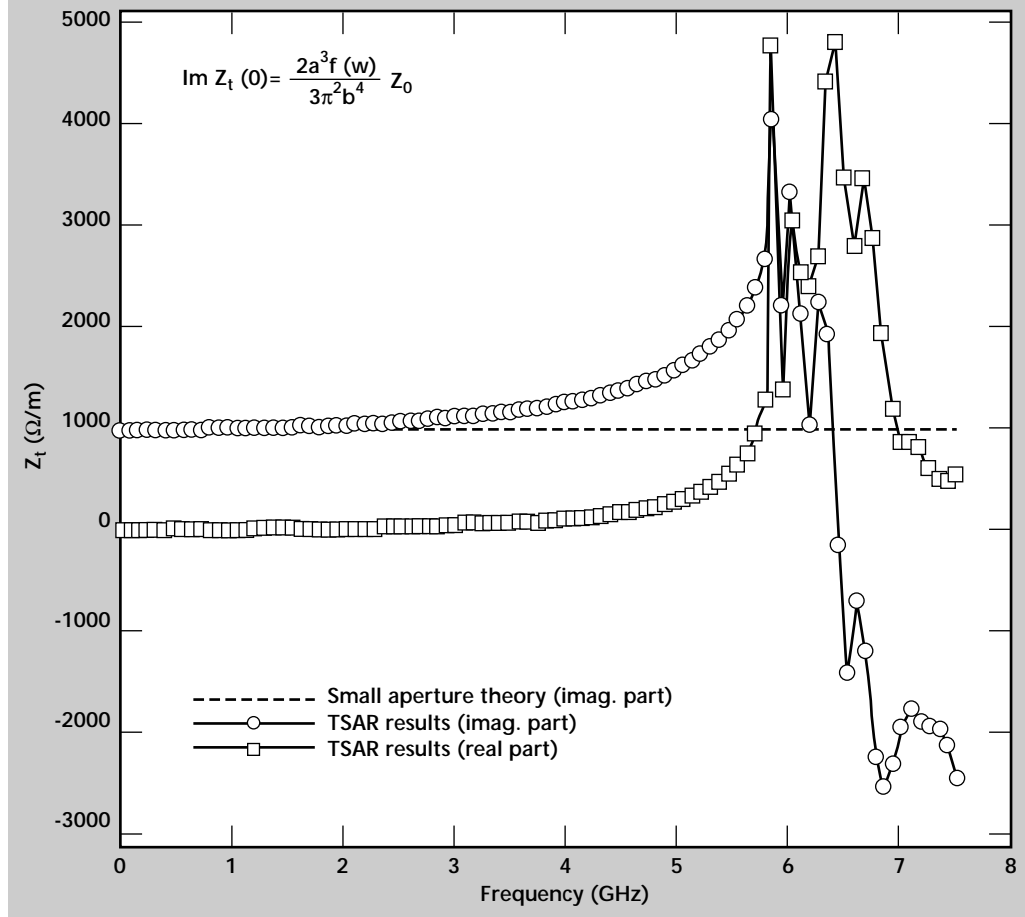
wakefield codes need to include the real prescribed electron beam trajectory to determine the wake properties of the structure.

Preliminary work in this area has already been applied to the TSAR cubical cell code with some success. However, proper interpretation of the results needs to be addressed to understand the wake physics in this regime. We are continuing work in this area by developing distributed impedance models of simple structures for complicated beam trajectories.

Acknowledgments

References

1. Poole, B. R., W. C. Ng, C. C. Shang, and G. J. Caporaso (1995), *Wakefield Modeling of Advanced Accelerator Components,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-ID-122235).

2. Ng, K-Y (1988), "Impedances of Stripline Beam-Position Monitors," *Particle Accelerators,* **23**, p. 93.

3. Gluckstern, R. L. (1992), "Coupling Impedance of a Single Hole in a Thick-Wall Beam Pipe," *Phys. Rev. A,* **46**, pp. 1106–1109.

4. Kurennoy, S. S. (1992), "Coupling Impedance of Pumping Holes," *Particle Accelerators,* **39**, pp. 1–13.

TIGER: An Object-Oriented Time-Domain Electromagnetics Simulation Code

David Steich, Jeffrey S. Kallman, J. Brian Grant, Steve Pennock, Niel A. Madsen, Gerald J. Burke, and David J. Mayhall
Engineering Research Division
Electronics Engineering

Brian R. Poole
Defense Sciences Engineering Division
Electronics Engineering

Karl Kunz
The Pennsylvania State University
State College, Pennsylvania

Hal R. Brand and Gary W. Laguna
Computer Applications Sciences and Engineering Division
Computation Organization

The TIGER (from "time domain generalized excitation and response") project began in FY-96. This report provides a description of TIGER's intended performance and the progress we have made.

Introduction

Lawrence Livermore National Laboratory (LLNL) has enjoyed a long history of time-domain leadership in the Computational Electromagnetics (CEM) community, with the development of codes such as AMOS, FDTD, TSAR, TSARLITE, and DSI3D. However, these codes are either research oriented (DSI3D), have limited applicability [AMOS (2.5-D), TSARLITE (2-D)], and/or are no longer regarded as state-of-the-art (TSAR, FDTD). Furthermore, many of these codes have multiple versions that were created over the years for specific applications requiring additional capabilities (such as dispersive TSAR, out-of-core TSAR, dispersive FDTD, near-to-far zone FDTD, advanced boundary condition FDTD, and anisotropic TSARLITE). These codes represent LLNL Engineering's ability to address programmatic time-domain linear electromagnetic modeling needs.

Without exception, all of these codes are extremely difficult to up-grade with the latest technology developments. For example, porting TSAR to a MPP platform, adding required communications to DSI to allow high-order advanced boundary conditions, and modifying AMOS to work for conforming meshes, are all formidable tasks.

At present, the incorporation of new physics into existing codes requires a conscious decision as to which specific code to enhance. Resources are rarely, if ever, available to place enhancements and new physics into several codes. The result is a series of codes with multiple versions built in isolation of one another, making it almost impossible to link them together to provide a larger set of capabilities. The existing problems associated with this collection of CEM codes in LLNL's engineering programs are magnified when difficult simulation requirements demand multiple technology enhancements simultaneously. These limitations will become increasingly debilitating as future modeling requirements become more complex.

In response to these problems, we developed the concept of the TIGER code, an MPP CEM physics kernel that would strengthen LLNL's leadership in linear time-domain CEM technology.

Due to the increasing complexity required to model realistic structures, simultaneous technologies must be suitably applied to reduce the computational burden. **Table 1** shows several such technologies and their potential improvement factors (that is, accuracy, required memory, and/or CPU improvements). Conforming grids eliminate staircasing errors associated with codes such as AMOS, TSAR and FDTD[1].

This improvement in accuracy can be thought of as a reduction in required number of unknowns, since for a given fixed accuracy a coarser grid can be used to represent a given geometry. The drawback is that conforming unstructured nonorthogonal grids often require roughly 30 times the memory overhead, compared to structured orthogonal grids.

However, by using hybrid structured/unstructured grids this high overhead can, to a large extent, be circumvented by judiciously using conforming grids only near appropriate surfaces. Running the code on MPP platforms allows much larger problems to be run much faster than on serial platforms.

Advanced Radiation Boundary Conditions (RBCs) can also reduce the computational effort considerably by allowing the truncation of the problem space to be placed much closer to the scatterer[2]. The actual improvement factors are problem-dependent.

To minimize computational requirements, this full set of technologies must be used simultaneously in a cooperative manner. Simultaneous incorporation of the above technologies in a code that includes advanced materials, physics, interface, and boundary conditions, each with disparate computation and memory requirements, is an ambitious task.

The vision of TIGER extends beyond state-of-the-art modeling of linear time-domain CEM modeling, and includes acoustic wave, elastic wave, and plasma physics modeling capabilities. To help meet the challenges of this task we are collaborating with the EIGER projects at LLNL and Sandia National Laboratories, and the Pennsylvania State University. Each of these collaborators offers much expertise that strengthens the overall effort.

Progress

During the past year we focused our attention on the design of the object-oriented infrastructure that defines the parameter space of the TIGER effort. Object-oriented codes allow flexibility far beyond conventional procedural-based codes. This greatly enhanced flexibility comes at the expense of the extensive up-front design required to build the object-oriented infrastructure. The better the design of the interfaces making up the object-oriented classes, the more flexible the code becomes. A successful design allows new physics capabilities to be included by simple (localized) code additions.

We have built multiple prototype versions of TIGER that test various aspects of the intended performance. We built a FORTRAN 77 quasi-object-oriented version of the code that runs on the CS-2 Meiko and the Cray T3D. This code tested MPI (Message Passing Interface) libraries, simple partitioning algorithms for structured grids, PACT/PANACEA software, and communication overheads associated with algorithms such as the near-to-far zone transformation.

The FORTRAN version works for rectangular structured-grid cells. It was determined that communication overhead for structured grids would be as much of a bottleneck as for unstructured grids even though less information must be communicated between partition boundaries.

Also, it was found that many of the advantages PACT/PANACEA software provides FORTRAN and C-based codes (that is, I/O management, variable dictionaries, and automatic restart capabilities) were not as useful in an object-oriented environment. For example, the need for variable

Table 1. Anticipated performance improvement factors using a given technology enhancement.

Technology enhancement	Improvement factor	Accuracy	Memory	CPU time
Conforming grids	up to 100 $\times$	√	√	√
Hybrid structured/ unstructured grids	up to 30 $\times$		√	√
Massively Parallel Processors (MPP)	up to 60 $\times$		√	√
Advanced Radiation Boundary Conditions (RBC)	up to 20 $\times$	√	√	√

dictionaries becomes practically obsolete in a C++ language that allows extensive information hiding and name space capabilities.

Also, at this time the PACT/PANACEA software does not support C++. For these reasons we decided not to pursue using PACT/PANACEA within the TIGER code. However, in the future we intend to interface with data and codes that use PACT/PANACEA software by providing filters that can translate to the PACT/PANACEA format.

Other prototype C++ versions of TIGER tested primarily the mesh management systems required to insulate the physics from the bookkeeping. Insulation of the physics from mesh-specific operations is the key factor in the extensibility, flexibility, and maintainability of the code.

Much effort during the past year has also been expended on issues of memory and execution time efficiency. A major goal of the TIGER project is to have overhead as close as possible to that of existing codes which are specialized for specific algorithms. In the past, object-oriented codes were not able to compete with "hand-coded" optimized FORTRAN or C codes. However, with the use of template expressions and traits[3,4] a fully objected-oriented C++ code can perform within a few percent of hand-coded software.

This year we also up-graded the Anastasia mesh generation package to allow BRLCAD (a popular army CAD/CAM code) solid models to be meshed in terms of rectangular brick elements. The modification of Anastasia will allow us to meet some short term deliverables. A much more sophisticated mesh generator will soon be required as TIGER becomes more functional.

Object-Oriented Approach

The object-oriented infrastructure we have begun building this year is shown in **Fig. 1**. The figure shows six management systems that will be built to insulate the core physics algorithms from the rest of the code, thus freeing physics programmers from having to worry about:
1) what type of mesh their algorithm will be applied to (mesh management);
2) whether the code will be run serial, parallel distributed, or shared (communication management);
3) the allocating and freeing of memory (memory management);
4) how to read in, store, and write out information (I/O management);
5) whether the user is attempting illegal tasks (parameter and constraint management); and

6) how to manage input parameters, runs, results, projects (database management).

Providing an object-oriented framework with long-term extensibility for each of these management systems will be accomplished by an emphasis on isolating concepts and their fundamental relationships. On the surface, this sounds insignificant but in actuality it is the key to extending the shelf-life of an object-oriented program.

A small portion of the mesh management system is displayed in **Fig. 1**. The mesh management abstraction allows different kinds of meshes to be used in an identical way by providing a common mesh interface that hides details. The structured mesh database that exists behind the interface consists of a series of functions, while the unstructured mesh database consists of a complex set of memory index arrays.

In numerical analysis, discretized equations are often used to approximate the actual equations. In TIGER's object-oriented approach, discretized operators are used to build these discretized equations (see lower portion of **Fig. 1**) that simulate the physics. In this way, equations can be thought of as a collection of operators applied to field quantities. These discretized operators manipulate field quantities that lie on the numerical mesh.

Manipulation of fields by these operators occurs through the mesh interface and allows the same operators to be used independent of mesh type. The mesh management system is a vital part of the overall TIGER effort, and has been the primary focus of our efforts this year. A major portion of this management system has been completed.

The memory management system organizes memory-specific allocations and deallocations in an efficient manner. Often in object-oriented codes the getting and freeing of memory can be a major portion of the overall execution time. This interaction with the operating system can degrade performance compared to "hard-wired" FORTRAN or C-based codes. Our memory management system grabs large blocks of memory from the system and dispenses required chunks of memory efficiently. A rudimentary memory management system has been built this year.

We have also begun work on the communication and parameter/constraint management systems as well. The I/O and database management systems are future tasks that will become important as TIGER becomes more functional.

The object-oriented operator-based approach of TIGER not only makes the code easily extensible and maintainable but also provides flexibility far beyond the conventional procedural approach. For example,
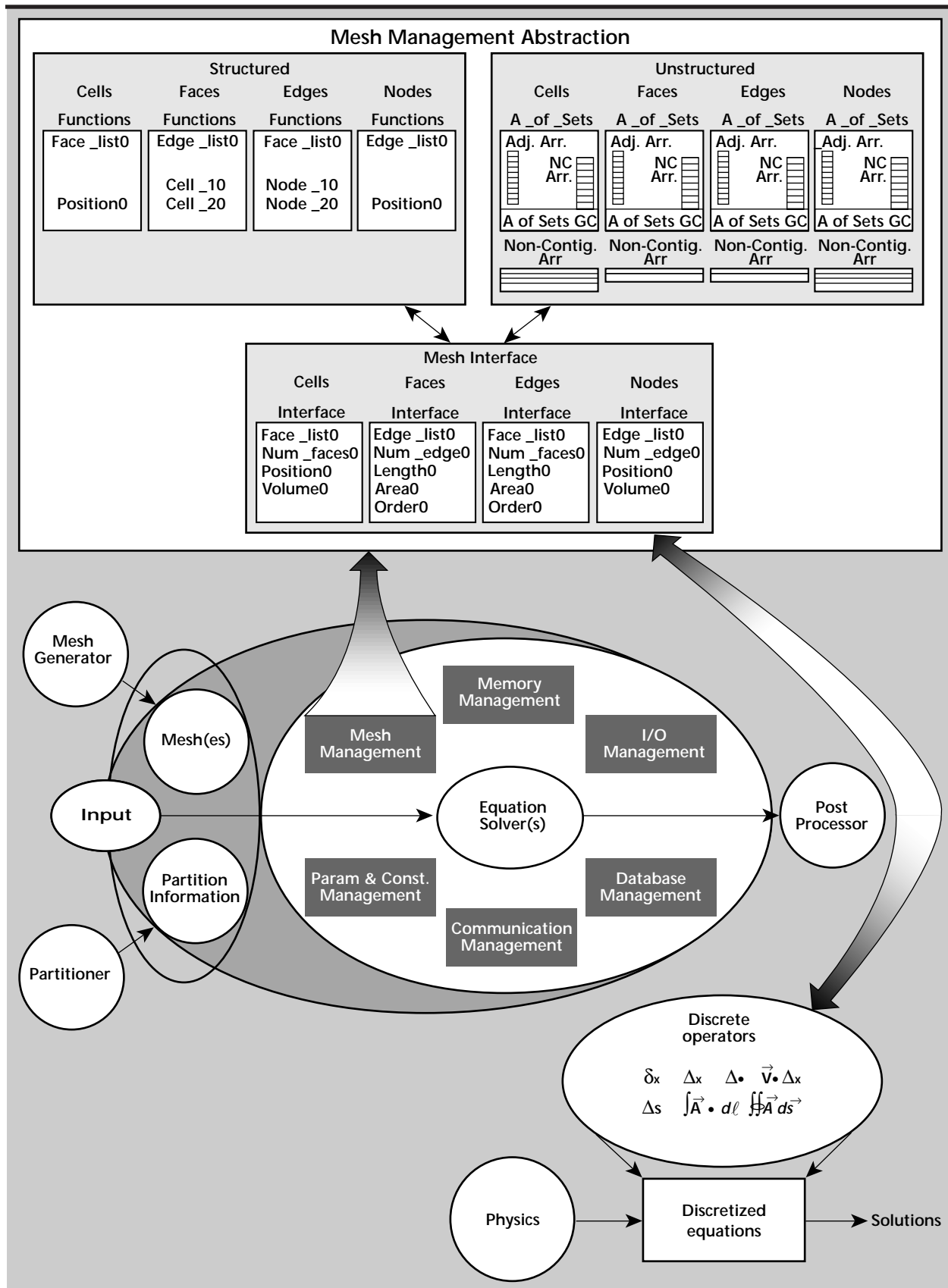
Figure 1. Insulating generalized physics from bookkeeping.

consider the implementation of RBCs, which are required to truncate the numerical mesh for all open scattering problems.

A pictorial representation of a procedural versus object-oriented approach is shown in **Fig. 2**. To write a code containing many possible RBCs which could be applied to many different mesh types typically would require the product of (# of RBCs) × (# of mesh types) implementations (see top portion of **Fig. 2**).

Using an object-oriented approach, the mesh and RBC equations can both be abstracted (see bottom portion of **Fig. 2**). A single implementation can then be constructed using the mesh interface and the abstracted set of operators to obtain the appropriate set of operations. The single "operator-based" equation is more complicated to write than one of the "procedural-based" equations, but the same implementation works for almost every mesh type and RBC type which can be cast into the abstracted forms (factors such as numerical stability, numerical dispersion, and stencil size may prohibit certain implementations). The key is to have the abstractions be as general as possible to encompass the largest set of capabilities. Progress on these abstractions is underway and is at the heart of the TIGER effort.

The mesh management system is responsible for the grid abstractions (**Fig. 3**). A very small subset of RBC types is shown on the left-hand side of **Fig. 3**. Each RBC type has multiple orders and

variations resulting in a plethora of possible RBC equations. Shown near the center of **Fig. 3** is a blow-up of a third-order Higdon boundary condition. Inside the matrix are 64 terms, with each term having a dozen or so sub-terms. Other more complicated boundary conditions require 4-D matrices involving hundreds of terms to describe the RBCs' coefficients. Shown in the lower portion of **Fig. 3** are simple operators **A**, **B**, **C**, that have been abstracted out of the RBC equations along with some operations that can be used with the operators, such as +, - , dot product, inverse.

Also shown in the lower portion of **Fig. 3** is the same third-order Higdon boundary condition written in terms of these operators. Note the significant reduction in the complexity of equation representation. The expression is exactly identical to the expression shown in the upper portion of **Fig. 3**. The approach is very general and even allows new boundary conditions to be invented simply by defining new operators or by concatenating them together (that is, a Liao-Higdon-Mur RBC).

The end result is a powerful technique that can greatly increase accuracy and reduce execution times required to solve problems. Shown on the right-hand side of **Fig. 3** are a second-order Mur and a third-order Taylor RBC. Shown are the results of a convergence study testing the number of cells required before a converged solution is obtained.
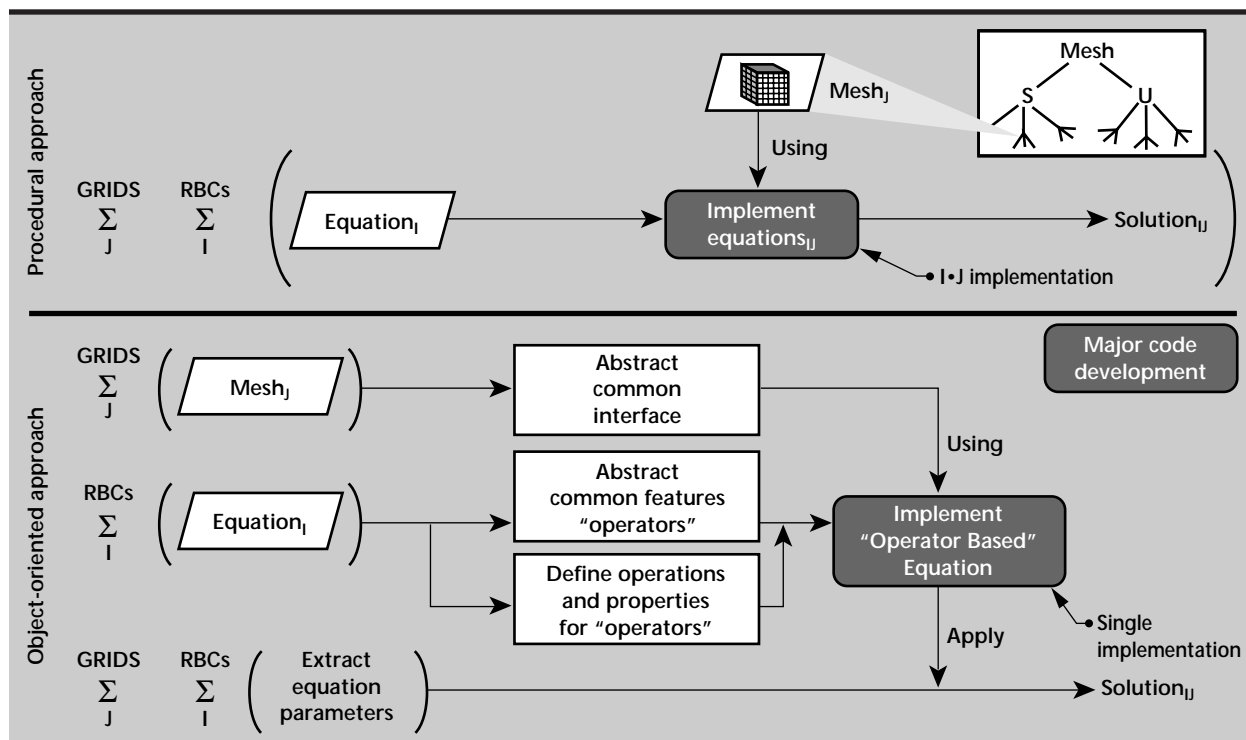
Figure 2. RBCs: Procedural versus object-oriented approach.

It is obvious that the Taylor boundary condition is fully converged already, using 80,000 cells. The Mur RBC, perhaps the most popular boundary condition used over the past 15 years, has still not converged, even after using over 1.2 million cells. The Taylor boundary condition was invented using the operator abstractions described above. We now have the capability to represent literally hundreds of RBCs using this research. We can model virtually every local RBC in the community today simply by changing parameters used to initialize the operators. (A new boundary condition called the Perfectly Matched Layer (PML) requires a slightly different set of operators but still fits into the above described approach.)

This operator-based approach to RBCs is cutting-edge technology, unique in the world.

Future Work

The focus of our collaborative efforts for FY-97 will be to continue to build the objected-oriented management class infrastructures that provide the required insulation between the physics and the bookkeeping. Specifically, the essential parts of the mesh, communication, and memory management systems are expected to be in place in FY-97. The mesh management system will eventually allow numerous mesh types (such as structured, unstructured, expanding, uniform, 2-D, 3-D, skewed, and nonorthogonal) to be abstracted into a common interface.

The communication management system will organize processor topologies, sending and receiving of data, and execution barriers required for shared
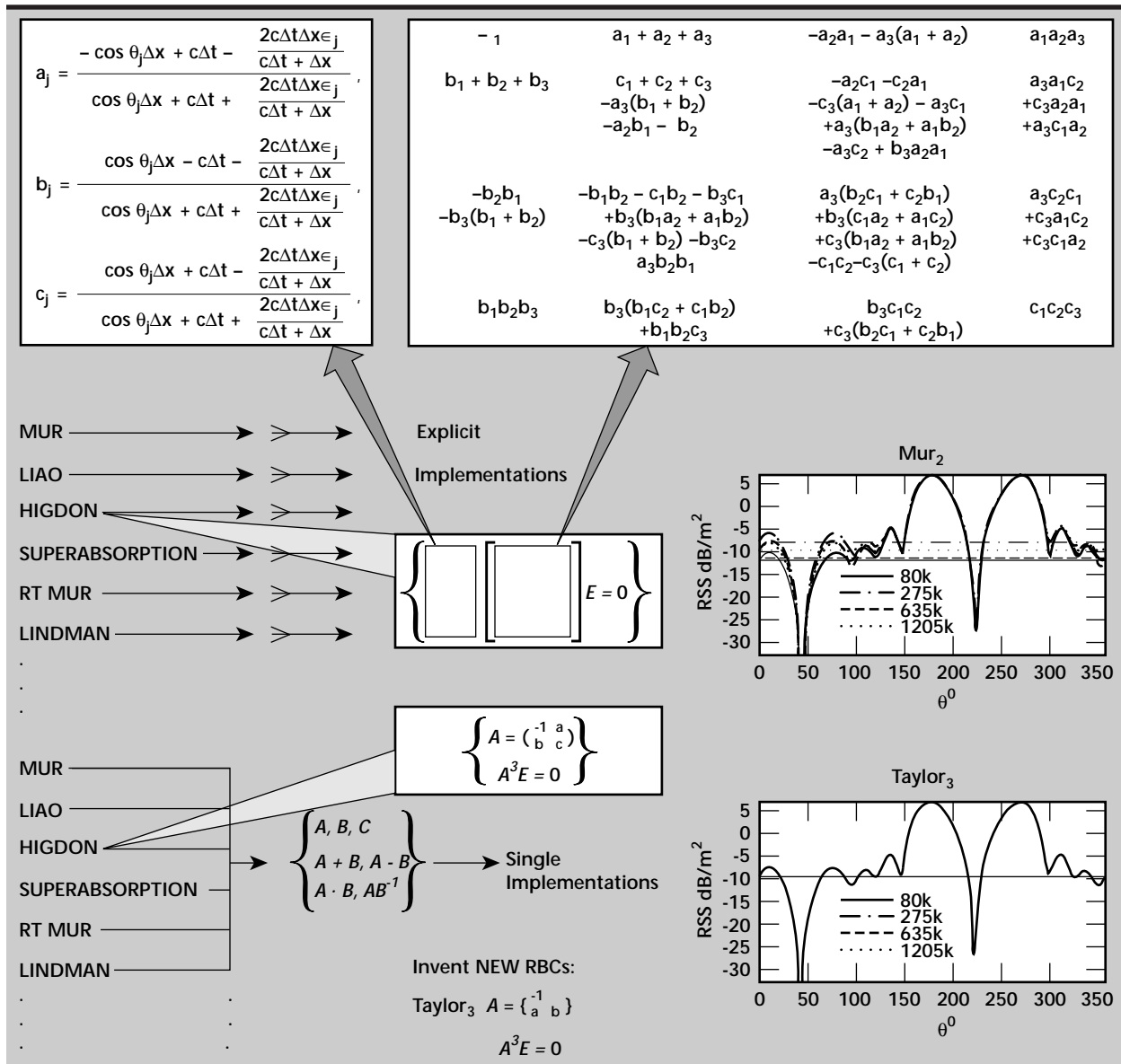
Figure 3. Operator-based approach to RBCs.

and distributed memory multiprocessor computer architectures. We will also continue building physics modules that use these management systems.

We will also research issues required for the simultaneous implementation of the technology enhancements shown in **Table 1**. The research areas are:

1) an efficient hybrid grid realization;
2) generalized advanced material algorithms for hybrid grids;
3) generalization of algorithms at grid interfaces that enforce conservation laws;
4) efficient MPP communication in the presence of these generalizations; and
5) efficient MPP load-balancing for hybrid grids.

In addition to the above research topics, we will develop the following new physics:

1) wakefield calculations for arbitrary beampipe cross sections;
2) advanced RBCs in the presence of Perfect Electrical Conductors (PEC);
3) advanced RBCs in the presence of complex sources;
4) advanced material modeling formulations for nonorthogonal grids; and
5) the simulation of a curved beamline source for complex geometries on hybrid grids.

We anticipate that TIGER will quickly surpass all of LLNL Engineering's linear time-domain EM code capability by the end of FY-97.

The first major use of the TIGER code will be in the modeling and design of accelerator and electron beam transport components for the Advanced Hydrotest Facility (AHF) which will be a major part of the Science Based Stockpile Stewardship (SBSS) Program. In AHF, a linear induction accelerator generates a high-current, relativistic electron beam for the generation of x rays for use in pulsed radiography applications. **Figure 4** is an illustration of the AHF facility with key accelerator components that will be modeled by the TIGER code.

Of prime importance in the design of high-current electron beam accelerator and beam transport systems is the understanding and control of the scattered electromagnetic fields associated with the passage of a relativistic electron beam through the system. These scattered electromagnetic fields are referred to as wakefields. The wakefields can lead to degradation of the beam with the resultant loss of beam quality at the target, or worse, can lead to beam breakup instabilities with total loss of beam transport.

Excellent results have been achieved with TSAR for the modeling of wakefields in 3-D structures (see "Three-Dimensional Electromagnetic Modeling of Wakefields in Accelerator Components," a separate report in this thrust area.) However, the full scale
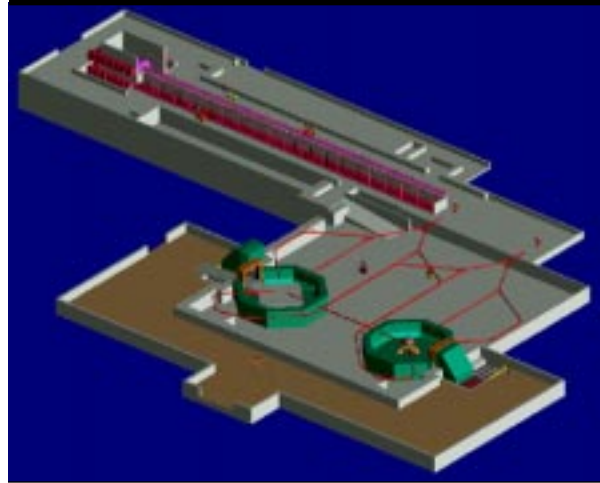
Figure 4. AHF facility, with four critical accelerator components that will be modeled with TIGER.

models will require a more robust code. Full 3-D modeling of complex structures has been limited due to the substantial computer resources required for a fully 3-D model. This is particularly true with TSAR, where its cubical cell-structured mesh significantly limits the development of realistic models. Development of TIGER using conforming unstructured meshes will allow a substantial increase in the geometric fidelity of the structures being modeled and, with its advanced boundary conditions, will allow a more realistic truncation of the computational problem space.

Acknowledgments

References

1. Madsen, N. K. (1995), "Divergence Preserving Discrete Surface Integral Methods for Maxwell's Curl Equations Using Non-Orthogonal Unstructured Grids," *J. Comput. Physics,* Vol. **119**, (1), June, pp. 34-45.

2. Steich, D. J. (1995), "Local Outer Radiating Boundary Conditions for the Finite-Difference Time-Domain Method Applied to Maxwell's Equations," Ph.D. dissertation, Pennsylvania State University.

3. Veldhuizen, T. (1995), "Expression Templates," *C++ Report Magazine*, Vol. **7**, (5), June.

4. Myers, N. (1995), "A New and Useful Template Technique: 'Traits'," *C++ Report Magazine*, Vol. **7**, (5), June.

Anisotropic Magnetic and Electric Material Models in Computational Electromagnetic Codes

Scott D. Nelson
Defense Sciences Engineering Division
Electronics Engineering

Gerald J. Burke and David J. Steich
Engineering Research Division
Electronics Engineering

We have demonstrated that the recursive-convolution solution for anisotropic and dispersive media yields accurate results for reflection from ferrite and plasma slabs up to a frequency limit set by the sampling interval. Several different forms of the up-date equations were considered which can result from different approximations in reducing the continuous equations to discrete forms.

Introduction

Magnetic and electric tensor material characteristics are central to a broad class of complex materials such as plasmas and crystals. Many devices in electromagnetics use external fields to bias a ferrite material or an electro-optic material. This bias then changes the field propagation characteristics, based on the direction of travel through the medium with respect to the direction of the applied field. The incorporation of anisotropic materials into Lawrence Livermore National Laboratory (LLNL)'s computational electromagnetic design codes allows for the expansion into these areas.

In recent years the finite-difference time-domain (FDTD) solution procedure developed by K. S. Yee has been extended to dispersive and anisotropic media to handle materials such as magnetized ferrites and plasmas. The constitutive relations for the dispersive material can be introduced into the solution through either a convolution integral or a differential equation. We have examined the convolution form, which uses a recursive update procedure, for use in design codes under development at LLNL.

The coupling of field components in anisotropic media can result in a substantial increase in complexity and evaluation time for the computational electromagnetic update equations. However, we note that the equations simplify greatly if the impulse response of the susceptibility is zero at time equal to zero, as is true of some materials. Some variations in the basic up-date equations, resulting from the approximations in going from continuous to discrete form, were examined, and results were compared for accuracy and stability.

A 3-D code was written to test these solutions, but was validated for the 1-D problem of reflection and transmission by a magnetized ferrite or plasma slab, since the exact reflection and transmission coefficients are available for this problem.

Progress

Material Formulation for Gyrotropic Media

The case of anisotropic and dispersive magnetic material will be used here as an example. The magnetic field up-date equation for such material is obtained by combining the Maxwell's equation

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \qquad (1)$$

and the equation relating $\mathbf{B}$ and $\mathbf{H}$ in convolution form

$$\mathbf{B}(t) = \mu_0 \left[\mathbf{H}(t) + \int_0^t \overline{\overline{\chi}}_m (t - \tau) \cdot \mathbf{H}(\tau) d\tau \right] \qquad (2)$$

where $\bar{\bar{\chi}}_m(t)$ is the susceptibility tensor representing the impulse response of the material. To eliminate **B** from **Eq. 1** the derivative of **Eq. 2** can be evaluated as

$$\frac{\partial}{\partial t}\mathbf{B}(t) =$$
$$\mu_o\left[\frac{\partial}{\partial t}\mathbf{H}(t) + \bar{\bar{\chi}}(t)\cdot\mathbf{H}(0) + \int_o^t\bar{\bar{\chi}}(\tau)\cdot\frac{\partial}{\partial t}\mathbf{H}(t-\tau)d\tau\right]$$

(3)

or as

$$\frac{\partial}{\partial t}\mathbf{B}(t) =$$
$$\mu_o\left[\frac{\partial}{\partial t}\mathbf{H}(t) + \bar{\bar{\chi}}(0)\cdot\mathbf{H}(t) + \int_o^t\frac{d}{dt}\bar{\bar{\chi}}(t-\tau)\cdot\mathbf{H}(\tau)d\tau\right].$$

(4)

Converting **Eq. 3** to discrete form with a pulse approximation of the integral leads to the magnetic field up-date equation

$$\mathbf{H}^{n+1/2} = \left[\bar{\bar{\mathbf{I}}} + \frac{\Delta t}{2}\bar{\bar{\chi}}_m(0)\right]^{-1}\left[\bar{\bar{\mathbf{I}}} - \frac{\Delta t}{2}\bar{\bar{\chi}}_m(0)\right]\cdot\mathbf{H}^{n-1/2}$$

$$-\Delta t\left[\bar{\bar{\mathbf{I}}} + \frac{\Delta t}{2}\bar{\bar{\chi}}_m(0)\right]^{-1}\cdot\Psi^n$$

$$-\frac{\Delta t}{\mu_o}\left[\bar{\bar{\mathbf{I}}} + \frac{\Delta t}{2}\bar{\bar{\chi}}_m(0)\right]^{-1}\cdot\nabla\times\mathbf{E}^n$$

(5)

where

$$\Psi^n =$$
$$\sum_{i=0}^{n-1}\left[\bar{\bar{\chi}}_m\big[(n-i)\Delta t\big] - \bar{\bar{\chi}}_m\big[(n-i-1)\Delta t\big]\cdot\mathbf{H}^{i+1/2}\right].$$

Starting with **Eq. 4** and representing $\partial\mathbf{H}(t)/\partial t$ with a central difference and **H**(t) with an average from times $(n-1/2)\Delta t$ and $(n+1/2)\Delta t$ leads to an equation similar to **Eq. 5**, but with Ψ^n replaced by $\Delta t\Psi'^n$ where

$$\Psi'^n = \sum_{i=0}^{n-1}\bar{\bar{\chi}}'_m\big[(n-i-\tfrac{1}{2})\Delta t\big]\cdot\mathbf{H}^{i+1/2}.$$

Somewhat different results can be obtained for **Eq. 5** from different interpretations of the pulse approximation of the integral, or in the equation derived from **Eq. 4** by using the value at the forward

time step rather than the average for **H**(n\Delta t). In fact, adding the same small quantity to both square-bracketed terms results in a second-order change in the product multiplying $\mathbf{H}^{n-1/2}$. Hence **Eq. 5** can be reduced to

$$\mathbf{H}^{n+1/2} =$$
$$\left[\bar{\bar{\mathbf{I}}} + \Delta t\bar{\bar{\chi}}_m(0)\right]^{-1}\cdot\left(\mathbf{H}^{n-1/2} - \Delta t\,\Psi^n - \frac{\Delta t}{\mu_o}\nabla\times\mathbf{E}^n\right)$$

(6)

by adding $(\Delta t/2)\bar{\bar{\chi}}(0)$ to both coefficients, or by subtracting the same quantity

$$\mathbf{H}^{n+1/2} =$$
$$\left[\bar{\bar{\mathbf{I}}} - \Delta t\bar{\bar{\chi}}_m(0)\right]\cdot\mathbf{H}^{n-1/2} - \Delta t\,\Psi^n - \frac{\Delta t}{\mu_o}\nabla\times\mathbf{E}^n.$$

(7)

When $\bar{\bar{\chi}}(0) = 0$, which occurs in materials such as Lorentz dielectrics, **Eqs. 5, 6,** and **7** all reduce to the simpler and easier to use form

$$\mathbf{H}^{n+1/2} = \mathbf{H}^{n-1/2} - \Delta t\Psi^n - \frac{\Delta t}{\mu_o}\nabla\times\mathbf{E}^n.$$

(8)

The evaluation of a single vector component of the product of a tensor with the curl of the field can involve 36 field values. However, this can be reduced to 20 values by combining and canceling terms. When the elements of $\bar{\bar{\chi}}(t)$ are represented as sums of exponentials Ψ^n can be evaluated from Ψ^{n-1} through a simple recursion relation, although additional averaging is needed for anisotropic media.

The results of solving **Eq. 5** for a pulse normally incident on a ferrite slab are shown in **Fig. 1**.

The magnitudes of the reflection and transmission coefficients for right-hand circular polarization are compared with the exact results, and the relative errors in the complex quantities are also shown. The relative errors from solving **Eqs. 5, 6,** and **7** are compared in **Fig. 2**.

While the differences in error are fairly small, **Eq. 5** appears to be more stable near the Courant limit, as shown by the comparison with **Eq. 6** in **Fig. 3**.

Equation 6 continued to show greater instability at late time when Δt was reduced to $\Delta\mathbf{x}/c\sqrt{3}$, while **Eq. 5** showed no significant instability at this time increment. The stability of **Eq. 7** was similar to that of **Eq. 6**.
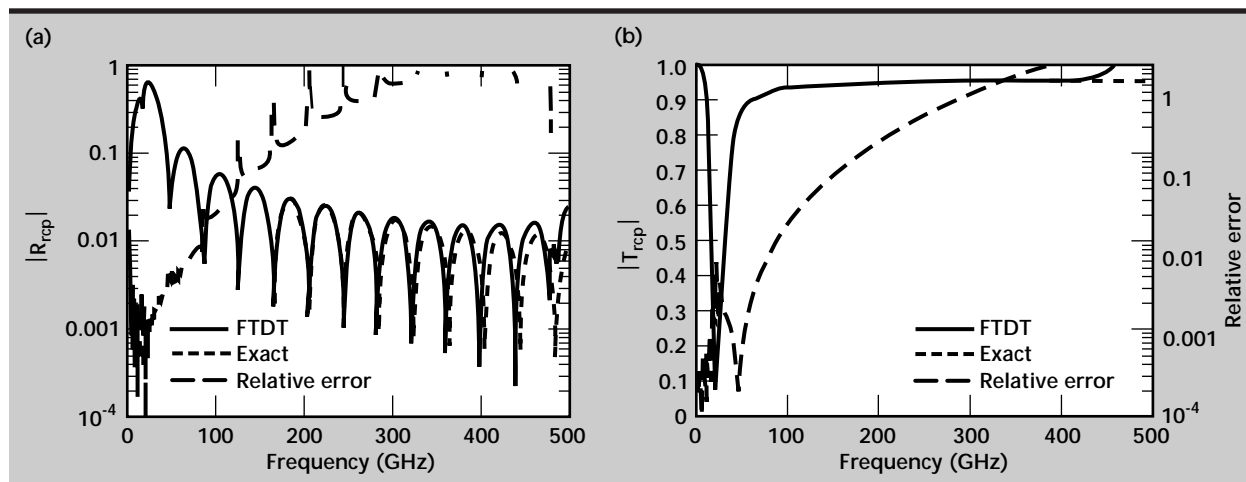
Figure 1. Magnitudes of the reflection (a) and transmission (b) coefficients for normal incidence on a ferrite slab from the FDTD solution, compared with the exact solutions. The relative errors in the complex quantities are also shown.
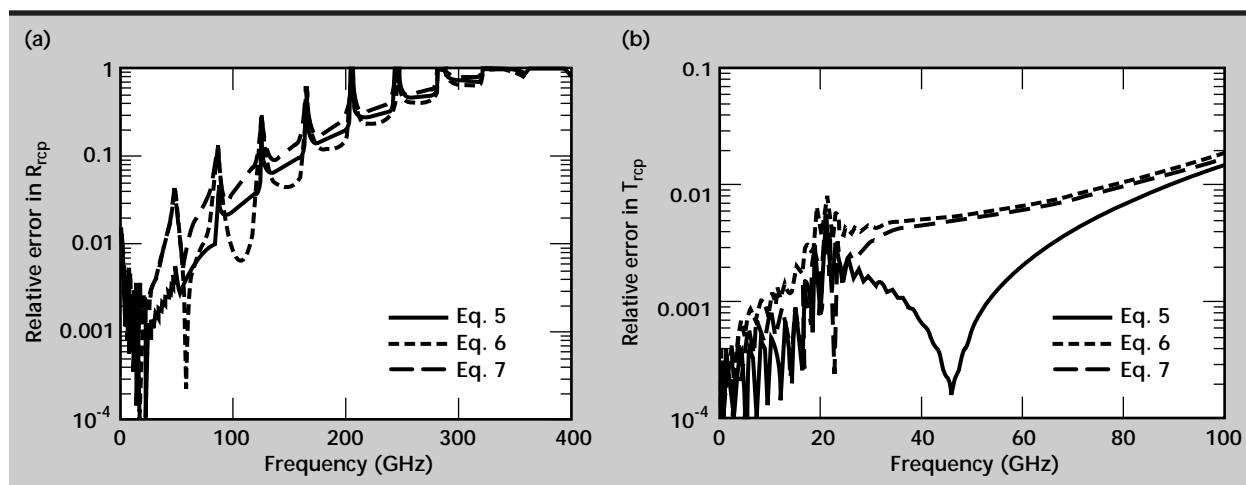
Figure 2. Relative errors in the complex reflection (a) and transmission (b) coefficients from solving Eqs. 5, 6, and 7 for normal incidence on a ferrite slab with $\Delta t = \Delta x / 2c\sqrt{3}$.
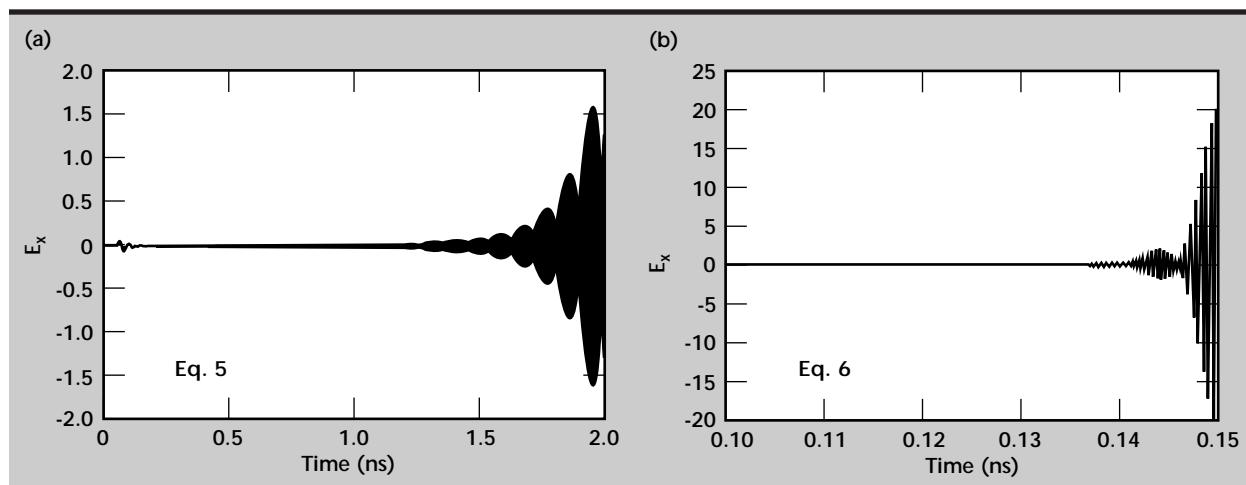
Figure 3. Electric field in front of the ferrite slab, showing instability of Eqs. 5 and 6 solved at the 1-D Courant limit of $\Delta t = \Delta x/c$.

Non-Dispersive Anisotropic Media

For a simple non-dispersive anisotropic medium ε_r and μ_r become constant tensors, so that the constitutive relations are

$$\mathbf{D} = \varepsilon_o \overline{\overline{\varepsilon}}_r \cdot \mathbf{E} = \varepsilon_o \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{9}$$

$$\mathbf{B} = \mu_o \overline{\overline{\mu}}_r \cdot \mathbf{H} = \mu_o \begin{pmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \\ H_z \end{pmatrix} \tag{10}$$

Substituting **Eqs. 9** and **10** into the basic up-date equations for **B** and **D**, the up-dated fields are obtained in terms of the inverses of the $\overline{\overline{\varepsilon}}_r$ and $\overline{\overline{\mu}}_r$ tensors as

$$\mathbf{H}^{n+1/2}(i,j,k) =$$

$$\mathbf{H}^{n-1/2}(i,j,k) - \left(\Delta t/\mu_o\right) \overline{\overline{\mu}}_r^{-1} \cdot \nabla \times \mathbf{E}^n(i,j,k) \tag{11}$$

and

$$\mathbf{E}^{n+1}(i,j,k) =$$

$$\mathbf{E}^n(i,j,k) + \left(\Delta t/\varepsilon_o\right) \overline{\overline{\varepsilon}}_r^{-1} \cdot \nabla \times \mathbf{H}^{n+1/2}(i,j,k). \tag{12}$$

The components of the vector equations **Eqs. 11** and **12** are still evaluated at the off-set locations of the Yee cell. However, the dot product requires all vector components of the curl at each location. The components that are not directly available from the central difference formulas must be obtained by averaging. The equation for $H_x\left(i, j + \frac{1}{2}, k + \frac{1}{2}\right)$ is

$$H_x^{n+1/2} =$$

$$H_x^{n-1/2} - \frac{\Delta t}{\mu_o} \begin{bmatrix} \mu_{11}^{-1}\left(\nabla \times \mathbf{E}^n\right)_x \\ +\mu_{12}^{-1} A_{xy}^H\left[\left(\nabla \times \mathbf{E}^n\right)_y\right] \\ +\mu_{13}^{-1} A_{xz}^H\left[\left(\nabla \times \mathbf{E}^n\right)_z\right] \end{bmatrix}$$

while

$$H_y\left(i + \frac{1}{2}, j, k + \frac{1}{2}\right) \text{ is}$$

$$H_y^{n+1/2} =$$

$$H_y^{n-1/2} - \frac{\Delta t}{\mu_o} \begin{bmatrix} \mu_{21}^{-1} A_{yx}^H\left[\left(\nabla \times \mathbf{E}^n\right)_x\right] \\ +\mu_{22}^{-1}\left(\nabla \times \mathbf{E}^n\right)_y \\ +\mu_{23}^{-1} A_{yz}^H\left[\left(\nabla \times \mathbf{E}^n\right)_z\right] \end{bmatrix}$$

and

$$H_z\left(i + \frac{1}{2}, j + \frac{1}{2}, k\right) \text{ is}$$

$$H_z^{n+1/2} =$$

$$H_z^{n-1/2} - \frac{\Delta t}{\mu_o} \begin{bmatrix} \mu_{31}^{-1} A_{zx}^H\left[\left(\nabla \times \mathbf{E}^n\right)_x\right] \\ +\mu_{32}^{-1} A_{zy}^H\left[\left(\nabla \times \mathbf{E}^n\right)_y\right] \\ +\mu_{33}^{-1}\left(\nabla \times \mathbf{E}^n\right)_z \end{bmatrix}$$

where A_{xy}^H yields a value of the argument at the H_x location by averaging from four surrounding H_y positions. The averaging for all components of the curl at a single location involves 36 field values. However, eight of these cancel and another eight can be combined in the dot product of the curl with a vector, leaving 20 unique field values. The complete result for a product such as $\overline{\overline{\mu}}_r^{-1} \cdot \nabla \times \mathbf{E}$, with each component of the resulting vector located at the correct Yee-cell location, can be obtained from a total of 36 unique values of **E**, since each value is used more than once. This compares with 108 if no attempt is made to cancel or reuse values.

Results

The solution for gyrotropic media was validated for a plane wave normally incident on a ferrite slab with the biasing magnetic field in the direction of propagation, along the z axis. This problem was chosen because the reflection and transmission coefficients for the slab are available in simple analytic form. In this case, a 3-D code was written, using the equations from the preceding section. Since the plane wave propagating through the ferrite in the direction of the biasing field splits into right-hand and left-hand circularly polarized waves with different propagation constants, the problem space was terminated in even-symmetry boundary conditions in both x and y boundary planes. First-order Mur boundary

conditions were used at the z boundaries, but for the results shown here the z extent of the problem space was made large enough to gate out the reflections, thus eliminating the boundary conditions as a source of error.

For the ferrite, biased with magnetic field along the z axis, the relative permeability tensor at a frequency ω is

$$\overline{\overline{\mu}}_r(\omega) = \begin{pmatrix} \mu(\omega) & jk(\omega) & 0 \\ -jk(\omega) & \mu(\omega) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1+\chi_{11}(\omega) & \chi_{12}(\omega) & 0 \\ \chi_{21}(\omega) & 1+\chi_{22}(\omega) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where

$$\chi_{11}(\omega) = \chi_{22}(\omega) = \frac{(\omega_o + j\omega\alpha)\omega_m}{(\omega_o + j\omega\alpha)^2 - \omega^2}$$

$$\chi_{12}(\omega) = -\chi_{21}(\omega) = \frac{j\omega\omega_m}{(\omega_o + j\omega\alpha)^2 - \omega^2} .$$

The parameters of the ferrite modeled here were

$$\omega_o = (2\pi) \cdot 20 \times 10^9 \, rad \, / \, s$$
$$\omega_m = (2\pi) \cdot 10 \times 10^9 \, rad \, / \, s$$
$$\alpha = 0.1$$

Results for this ferrite with $\Delta x = 75(10^{-6}) m$ and $\Delta t = \Delta x / (2c\sqrt{3})$ with 6000 time steps are shown in **Fig. 1**. The source was a Gaussian-pulse plane wave with full-width-half-max equal to 0.001 m. Since the solution is uniform in the x and y directions the problem was solved with 3 cells in x and y and 4000 cells in z, to eliminate the radiating boundaries, and the ferrite filled 50 cells for a thickness of 0.00375 m.

Conclusions

The recursive-convolution solution for anisotropic and dispersive media was seen to yield accurate results for reflection from ferrite and plasma slabs up to a frequency limit set by the sampling interval. Depending on the application, the results shown might be considered usable up to about 300 GHz, which corresponds to about 13 cells per wavelength. Results at still higher frequencies might be usable when a time delay or frequency shift due to dispersion can be tolerated. The deconvolved results became unstable not far beyond the upper limits of the plots.

The main difficulties in implementing the solution for anisotropic and dispersive media are evaluating the product of a tensor with the curl of a field, and the tensor multiplying the previous values of the field that is being up-dated. The number of values needed in evaluating the product with the curl can be reduced by combining and canceling terms, but that can restrict generality of a code in the treatment of mesh types. The product of a tensor with the field being up-dated presents bookkeeping problems that can be circumvented, at the expense of added storage, by computing the new values in a separate array.

Several different forms of the up-date equations were considered which can result from different approximations in reducing the continuous equations to discrete form.

Nodular Defect-Induced Field Enhancements in High-Power Multilayer Optical Coatings

Nicole E. Molau and Clifford C. Shang
Engineering Research Division
Electronics Engineering

Hal R. Brand
Computer Applications Sciences and Engineering Division
Computations Organization

Mark R. Kozlowski
Inertial Confinement Fusion
Laser Programs

In FY-96, we completed the 2-1/2-D electromagnetic modeling of rotationally-symmetric nodular defect geometries in high-power optical high-reflectance coatings.

Introduction

Advances in the design and production of high damage threshold optical coatings for use in mirrors and polarizers have been driven by the design requirements of high-power laser systems such as the proposed 1.8-MJ National Ignition Facility (NIF) and the prototype 12-kJ Beamlet laser system at Lawrence Livermore National Laboratory (LLNL). The present design of the NIF will include 192 polarizers and more than 1100 mirrors. Currently, the material system of choice for high-power multilayer optical coatings with high damage threshold applications near 1.06 μm is e-beam deposited HfO_2/SiO_2 coatings. However, the optical performance and laser damage thresholds of these coatings are limited by micron-scale defects and insufficient control over layer thickness.

In multilayer dielectric stack mirrors, the predominant surface defects are micron-scale domes associated with the classic nodule defect. These defects are initiated at seed particles that are either present on the substrate or deposited during the film deposition process. The seed particles can arise from several sources: contamination within the vacuum system, substrate contamination, or particles ejected from the coating source material itself. Nodular defects are produced during film deposition as a result of shadowing effects at seed particles during the growth process. These defects give rise to local electric field enhancements which lead to "hot spots" in the coatings since the heat generation is proportional to the square of the electric field, and these rapidly expanding regions create tensile stresses which may result in mechanical failure of the coatings.

In FY-96, we completed the 2-1/2-D electromagnetic (EM) modeling of rotationally-symmetric nodular defect geometries in high-power high-reflectance (HR) optical coatings. Twenty-three different defect geometries were studied to investigate the influence of nodule size, depth, and material composition on electric field enhancements within the structure.[1,2] The heat generation rates for these defect geometries were computed based on the EM simulations and coupled into the thermal-mechanical model for each geometry.[2] This report will summarize the 2-1/2-D EM modeling results; the thermal-mechanical modeling results are reported elsewhere.[3]

Progress

In FY-96 we completed the 2-1/2-D EM modeling (assuming normal-incidence illumination) of 23 different rotationally-symmetric nodular defect geometries similar to that shown schematically in **Fig. 1**. The 2-1/2-D finite-difference time-domain (FDTD) EM code, AMOS[4,5] was used to calculate the steady-state electric field profiles throughout each rotationally-symmetric defect geometry, by solving

the Maxwell equations and applying the appropriate boundary conditions[6] and material models. AMOS generates a 2-D solution in the r-z plane and assumes an explicit φ-variation to achieve the full 3-D solution for the rotationally-symmetric object.

These steady-state electric field profiles were used to calculate the heat generation distribution (σE^2 in W/m^3 with σ, the electrical conductivity in S/m and E, the steady-state electric field in V/m) throughout each geometry. This heat generation distribution was then coupled into the thermal-mechanical model as the initial conditions, resulting in a full electro-mechanical model of nodule failure.

The steady-state electric field profile everywhere within a 2-D r-z cross-section (for φ = 0°) of a typical nodular defect geometry is shown in **Fig. 2**. In this case, the nodular defect geometry consists of a spherical HfO_2 seed of diameter, d = 0.73 μm; depth, T = 1.97 μm; and nodule dome diameter, D = 3.392 μm. The source was assumed to be a normally-incident, sinusoidally-varying, linearly-polarized, plane wave (λ_o = 1.06 μm) with an average input power density of 1 GW/cm^2. As is illustrated in **Fig. 2**, for rotationally-symmetric nodular defect geometries under normal-incidence illumination, the peak electric fields and highest electric field enhancements are typically located along the axis of symmetry.

The electric field standing wave (EFSW) pattern along the axis of symmetry (r = 0 μm) of the nodule and the EFSW pattern along a line parallel to the axis of symmetry but at a radius of 5 μm are also shown in **Fig. 2**. The EFSW pattern along the axis of symmetry shows several localized electric field enhancements within the HR coating, one in the SiO_2 overcoat and one in the HfO_2 seed, while the EFSW pattern for the line at r = 5 μm looks very much like that of the perfect HR stack.

This is to be expected since, for radii larger than the defect dome radius, D/2 = 1.696 μm, the nodule defect geometry rapidly returns to the perfect stack geometry. By comparing the EFSW pattern along the axis of symmetry to that along the line at r = 5 μm (that is, to the EFSW pattern of a perfect HR stack), we see that the peak electric field enhancement for this particular geometry is approximately a factor of 2.2. This peak electric field enhancement is determined by comparing the peak electric field in the lossier material (that is, in one of the HfO_2 quarter-wave layers or in the HfO_2 seed) in the nodule defect geometry to the peak electric field in the lossier material (that is, in the first HfO_2 quarter-wave layer) in the perfect HR stack. Note that a peak electric field enhancement of 2.2 corresponds to an enhancement in the heat generation (σE^2) of a factor of 4.8 since the heat generation is proportional to the square of the electric field.

Figure 3 shows the heat generation distribution (in $W/\mu m^3$) that corresponds to the electric field profile shown in **Fig. 2**, assuming electrical conductivities of 1.1956 S/m and 0.3366 S/m for HfO_2 and SiO_2, respectively. Since the heat generation is calculated as σE^2, it is evident that the peaks in the electric field profile that lie in the lossier material (the material with the higher electrical conductivity, σ) will be emphasized in the heat generation profile, while the electric field peaks in the less lossy material will be de-emphasized in the heat generation profile. Typically, the electric field enhancements in the lossier material (that is, in the HfO_2 layers) are of the most importance since they will lead to "hot spots" in the heat generation (σE^2) distribution and will therefore control the damage threshold of the HR coating.

Several conclusions can be drawn from the 2-1/2-D EM modeling results for the matrix of 23 different nodular defect geometries. It was found that the dependence of the electric field distribution, the electric field enhancement, and the heat generation on the nodular defect geometry (for example, seed diameter, seed depth, and seed material composition) is, in general, quite complex. Variations in the nodular defect geometry not only

Figure 1. 2-D cross-sectional schematic of a typical rotationally-symmetric nodular defect geometry containing a seed of diameter, d, embedded some depth, T, into the HR stack, with parabolic sides defined by the relation, $D = (8*T*d)^{1/2}$.

affect the peak value of the electric field in the geometry, but also the electric-field profile itself. As the geometry parameters are varied, the locations of the "hot spots" migrate within the geometry.

It was determined that the highest electric field enhancements occur along the axis of symmetry, with the highest electric field enhancement of 6.4 reported for the largest, most shallow HfO_2 seed (d = 1.5 μm and T = 1 μm), and the lowest electric field enhancement of 1.3 reported for the smallest, most deeply embedded HfO_2 seed (d = 0.25 μm and T = 2.94 μm), corresponding to enhancements in the heat generation by factors of 41 and 1.7, respectively. For small seeds (d $\leq$ 1 μm), the important factor influencing electric field enhancement is the seed depth, with an increasing seed depth leading to a decreasing electric field enhancement.

On the other hand, for deeply embedded seeds (T $\geq$ 2.94 μm), the important factor influencing the electric field enhancement is the seed size, with an increasing seed diameter leading to an increasing electric field enhancement. Trends for some of the other nodular defect geometries are not as clear, indicating that perhaps other complex effects such as resonance or other higher-order EM interactions may be the dominant influences on electric-field enhancements for those cases. These results predict that, as a general rule-of-thumb, geometries containing large, shallow seeds lead to the largest electric-field enhancements (or, equivalently, to the highest heat generation) while geometries containing small, deep seeds lead to the smallest electric field enhancements (or, equivalently, to the lowest heat generation).
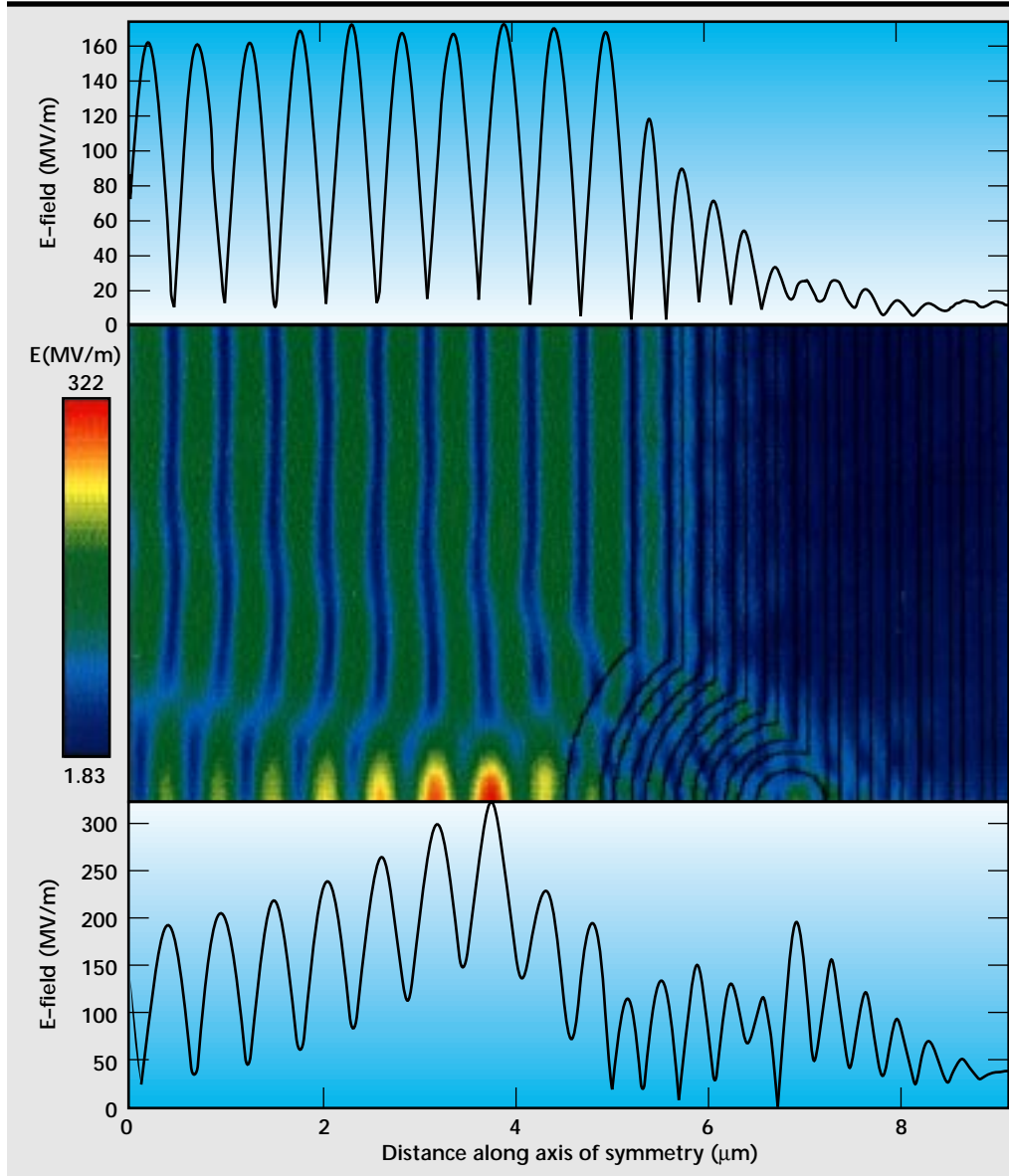
Figure 2. The steady-state electric field profile everywhere within a 2-D r-z cross-section (for $\phi = 0°$) of a typical nodular defect geometry, consisting of a spherical HfO_2 seed of diameter, d = 0.73 μm and depth, T = 1.97 μm. Also shown is the EFSW pattern along the axis of symmetry (r = 0 μm) of the nodule as well as the EFSW pattern along a line parallel to the axis of symmetry but at a radius of 5 μm.

Figure 3. The heat generation distribution (σE^2, in $W/\mu m^3$) that corresponds to the electric field profile shown in Fig. 2, assuming electrical conductivities of 1.1956 S/m and 0.3366 S/m for HfO_2 and SiO_2, respectively.
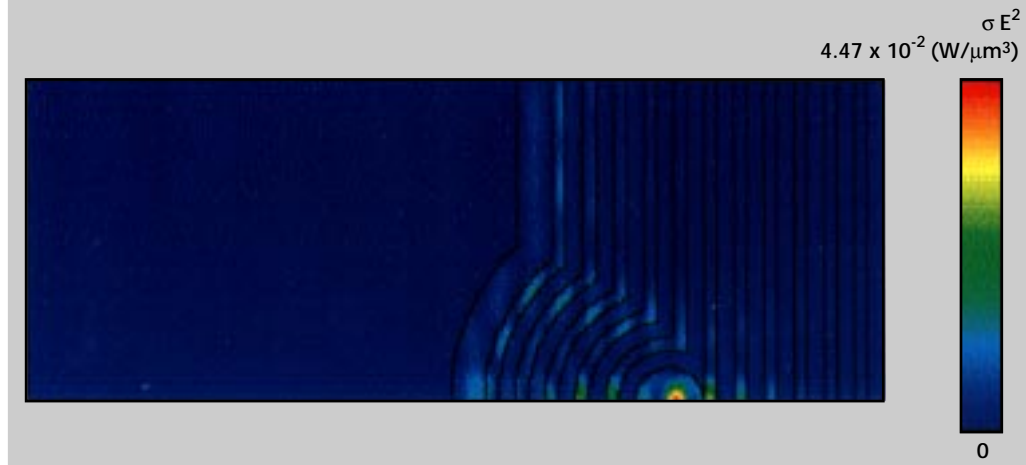
These conclusions are in close agreement with earlier modeling work reported by DeFord,[7,8] as well as with experimental work reported by Kozlowski.[8] Kozlowski found that nodular defects with large dome heights (particularly those above 0.6 μm) were the most susceptible to laser damage, a result supported by our conclusion that large, shallow defects lead to the highest electric field enhancement and, therefore, the highest heat generation.

Future Work

In the future, the EM modeling of nodular defects in high-power optical coatings could be extended to the fully 3-D regime, allowing for both generalized sources (arbitrary angle-of-incidence and arbitrary polarization) and generalized defect geometries (asymmetric geometries). Once available, the 3-D frequency-domain EM code, EIGER,[9] would be well-suited to perform this 3-D EM analysis.

Coupled engineering codes are being applied to the analysis and design of high-average-power microwave components for Accelerator Production of Tritium (APT).

Acknowledgments

References

1. Molau, N. E., H. R. Brand, M. R. Kozlowski, and C. C. Shang (1996), *2-1/2-D Electromagnetic Modeling of Nodular Defects in High-Power Multilayer Optical Coatings*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-ID-124808).

2. Molau, N. E., C. C. Shang, F. Y. Genin, M. R. Kozlowski, H. R. Brand, and R. H. Sawicki (1995), *Electro-mechanical Modeling of Nodular Defects in High-Power Multilayer Optical Coatings*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-95), **1**, p. 7.

3. Shapiro, A. B., D. R. Faux, T. A. Reitter, and R. A. Riddle (1996), "Characterization of Laser-Induced Mechanical Failure Damage of Optical Components," *Engineering Research, Development and Technology FY96 Thrust Area Report*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-96).

4. DeFord, J. F., G. D. Craig, and R. McLeod (1989), "The AMOS (Azimuthal Mode Simulator) Code," *Proceedings of the 1989 Particle Accelerator Conference, Chicago, Ill.*, Mar. 20–23, p. 1181.

5. Shang, C. C., and J. F. DeFord (1990), "Modified-Yee Field Solutions in the AMOS Wakefield Code," *Proc. 1990 Linear Accelerator Conf., Albuquerque, New M.,* September.

6. Mur, G. (1981), *IEEE Trans. Electromagn. Compat.,* **EMC-23**, p. 1073.

7. DeFord, J. F., and M. R. Kozlowski (1993), "Modeling of Electric-Field Enhancement at Nodular Defects in Dielectric Mirror Coatings," *Laser-Induced Damage in Optical Materials: 1992,* (SPIE) Vol. **1848**, pp. 455–473.

8. Kozlowski, M. R., J. F. DeFord, and M. C. Staggs (1993), "Laser-Damage Susceptibility of Nodular Defects in Dielectric Mirror Coatings: AFM Measurements and Electric-Field Modeling," *Proceedings of Second International Conference on Laser Ablation, Knoxville, Tenn.,* April 19–22.

9. Sharpe, R. M., J. B. Grant, W. A. Johnson, R. E. Jorgenson, D. R. Wilton, D. R. Jackson, and J. W. Rockway (1995), *EIGER: A New Generation of Computational Electromagnetics Tools*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-95), **1**, p. 33.

EIGER: Electromagnetic Interactions Generalized

Robert M. Sharpe and J. Brian Grant
Engineering Research Division
Electronics Engineering

William A. Johnson and Roy E. Jorgenson
Sandia National Laboratories
Albuquerque, New Mexico

Donald R. Wilton and Nathan J. Champagne
University of Houston
Houston, Texas

John W. Rockway
Naval Command and Control Ocean Surveillance Center
San Diego, Calif.

New software engineering methods, specifically, object-oriented design, are being used to abstract the key components of spectral analysis methods so that the tools can be easily modified and extended to treat new classes of problems. This software design method also yields a code suite that is more easily maintained than standard designs. This report gives a brief update and highlights planned future developments.

Introduction

The EIGER (Electromagnetic Interactions Generalized) development project is a multi-institutional collaboration that is bringing a variety of spectral domain analysis methods into a single integrated software tool set. Staff from the Sandia National Laboratories, the Navy (NCCOSC/NRaD), Lawrence Livermore National Laboratory, and the University of Houston are collaborating to develop this package which has an unparalleled ability to take general purpose analysis methods and optimize the approach used for specific applications.

The object-oriented design method enables the abstraction of the basic underlying components of computational electromagnetics to yield a suite of tools with unprecedented flexibility. The key components of the numerical analysis and their roles are:

1) elements, which are used to describe the geometry,
2) basis (expansion) functions, which interpolate the unknowns (for example, fields) locally, and
3) operators, which are the underlying formulation of the physics used to propagate the energy or enforce fundamental principles.

By contrast, in standard design procedures entire codes are developed around a single element with a specialized basis function for a specific operator. Although such tools can be used effectively to model large classes of problems, it is often very difficult, if not intractable, to extend the tools beyond their initial design. Overcoming this limitation is one of the most compelling goals of our project. Indeed, the applicability of EIGER is significantly broadened as we cast a variety of analytic treatments (Green's functions) into a form compatible with the numerical procedures in EIGER.

The following sections briefly describe this year's efforts in several key areas.

Progress

Elements

Elements are the basic building blocks that are used to describe a given geometry for numerical computation. Elements are typically the output of a commercial mesh generation package that has discretized a solid CAD model into pieces that are

amenable to numerical computation. A pre-processor then reads these elements and assigns such features as material characteristics and excitation parameters to generate the actual input for the physics code.

Our development team has been using two commercial CAD and mesh generation packages to generate the element descriptions for a given problem. Specifically, MacNeal Schwendler's CAD tool PATRAN and SDRC's IDEAS are being used to feed the electromagnetics analysis tools. Even though these packages were primarily designed to generate computational meshes and information for mechanical engineering analysis, they have proven flexible enough to label and group information associated with a mesh such that EIGER can easily take advantage. IDEAS advanced grouping functionality has proven especially useful for generating compute meshes for complicated geometries.

The EIGER suite can treat both 2- and 3-D problems in a single integrated tool set. Initially, emphasis was placed on the requirements for modeling surface physics. Thus, for 2-D geometries, combinations of linear segments or "bar" elements are currently used to describe the surface problem. Likewise, for 3-D geometries, both linear triangles and rectangles can be combined to render accurate surface models. In addition to the basic elements for modeling 3-D surfaces, wire segments are also incorporated into EIGER. Currently, at the University of Houston a scheme is being implemented to use higher-order surface elements to more accurately resolve the geometry of curved objects.

The code structures have been designed to easily incorporate volumetric elements directly into the analysis framework. In fact, the 3-D surface elements are all that are required for 2-D volumetric analysis. Currently, general 3-D volumetric elements such as tetrahera, hexahedra, and pentahedra are being added to the code suite to address general finite element problems.

Basis Functions

Basis (or expansion) functions are used to locally interpolate an unknown quantity in many numerical methods. These unknowns may physically represent surface currents or fields for dynamic operators, or potentials for statics. Both constant and linear basis functions are currently available for each of the elements described above. The characteristic integrals (for integral equation solutions) that arise in many problems require that the potentials, and their gradients due to source distributions that are represented by these basis functions, must be efficiently computed.

Extensive efforts have enabled most of these calculations to be cast into the same basic form, thus simplifying their integration into a general purpose code. In conjunction with the efforts to add higher-order elements, extensions are currently under investigation to use higher-order (smoother) basis functions to enable a more efficient numerical solution for electrically large problems. In addition, we have developed a new formulation for singular basis functions that will incorporate known edge conditions and other known local variations directly into the numerical procedure.

Operators

An electromagnetic operator is a mathematical construct that relates the field at a point to the sources that produced the field. Since previous efforts focused initially on incorporating surface physics, emphasis was placed on integral operators, which are among the most efficient classes of operators for enforcing this type of physics. These operators explicitly propagate a field between locations. The numerical solution of these operators is known as the method of moments. Specific integral operators are often given specific names that are related to boundary conditions that are applied and the manner in which they are enforced. A variety of integral operators have been incorporated into EIGER to treat the following boundary conditions: perfect electric conductors, conductors coated with complex impedance materials, and homogeneous dielectrics (penetrable materials). Specifically, these operators are:

1) the Electric Field Integral Equation (EFIE);
2) the Magnetic Field Integral Equation (MFIE);
3) the Combined Field Integral Equation (CFIE); and
4) general Homogeneous Lossy Dielectrics (PMCHW).

The basic formulation for the integral operators has been generalized and simplified so that all of the above operators are generated by simply taking linear combinations of basic coupling or interaction operators. This powerful abstraction means that a variety of other boundary conditions can easily be added to the code simply by choosing the appropriate set of coefficients (that is, equivalent aperture formulations, and alternative dielectric treatments).

This year, efforts were begun to add the types of finite-element operators that are required to solve partial differential equations. Rigorous treatments for 2-D finite element solutions of both electrostatic and magnetostatic problems were incorporated for the analysis of micro-electro-mechanical devices (MEMs). Specific numerical treatments were

designed to allow the termination of the finite element volumetric mesh with the appropriate surface mesh using an integral operator. Currently, efforts are underway to extend these same capabilities to the fully 3-D case.

Green's Functions

A Green's function can be thought of as a characteristic solution of a problem for an ideal infinitesimal source. Ideally, this solution directly satisfies as many of the boundary conditions as possible in an analytic fashion, thus reducing the numerical requirement for a given problem. One of the most promising features of the EIGER development is the ability to directly incorporate a variety of Green's functions into the solution. In addition to the standard homogeneous free-space type Green's functions, this year a completely unique capability was added to EIGER by using a multi-layered media Green's function that can treat arbitrarily-shaped objects that are penetrable. These extensions allow applications to be addressed in a variety of areas such as multi-layered electronic circuit boards, high-powered optical mirrors, and geophysical problems.

In **Fig. 1** the scattering from a dielectric cube (ε_r = 2.9) immersed in a half space with a dielectric constant of 3.0 is shown. The results are compared with a Born approximation and the agreement is seen to be quite good.

Solvers and Post-Processing

An end user of EIGER has a choice of basic numerical solutions of linear systems by either a Gaussian elimination method or a bi-conjugate gradient method. The later method is extremely useful for large systems of equations. Solvers that

are tuned for the sparse systems that arise from a finite element solution are also planned for incorporation into the code.

Post-processing capabilities currently include the calculation of near or far fields from the basic current solutions. **Figure 2** shows the geometry of a commercial broadband horn made by AEL which is designed to operate between the frequencies of 1 and 10 GHz. Multiple compute meshes were generated for the geometry, ranging from 2000 to about 8000 unknowns.

Figure 3 shows the 3-D radiation pattern for the antenna at a frequency of 5 GHz. This pattern consists of roughly 16,000 field points.

Figure 4 shows the radiation pattern at a frequency of 10 GHz. The maximum directivity is used to calculate a directive gain for the antenna. **Figure 5** depicts the calculated gain vs frequency as compared to measured data for the structure.

Future Work

Each of the areas outlined above will continue to evolve and extend to reach broader classes of problems with a greater diversity of tools. Specific examples of future development are briefly noted here.

As mentioned previously, we are extending the element modules of the code to include higher order elements. These elements can be of arbitrary order, and will more accurately resolve the local

Figure 1. The scattered far field in the ϕ = 0° plane from a small dielectric cube (ε_b = 3.0ε_0). The plane waves are incident from θ^i = 60° and ϕ^i = 0°.
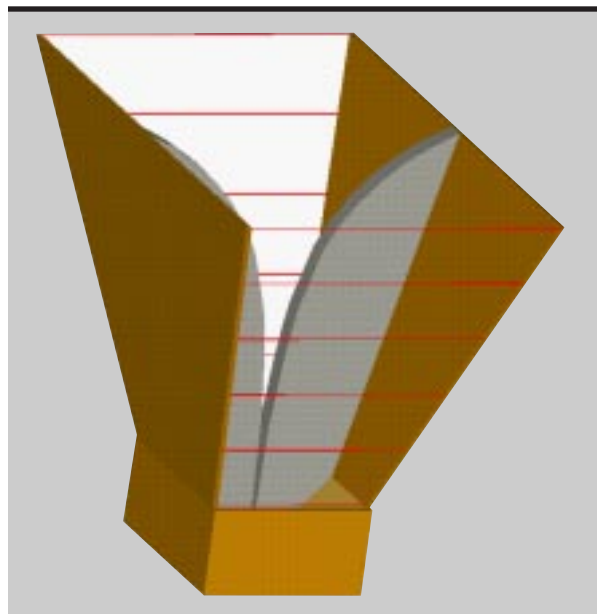
Figure 2. CAD representation of a commercial flared notch horn antenna made by AEL.
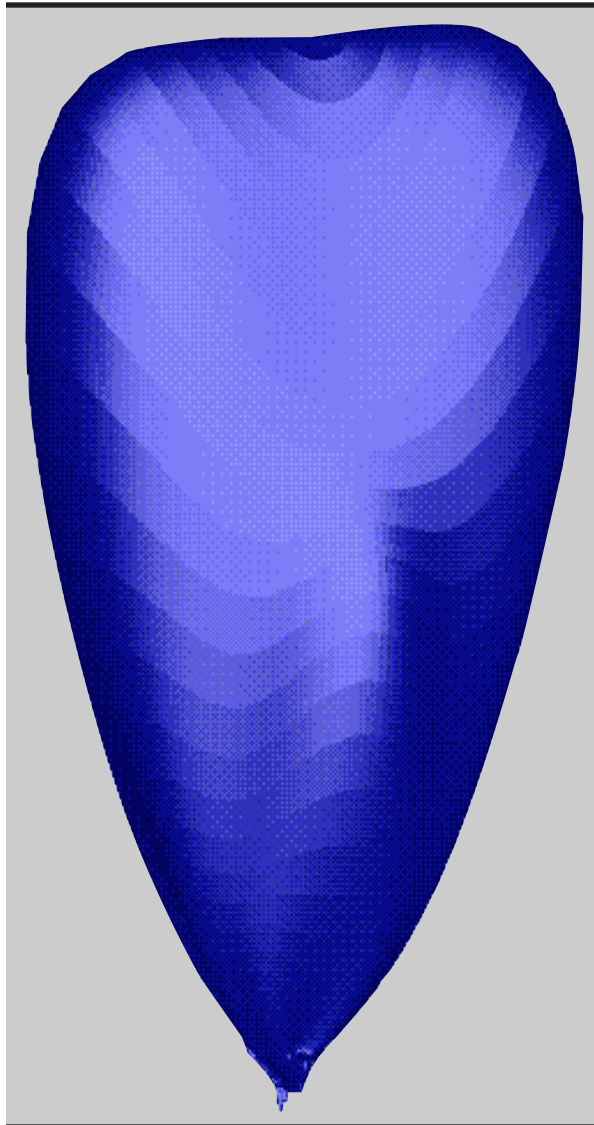
Figure 3. *Three-dimensional radiation pattern for the AEL horn at 5 GHz.*
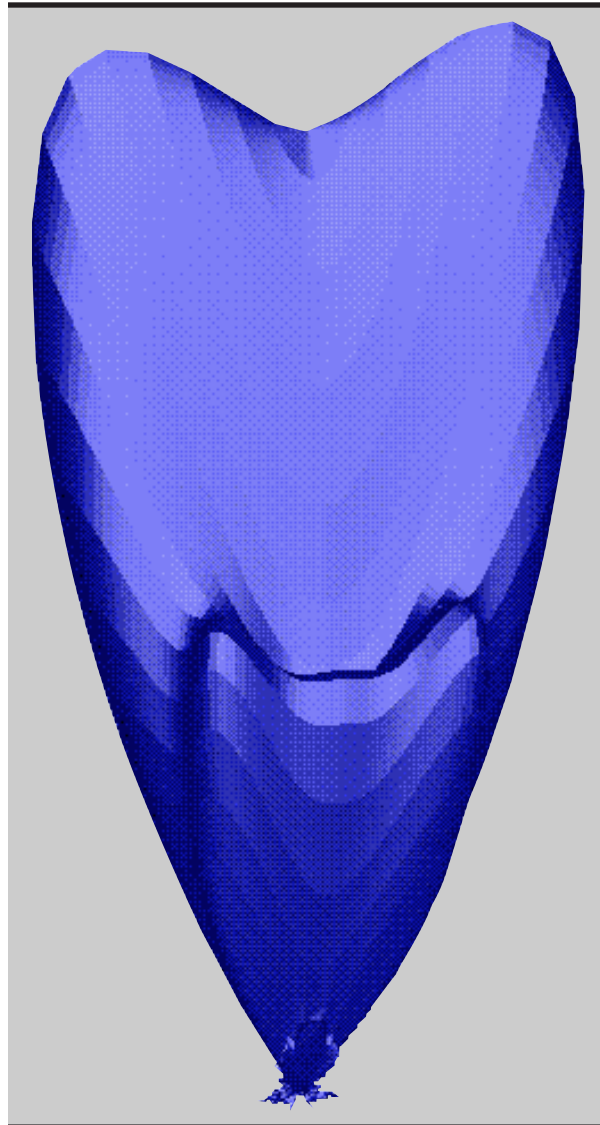
Figure 4. *Three-dimensional radiation pattern for the AEL horn at 10 GHz.*

Figure 5. Antenna gain (directivity at bore sight) vs frequency. Both measurements and experiments are shown.
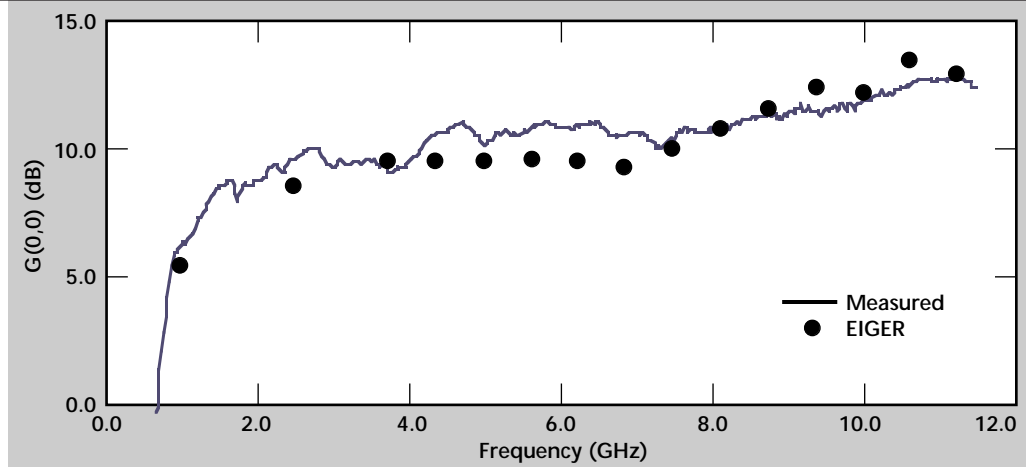
variations of the geometry. Also, volume elements will be incorporated for treating inhomogeneous material regions.

The combined versatility of elements and basis functions suggests multiple combinations and approaches to solve models, possibly overwhelming the modeler with choices. However, this versatility also provides the flexibility to tackle portions of the model with the most appropriate combinations. During this past year, key steps were taken to address the increased complexity of effectively using a code suite such as this. Efforts will continue to develop and extend the pre-processor interface to the physics tools for more efficiency. This activity is being performed in conjunction with the TIGER effort, also described in this publication.

Additional integral operators will continue to be added to the EIGER suite. In addition, the differential operators, which are the underlying mathematics, are being extended from 2-D to 3-D.

Extensions to the Green's functions in the code are also underway. General symmetry treatments have been added to the analysis code and are being incorporated into the post-processor. In addition, an effort to incorporate general periodic analysis, that was previously developed for treatment of cavities and array analysis is also underway.

Another key area for the coming year is the efficient migration of this tool set to advanced computer architectures. Both symmetric multiprocessor (SMP) and massively-parallel processor (MPP) systems are targeted, and work is being done on both the physics kernel and the associated numerical linear algebra routines.

Acknowledgment

Laser Drilling Modeling

David J. Mayhall and Jick H. Yee
Engineering Research Division
Electronics Engineering

We have modified an existing, one-dimensional, laser drilling computer code to run on an IBM-compatible PC, investigated some of the code's behavior, estimated the pressure pulse transmitted to two different solid materials due to laser ablation, and modestly enhanced the code's user interface.

Introduction

Laser drilling and laser shock generation are being explored to an increasingly greater extent for various manufacturing processes, such as circuit-board machining, metal surface hardening, and parts forming in powder metallurgy. In these endeavors, potential increases in efficiency and reductions in expense result from understanding the complex physical and chemical processes between the laser light and the materials upon which it impinges and ablates. Sufficiently accurate computational modeling of the laser ablation of materials should yield improved understanding of the processes involved.

As a step in the computer modeling of laser ablation for manufacturing, we have embarked on a laser drilling modeling project. The objectives of this project are 1) to convert a one-dimensional, laser drilling computer code, DRILL[1], which was written in Hewlett-Packard (HP) FORTRAN on an HP workstation with a UNIX operating system, to run on an IBM-compatible, Pentium-class PC; 2) to benchmark PC run results against HP workstation

results; 3) to improve the code's user-related graphics capabilities; and 4) to improve the physics models in the code.

Progress

We transferred a version of the DRILL computer code to a Gateway 2000, model P166, 166 MHz Pentium PC and modified it to run under the Windows 95 operating system with Microsoft Power Station FORTRAN 90. After running a problem for atmospheric air over solid aluminum with supplied input and data files to completion at 5 µs, we benchmarked the calculated results against the results for the same input and data files from an HP series 700 workstation. The agreement between the results was excellent.

We then performed a set of computer runs by varying the peak laser incident pulse intensity and the pulse length. With the initial 120-ns-long laser pulse and spherical vapor expansion into the air, the code ran to completion for peak laser intensities of 3×10^8 and 1×10^9 W/cm^2. At 5×10^9 W/cm^2 of peak incident laser intensity, the code suffered a

Solid	Liquid	Vapor	Compressed gas	Ambient gas

Solid-melt interface   Ablation surface   Contact surface   Shock surface
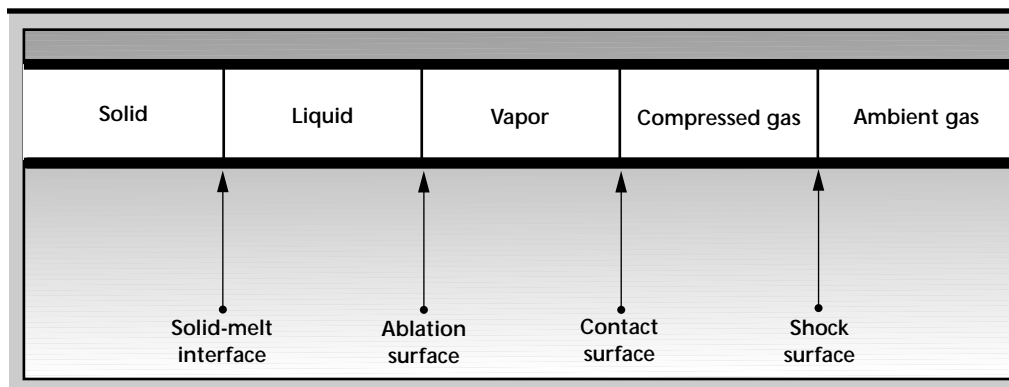
Figure 1. Schematic arrangement of the evolved material regions.
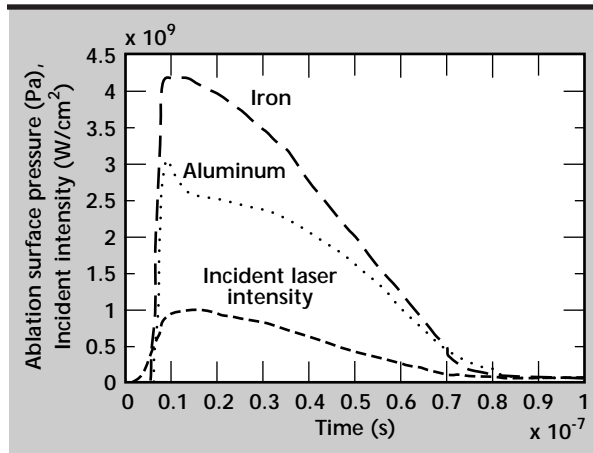
Figure 2. Calculated ablation surface pressure vs time for aluminum and iron solids at peak incident laser intensity of 1×10^9 W/cm^2.

Figure 3. Ablation surface pressure for iron at 5×10^9 W/cm^2.

dramatic slowdown in integrator time step. It ran for several hours, but could not proceed past about 18.6 ns in simulation time. At 5×10^9 W/cm^2, with the original laser pulse truncated to 20 ns, the code ran to completion. At 1×10^{10} W/cm^2, with the truncated incident pulse, the integrator in the code was unable to achieve corrector convergence beyond 6.98 ns into the simulation. The reason for the inability to obtain solutions at the higher incident laser intensities is not yet known.

Several runs with atmospheric air over iron were investigated next, with spherical vapor expansion into the air. The code ran to completion with the initial incident laser pulse at 1×10^9 and 5×10^9 W/cm^2 of peak intensity, but encountered the time-step slowing down problem at about 26.0 ns into the simulation, with a peak incident intensity of 1×10^{10} W/cm^2.

When a sufficiently intense laser pulse impinges on most solid materials, it melts the surface of the material. If the incident pulse continues, the melted, liquid layer expands into the solid. As this occurs, the liquid surface vaporizes, and a vapor front expands into the ambient gas over the solid. The ambient gas is then compressed, and a shock surface is generated. This array of material phases is schematically portrayed in **Fig. 1**, which has been

adapted from Reference 1. The time-varying thicknesses of the regions beyond the solid are usually on the order of microns.

The pressure/time history at the ablation surface is one calculated quantity of interest to the developers of some manufacturing processes. This pressure/time history is an estimate of the pressure pulse transmitted into the solid. **Figure 2** shows the calculated ablation surface pressure/time history for aluminum and iron solids at a peak incident laser intensity of 1×10^9 W/cm^2. The time history of the incident laser pulse is also shown. **Figure 3** shows the ablation surface pressure for iron at 5×10^9 W/cm^2 of peak incident intensity. These peak pressures agree roughly in order of magnitude with those reported in the literature on experimental investigations for metal ablation.

We next added comment lines to the input file and the source output routine to describe briefly the input parameters and output arrays for a code user. These modifications allow a user to more easily know what physical quantities the input parameter and output array names represent.

References

1. Boley, C. D., and J. T. Early (1994), "Computational Model of Drilling with High Radiance Pulsed Lasers," *Proc. ICALEO '94, Orlando, Fla.*, pp. 499–508.

Gerald L. Goudreau, Thrust Area Leader

The Computational Mechanics thrust area has continued to sponsor investigations into the solid, structural, and fluid mechanics and heat transfer underlying the state-of-the-art computational software used in engineering analysis in support of programs at Lawrence Livermore National Laboratory (LLNL).

The scale of capability spans the office workstation and its continued expansion of interactive and graphic capability, to "departmental" or building compute servers, and finally, central or national massively parallel supercomputing. Our past breadth in technology transfer to industry has narrowed, and is focused toward funded activity.

The Department of Defense High Performance Computing and Modernization Program is funding the LLNL ParaDyn Project to port DYNA3D to their major resource centers. The DOE Defense Program's ASCI Blue acquisition holds ParaDyn as one of its prototype applications. In addition, the Federal Aviation Agency is funding LLNL to provide DYNA3D (and eventually ALE3D) to the aircraft industry as a tool kit to address debris damage and mitigation assessment.

This past year has been the first full year that the Methods Development Group (MDG) has been under the leadership of P. Raboin, who has written a survey article on code activities, including general DYNA, NIKE, and TOPAZ enhancements, as well as mesh generation and graphic post-processing tools. C. Hoover presents an article on the overall ParaDyn Project, with her broad team and many activities. J. Lin and M. Puso each have articles on new rigid body material options just added to DYNA3D and NIKE3D, respectively. All of these activities are within the MDG, and together with P. Raboin's article will be expanded into a special code capabilities document.

Additional articles this year include the effort led by A. Shapiro to characterize the laser-induced mechanical failure and damage of optical components; creative use of a suite of heat transfer codes to perform a thermal evaluation of the nuclear weapons pit storage facility at Pantex, by B. Kornblum; radiation-induced thermal stresses in high-power flashlamps, by J. Maltby; and finally, the simulation of turbine fragment containment and mitigation of aircraft engines, by G. Kay and E. Zywicz.

Computational Mechanics

2

2. Computational Mechanics

Finite-Element Code Enhancement

Peter J. Raboin
Defense Technologies Engineering Division
Mechanical Engineering

The DYNA, NIKE, TOPAZ, and GRIZ family of finite-element codes provide Lawrence Livermore National Laboratory (LLNL) programs (50-60 analysts) with needed dynamic, structural and thermal analyses. Our broad service charter is to provide rapid and highly skilled user support and has resulted in over 75 improvements to these codes in FY-96. Material switching capabilities in DYNA3D and new NASA matrix solvers in NIKE3D have increased the execution speeds of these codes by factors of over 1000, and have permitted heretofore impossibly complex and large finite-element model analyses. Visualization capabilities are expanding with promising developments in the areas of parallel processing, ASCI-compliant standards and plot database flexibility. Exciting new advancements in computer hardware (DEC Alpha and IBM SP2) are quickly being matched with code portings to these new platforms. Web-browsing capabilities for our software manuals were demonstrated for DYNA3D and MAZE this year, with more manuals to follow. Finally, the Code Collaborators Program has been strengthened with tougher standards for admission, renewals and documentation.

Introduction

There are diverse and numerous year-end deliverables for the Methods Development Group (MDG) at LLNL, because of its primary function as a code support, code development, and research organization. Each year within MDG, there are specific projects to add larger code capabilities to our software. For example, the ParaDyn project is adding parallel capabilities, and another project adds rigid body capabilities to DYNA3D and NIKE3D.

In FY-96, our code enhancement work consisted of five projects: DYNA, NIKE, TOPAZ, pre- and post-processing and rezoning. This report describes some of the general code up-grades, gives short descriptions of the major accomplishments for each project, and provides a Code Collaborators Program status update.

A more detailed listing of the specific code enhancements for FY-96 is given in an appendix.

Progress

General Up-grades

During FY-96, up-grades in computer equipment and operating systems had a broadening effect on the base of MDG code development. Sun Solaris operating systems were added across LLNL, and new DEC Alpha machines were purchased for open and classified computing. The MDG codes were ported to support these changes.

At present, our serial codes support the following platforms: Cray, DEC Alpha, HP, IBM RS6000, SGI, and Sun. We support multiple hardware and operating system variations for several of these machines. Executables for these codes are maintained on both the open and classified computing systems. The supported parallel machines are Meiko, Cray T3d, and IBM SP2. Concurrent Version System (CVS) version control is used to provide single-source code tracking and to facilitate multiple code developer activities.

During the last months of FY-96, the MAZE and DYNA3D manuals appeared for the first time as Web HTML documents. This was accomplished with a software package that converts FrameMaker documents to HTML. Formal requests have been made to provide NetScape on LLNL's classified system so that these manuals may be interactively viewed. It is our intention to provide up-to-date manuals via this medium for all of the MDG codes on the unclassified and classified systems.

DYNA

Mixed-time integration methods, contact algorithms and user-requested code improvements were

important priorities in this year's MDG DYNA3D development. Two rigid body features, improved rigid body joints and deformable/rigid material switching were also added. The implementation strategy for mixed-time integration is laid out. Element/node partition work is in progress. The sub-cycling scheme will be implemented with selected contact types in the coming year.

Two contact capabilities were improved: tying SAND into type-8 contact, and adding automatic contact velocity damping for Lagrange projections. User-requested DYNA improvements aided the AT-400A (**Fig. 1**), W56 Transportation Safety Risk Assessment (TSRA), FL Container, and DOD penetrator studies in the Defense Technologies Engineering Division (DTED). Work performed for Energy Manufacturing and Transportation Technology (EMATT) by the New Technologies Engineering Division (NTED) was helped with features added to assist the VM and Composites Cooperative Research and Development Agreements (CRADAs) and the Boeing project. Another important accomplishment this year was the preparation and presentation of a DYNA3D training course.
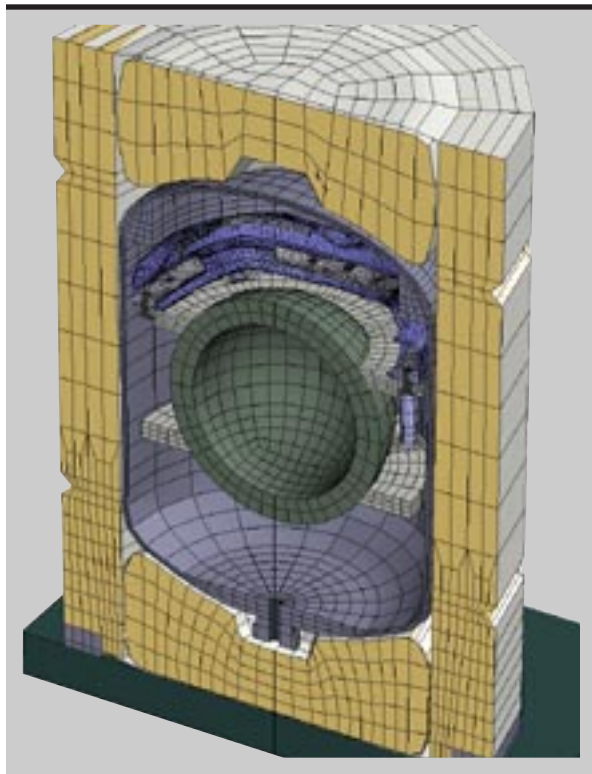
Figure 1. The AT400A transportation container, experimentally and analytically dropped, crushed, and cooked to validate and prove its functionality as a transportation container for nuclear weapon components.

NIKE

For NIKE3D development, the top priorities were up-grades to the structural elements, improved linear solvers, double precision, a short effort in NIKE-DYNA-NIKE linkages, and user-requested code improvements. These goals were largely achieved. There are now concrete shell and beam models for structural elements, user-defined integration for concrete and windshield composites, and a new membrane element for metal forming.

Two new linear solvers were added and a double precision version of NIKE3D was completed for workstation use. A linkage capability was also demonstrated and used to motivate the element technology Laboratory-Directed Research and Development (LDRD) proposal. Two Lasers Program projects benefited this year from NIKE3D development: a NIKE3D coupling to thermo-optical diffraction calculations and another NIKE3D coupling to a commercial multiple-optic beam path calculation code (TSO). Special assistance was provided for the Boeing CRADA (in NTED) with additions of the membrane element and added capabilities for the Super Plastic Forming pressure scheduler. The FL container vibration problem (in DTED) benefited from close user support and two new capabilities: NASA's variable band solver (VBS) and an auto-time-stepping algorithm for dynamic vibration problems (**Fig. 2**).

TOPAZ

Based on programmatic requests (from DTED) and the prevalence of computer workstations, a double precision version of TOPAZ3D was created this year. This work, which affects data structures and I/O routines, was made compatible with the new double precision version of NIKE3D. This was done to simplify the future integration/up-grade of these codes into the coupled NITO3D code. This latter code was ported in FY-96 from the workstation to the Cray J-90 supercomputer.

FACET was also up-graded in FY-96 by implementing the more accurate axisymmetric radiation view factor algorithm from the GLAM physics code. GLAM was written in the late 1970's on the CDC7600 using LRLTRAN.

Pre- and Post-Processing

The GRIZ post-processor was our top priority in FY-96. Its development is central to the ParaDyn effort and to our goals in supporting the new LLNL

ASCI Program. We are a member of the ASCI Common Data Format Working Group, the goal of which is to create a common (portable) scientific data file format/Application Program Interface (API) to support ASCI work. In conjunction with our goals for GRIZ and our participation in ASCI, the mesh input/output (Mili I/O) library received considerable development this year as a replacement to the fixed TAURUS database. The Mili I/O library has both C and Fortran interfaces that will be used by the analysis codes and GRIZ. Completion of the Mili library interface is now proceeding, in conjunction with the conversion of GRIZ to a mesh I/O based (Mili) application.

For pre- and post-processing in general, we have completely revamped three manuals, for GRIZ, THUG, and MAZE. Running GRIZ on Sun workstations was finally accomplished this year. The latest Ultra-Sparc workstations have a hardware graphic OpenGL capability, and the older Sparc workstations can use Mesa, a freeware version of OpenGL. GRIZ is now functional on all of the MDG supported platforms and its use by the analysts is increasing. The addition of THUG as a specialized time-history graphical program has bridged all of the capabilities of TAURUS.

GRIZ and THUG supported important programmatic and research work this year with new time-history filtering features for the AT400A impact analyses, advanced iso-surface visualization capabilities for the Array of Conventional Explosives (ACE) project and unique feature support of the Large Eddy Simulation (LES) turbulence modeling LDRD.

Minimal support was maintained for Orion and Ingrid in FY-96. A low level of support is anticipated for MAZE in FY-97. With a new MAZE manual and the addition of a repertoire of region commands, the FY-97 feature requests for MAZE are limited. Animation activities were in strong demand this year (21 high quality videos were produced). MDG continues to support a modest video creation capability, and new animation capabilities were added with the purchase of advanced visualization capabilities through WaveFront.

Rezoner

A new 3-D rezoning task was initiated this year with three goals: 1) graphical monitoring of the analysis results while the code is running; 2) interactive rezoning between analysis iterations; and 3) automatic rezoning between iterations, either through scripting or error detections. This was to be a multi-year phased project, but a departure of

key personnel in FY-96 has forced suspension until suitable replacements are identified. Completed this last year is a software development design plan that investigated user requirements and compatibility between analysis codes, GRIZ, and the rezoner requirements. We completed work on interfaces for 3-D interactive rezoning and also demonstrated a prototype capability to display mesh quality measures (skewness and aspect ratio) in GRIZ.

Code Collaborators Program

The Code Collaborators Program this year saw the addition of 32 new collaborators (32 renewals) and the dropping of old collaborators who had
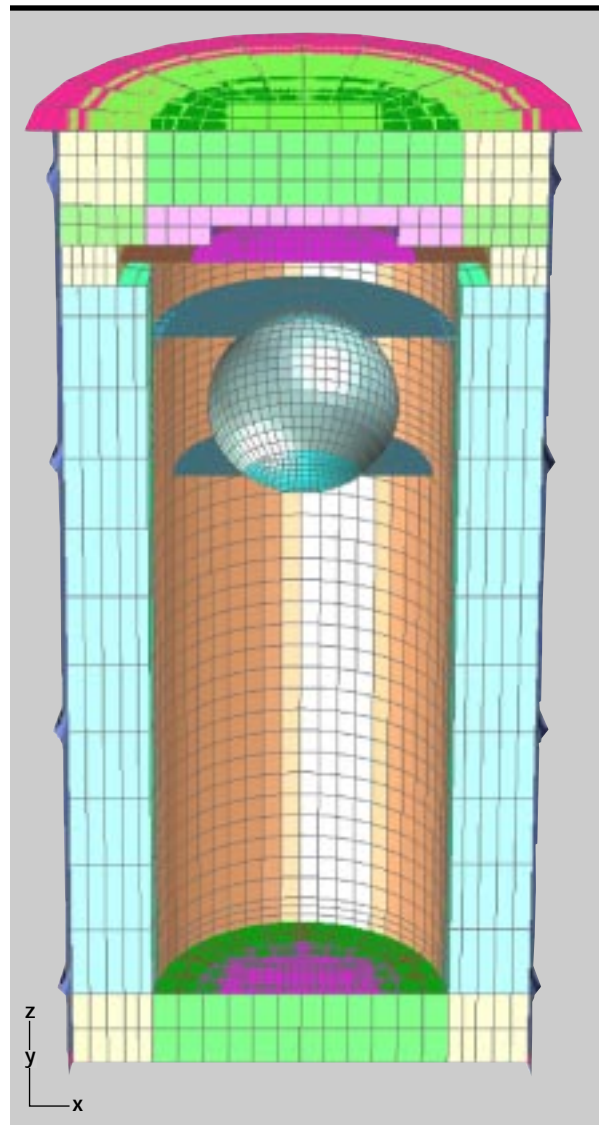
Figure 2. The FL container, subjected to extensive vibration analyses with NIKE3D to study its overall mechanical response and the structural integrity of certain welded joints.

shown no participation in the last 18 months. The program currently has 152 collaborators, consisting of 54 academic institutions, 72 industrial partners, 21 government organizations, and 5 associates who have legal partnerships with LLNL. Following DOE requirements, the Energy Science and Technology Software Center (ESTSC) was up-dated this year with copies of our most recent "stable" software. The ESTSC meets the needs of those interested in obtaining our codes for analysis, thereby decreasing the demands placed on the program, while retaining those organizations who truly pursue code development.

The purpose of collaboration remains the same: we wish to advance the development of MDG codes through the mutual sharing of software developments, bug fixes, and general capability enhancements. Several notable code contributions were obtained this year. A new NASA solver was incorporated into NIKE3D. This solver was essential in solving a FL container vibration problem for DTED. It provided a 6500% code speed-up using four times less memory. Logicon gave us a new material model (#45) which is an enhancement of model #16 for geologic materials. The Defense Evaluation Research Agency in the UK provided complete copies of their source up-dates as well as their documentation for ISO 9003 software quality assurance registration of DYNA3D software.

There remains a back-log of unincorporated collaborator contributions that are prioritized by the LLNL community needs. This program remains valuable for the capability improvements and bug fixes we receive.

Summary

The MDG group activities affect a large user base of 50 to 60 analysts who are principally in the DTED, NTED and Laser divisions of Mechanical Engineering at LLNL. In FY-96, the MDG code development projects were aligned with the codes that require the most support and receive the greatest use. Most interactions between the code developers and the users are brief, being either quick feature enhancements or bug reports with subsequent fixes. Not shown in this report are the day-to-day user interactions of instruction, model debugging, and problem brainstorming which often go into solving difficult solid mechanics and heat transfer problems. The Code Collaborators Program has continued this year with tougher participation requirements and an increased reliance on the ESTSC for satisfying demands for "public domain" versions of DYNA, NIKE, and TOPAZ software.

Future Work

Since the numerous code capability improvements undertaken by MDG in FY-96 were motivated by the user community, a prediction of future work is difficult.

What is known is that the Cray J-90's will get an operating system up-grade, which means that Fortran 90 modifications are needed for all of MDG codes. Element technology research will provide greater mid-analysis switching capabilities between DYNA3D and NIKE3D, compatible DYNA3D/NIKE3D element formulations and more stable element formulations.

Parallel developments undertaken by the ParaDyn project will deliver more computing horsepower to the analyst than ever before, so more MDG user support is expected in this area. DYNA3D work continues on sub-cycling, contact improvements, and rigid body joint failure. NIKE3D work will improve implicit dynamics (HHT integration), rotational inertia for shells, and beams, and modal superposition. Incorporation and merging of the GEMINI capabilities into NIKE3D is also planned. TOPAZ3D expects to improve its shell element heat transfer capabilities as well as implement some of the new matrix solver routines. Continued GRIZ development is focused on Mili, parallel implementations, and a substantial user request list.

Web browsing for our manuals is seen as a necessary task to bring MDG documentation up to current technology standards. Finally, more effort is being directed to quality assurance improvements in our conduct of code development.

Appendix

The lists in this section are for individual codes that received Computational Mechanics Thrust Area funding in FY-96. Items indicated by ◊ are purely LLNL activities; those marked with an asterisk (*) are external collaborator contributions.

DYNA3D

New and Improved Features
- ◊ Material switching
- ◊ Revolute joints
- ◊ Dynamic memory allocation
- ◊ Velocity correction to Lagrangian contact algorithm
- ◊ Mass proportional damping switch (on/off)
- ◊ Command line file sizes
- ◊ Volumetric strain failure criterion added to material type 11

◊ Composite material orientation
∗ Material model #45: improved material type 16
◊ Material model #46: fully anisotropic elasticity
◊ Multiple type-2 load curves
∗ Type-8 slide surfaces in SAND
◊ Additional NIKE3D to DYNA3D stress initialization options
◊ One-step plane stress elastoplasticity algorithm added
◊ Bathe-Dvorkin and fully-integrated YASE shell element formulations
◊ Slave node radii in type-5 contact—discrete nodes impacting a surface
◊ Model #22 minimum time-step size deletion criteria, plus additional failure options
◊ Multiple groups of materials for Rigid/Deformable Material Switching
◊ Breakable Rigid Body Joints with activation time and acceleration failure criterion
◊ User-defined initial velocity superposition for rotational motion
◊ Suppression of SAND failure information in printout

Bug Fixes
◊ Internal energy for discrete spring/damper elements
◊ Hughes-Liu shell element
◊ Rigid Body Joints error in initial position differences
∗ Vectorized auto contact
◊ Tie-Breaking Shell Slidelines
◊ Deformable/Rigid Material Switching
◊ Thick shell elements
◊ Equation of state 11

DYNA2D

Bug Fixes
◊ TOPAZ2D plotfile compatibility
◊ Double precision bugs in common blocks

NIKE3D

New and Improved Features
◊ Makefile
◊ Viscoelasticity material model
◊ 24-point limitation for user-defined beam and shell integration points eliminated
◊ Concrete damage and cyclic steel plasticity beam model
◊ Double precision
◊ Compressed row format storage of stiffness matrix

∗ Cray Research Inc. VBS solver
∗ SGI and HP Sparse Solvers
◊ Additional time-history and incompatible modes output
◊ Membrane element
◊ Bisection (line search) algorithm for SPF pressure scheduler
∗ NASA VSS solvers
◊ Dynamic memory allocation
◊ Auto-time-stepping algorithm, called AUTV, to solve dynamic vibration
◊ Improved tangent stiffnesses for contact
◊ Single pass contact and improved search algorithms

Bug Fixes
◊ Thermomechanical double precision
◊ Memory holes in dynamic memory allocation for HP workstations
◊ File size mismatches with TOPAZ3D
◊ Slide surface memory conflicts with rigid body data structures
◊ New FL input format

NIKE2D

New and Improved Features
◊ GRIZ plot format
◊ Interference fits for slide lines

Bug Fixes
◊ Double precision load curves
◊ Zero load application error

TOPAZ

New and Improved Features
◊ Added logic to reduce the time step if the stiffness matrix is non-positive definite
◊ Replaced variable time step coding in TOPAZ3D with more robust coding from TOPAZ2D
◊ Double precision

Bug Fixes
◊ Corrected bug in dump/restart subroutines—the number of words in a common block was incorrectly defined.
◊ Corrected bug in slide line logic (a crash occurred when the user specifies 0 for the contact conductance). This is now captured and the slide line is ignored.
◊ Corrected bug in dump/restart logic (dt for the last time step may be 0 or very small so that the end time is hit exactly).

◊ Fixed bug in reading an exchange factor text file. The exchange factor matrix is non-symmetric. The previous coding read a symmetric view factor matrix.

◊ Unresolved FEM formulation problem. A NIF problem showed an anomalous negative temperature wave moving through the mesh. No bug could be found in TOPAZ. The commercial code ANSYS gave the same anomalous results.

GEMINI

New and Improved Features
◊ Composite nodal damping
◊ Workstation version
◊ Repetitive load case generation

Bug Fixes
◊ Corrected bandwidth minimization for shells having composite nodal damping

GRIZ

New and Improved Features
◊ Updated GRIZ manual to 2.0
◊ Unit conversion via scale and intercept
◊ Material manager material selection picking
◊ "tellpos" node/element position queries via command line or mouse picking
◊ Refined color legend and scaling procedure
◊ GRIZ under HPUX using Mesa on an HP workstation
◊ GRIZ on the Dec Alpha
◊ Automatic length determination for floating point displays in time-history plots
◊ Mesh bounding box up-dates
◊ Mode-dependent cursor shapes
◊ Dynamic linking to user-defined libraries of time-series transformation functions
◊ Multiple independent sets of particle traces
◊ Save/restore of particle traces
◊ Disable particle traces for selected materials
◊ Save vector field values
◊ Node-base vector fields

Bug Fixes
◊ Material manager GUI
◊ Interpolated time display
◊ Y-axis labeling for time-history plots
◊ Time-history re-calculation after a change in the basis, reference surface, or strain variety
◊ Ensure that edge lines are rendered on top of coincident polygons
◊ TOPAZ3D plotfile database compatibility

Mili I/O

New and Improved Features
◊ Mili allows per-subrecord data organization so that either object-ordered or results-ordered data structures can be used.
◊ Mili has a memory management layer for facilitating the traversal and re-organization of internal data structures.
◊ Mili can detect an "active" database and is able to operate on a growing database.
◊ Mili has added the data structures (hash tables) to support derived and primal results and GRIZ is similarly supporting dynamic menu creation.

THUG

New and Improved Features
◊ Interpreted "sum" and "diff" commands for derived results
◊ thd2s utility added to convert Cray to Workstation data formats
◊ Dynamic linking to user-defined libraries of time-series transformation functions
◊ Nodal data exclusion flag for time-history database
◊ Incorporated Low-Pass Butterworth digital filter
◊ Up-dated the THUG User Manual
◊ Completed THUG software distribution for the Code Collaborators Program

MAZE

New and Improved Features
◊ Expanded region command capabilities
◊ Completed a new User's Manual
◊ New slideline types supported
◊ Equation-of-state modifications
◊ Region node number assignment to a parameter symbol

ANIMATION

◊ Titling modifications to "Advanced Hydro Test Facility" (J. Pastrnak)
◊ "Pulse-Induced Deformation in a Laser Slab" (M. Rotter)
◊ Shortened "Computational Biomechanics" (for G. Goudreau)
◊ "Large Eddy Simulation of Flow Over a Backward Facing Step" (B. Kornblum and R. McCallen)
◊ "GM Casting Rework" (P. Raboin)
◊ "ParaDyn Benchmark Applications" (C. Hoover and T. Degroot)

◊ "The Role of Animation Production in Visualization" (M. Loomis)

◊ "High-Bypass Commercial Aircraft Engine Simulation" (E. Zywicz)

◊ "Vector Carpets for Irregular Grids" (D. Dovey)

◊ "Earthquake Simulation Compilation" (D. McCallen and T. Chargin)

◊ "Finite Element Modeling of Lower Extremities" (S. Perfect)

◊ "TEAM Virtual Manufacturing: Sheet Metal Forming" (T. Lee)

◊ "LLNL High Performance Computing" (G. Goudreau, J. Peters (NASA))

◊ "Large Eddy Simulation of Turbulent Separating Flow" (B. Kornblum and R. McCallen)

◊ "Computational Biomechanics at LLNL" (K. Hollerbach and D. Schauer)

◊ "Spin-Forming of Uranium Components" (P. Raboin)

◊ Presentation graphics (transparency and paper hardcopy) (for H. Louis)

◊ Copies of "Seismic Response of the 24-580-980 Interchange" (D. McCallen)

◊ "Array of Conventional Explosives In Use Against a Deep Target" (D. Badders and M. Shannon)

◊ "Laminar Backward-Facing Step Flow" (B. Kornblum)

◊ "Arrays of Conventional Explosives in Use Against a Deep Tunnel" (D. Badders)

◊ Presentation graphics (transparency and paper hardcopy) (for G. Bernal)

◊ "DYNA3D Model of AT400A Drop Tests" (D. Badders)

◊ "OAB-LTAB Material Flow Animation" (M. McDaniel)

Parallel Algorithm Research for Solid Mechanics Applications Using Finite Element Analysis

Carol G. Hoover and Daniel C. Badders
Defense Technologies Engineering Division
Mechanical Engineering

Anthony J. De Groot and Robert J. Sherwood
Engineering Research Division
Electronics Engineering

The goal of this project is parallel algorithm development for both explicit and implicit finite element applications in solid and structural mechanics. New capabilities developed this year include the development of software modules for multiple partitioning of the finite element meshes, efficient parallel implementation of mechanics algorithms in DYNA3D, and production simulations of underground shock-structure interactions with over a million elements. Performance results for ParaDyn, the parallel version of DYNA3D, demonstrate parallel scalability and efficiency for several benchmark applications and production problems.

Introduction

This research in parallel algorithms is targeted for the two nonlinear, large deformation solid mechanics programs, DYNA3D and NIKE3D. DYNA3D[1] is an explicit finite-element program for analyzing the transient dynamic response of 3-D solids and structures. The element formulations available include 1-D truss and beam elements, 2-D quadrilateral and triangular shell elements, and 3-D continuum elements. The thirteen contact-interface algorithms provided in DYNA3D include state-of-the-art algorithms for modeling arbitrary contact for large deformations, frictional sliding, single surface contact, and contact with failure. Rigid materials provide added modeling flexibility.

NIKE3D[2] is a fully implicit finite element program for analyzing the finite-strain, static response of 3-D solids and structures. Spatial discretization includes solid, shell, beam and truss elements with 8-point, 4-point and 2-point integration for these element types. Four contact interface algorithms permit gaps, frictional sliding, and mesh discontinuities along material interfaces. The resulting system of simultaneous linear equations is solved either iteratively using an element-by-element method, or directly by a factorization method. Nonlinear solution strategies include full-, modified- and quasi-Newton methods.

DYNA3D (ParaDyn) uses a successful strategy for parallel implementation of an explicit finite element method[3,4] based on dividing the finite element mesh among the processors and executing ParaDyn on the individual subdomains. This is a message-passing parallel model with nodal point data identified as the shared data between processors. The challenging aspects of the parallel algorithm development are the optimal partitioning of the finite element mesh and the treatment of the interfaces at material boundaries (contact algorithms). Reference 4 describes a benchmark application in crashworthiness simulations.

Current research[5] in graph partitioning algorithms provides a means for developing automated tools for partitioning finite element meshes. It is now possible to develop multiple partitions for domains based on a physical partitioning of the problem as well as for treating boundary conditions such as contact. Software development for automating multiple partitioning methods is described in a section below. These methods can be extended into parallel form as needed for future large applications.

Parallel computers today are characterized by their memory hierarchy as well as by the number of processors on the full system. Typically, the number of processors varies from as few as 8 to as many as 4,000. For computers with a large number of processors, the memory is always distributed with an

interconnect network between individual nodes. The nodes may include multiple processors that share cache and secondary memory. The parallel programming model in this case is generally message-passing.

The smallest parallel computers are typically a single node with processors sharing the memory by means of a high-speed bus interconnect. Although in the past the parallel programming model for these computers has been shared memory, the emergence of the Message Passing Interface (MPI) standard has provided shared memory computer vendors the motivation and impetus to provide this programming model as well. Future parallel computers are expected to include shared memory nodes with an interconnect network. The implementation of the MPI standard in ParaDyn in this last year, along with optimal mesh partitioning, will enable efficient problems on future parallel computers regardless of their architecture.

Research leading to parallel algorithms for NIKE3D is just beginning. The plan for implementing parallel algorithms in NIKE3D is to develop first a shared memory version of linear solvers for an 8- to 10-fold increase in problem performance for a minimum amount of reprogramming effort. This will enable longer time simulations for applications such as spin-forming in the near future. Current research efforts are being directed toward message-passing solution techniques that will enable the larger size NIKE3D applications on the next generation massively parallel computers.

Progress

The most notable accomplishments in FY-96 have been (1) the simulation of problems with over a million elements; (2) simulations of large underground shock/structure problems with localized contact; and (3) the development of parallel algorithms to implement the mechanics capabilities needed for production applications and parallel benchmark tests.

Parallel Applications and Performance

One of our most challenging goals has been to demonstrate a production ParaDyn calculation with up to a million elements. The calculation shown in **Fig. 1** is one of a series of underground blast calculations simulated on a 128-node Meiko CS-2, and more recently on a 256-processor IBM SP-2. The problem sizes vary from a million elements to several million elements, and demonstrate excellent scaling for the Meiko CS-2, the IBM SP-2, and the Cray Research Incorporated T3D, as shown in **Fig. 2**. A series of performance studies using 128 Meiko nodes with two processors per node resulted in roughly 50% improvement over the results using one processor per node. The absolute performance on the SP-2 using 128 processors is .5 μs per element cycle. This is roughly 50 times faster than a single processor on a Cray Y-MP.

Figure 1. Underground blast simulation modeled with over seven million degrees of freedom.
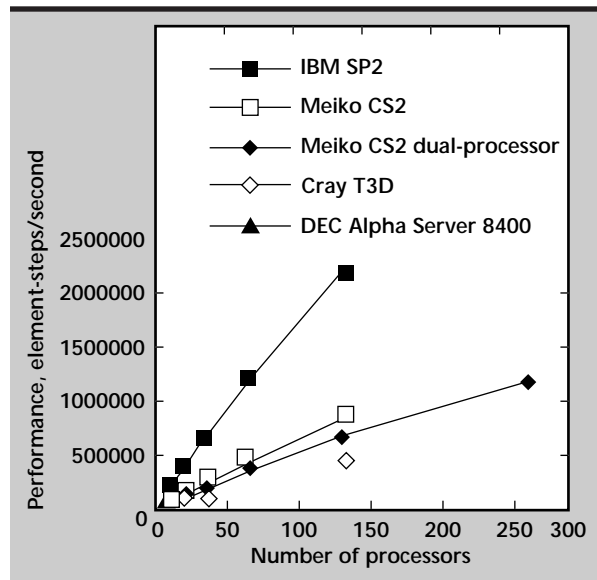
Figure 2. Performance of parallel computers. The performance measurements are based on automatic optimization provided by the compilers. No machine-dependent changes were made to the ParaDyn source program.

Last year a collaborative activity with researchers at the Waterways Experiment Station in Mississippi was initiated through the Department of Defense High Performance Computing and Modernization Program. **Figure 3** is the first of several production benchmarks problems which are deliverables to demonstrate parallel performance. This benchmark is a simulation of an underground shock/structure interaction. The contact surfaces between the soil and the structure were modeled with the nodes-in-processor contact algorithm.[3,4] The contact included one tied surface and five surfaces modeled with a symmetric penalty algorithm. New preprocessing software, described below, was developed to partition the contact surfaces.

Validation of results on parallel computers is an important component of this project. **Figure 4** illustrates a benchmark application that is a robust test of the parallel automatic contact algorithm and that provides an interesting physical system for studying nonlinear phenomena and fractal power laws. The dynamics of the crumpled surface shown in the figure is modeled by compressing a thin sheet with six moving stonewall boundaries. This simulation is run for many sound traversal times and the calculation time is significantly affected by the contact search and force calculations. The large number of contacts at high compression make this a severe test for verifying that all contact points are located and that parallel implementation is correctly summing the contact forces on all processors with shared nodes.

The design of production parallel runs on a routine basis requires characterizing the performance of individual parallel algorithms in ParaDyn to provide guidelines to the analyst for selecting an optimal number of processors for the algorithms used in a particular application. Current performance results show that problems composed primarily of solid elements will perform efficiently with roughly 1000 elements allocated in a processor. In contrast, applications dominated by shell elements, such as the simulations of automobile crashworthiness, perform efficiently with as few as
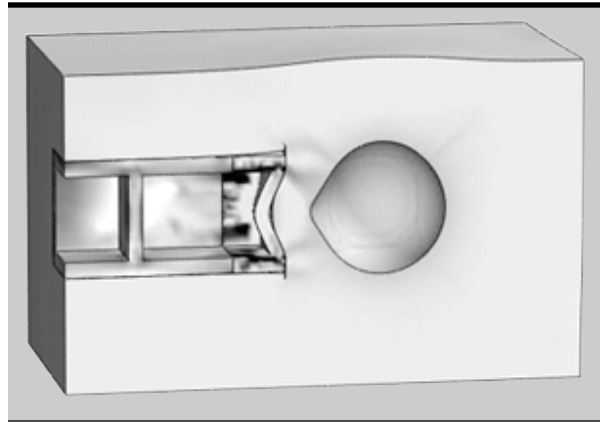
Figure 3. Underground shock/structure interaction at the end of the calculation. The right wall and the right portions of the back wall, ceiling and floor are a reinforced concrete material. The maximum pressure is shown with the darkest shading and the minimum pressure is white. Maximum damage occurs at the center of the right wall.
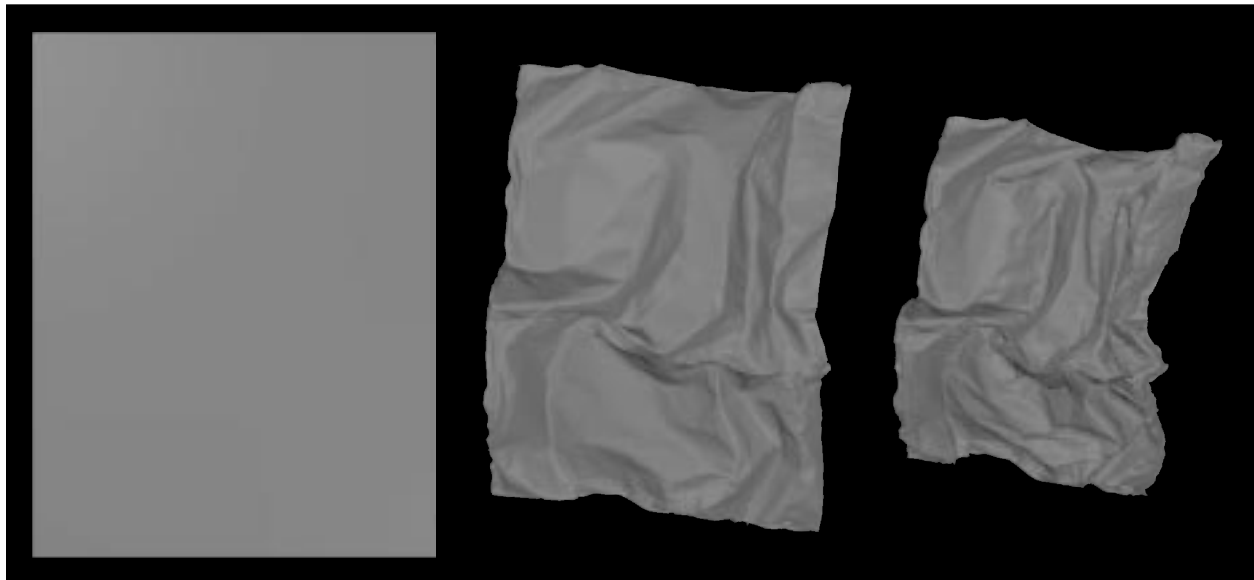
Figure 4. Thin square sheet of shell elements compressed by six moving stonewall boundaries which form the faces of a cube. The boundaries are moving at a speed slow compared to the sound speed of the material so that the initial pressure is small compared to the yield strength. The nonlinear folding of the material provides a very robust test of the contact algorithms.

250 elements in a processor. The nodes-in-processor algorithm used for the 21 contact surfaces in the crashworthiness simulation described in Reference 4 can become inefficient if the partitioning of the elements produces a significant number of shared nodes in each processor. We are adapting the preprocessing software to prevent this. In particular, contact surface pairs can be allocated fully within one processor or a subset of the processors. Progress described below on the preprocessing software is a major step toward reaching this result.

ParaDyn Algorithms

An important emphasis of the work this year has been the implementation of mechanics capabilities needed for parallel production applications. Parallel and serial versions of DYNA3D and ParaDyn have been merged into a single source. Parallel versions of the Hughes-Liu shell, applied loads (pressure loads, nodal forces, base accelerations), non-reflecting boundary conditions and rigid body mechanics

have been completed. Work is in progress on the parallel automatic contact algorithm to complete the implementation of new features such as material inclusion/exclusion and boxes defining the portion of the grid for automatic contact. The new multiple automatic contact algorithm requires the redesign of the parallel data structures and will be continued into next year.

The parallel rigid body algorithm is designed to minimize communication for rigid body motion through the use of efficient parallel global sums and is implemented as follows. The rigid material option in DYNA3D (material type 20) transforms the finite elements in the material into the list of nodes representing a rigid body and the physical parameters associated with the rigid body: center of mass coordinates, body velocities and accelerations, the moment of inertia, body forces and constraints. The parallel algorithm for rigid bodies assigns the nodes in a rigid body to processors based on the partitioning for the elements. For a shared node, the processor selected to include the node in its rigid body sums is the processor with the lowest logical processor number.

Figure 5. Load balancing achieved with multiple partitions. A partitioning of the boundary elements separate from the remaining elements on the mesh provides a method for load balancing a calculation. (a) Two partitions are used for a penetration mechanics problem assigned to the four processors. The elements in the penetrated material are allocated to four processors and the penetrator elements are allocated to the first two processors. The contact surface boundary conditions are allocated to the first two processors. (b) A fluid/structure interaction calculation uses three partitions, one for the fluid, another for the submarine hull, and a third for the surface elements between the fluid and structure.

(a)

Contact surfaces in processors 1 and 2

Element partition in all four processors

(b)

Fluid + submarine + boundary
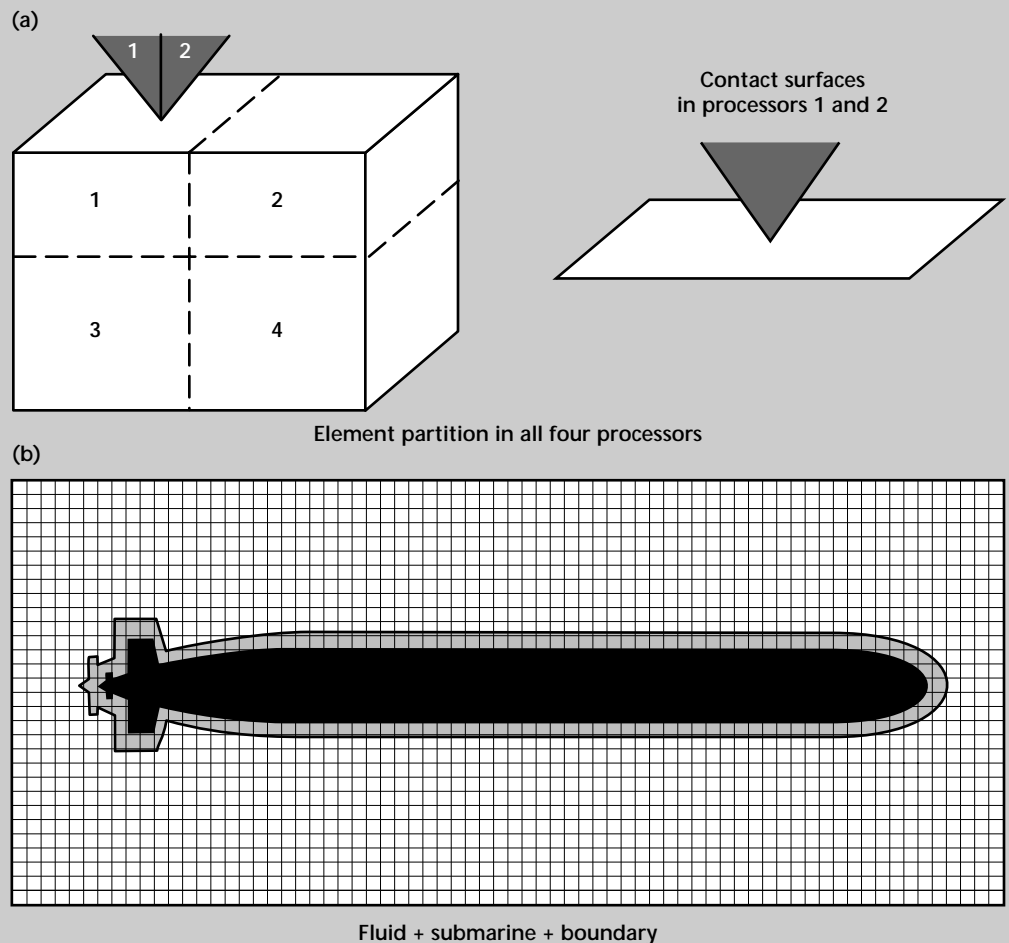
For rigid materials communication during the initialization is used to assign the shared nodes to the lowest processor and to search over individually constrained nodes to evaluate body constraints. Global sums are used to compute body parameters so that all processors know these values for every rigid material. The time integration for the rigid nodes, including shared nodes, is carried out knowing the updated rigid body velocity. The advantages of this algorithm are that the partitioning of rigid materials is the same as that for deformable materials and the partitioning can be weighted by the relative cost of the calculations for rigid body motion and deformable mechanics.

Multiple Partitions for Load-Balanced Calculations

Flexible partitioning software is needed to provide multiple partitions based on the physical characteristics of the problem and the need to load-balance a parallel calculation. **Figure 5** illustrates this idea. **Figure 5a** shows a problem using two partitions, the first for the elements and the second for the contact surfaces. The elements are partitioned using an algorithm selected from the METIS software. The contact surfaces are allocated to processors so that each sliding interface is wholly contained in one processor or a subset of the processors.

The second example (**Fig. 5b**) illustrates the partitioning of a mesh for the PING program, a structural acoustics program coupling a wave equation solver in the fluid with an implicit structural analysis of the submarine. The computer time required to impose the boundary condition between the submarine hull and the fluid represents a significant fraction of the total computational time.

To load-balance this calculation it is important to distribute the elements on the boundary among all of the processors. This can be accomplished using three partitions, one for the elements composing the boundary and two others for the elements in the remainder of the problem. Merging the partitions together produces the desired effect of load-balancing the expensive calculations at the boundary with the calculations in the fluid and structure.

The preprocessing software for partitioning a mesh is illustrated in **Fig. 6**. New programs developed in this last year are shown in shaded boxes. These programs automate the partitioning of the boundary conditions and handle all of the special cases (for example, the local contact surfaces and tied node sets). The new programs were developed to automatically treat the sliding interface partitioning

for the underground shock/structure interaction problem described previously. In this problem the elements that form the sliding interfaces between the soil and the structure are all allocated to one processor. The remaining elements in the problem are partitioned using METIS and assigned to other processors. The partitioned results from METIS are merged with elements assigned to the processor for contact calculations using the MERGECOL program. The merged result is used as input to the PFGEN program to produce the partitioned file used as input to ParaDyn.

The following is a brief description of each of the programs used in the partitioning software. Additional details are found in References 3 and 6.

SNPGEN parses the DYNA3D input file and extracts the slide surface data and other boundary data needing special treatment in the partitioning of the mesh. The output from SNPGEN is a file containing Special Nodal Point sets. These sets of nodes, one for each boundary condition, are subsequently input to PFGEN to be allocated to processors based on specific rules. One example of a rule is the node-in-processor contact algorithm.

DUALGEN reads the list of nodes and elements in the DYNA3D input file and uses efficient algorithms to generate the dual mesh. The dual mesh is the mesh of connected elements. DUALGEN additionally generates a file of Special Element sets using the
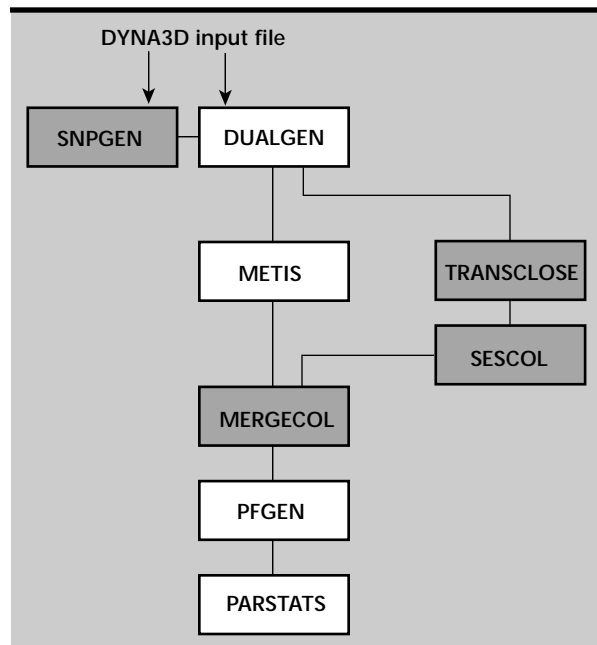
Figure 6. Preprocessing software, designed as a collection of tools to be used with several analysis programs. Each tool can be used separately for generating partitions for specific problem meshes, boundary conditions, and analysis programs.

output of SNPGEN. The Special Element sets include all elements which have one or more of their nodes in the corresponding Special Nodal Point set.

TRANSCLOSE combines Special Element sets if the sets share one or more elements. Duplicate elements are removed in the combined set.

SESCOL is a program planned for the future. It will assign processors to the Special Element sets developed by TRANSCLOSE.

MERGECOL combines the results from the METIS partitioning with the results from the TRANSCLOSE and SESCOL programs to produce the processor assignment list for all elements.

PFGEN uses the processor assignment list to produce the partition file used as input to ParaDyn.

Database Development for Parallel Applications and Visualization

The results of an analysis with DYNA3D and NIKE3D are written into binary databases for subsequent visualization with the post-processor, GRIZ. Efficient use of on-line and off-line storage and networks strongly suggests limiting the size of the databases as much as possible for large runs. This is especially true for a dynamic analysis in which animations provide optimal understanding and interpretation of the results.

Mili, a new mesh input/output library, provides a self-describing database that has the desirable feature that the content of the database can be limited by providing templates in the analysis programs. Furthermore, Mili provides cross-platform database portability. This allows targeting of the database for the desired post-processing environment. Requirements for databases generated in a parallel calculation naturally fit into the grid object design provided with the Mili design. The Mili library development for writing the databases has been completed in this last year. The modifications needed in GRIZ for reading the Mili database will be continued into next year along with a parallel implementation for GRIZ.

Future Work

Our future work will focus on three significant activities: (1) message-passing research for NIKE3D; (2) parallel contact algorithms for both NIKE3D and DYNA3D; and (3) parallel visualization.

Linear solver research for NIKE3D is being planned as a collaborative activity with the Center for Applied Science and Engineering at the Lawrence Livermore National Laboratory (LLNL).[6]

Mathematical research is being conducted in both parallel iterative solvers and direct solvers using a message-passing interface (MPI). Improved pre-conditioning methods are being investigated for iterative solvers and several methods for direct solvers are being designed and written.[7]

We will be designing parallel versions of penalty contact algorithms in DYNA3D that allow the contact surfaces to be partitioned over multiple processors. Lagrangian methods are not conveniently or efficiently treated in this way as coordinate and velocity values at an advanced timestep are required during the calculation of the constraint condition. Parallel techniques developed for contact in DYNA3D provide the experience needed to design parallel contact for NIKE3D.

We will continue to work on the parallel automatic contact algorithm in DYNA3D to incorporate new features such as boxes and multiple automatic contact.

Parallel visualization is essential for the large applications being simulated on the massively parallel computers planned through the Accelerated Strategic Computing Initiative (ASCI) at LLNL. We will be investigating and pursuing client/server models running GRIZ on a small number of processors on the parallel machine and displaying results on an analyst's workstation.

Acknowledgments

References

1. Whirley, R. G., and B. E. Engelmann (1993), *DYNA3D: A Nonlinear, Explicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics—User Manual, Rev. 1*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-107254).

2. Maker, B. N. (1995), *NIKE3D: A Nonlinear, Implicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics—User Manual, Rev. 1*, Lawrence Livermore National Laboratory, Livermore Calif., (UCRL-MA-105268).

3. Hoover, C. G., A. J., De Groot, J. D. Maltby, and R. J. Procassini (1994), *ParaDyn-DYNA3D for Massively Parallel Computers,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-53868-94).

4. Schauer, D. A., C. G. Hoover, G. J. Kay, A. S. Lee, and A. J. De Groot (1996), *Crashworthiness Simulations with DYNA3D*, Transportation Research Board, Washington, D.C., (Paper No. 961249).

5. Karypis, G., and V. Kumar (1995), *METIS: Unstructured Graph Partitioning and Sparse Matrix Ordering System*, University of Minnesota, Department of Computer Science, Minneapolis, Minn., (http://www.cs.umn.edu/~karpis).

6. Procassini, R. J., A. J. De Groot, and J. D. Maltby (1994), *PARTMESH: Partitioning Unstructured Finite Element Meshes for Solution on a Massively Parallel Processor—User Manual*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-118774)**.**

7. Baddourah, M. (1996), National Energy Research Supercomputer Center, Lawrence Berkeley National Laboratory, Berkeley, Calif., private communication.

Characterization of Laser-Induced Mechanical Failure Damage of Optical Components

Arthur B. Shapiro and Thomas A. Reitter
New Technologies Engineering Division
Mechanical Engineering

Douglas R. Faux and Robert A. Riddle
Defense Technologies Engineering Division
Mechanical Engineering

The goal of this research is to quantify by numerical calculations and discriminating experiments the effects of surface and subsurface defects on damage initiation and growth in high power laser optical components. The defects include absorbing spots (for example, surface particulate contamination) and surface damage regions (for example, microcracks and voids) which are present due to environmental exposure and fabrication processes. The damage initiation process is described in terms of the stress waves and thermal stresses that develop when intense laser radiation is absorbed by foreign material attached to the optical surface, producing rapid local heating and material evaporation and ablation. The failure and damage growth in the optical components is described using the concepts of fracture and continuum mechanics. Understanding the precursor to damage will lead to surface cleanliness and surface finish design requirements.

Introduction

We investigated the cause of mechanical failure of two optical components due to laser radiation: (1) the Beamlet spatial filter lens, and (2) multilayer dielectric coated mirrors.

Spatial Filter Lens

The Beamlet spatial filter catastrophic failure caused $128,000 damage to the laser and one month of down-time. Experimental evidence indicates that particulate contamination on the lens surface was the precursor to failure (**Fig. 1**). For NIF to set cleanliness and surface finish requirements, we must determine the particulate material and size that results in critical damage morphologies after laser irradiation.

Particulate contamination on optics can act as an obscuration in the beam or as a local initiation site for laser damage. An on-going experimental program, initiated in FY-96 and funded by NIF, is looking at damage morphologies on optical surfaces contaminated with particles of well-controlled size and composition. The experimental effort is focusing on three main types of contamination: organic
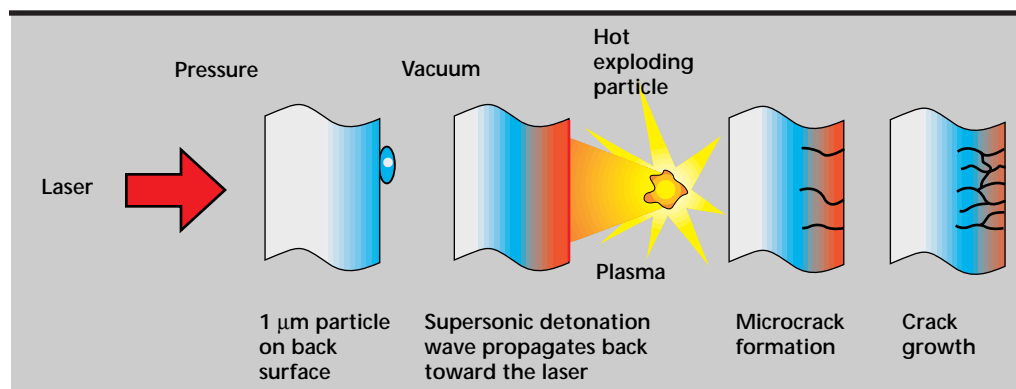
Figure 1. Schematic of the formation of pits/faults in glass as various types of surface debris explode on the first shot, and the evolution of these pits/faults into propagating cracks with subsequent shots.

Laser → Pressure | Vacuum | Hot exploding particle | Plasma

1 µm particle on back surface | Supersonic detonation wave propagates back toward the laser | Microcrack formation | Crack growth

(for example, skin and fibers), inorganic (for example, glass chips), and metals (for example, Al). The work is examining the growth of the plasma plume and the initiation of damage when the optic is in air, vacuum, and dry nitrogen.

To improve our understanding of damage initiation and growth, a modeling effort, complementing the experimental work, was also initiated in FY-96. The modeling group is a multidisciplinary team with representatives from NIF, Physics, Lasers, and Mechanical Engineering. Physics is responsible for formulating theoretical models of the problem with emphasis on plasma physics, laser-matter interaction, and radiation transport. The Lasers Program is using LASNEX to model the laser interaction with surface particulate, plasma formation, and calculation of the magnitude of the surface pressure pulse on the optic. Mechanical Engineering is using DYNA to investigate the optic material response to the surface pressure pulse calculated by LASNEX. Additionally, Mechanical Engineering is using NIKE and TOPAZ to model crack growth due to cyclic thermal-stress loading.

Multilayer Dielectric Coated Mirrors

This work is a continuation of work started in FY-95. The system output fluence and pulse shape of high peak-power laser systems, such as the prototype
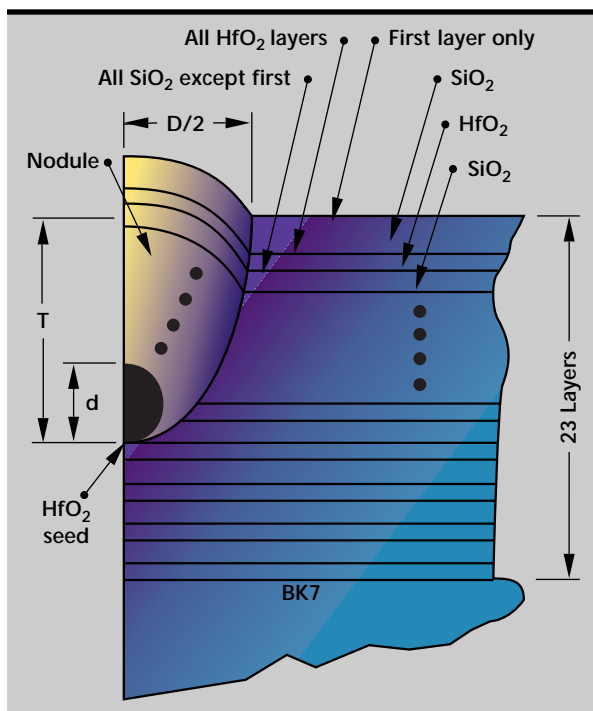
Figure 2. Schematic of a nodule defect. A nodule defect is created when a "seed" particulate is deposited on the substrate or thin film layer during the multilayer coating process.

Beamlet Laser and the proposed NIF, are influenced by the damage threshold of the multilayer coatings used in the high-power optical components such as mirrors and polarizers. In multilayer dielectric coated mirrors, the predominant surface defects are micron-scale domes associated with the classic nodule defect. These defects are initiated at seed particles that are either present on the substrate, or that are deposited during the film deposition process (**Fig. 2**). Upon laser illumination, these defects give rise to local electric field enhancements which lead to "hot spots" in the coatings. These rapidly expanding regions create tensile stresses, which may result in mechanical failure of the coatings.

During FY-96, we performed electro-thermal-mechanical modeling of rotationally symmetric nodular defect geometries, assuming normal incidence illumination. We studied the influence of defect geometry (for example, nodule size, depth, and material composition) on the location and magnitude of the "hot spots" and thermal stresses.

Progress

Spatial Filter Lens

Material Dynamic Response. Our effort has focused on using DYNA to model the optic material response to a surface pressure or energy deposition loading calculated by LASNEX. This is not a routine modeling effort, due to the small time scale of the laser pulse (5 ns) and the magnitude of the pressure pulse (15 GPa). Additional calculations have been performed assuming an aluminum particle on the surface of the optic being irradiated by the laser.

A literature review was conducted to find information on the dynamic properties and failure modes of brittle materials, concentrating on SiO_2 and other forms of glass. For brittle materials the mechanical properties vary drastically from virgin material to damaged material. This necessitates a material model that accounts for rate-dependent cumulative damage and associated material property degradation. Many damage models exist for brittle materials, ranging from pure phenomenological models to ones that incorporate damage nucleation and the growth and coalescence of microcracks. At this point it is not clear which model would be best for our particular case. It may require combining certain features of existing models with new features to accurately predict SiO_2 damage based on the results of experiments.

If a metal particle on the front side of an optic is hit with a laser, a pressure pulse due to the laser light absorption on the surface of the particle will be

transmitted through the particle. Spallation of the particle may occur if the particle is not flush against the optics surface. The spallation is due to the crossing of two rarefaction waves, one from the front of the particle when the loading falls off, and the other generated from the rear surface when the incident shock reflects back into the material. If the tensile conditions are sufficient in magnitude and time application, they can lead to the formation of a spall plane. Spallation calculations were performed using a Tuler-Butcher time-dependent failure model to predict the amount of particle debris that may be deposited onto the surface of the SiO_2 from laser illumination on the front surface. **Figure 3** shows the peak pressure of the Gaussian pulse at various pulse widths required to initiate spallation. The pressure required to cause spallation at pulse widths less than 5 ns increases significantly, indicating the need for a time derivative failure model.

An Al particle on the front surface of the optic was modeled with DYNA2D. The Al particle was placed in 20° half-angle contact with a SiO_2 substrate. A 150-kbar, 4-ns Gaussian pressure pulse was applied to the surface of the sphere with a $\cos^2\theta$ pressure distribution. The Steinberg-Guinan constitutive law, Mie-Grüneisen equation of state and the Tuler-Butcher failure criteria for damage were used to model the behavior of the Al particle. The SiO_2 substrate was modeled with an elastic plastic constitutive model and a Mie-Grüneisen equation of state. Failure in the substrate was modeled with a simple spall criterion. Failure in the

substrate occurs when the maximum principle stress reaches a tensile limit of 6.5 kbar. When failure occurs, the deviatoric stresses are set to zero and the element can only handle compressive pressures. A complete brittle damage model is not used for the SiO_2 at this time, which should account for cumulative damage, time-dependent failure and compressive failure. **Figure 4** shows the cumulated damage in the Al particle and SiO_2 substrate after 50 ns. These calculations begin to address the association of surface contaminants with the initiation of damage sites on the SiO_2 optic.

Crack Growth Dynamics. After the damage starts, the cracks can grow until they reach a detectable size, at which time the optic component is removed from service before it can fail catastrophically. The cracks may grow by a number of mechanisms: (1) thermally-induced crack growth due to time varying stresses; (2) crack growth due to the propagation of stress waves from rapidly heated

Figure 3. *Maximum pressure of a Gaussian pulse required to cause spallation in a 1-D plane strain analysis of an Al bar. The pressure required to cause spallation at pulse widths less than 5 ns increases significantly, indicating the need for a time-derivative failure model.*
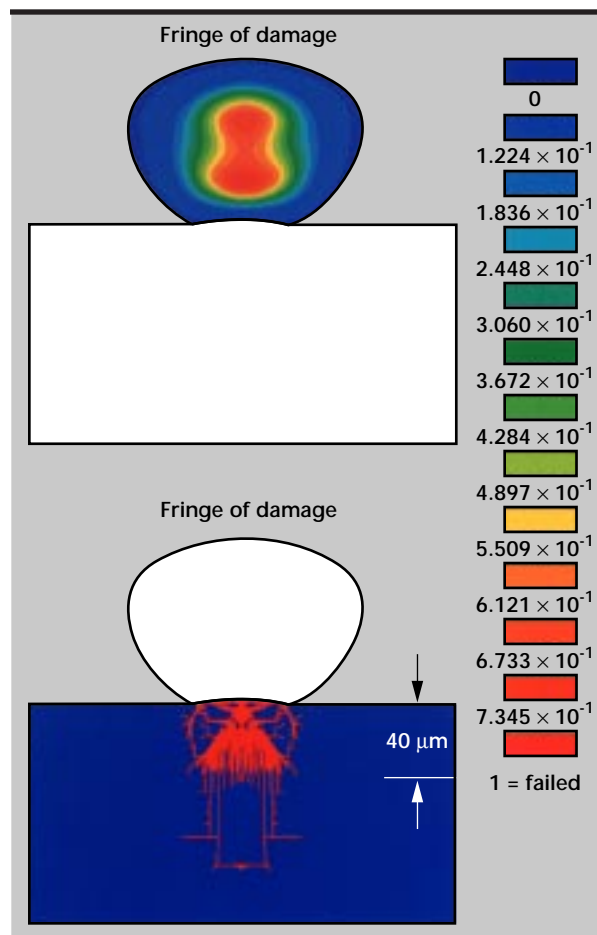
Figure 4. *Cumulative damage in an Al particle and SiO_2 substrate after 50 ns. Failure in the substrate, using a simple spall criterion, occurs when the maximum principle stress reaches a tensile limit of 6.5 kbar.*

defect sites near the surface; and (3) crack growth due to plasma formation in existing surface cracks. The damage from these mechanisms cause both optical and mechanical changes in performance.

For thermally-induced crack growth there are critical length scales relating the size of the absorbing defect, the crack size, and the distance between the hot spot and the crack. Beyond a certain separation of the hot spot and the crack there is not enough energy in the laser pulse to initiate crack growth. This envelope of thermal stress damage has started to be quantified by thermal cracking calculations.

Near the absorbing defect, the optical material experiences large temperature changes that induce large residual stresses and cracking upon cooldown. Thermally-induced crack growth in these regions must depend on the time- and temperature-dependent viscous response of the silica constitutive model. This constitutive model, and the analytical tools to measure the propensity for cracking under such conditions are in the process of development.

In this first series of analyses, the crack is 50 μm beneath the surface and 100 μm in length. The orientation of the crack is parallel to the surface. The laser energy flux of 1.85×10^9 W/m^2 is deposited over a spot size of 75 μm. **Figure 5a**
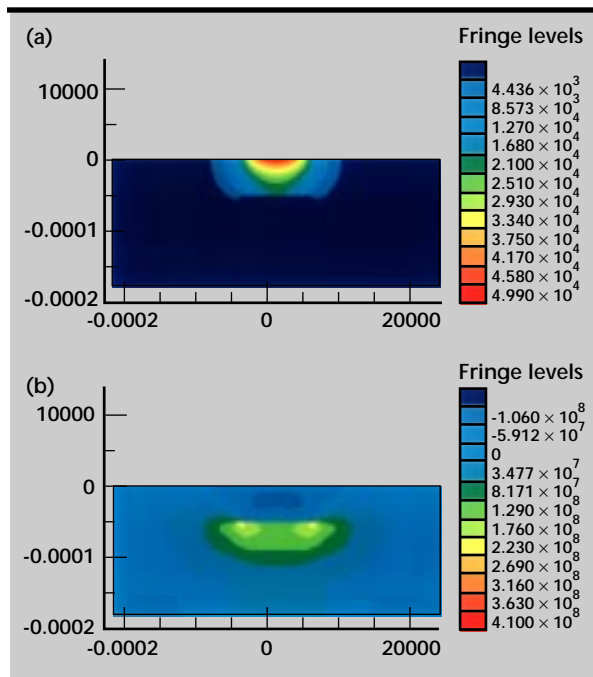
Figure 5. Temperature field (a) and stress state (b) near a subsurface crack. The crack acts as a barrier to thermal diffusion, forcing the heat flow around the ends of the crack, and thereby causing thermal gradients and singular stresses.

shows the temperature contours near the crack at 2 ms. The crack acts as a barrier to thermal diffusion, forcing the heat flow around the ends of the crack, causing thermal gradients and singular stresses. **Figure 5b** shows contours of the maximum principal stress concentrating around the crack tips with a pattern characteristic of mixed mode I and II loading. The energy available for crack advance at the crack tip as a result of these loading conditions is 14.0 J/m^2. This is a very small fraction of the energy absorbed at the surface $(3.7 \times 10^6$ J/m^2), but large enough to initiate cracking in the material based on adding the mode I and mode II stress intensity factors. The root mean square addition of the mode I and mode II stress intensity factors may not be the correct method of predicting mixed mode failure, but an experimentally verified criterion for failure of brittle materials under mixed mode loading is not available.

The importance of this flux level is that it represents a critical value for the onset of crack growth. The thermal stresses represent only one part of the total loading on cracks near a laser surface, but a part which may be particularly important combined with the presence of absorbing optical surface contamination.

Multilayer Dielectric Coated Mirrors

The objective of this task was to identify the precursor and characterize the catastrophic delamination of multilayer dielectric coated mirrors under very high laser fluences. We focused our attention on the nodule defect (**Fig. 2**). A nodule defect is created when a "seed" particulate, which is usually hafnia because the silica evaporates more continuously, is deposited on the substrate or a layer during the multilayer coating process. Subsequent layers then produce a bump in the silica overcoat. The distortion represented by the nodule causes non-uniformities in the electric field that lead to local hot spots.

During FY-96, we performed electro-thermal-mechanical modeling of rotationally symmetric nodular defect geometries, assuming normal incidence illumination.

The electromagnetic code AMOS was used to calculate the E-field enhancement due to the nodule defect. The E-field is enhanced by a factor of 2 to 4 in or near the defect. The thermal energy deposition scales as E^2. This creates a local hot spot causing compressive stresses within the region and tensile stresses around it. The heat transfer code TOPAZ2D was used to model the thermal effects due to the electromagnetic energy deposition and

NIKE2D was used to predict the resulting thermal stresses. We believe these tensile stresses are what caused the coating failure.

We looked at six nodule defect models representing a range of seed defect sizes and depths of burial (see **Fig. 2** and **Table 1**). **Table 1** gives the thermal and mechanical results for the six models. The third geometry gave the highest temperatures and stresses. However, after many computer simulations investigating thermal-stress levels caused by various size and location nodule defects, we have concluded that the stress levels are just too low to be concerned about as a failure mechanism. The analyses showed that geometry differences alone could not be used as a criterion for failure due to thermal stresses.

We also investigated models which included residual stresses and cracks from the manufacturing process. The above models assume that the initial state of the multilayers is stress-free. However, an intrinsic residual stress exists in the multilayers due to the coating process. We were able to calculate the residual stress for a single layer on a substrate, but could not "dial in" the intrinsic stresses for a multilayer. In fact, no one knows the appropriate intrinsic stresses for the multilayers of interest.

The next effort involved putting a 0.4-μm radius crack between the first hafnia and the second silica layers. We used the J-integral method in NIKE2D to calculate the stress intensity factor for delamination. The thermal gradients produced by the reduced conduction across the crack were insufficient by orders of magnitude to produce the critical stress intensity factor for delamination. Only pressure in the crack exceeding 100 MPa is capable of producing stress intensity factors large enough for crack growth. These porous multilayers readily take in water vapor from the air. If one assumes condensation to liquid water and subsequent heating, pressures of the right order of magnitude for crack growth are possible.

The calculated prediction of failure in sub-micron multilayer coatings is a very difficult problem. While the work reported here has not been successful at identifying failure mechanisms, it has at least highlighted some requirements for future modeling activities, including:

1) obtaining reliable material properties for thin films of silica and hafnia;
2) defining failure criteria for thin films and their interfaces, such as bonds;
3) developing methods for putting intrinsic stresses into the layers as initial conditions; and
4) developing an ability to model stress due to phase change of materials and trapped water.

Future Work

Our focus for FY-97 will be a continuation of the calculations investigating the spatial filter material response due to stress waves caused by the formation of surface plasma plumes, and calculations investigating crack growth and fracture. We are expecting NIF to continue funding the experimental effort and the LASNEX calculations in FY-97. Continuation of the LASNEX calculations are essential to our effort because they define the boundary condition to be used in DYNA.

The 1-D plane strain calculations supplied an understanding of the stress state in the SiO_2 due to pressure/thermal/energy deposition loading. However, the damage observed in the optics is of a 2-D nature. Two-dimensional and possible 3-D calculations are required to understand and adequately model the observed damage. The merging of the LASNEX and DYNA calculations is required to properly model the vapor-liquid-solid interfaces observed on the surface of the SiO_2 when modeling rear surface particles. An extension of the 2-D modeling will be the incorporation of material failure via one or a combination of several brittle

Table 1. Results from investigation of the six geometries in Fig. 2. Local hot spots resulting from the E-field enhancement due to the nodule defect cause tensile stresses around the defect. The thermal stresses alone are not sufficient to cause failure at NIF operating conditions.

Geometry	d	T	D/2	Seed Material	Tmax (K)	Stress (MPa) $\left[\sigma_{max}^{prin} - \sigma_{min}^{prin}\right]_{max}$
1	0.730	1.9723	1.6969	HfO_2	469.6	25.1
2	0.250	1.9723	0.9931	HfO_2	470.4	36.4
3	1.500	1.9723	2.2205	HfO_2	1000.5	136.0
4	.730	1.0087	1.2135	HfO_2	636.4	58.3
5	.730	2.9359	2.0704	HfO_2	408.3	25.7
6	.730	1.9723	1.6969	SiO_2	576.8	34.9

damage models. Failure in this sense is termed bulk material failure, including microcracking, versus explicitly modeling a crack and calculating propagation rates, which will be addressed separately.

A sizable fraction of failures are associated with crack growth. Crack extension from an existing crack can be predicted based on a critical energy release rate (a J-integral value), which in turn is related to a critical stress intensity factor. A J-integral capability has been implemented for use with NIKE2D and TOPAZ2D. This capability will be implemented in the 3-D codes, NIKE3D and TOPAZ3D. Additionally, the J-integral capabilities will be implemented in the DYNA codes to model crack propagation resulting from dynamic mechanical events. The J-integral capability will allow us to predict dynamic, static, and thermal loads which would cause crack growth and failure. The J-integral predictive capability will be validated with the NIF-funded experimental effort.

With a better understanding of the physics behind the observed damage, one hopes to be able to develop a predictive capability which can be used to establish critical flaw sizes to prevent catastrophic optics failure. The multidisciplinary team represents a technically potent attempt to solve the laser damage issues, with every facet of the problem represented in the integrated team.

New Rigid Body Features in DYNA3D

Jerry I. Lin
Defense Technologies Engineering Division
Mechanical Engineering

Anthony S. Lee
New Technologies Engineering Division
Mechanical Engineering

We have implemented two new rigid body features in the explicit finite element code DYNA3D. The first is the modified Rigid Body Joints that provides a tighter bond between rigid bodies with reasonable integration time increments. The second feature is Material Switching, a capability that allows groups of materials to be changed from deformable to rigid then back at user-specified times.

Introduction

As DYNA3D is increasingly applied to analyzing mechanical systems and impact simulations, adding new functionalities involving rigid body mechanics/kinematics becomes more important. One such example is the capability to model joints, such as revolute, universal or spherical joints in complicated machinery or suspension systems. The Rigid Body Joints feature in DYNA3D[1] did provide the necessary modeling tool. However, ideal bonding conditions could only be achieved at the expense of requiring very small integration time steps. This deficiency occurs because the penalty stiffness between adjoining rigid bodies was considered a function of the time step.

A new algorithm based on a rigid body mechanics approach is implemented to overcome this shortcoming. The new method is always stable with the time step, limited by the Courant stability condition, and the bonding force between two adjoining rigid bodies is calculated according to their masses and current distance.

In some impact applications of DYNA3D, "freezing" the entire model or part of the model as a rigid body during a specific period of time appears to have no adverse effect on the overall simulation, but does offer a significant economic advantage. One example is the simulation of multiple ground impacts of an object during a drop test. Depending on the model characteristics, such as the material properties and model geometry, and the kinematic parameters, such as the drop height, angle and initial velocity, the oscillatory deformation occurring between impacts may be secondary compared to the gross deformation inflicted by ground impacts. Thus, treating the whole or part of the model as a rigid body between impacts could save considerable CPU time.

The other example is a vehicle impacting a roadside safety structure such as a highway guardrail at high speed. In this case, the post-impact vehicle trajectory is of particular interest because of its likely interference with the traffic flow or rollover. Since the overall vehicle motion is far more important than the component deformation in the vehicle redirection stage, it is desirable to have the finite element mesh treated like a rigid body. The new Material Switching feature in DYNA3D allows individual components (material groups) to be switched back and forth between any deformable and rigid material types at user-specified times.

Progress

Rigid Body Joint Algorithm

There are six types of joints, namely spherical, revolute, cylindrical, planar, universal and translational, currently available in DYNA3D. These joints are defined by 1 to 3 pairs of nodes, depending on

the particular relative movement between adjoining rigid bodies permitted by each joint type. Each pair of nodes consists of a node from each adjoining body, and a force between this pair of nodes is calculated every time step to prevent the pair from separating. The direction of the forces imposes constraints on the relative motion between the pairs of nodes, which in turn creates the desirable motion characteristics for each joint type, whereas the magnitude of the force determines whether the adjoining bodies are tied together.

In the original joint algorithm, with maintaining the global stability in mind, the force magnitude is determined by the time step, and reducing the time step is the only means to increase the stiffness associated with the penalty spring. The new algorithm instead takes both the time step and masses of the adjoining rigid bodies into consideration to achieve the desired stability.

Without the presence of a bonding force, a gap between a pair of nodes that defines the Rigid Body Joint will be created during a time step Δt. To close this gap, a penalty spring that will impose sufficient force needs to be added. Let us consider two isolated rigid bodies of mass m_1 and m_2, respectively. The distances d_1 and d_2 these rigid bodies would be moved over a time span of Δt by a force f can be expressed as

$$d_1 = \iint \frac{f}{m_1} dt^2 = \frac{1}{2} \frac{f}{m_1} \Delta t^2 \tag{1}$$

and

$$d_2 = \iint \frac{f}{m_2} dt^2 = \frac{1}{2} \frac{f}{m_2} \Delta t^2. \tag{2}$$

To find the force necessary to close the gap, d, between a pair of nodes defining the joint, we combine **Eq. 1** and **Eq. 2** to yield

$$d = d_1 + d_2 = \frac{f\Delta t^2}{2} \left(\frac{1}{m_1} + \frac{1}{m_2} \right), \tag{3}$$

and

$$f = \frac{2m_1 m_2}{(m_1 + m_2)\Delta t^2} (d_1 + d_2). \tag{4}$$

The penalty spring stiffness, K_p, can then be obtained by making $d_1 + d_2 = 1$ in **Eq. 4**:

$$K_p = \frac{2m_1 m_2}{(m_1 + m_2)\Delta t^2}. \tag{5}$$

If the rigid bodies are completely isolated from the rest of the model, K_p will tie them perfectly. However, that is not the case for most problems. A scale factor that provides users with provision to tune the penalty forces is made available in case forces of greater or less magnitude are needed.

A vehicle suspension system is used to demonstrate the effectiveness of this new algorithm. **Figure 1** shows the initial stage of the suspension system going up a ramp. Two simulations, one using the original joint algorithm, the other the new joint algorithm, are made, and the results are shown in **Fig. 2** and **Fig. 3**, respectively. Several parts are modeled as rigid bodies and connected by the Rigid Body Joints. One of the critical joints is the revolute joint, placed between the wheel hub and the spindle, through which the loads are transmitted between the vehicle frame and the wheels. With the backing plate removed, the disjoining of the wheel hub and the spindle can be clearly seen in **Fig. 2**. This is a direct consequence of insufficient penalty force to hold the joint in place. The time step needs to be reduced to a small fraction of the calculated time step to generate enough forces to keep the parts together. As shown in **Fig. 3**, when the new algorithm is used with a

Figure 1. A vehicle suspension system going up a ramp.

Backing plate

Lower control arm
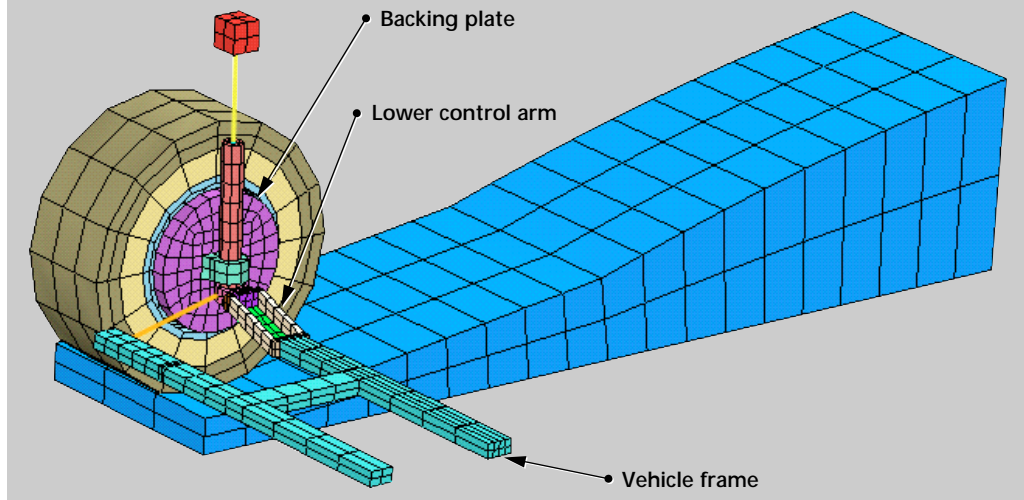
Vehicle frame

default time step, the wheel hub and the spindle remained connected, and the motion of the vehicle frame dictates a realistic wheel position.

Time Stability

From a stability standpoint, the new Rigid Body Joint can be simplified as two masses, m_1 and m_2, connected by a spring with spring constant K_p. The frequency of such a system, ω, can be expressed as

$$\omega = \sqrt{\frac{(m_1 + m_2)K_p}{m_1 m_2}},$$

$\qquad$ (6)

and the stability limit is in turn defined by

$$\Delta t_c = \frac{2}{\omega} = \sqrt{\frac{4 m_1 m_2}{(m_1 + m_2)K_p}}.$$

$\qquad$ (7)

Substituting **Eq. 5** into **Eq. 7** yields

$$\Delta t_c = \sqrt{2}\ \Delta t$$

$\qquad$ (8)

which indicates that the current time step Δt is always within the stability limit imposed by the penalty spring.

Deformable/Rigid Material Switching

This feature lets users designate lists of materials as material groups. Upon activation of the Deformable/Rigid Material Switching feature, DYNA3D collects the materials in a group and merges them to form a rigid body for the subsequent calculation. When off-time is reached, this rigid body is disassembled and the constituents are returned to their original material types. The boundary conditions for the collective rigid body are determined by superimposing the boundary conditions on each individual constituent. During the rigid phase, all elements are "frozen," that is, the element

quantities, such as stresses, strains, strain energy, and other history-dependent variables, are kept unchanged, whereas the nodal quantities, such as velocities and displacements, are updated by rigid body mechanics/kinematics.

This approach perfectly represents the element states for structural elements since stresses and strains are measured with respect to the element coordinate system. For continuum elements, because their stresses and strains are expressed in terms of global Cartesian coordinates, keeping them constant during rigid body motion could be incorrect if finite rotation is involved.

To maintain the stresses and strains with respect to the right reference frame for the continuum elements at all times, second-order tensor transformations need to be done at every time step. Making such a time consuming effort for bookkeeping purposes during the rigid phase is uneconomical and does not affect the mechanics in any way. Instead, the code calculates the rigid body's principal axes by solving an eigenvalue problem at the activation time, and incrementally updates them to account for the motion the rigid body is undergoing. The continuum element stresses and strains are then transformed to the right configuration according to the accumulative rotation of these principal axes at the end of the rigid phase.

The economical advantages of this feature are three-fold:
(1) the burdensome element calculations are bypassed during the rigid phase;
(2) through careful selection, contact checks can be reduced to a minimum during the rigid phase; and
(3) greater time steps could be used during the rigid phase since accuracy, not stability, is the deciding factor in selecting a time step for rigid materials.
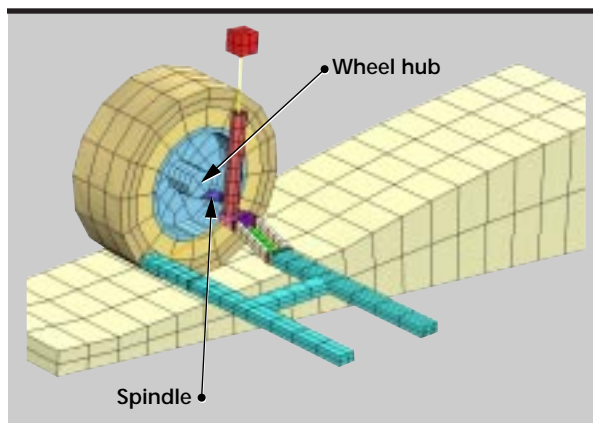
Figure 2. Final configuration, with backing plate removed, using the original Rigid Body Joints.
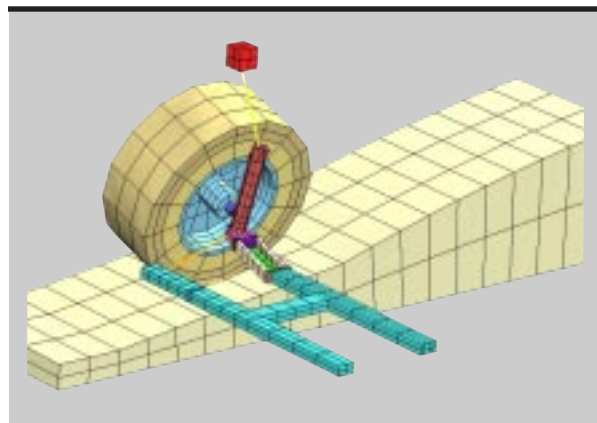
Figure 3. Final configuration, with backing plate removed, using new Rigid Body Joints.

The applications of this feature identified to date include multiple-impact simulation, post-impact trajectory analysis, post-event visualization and partial stress initialization of large models. The simulation of a cylinder with initial velocity impacting into a rail is chosen as an example for this feature. Because of double symmetry, only a quarter of the cylinder is modeled. Though the actual impact between the cylinder and the rail is through the Rigid Wall feature in DYNA3D, a block representing the rail is added for demonstration purposes. **Figure 4** shows the finite element mesh for this model.

Additional details for this model can be found in the DYNA3D Example Problem Manual.[2] Two runs were made, with the Material Switching feature turned off in run 1 but turned on in run 2. The total simulation time is 15 ms, and the entire model was changed to a rigid body at the half-way point in the second run. As expected, the CPU time for the second run is about 50% of that for the first run. The final deformation, kinetic energy history and momentum history in the impact direction are all almost identical for the two runs, which justifies the use of the Material Switching feature. The final deformation for run 2 is shown in **Fig. 5**.

We also recorded the displacements at node 205, a point at the center of the cylinder and in contact with the rail at the beginning, for both runs in **Fig. 6** to monitor the distance between the cylinder and the rail. It is of interest because it provides estimates of when the subsequent impacts would take place. An agreement with run 1 is essential for run 2 in successfully simulating the subsequent impacts.

When the switching feature is used, the cylinder needs to be returned to its original material types before the next impact takes place. Although the results appear promising, the study is not complete. Because of the high initial velocity, the simulation time is not long enough to let gravity bring the cylinder down for a second impact. A lower initial velocity will be used in our future studies of this problem.

Future Work

Rigid Body Joints

We are undertaking the task of making Breakable Rigid Body Joints. The Federal Highways Administration (FHWA) has requested this capability

Figure 4. Finite element mesh for a cylinder hitting a rail.
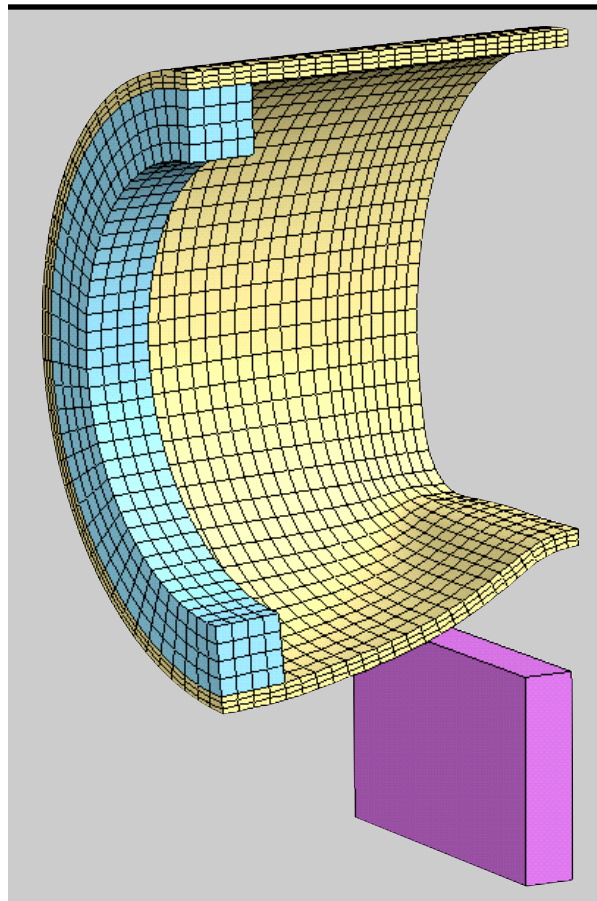
Figure 5. Final deformation of a cylinder hitting a rail.
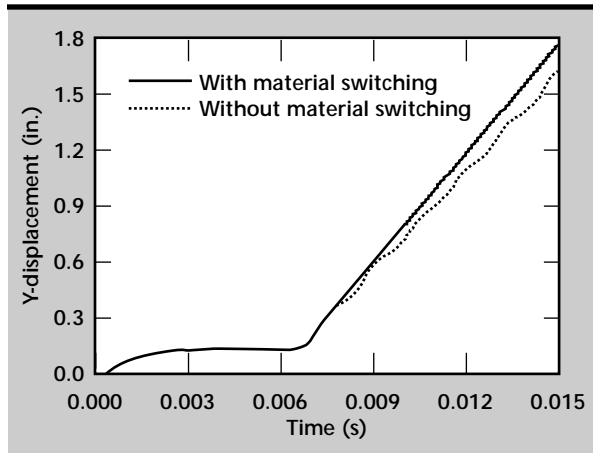
Figure 6. Distance between the center of the cylinder and the rail.

to simulate the disintegration of the vehicle parts in the event of a severe impact. Possible failure criteria for the joints are relative velocity or acceleration between the two adjoining parts.

Deformable/Rigid Material Switching

At this time, contact check can be completely eliminated when the entire model is switched to rigid. For more general cases involving partial model switching, we need to design and implement algorithms to identify the contact interactions between the converted rigid parts during the rigid

phase. One possible way is cross-checking contact definitions and switching material groups at the initialization phase and forgoing a specific contact definition only if all entities involved are switched to rigid.

In addition to the cylinder impact problem, more quantitative studies are scheduled. Boeing Commercial Aircraft is using this feature along with the Dynamic Relaxation feature in DYNA3D to add initial stress in a large scale model. With the possible interest from FHWA, a car/guardrail angle impact simulation model is being prepared. This ability to predict the post-impact vehicle motion and trajectory will be the focal points for the latter case. This feature can also be used to add the post-impact motion of the container to visual display.

References

1. Whirley, R. G., and B. E. Engelmann (1993), *DYNA3D: A Nonlinear, Explicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics - User Manual,* University of California, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-107254).

2. Lovejoy, S. C., and R. G. Whirley (1990), *DYNA3D Example Problem Manual,* University of California, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-105259).

Implicit Rigid Body Static and Dynamic Analysis with NIKE3D

Michael A. Puso
New Technologies Engineering Division
Mechanical Engineering

In this work, special treatment is given to the rigid body formulation for implicit structural analysis. The inertial contribution to the coupled deformable-rigid body equations is first formulated and implemented in the implicit finite element code NIKE3D. Special treatment is then given to the non-symmetric tangent matrix that results from the linearization of these equations. For statics problems, the coupling of the rigid body to the flexible finite elements can cause a large bandwidth. A special equation-reordering technique is used to alleviate this problem.

Introduction

In finite element analysis, the rigid body idealization can be exploited for many different problems, both static and dynamic. The die in a metal-forming problem can often be treated as rigid. This would usually be a statics application. In dynamics, the rigid body assumption can be applied to the stiff sub-systems of flexible structures when only the low frequency response is desired. Examples of this are vibration analysis of transportation containers, suspension systems and space structures.

The rigid body formulation in an implicit finite element setting is different in a number of ways from that in an explicit setting, as would be expected. For example, in an explicit code, attention is paid to the control of the critical time step when penalty methods are used to attach rigid bodies. In an implicit code, the penalty method does not impose such a time step restriction. In an explicit formulation, solution of the equilibrium equations is much simpler than that in the implicit formulation. The solution of the non-linear equations is the primary issue in the implicit formulation.

In this work, the equations of motion for coupled rigid-deformable finite elements are formulated and methods of solution are developed to solve the problems efficiently.

Progress

The rigid body material model has been available in NIKE3D for some time for statics problems only. In this implementation, a part is made rigid by specifying its material to be type 20. The kinematics of this rigid part are then described by three translational and three rotational degrees of freedom. Integration of the incremental rotations as provided by the non-linear solution scheme is made by use of the exponential map.

We used a version of Newmark's method for the integration of the rotation rates in the body frame. In formulating the inertial contribution of the rigid bodies to the finite element equations, it is recognized that the inertial tangent matrix is non-symmetric. This presents a major hurdle in the implementation, since full scale non-symmetric linear solution techniques are far too costly. No valid symmetrized approximation to the tangent matrix is apparent. Furthermore, it is not clear whether the symmetric stiffness updates made by the BFGS method to a symmetrized approximation of the stiffness matrix can be successful in handling the problem. An incomplete tangent matrix could diminish some of the savings rigid bodies offer by causing additional equilibrium iterations.

We have performed a special decomposition of the global stiffness matrix into symmetric and non-symmetric parts, so that independent factorization is made of the two parts by a symmetric and non-symmetric solver. The method is slightly more expensive (by about 10%) than the symmetric solver, due to additional matrix multiplications, but is worth the effort for the dynamics problems.

The rigid body idealization in implicit finite element analysis is attractive, due to the reduction in the number of degrees of freedom in the problem that results. In explicit codes such as DYNA3D, the main economical advantage is derived from the reduction in constitutive evaluations. In an implicit

code the reduction of constitutive evaluations is helpful, but the main advantage of rigid parts is derived from their potential to reduce the size of the stiffness matrix and the time to factorize it.

Although rigid materials will reduce the number of degrees of freedom of a large part with thousands of degrees of freedom to six, it often causes an extremely large amount of coupling between non-adjacent deformable elements that are attached to the rigid body. At times this causes the rigid bodies to make the problem less efficient. This increased bandwidth is due to the inability of the bandwidth minimizer to handle the large amount of coupling caused by the rigid part. A multiple minimum-degree reorder technique with a sparse direct solver is used in lieu of the bandwidth minimization/skyline solver scheme. It is shown to be unaffected by the large amount of coupling and recovers the efficiency provided by rigid bodies.

Rigid Body Dynamics Formulation

In the coupled deformable-rigid problem, degrees of freedom x are segregated into those that lie within the flexible mesh x_f and those that lie on the interface of the rigid body x_r. Furthermore, rigid interface degrees of freedom can be computed from the current center of rotation of the rigid body X_o and the current rigid body position vector R such that:

$$x = \begin{Bmatrix} x_f \\ x_r \end{Bmatrix} = \begin{Bmatrix} x_f \\ X_o + R \end{Bmatrix} \tag{1}$$

The finite element nodal forces can be treated similarly:

$$f = \begin{Bmatrix} f_f \\ f_r \end{Bmatrix} = \begin{Bmatrix} f_f^I + f_f^\sigma \\ f_r^I + f_r^\sigma \end{Bmatrix} \tag{2}$$

where f^I is the inertial contribution. To get the equations of motion using the generalized coordinates, the internal virtual work is computed:

$$f \cdot \delta x = f \cdot \begin{Bmatrix} \delta x_f \\ \delta X_o + R \times \delta \theta \end{Bmatrix}$$
$$= f_f \cdot \delta x_f + f_r \cdot \delta X_o + M_r \cdot \delta \theta \tag{3}$$

where θ is the incremental rotation and $M_r = f_r \times R$ is the moment. The forces and moments, f_r and M_r, at the rigid body interface are in equilibrium with the body forces and moments, f_o and M_o, applied at the center of rotation and the rigid body inertia such that:

$$m\ddot{X}_o - m(\dot{\omega} \times d + \omega \times \omega \times d) = f_r + f_o$$
$$-d \times m\ddot{X}_o + I\dot{\omega} + \omega \times I\omega = M_r + M_o \tag{4}$$

where m is the mass, ω is the rotational velocity, I is the tensor of moments of inertia, and d is the distance from the center of rotation to the center of mass of the rigid body.

Currently in NIKE3D, Newmark's method[1] is used to integrate the finite element accelerations. To maintain consistency with the flexible portion of the mesh, Newmark's method is chosen over midstep formulations[2] to integrate the rigid body transla-tional and rotational accelerations. This version integrates the body frame rotational velocity W and acceleration A. The body frame coordinates are rotated to the inertial frame via the rotation matrix Λ. By this transformation $\omega = \Lambda W$. The rotational inertia terms in **Eq. 4** are transformed to the body coordinates to accommodate the integration scheme, yielding:

$$m\ddot{X}_o - m\Lambda(A \times d^o + W \times W \times d^o) = f_r + f_o$$
$$-\Lambda d^o \times m\ddot{X}_o + \Lambda(JA + W \times JW) = M_r + M_o \tag{5}$$

The following is Newmark's method modified for integration of the rotation rates:

$$\Lambda_{n+1} = \Lambda_n \exp(\Theta_n)$$
$$\Theta_n = \Delta t\, W_n + (\Delta t)^2\left[\left(\frac{1}{2} - \beta\right)A_n + \beta A_{n+1}\right]$$
$$W_{n+1} = W_n + (\Delta t)\left[\left(\frac{1}{2} - \gamma\right)A_n + \gamma A_{n+1}\right] \tag{6}$$

Where exp $(\cdot)$ is the exponential map operator. The algorithm given by **Eq. 6** is substituted into **Eq. 5**, and then linearized to get the rigid body contribution of the tangent matrix k_{rr}. As it turns out, the inertial part of this tangent matrix is non-symmetric even when the center of rotation is at the center of mass.

We developed a method to minimize the work in solving the non-symmetric system. In view of **Eqs. 1** to **3**, the linear system of equations can be segregated as follows:

$$\begin{bmatrix} k_{ff} & k_{f\tilde{r}} \\ k_{f\tilde{r}}^T & k_{\tilde{r}\tilde{r}} \end{bmatrix}\begin{Bmatrix} \Delta x_f \\ \Delta x_{\tilde{r}} \end{Bmatrix} = \begin{Bmatrix} f_f \\ f_{\tilde{r}} \end{Bmatrix} \tag{7}$$

where

$$\Delta x_{\tilde{r}} = \begin{Bmatrix} \Delta X_o \\ \Delta \Theta \end{Bmatrix}, \quad f_{\tilde{r}} = \begin{Bmatrix} f_o \\ M_o \end{Bmatrix}.$$

Matrix condensation is performed on **Eq. 7** to yield the uncoupled equations:

$$\left[k_{\tilde{r}\tilde{r}} - k_{f\tilde{r}}^T k_{ff}^{-1} k_{f\tilde{r}}\right]\{\Delta x_{\tilde{r}}\} = \{f_{\tilde{r}}\} - k_{f\tilde{r}}^T k_{ff}^{-1} f_f \tag{8}$$

$$\left[\boldsymbol{k}_{ff} \right] \left\{ \Delta \boldsymbol{x}_f \right\} = \boldsymbol{f}_f - \left[\boldsymbol{k}_{\tilde{r}f} \right] \left\{ \Delta \boldsymbol{x}_{\tilde{r}} \right\} . \qquad (9)$$

In this method, the rigid body degrees of freedom are solved first from **Eq. 8** with a non-symmetric solver and substituted into **Eq. 9**. The inverse of $\boldsymbol{k}_{ff}$ never needs to be calculated explicitly. Instead, $\boldsymbol{k}_{ff}$ is factorized and then successive back solves are made to form the Schur complement matrix and right hand side in **Eq. 8**. Typically, $\boldsymbol{k}_{\tilde{r}}$ is small (especially when compared to $\boldsymbol{k}_{ff}$) such that **Eq. 8** and **Eq. 9** take roughly 10% to 15% longer than the similar size symmetric linear solve for typical problems.

Rigid Body Dynamics Implementation

Bricks, shells, beams and discrete elements can be treated as rigid by specifying the element material type to be 20 in the NIKE3D[3] input deck. NIKE3D calculates the moments of inertia, and by default uses the center of mass as the center of rotation. Alternate moments of inertia and centers of rotation along with rigid body initial velocities can be specified on the material cards. Nodes and facets can also be defined as rigid by specifying the nodal list on the appropriate cards. Rigid bodies can be fastened together by use of connecting penalty springs.

Rigid Body Dynamics Applications

A simple top (**Fig. 1**) is given an initial rotation of 1 rad/s about its axis of symmetry 3(I_3 = 4.0) and along the 1 axis (I_1 = 5.22), and allowed to rotate freely. Since this is torqueless motion, the angular momentum must remain constant. Due to the axis of symmetry, the top will precess about some preferred axes. Solution of the Euler equations shows that the body frame angular velocity vector will rotate about the symmetry axes at a uniform rate, with a period given by

$$T = \frac{2\pi I_1}{\left(I_1 - I_3 \right) \omega_3^o} = 26.85 \; s \; . \qquad (10)$$

The angular velocities computed by NIKE3D, along with the exact answers, are plotted in **Fig. 2**. The rates ω_1 and ω_2 predicted by NIKE3D are nearly exact, while ω_3 is exact. A more complicated motion results when the top is fixed about its pointy end and is given an initial rotational ω_3^0 = 1000 rad/s with a gravitational force acting through its center of mass, causing a moment about its center of rotation. **Figure 3** shows the x displacement of the center of mass as the top rotates. The high frequency response in **Fig. 3** is due to the top's precession, while the low frequency response is due to nutation.

A potential application for rigid body dynamics is the vibrational response of a transportation container. The FL transportation container (**Fig. 4**) has concentric capped cylinders that are separated by foam. The payload rests on plates welded to the inner cylinder. During transport, the outer canister is tied down by guy wires and is relatively rigid. The inner cylinder is also relatively rigid compared to the foam. The low frequency response is dominated by the behavior of the foam. It is then reasonable to make the two cylinders and the payload rigid.

Figure 1. Top shown in body coordinate system 123 and inertial coordinate system xyz.
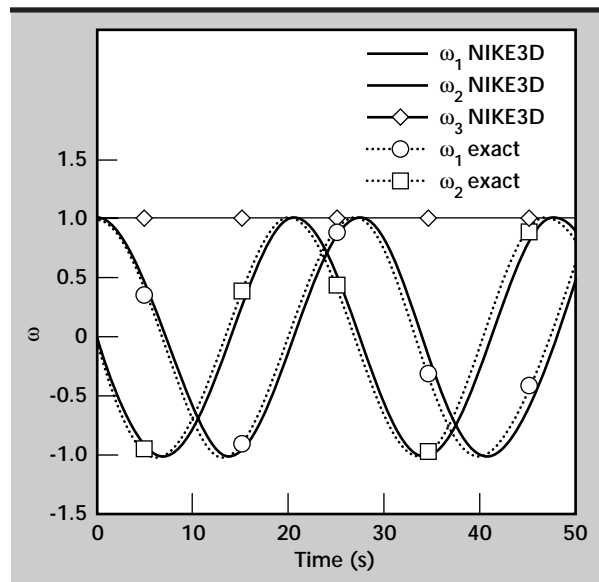
Figure 2. Body frame rotational velocities for precessing top.

A vibrational base acceleration (**Fig. 5**) was applied to the cylinder and the response was modeled with NIKE3D, with all flexible elements as one case, and the cylinders and payload rigid as another case.

This amounts to making 6260 of the 7032 shells rigid. The only shells that are flexible are the plates supporting the payload. The horizontal displacement of the payload is shown for both cases in **Fig. 6**. The fully

Figure 3. Plot of x-displacement of the center of mass for top spinning under the influence of gravity in the z direction.
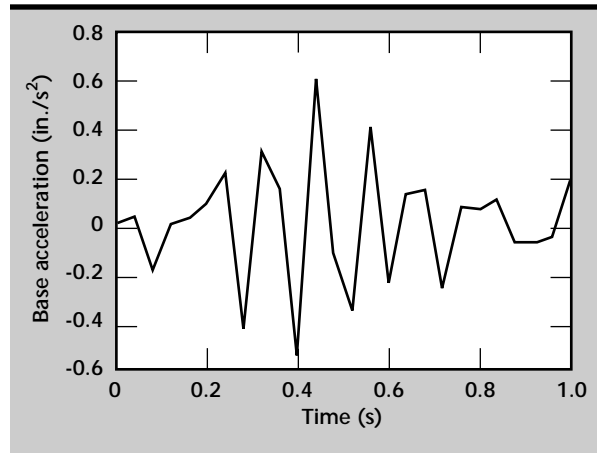
Figure 5. Base acceleration input for NIKE3D model of FL container.
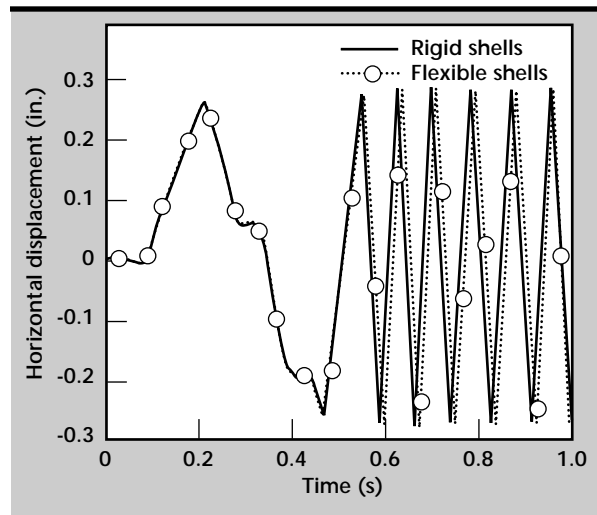
Figure 6. Horizontal displacement of payload in FL container due to base acceleration.

Figure 4. FL shipping container.
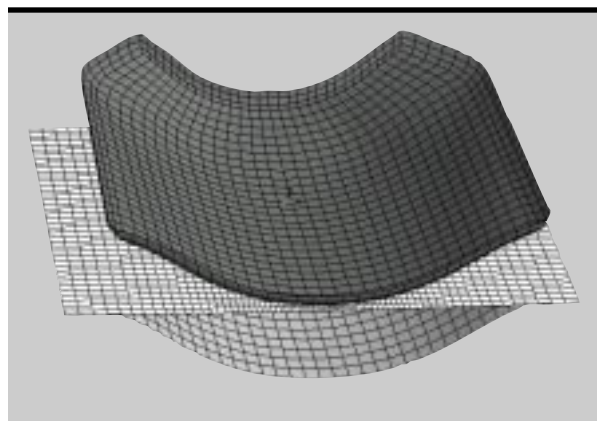
Figure 7. Sheet metal stamping.

Table 1. Results from NIKE3D implementation.

	Avg. equation bandwidth	Avg. # of words in stiffness matrix	Avg. time per factorization	Total run time
Case 1. GPS	260	5.6 million	22.0 s	59 min
Case 2. GPS	330	6.6 million	36.0 s	62 min
Case 3. MMD	N/A	1.8 million	6.1 s	29 min

flexible model on the DEC Alpha 2100 took about 31 h with 0.001-s time steps to model the problem out to 1 s. The rigid model took 18 h to complete.

Equation Reordering

As mentioned, using rigid bodies will reduce the number of degrees of freedom but can cause the equation bandwidth to increase drastically, since all the nodes attached to the rigid body are now coupled. Furthermore, GPS[4] bandwidth minimization is shown to be ineffective in minimizing the bandwidth with such large amounts of coupling. For statics problems the tangent stiffness matrix is symmetric so the decompositions given by **Eqs. 8** and **9** are not necessary. On the other hand, such a decomposition could reduce the large bandwidth. Instead of doing such a factorization, we used a different reordering scheme, called the multiple minimum degree, or MMD.[5] In the GPS scheme all the terms in the stiffness matrix are reordered in an attempt to push them toward the diagonal to minimize the matrix profile, and thus the storage. With the MMD scheme, reordering is done to minimize the fill in the factorized stiffness matrix. The MMD method tends to push off diagonal terms to the corners of the stiffness matrix.

Implementation and Application of MMD Reordering

The MMD method is used with the sparse direct solver and is specified by selecting option #3 for solver type.

An example is given where rigid bodies do not provide any efficiency using GPS reordering. The mesh shown in **Fig. 7** is a NIKE3D model of a sheet metal stamping. The upper die is displaced downward, stamping the sheet into the lower die. Both the upper and lower die can be considered rigid. Three different cases were run with NIKE3D: 1) all flexible elements; 2) both dies rigid with GPS reordering; and 3) both dies rigid with MMD reordering. The results are shown in **Table 1**.

In Case 1, all the degrees of freedom are fixed on the bottom die, forcing it to be rigid.

As seen from **Table 1**, the rigid body dies do not provide any savings over an all-flexible model. This is due to the top die's not being fixed and free to displace. Its degrees of freedom are therefore coupled with every node on the face of sheet, due to contact. The bottom die is fixed, representing an essential boundary condition whose degrees of freedom don't appear in the stiffness matrix and therefore don't increase the bandwidth. In an analysis where there are only fixed dies, this problem doesn't appear.

Future Work

Currently, only faceted contact is performed with rigid bodies. Contact that exploits the analytical surfaces typical of rigid bodies would provide smooth contact surface to better represent rolling cylinders or other types of rigid body contact. These smooth surfaces would also facilitate convergence of the non-linear equation solving. In addition, rigid bodies can only be connected using fle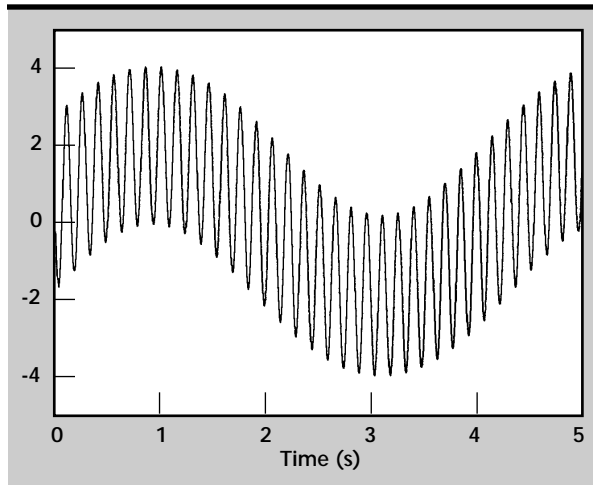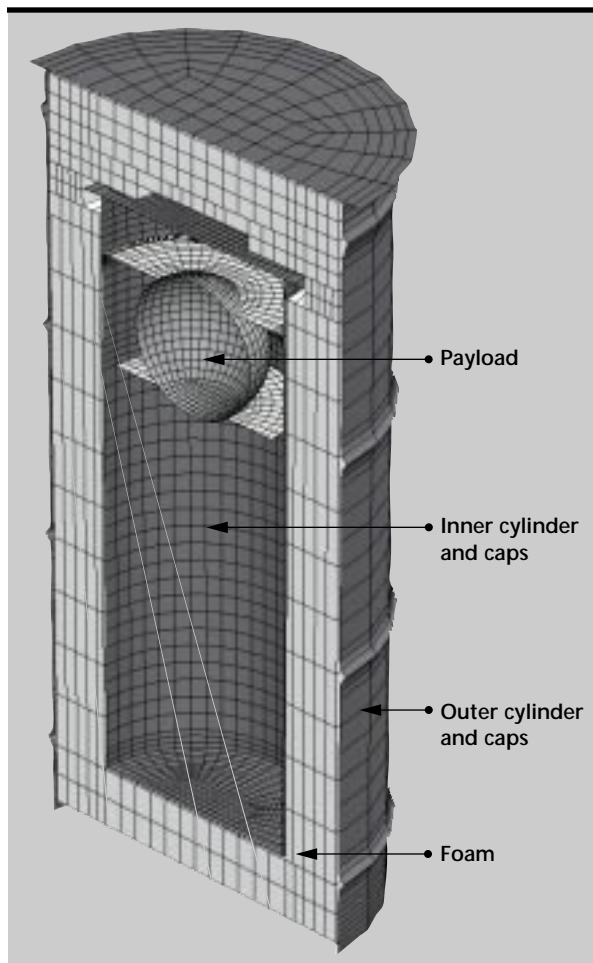xible elements such as penalty springs. The Lagrange multiplier method would provide a better solution over the penalty spring method to couple rigid bodies.

References

1. Simo, J. C, and L. Vu Quoc (1988), "On the dynamics of space rods undergoing large motions—a geometrically exact approach," *CMAME*, **66**.

2. Simo, J. C., and K. K. Wong (1991), "Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and moment," *IJNME*, **31**.

3. Maker, B. N. (1995), *NIKE3D: A nonlinear, implicit, three dimensional finite element code for solid and structural mechanics—users manual, Rev. 1*, University of California, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-MA-105268).

4. Gibbs, N. E., W. G. Poole Jr., and P. K. Stockmeyer (1976), "An algorithm for reducing the bandwidth and profile of a sparse matrix," *SIAM J. Num. Anal.*, **13**.

5. George, A., and J. H. Liu (1981), *Computer Solutions of Large Sparse Positive Definite Systems*, Prentice Hall, Englewood Cliffs, N.J.

Thermal Evaluation of Pantex Pit Storage

Barbara T. Kornblum
New Technologies Engineering Division
Mechanical Engineering

Salvador M. Aceves
Applied Research Engineering
Mechanical Engineering

We are modeling the heat transfer and air flow around weapon pit storage containers in Pantex magazines to predict container temperatures. This thermal-fluid mechanics problem presents a significant computational challenge because it involves transient, 3-D natural convection and thermal radiation around a large number of interacting containers with various heat generation rates. We synthesized a modeling procedure for arrays of pit storage containers, and applied the procedure to simulate a recent Lawrence Livermore National Laboratory (LLNL) experiment involving a group of 20 containers. Our calculated container temperatures compare well with data from the experiment, thus validating the modeling procedure. We performed parameter studies to determine the sensitivity of the convection and radiation heat transfer modes, and also demonstrated how the modeling procedure can be applied to an entire Pantex magazine holding 252 pit storage containers.

Introduction

This report investigates potential pit storage configurations that could be used at the Mason and Hanger Pantex Plant. The study uses data from a thermal test series performed at LLNL that simulated these storage configurations. The heat output values used in the LLNL test series do not represent actual pits, but are rounded numbers that were chosen for convenience to allow parameter excursions.

Pits from dismantled nuclear weapons are currently stored in magazines at the Pantex plant in Amarillo, Texas. The pits generate heat due to natural radioactive decay, and therefore may overheat if not cooled properly. Pits are carefully packaged for both structural integrity and fire safety inside cylindrical containers that are stacked closely together in columns inside the magazines. However, the internal container packaging and the compact storage arrangement also thermally insulates the pits, making the problem of cooling more difficult. It is important to keep weapon pits from overheating for safety and reliability.[1]

The Pantex pit storage magazines are passively cooled by natural convection and thermal radiation, subjecting them to the diurnal cycle as well as seasonal temperature variations. Heat generated inside the magazine can exceed several kilowatts, raising the internal magazine temperature well above the ambient outdoor temperature. Natural convection cooling occurs because a draft is created by the heat generated inside the magazine, pulling air in through intake vents near the floor while exhausting air out through a ventilation stack in the roof. Thermal radiative exchange occurs between the containers and the surroundings. A full understanding of the thermal transport characteristics of the magazine is desirable to predict pit temperatures under a variety of conditions.

The goal of this work is to develop a robust modeling procedure for analyzing large arrays of pit storage containers. Full-scale discretization of an entire magazine is not possible, due to the complicated geometry of the array of containers, the complexity of the air flow, and the physical limitations of today's computers. In the past there have been both computational and experimental efforts to

analyze Pantex storage magazines,[1,2,3] but they did not demonstrate the ability to extrapolate away from the known data envelope.

Here we present a different approach that links together two separate computational methods: a detailed fluid mechanics analysis on a few containers, followed by a lumped-parameter thermal model of an array of containers. With this technique, we are predicting container surface temperatures based on the rate of heat generation inside the container. Detailed documentation of our work is presented in Reference 4. We validated our modeling approach by comparing predicted temperatures with data from a recent experiment (performed at LLNL) in which container temperatures were monitored in a controlled, steady-state environment.[2] Finally, with some modifications to our model, we demonstrated how an entire Pantex magazine can be analyzed.

Progress

We used the commercial finite-element fluid mechanics program FIDAP[5] to model the air flow around columns of heated cylindrical containers to obtain natural convection coefficients for the containers. Thermal radiation coupling between containers was calculated using traditional view factor techniques[6] and a Fortran program written specifically for this project. The convection and radiation data were transformed into resistances and used in lumped-parameter modeling, using the commercial

Figure 1. Schematic of lumped-parameter model.

finite-difference code SINDA,[7] to calculate storage container temperatures. The lumped-parameter resistance network is shown schematically in **Fig. 1**.

Two problem geometries were analyzed: a group of 20 storage containers inside a tent, and a Pantex magazine. The tent is a sub-set of a magazine storage configuration, in which four columns of storage containers were equipped with mock pits producing internal heat generation, instrumented with thermocouples, and monitored in steady-state tests. We selected a data set (Test Run 5, Reference 2) from that experiment for validation of our computational work. Temperature data from the experiment is plotted in **Fig. 2** along with calculated results for the tent model. These results validate the model because the calculated and measured temperatures agree within the range of the data spread.

To analyze the magazine problem with 252 containers we again chose a group of 20 containers, as in the tent, but made modifications to the model (indicated on **Fig. 1**) that adapted it to the highly insulated magazine geometry. This approach allowed a direct comparison of the magazine results with the validated tent results, shown in **Fig. 3**. As expected, the magazine container temperatures are higher than the tent temperatures because the tent radiates to the building while the magazine is insulated by a thick soil berm.

Fluid Mechanics Analysis

A detailed fluid mechanics analysis of the natural convection around the storage containers was required to obtain heat transfer coefficients for the lumped-parameter thermal model. The container diameter is 0.5 m, and the spacing between neighbors is 0.025 m. This close spacing causes the thermal boundary layers of neighboring containers to overlap, and presents a challenging fluids problem because the Rayleigh Number characterizing the flow regime is on the order of 10^8. Empirical correlations for cylinders with this compact spacing are not available in the literature because very close spacing yields poor heat transfer and thus is not used in common design practice. We were, however, able to confirm the order of magnitude of our calculated results by using a correlation for a single cylinder.[8]

We used FIDAP, with a turbulence wall model, and a 2-D approximation of the containers to obtain the necessary coefficients. We found that the coefficients decrease slowly with both decreasing heat generation rate and decreased spacing between containers. Although the coefficients are actually a function of position around the cylindrical containers, we extracted only average coefficients from
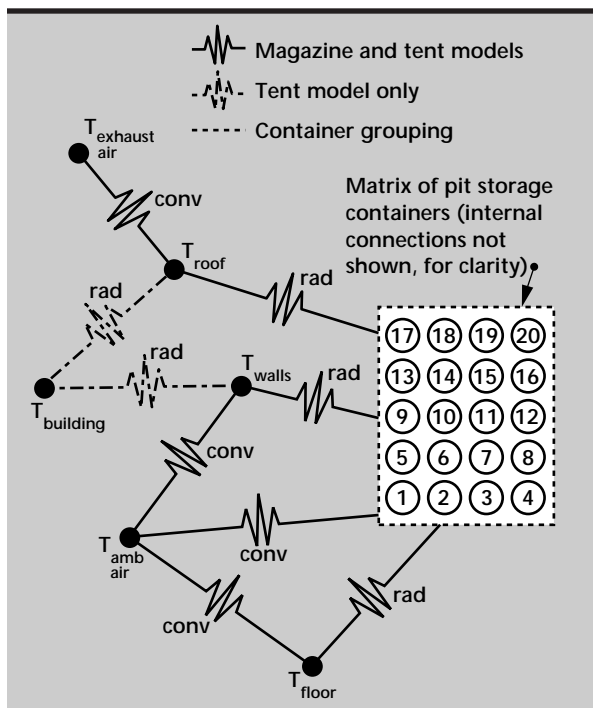
FIDAP for subsequent use in our lumped-parameter models. To quantify the possible error introduced by our modeling assumptions, we lowered the coefficients by 50%, but found that the container surface temperatures rose by only about 1 °C, as **Fig. 2** shows.

Thermal Radiation Analysis

Including thermal radiation in the analysis was found to be extremely important in obtaining correct container temperatures. Without radiation, the calculated surface temperatures were 4 to 6 °C higher than the corresponding experimental data (see **Fig. 2**). To include radiation, the lumped-parameter network code required radiation resistances. To determine these resistances we formulated and solved the radiosity matrices, which relate heat flux to the fourth power of temperature. View factors for each container in the array were computed using the crossed-string method and view factor algebra.[6] A total of 253 radiation resistances resulted from this process.

Two different types of storage containers were used in the tent experiment, so this necessitated the use of two different surface emissivities (0.85 for AT-400A and 0.96 for AL-R8) in the calculations. However, we found only 0.1 °C difference in the calculated temperatures when we changed both emissivities to 1.0. This important result shows that assuming black-body radiation is sufficiently accurate to characterize the radiation exchange and will greatly simplify future calculations.

Lumped-Parameter Model

The lumped-parameter approach uses an electrical resistance network analogy to model the response of a thermally interacting system. In **Fig. 1**
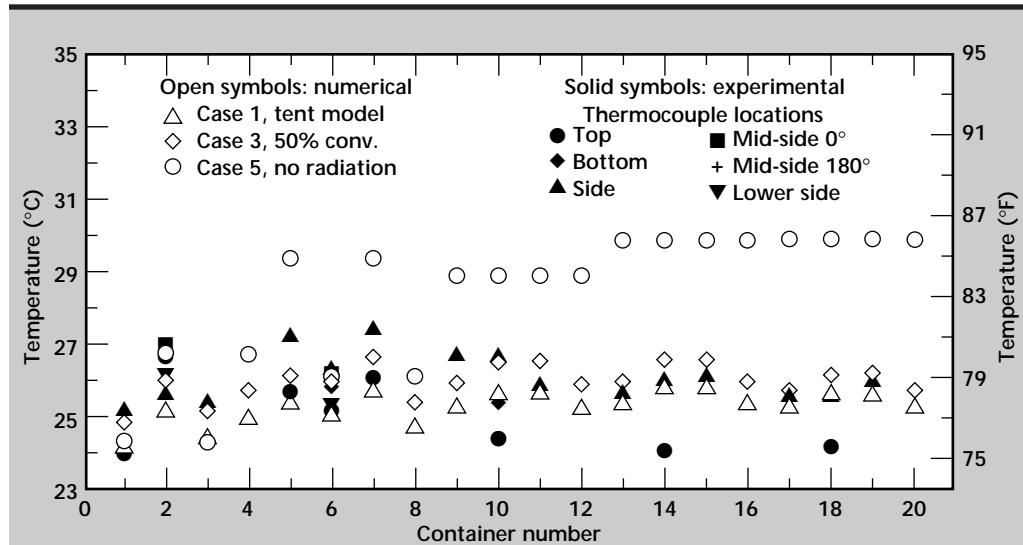
Figure 2. Parameter study of tent model and comparison with experimental data.
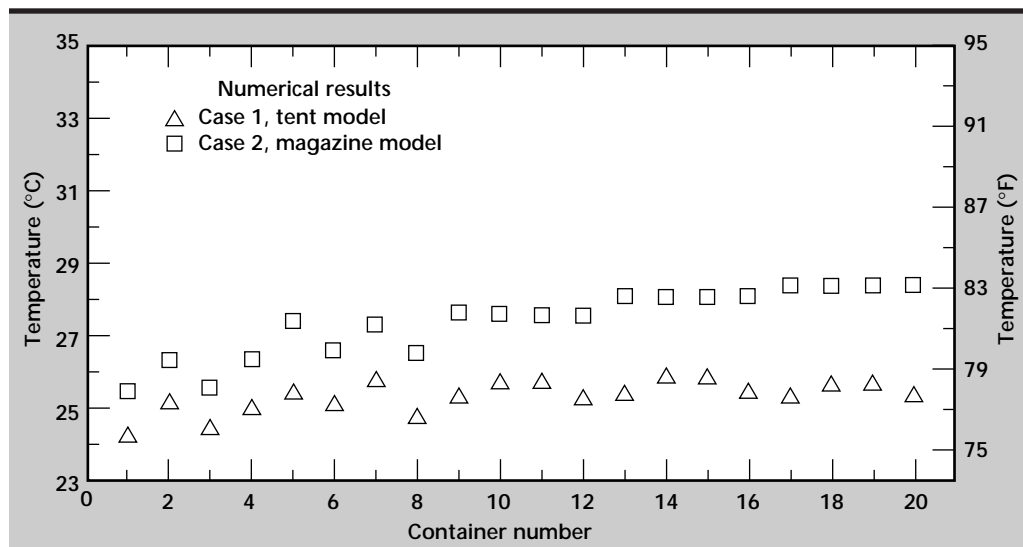
Figure 3. Comparison of tent and magazine model results.

we show the network that includes the array of 20 containers and the surrounding components. Each container is connected by radiation with the other 19 containers, but for clarity these connections are not shown in the figure. The convection coefficients and radiation exchanges within the system were translated into thermal resistances for the SINDA runs. **Table 1** lists the values of heat generation[2] which were used in both tent and magazine models. **Table 2** lists temperature results for the walls, roof, and floor, as well as temperature boundary conditions used in SINDA.

Determination of the temperature of the hot air flowing up the stack requires a calculation of the air flow rate through the magazine. This involved estimating the mass flow rate through the ventilation stack as a balance between the buoyancy forces, which tend to increase the air velocity; the pressure drops in the stack and intake vents, which tend to reduce the air velocity; and the kinetic energy of the air flowing out through the stack. This calculation was an iterative process that resulted in an air flow rate of 0.17 kg/s or 1.5 m/s. While this air flow rate is reasonable for the idealized models treated here, it may be sensitive to wind speed and direction and will also depend on the heat load within the magazine.

Future Work

We recommend establishing direct air flow rate measurements in the ventilation stacks at Pantex magazines to create a data base on this parameter for improved accuracy in future modeling efforts.

Performing additional studies with our lumped-parameter model will yield further understanding and confidence in this modeling approach. The next logical steps could include the following additions to the model: an array of more than 20 containers, stratification of the air inside the magazine, and dividing each container into more than one lumped mass.

Using commercial software limited our capabilities in the fluid mechanics computation of convection heat transfer coefficients, but it was the most economical way to proceed within the budget of the project. If a more detailed study of

Table 1. Container heat loads used in lumped-parameter model, taken from Test Run 5, Reference 2.

Cont. #	Load (W)	Cont. #	Load (W)	Cont. #	Load (W)	Cont. #	Load (W)
17	12.0	18	12.0	19	12.0	20	12.0
13	12.0	14	12.0	15	12.0	16	12.0
9	12.0	10	12.0	11	12.0	12	12.0
5	15.0	6	7.5	7	15.0	8	7.5
1	5.5	2	15.0	3	5.5	4	15.0

Table 2. Calculated temperatures of surrounding components and boundary conditions from SINDA.

Component	Tent model	Magazine model	Reduced convection	No radiation
T_{floor}	24.1	25.8	24.5	22.9
T_{walls}	23.3	25.6	23.5	22.9
T_{roof}	24.4	28.7	24.3	28.1
$T_{amb\ air}$ (BC)	22.9	22.9	22.9	22.9
$T_{exhaust\ air}$ (BC)	33.2	33.2	33.2	33.2
$T_{building}$ (BC)	22.9	n/a	22.9	22.9

the convection is deemed necessary, we recommend proceeding with an LLNL code such as HYDRA[9] so that the approach can be customized for this specific application.

Acknowledgments

The two-tiered computational approach was originally conceived by A. Shapiro and C. Landram. We thank them, along with J. Maltby and R. Martin for valuable discussions held during the progress of this project. We are grateful to L. Dobson for graphical artwork and skillful typing.

References

1. Buntain, G., K. Matsumoto, R. Steinke, J. Squires, P. Fuhrman, M. Lowrey, R. Rabie, W. Gregory, R. Bailey, T. Vidlak, P. Foster, and A. Duncan (1995), *Pit-Storage Monitoring*, Los Alamos National Laboratory, Los Alamos, New M., (LA-12907).

2. Fuhrman, P. W., private communication.

3. Gregory, W. (1995), *Thermal and Airflow Analysis of Pantex 21A Stage Right Configured Modified Richmond Magazine,* Los Alamos National Laboratory, Los Alamos, New M., (TSA8-LA-19).

4. Aceves, S. M., and B. T. Kornblum (1996), *Thermal Analysis of Simulated Pantex Pit Storage*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-ID-125467).

5. Fluid Dynamics International, Inc. (1993), *FIDAP User's Manual Set Fluid Dynamics Package Reference Documentation Version 7.52,* Evanston, Ill.

6. Siegel, R., and J. Howell (1972), *Thermal Radiation Heat Transfer*, McGraw-Hill, Inc., New York, N.Y.

7. Cullmore and Ring Technologies (1995), *SINDA/FLUINT User's Manual*, Littleton, Colo., September.

8. Incropera, F., and D. DeWitt (1981), *Fundamentals of Heat Transfer*, John Wiley & Sons, New York, N.Y., p. 447.

9. Christen, M. A. (1995), *HYDRA, A Finite Element Computational Fluid Dynamics Code*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-121344).

R adiation-Induced Thermal Stresses in High-Power Flashlamps

James D. Maltby, Virginia C. Garcia, and Barbara T. Kornblum
New Technologies Engineering Division
Mechanical Engineering

Accurate calculations of dynamic stresses in pulsed xenon flashlamps, an important lasing component in high-powered glass lasers, will lead to understanding the mechanisms of failure within flashlamp envelopes. The understanding of failure will, in turn, point to more efficient, higher quality, failure-resistant flashlamps. Dynamic stresses in a flashlamp envelope result from two processes: transient heating of the quartz envelope due to radiative absorption, and the pressure rise of the xenon plasma. These stresses were calculated using the 3-D finite-element analysis codes TOPAZ3D, NIKE3D, and DYNA3D. These calculations build upon previous studies of dynamic stresses in envelopes and suggest a possible mechanism of failure due to an extreme temperature rise on the inside surface during the pulse.

Introduction

Flashlamps are the laser pumping source for the NOVA and NIF lasers at Lawrence Livermore National Laboratory. Failure of flashlamps during operation can mean a catastrophic explosion which would severely damage surrounding laser components (such as the amplifier glass). Understanding the mechanisms of flashlamp failure can prevent explosions while increasing flashlamp efficiency and improving flashlamp design.

There are two types of flashlamp failure. The first type is due to fatigue from normal use. This study focuses on the second kind of failure: one-shot failure due to pulsing the flashlamp at an extremely high energy level. This high energy is known as the "explosion energy," or the energy required to explode a flashlamp. The formula for the explosion energy, E_x, is

$$E_x = 20,000 \cdot l \cdot d \cdot \sqrt{t}$$

where l is the arc length of the flashlamp, d is the diameter, and t is the pulse length[1]. This equation is an empirical relationship dating back to Harold Edgerton (1920), the inventor of flash photography. From the explosion energy, the explosion fraction is defined as:

$$f_x = \frac{E}{E_x}$$

where E is the electrical energy delivered to a flashlamp. The lifetime of a flashlamp has been empirically correlated to a simple function of explosion fraction, so understanding of the single-shot explosion mechanism has implications for flashlamp durability.

While a good empirical relationship for the explosion energy is known, the physical failure mechanism is poorly understood. In this project we performed a detailed radiation/conduction/thermal stress analysis of a typical NOVA flashlamp at several explosion fractions. A number of analytical approaches were used to evaluate the optimum method.

Progress

Analytical Approach

This study is an expansion of research on dynamic stresses in the envelopes of pulsed xenon flashlamps. Erlandson[2] used TOPAZ2D to calculate the temperature distribution in the flashlamp envelopes. Analytic formulae were used to approximate the dynamic stresses due to the envelope temperature distribution and the xenon pressure rise individually. Since that time, experimental data show that the transmissivity of cerium-doped quartz decreases with increasing temperature. This suggests the possibility that as the envelope heats up, the quartz would absorb more

energy, leading to increased temperature rise, leading to higher absorption. This runaway couple could potentially destroy the lamp.

To investigate this possibility, a detailed coupled analysis using TOPAZ3D and NIKE3D was performed. To check the validity of NIKE3D on such a rapid transient problem, DYNA3D was used to check a single case. If the initial analysis showed that it was warranted, a multiple wavelength band coupled radiation/-conduction/thermal stress analysis using TRIM3D was at the ready, though it proved not to be necessary.

Flashlamp Description and Boundary Conditions

A schematic representation of a NOVA flashlamp is shown in **Fig. 1**. The flashlamp arc length is 112 cm and the inner diameter (or bore diameter) is 1.5 cm. The flashlamp envelope is 2.5 mm thick and made of cerium-doped quartz glass. The cerium doping absorbs UV light in the range 0.25 to 0.35 to prevent the solarization of surrounding laser components. The envelope is filled with xenon gas at a fill pressure of 0.04 MPa (300 Torr). Under normal operating conditions, the electrical energy delivered to a flashlamp is 11.4 kJ for a pulse length of 943 ms, which gives a conservative explosion fraction of 0.19. During the pulse, dynamic stresses in the flashlamp envelopes are caused by the temperature and pressure rise of the xenon plasma and the heat absorption due to UV light within the envelope.

Figure 2 shows the mesh model of the flashlamp envelope. Only a piece of the envelope is modeled, taking advantage of symmetry and the extremely long, thin cylindrical characteristics of the envelope (so that end effects can be ignored for short pulse lengths). The mesh is graded at the inner and outer surfaces to capture steep temperature gradients and pressure distributions.

Conditions within the tube were derived using empirical formulas from Erlandson's earlier study of dynamic stresses, mentioned above. The rise in xenon pressure during the pulse is modeled as a uniform pressure on the inside surface of the envelope. The pressure profile during the pulse rises rapidly during the first 100 ms, reaches a peak at around 250 ms, then trails off gradually. A surface heat flux boundary condition on the inner surface of the tube was used to model the effect of the interaction between the xenon plasma molecules and the envelope as well as the absorption of UV wavelengths below 0.12. At these wavelengths the quartz is effectively opaque.

The profile for the surface heat flux boundary condition is similar to the pressure pulse. Additional UV radiation is absorbed within the envelope, due to the cerium doping. For the initial analysis, a single absorption coefficient independent of temperature was used. This volumetric heat absorption is a function of time and position. From these boundary conditions, temperatures and stresses within the envelope are calculated.

Heat Transfer Results

The temperature distribution of the flashlamp envelope was calculated using the 3-D finite-element heat transfer code TOPAZ3D. At the design explosion fraction of 0.19, the peak temperature is about 1150 K on the inside surface (the initial temperature is assumed to be 300 K), as shown in **Fig. 3**. This temperature rise is primarily due to the surface heat flux, and is concentrated within a thin layer near the surface. The volumetric heat generation due to

Figure 1. Schematic of flashlamp.
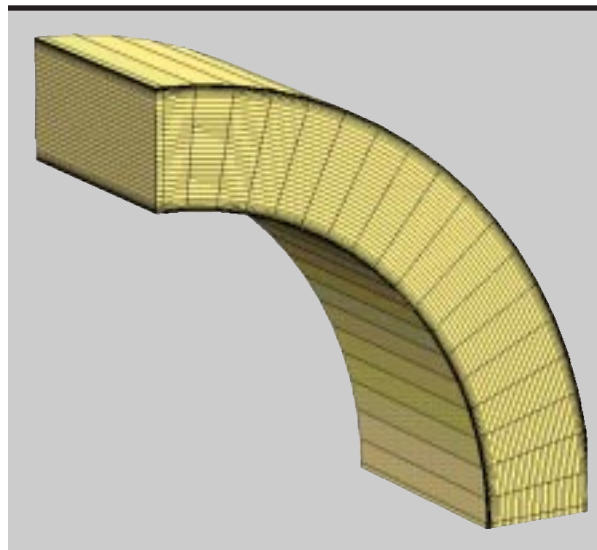
Figure 2. Finite-element mesh model of a piece of the flashlamp.

cerium doping results in a moderate temperature rise of about 30 K, distributed throughout the envelope. Though each of the two thermal loads contributes about the same total amount of heat to the envelope, the volumetric heat generation results in negligible additional thermal stress because of the weak gradients and small temperature rises.

This would imply that the cerium doping does not materially contribute to flashlamp failure, and thus a more detailed multiband analysis of cerium absorption was not deemed to be necessary.

Figure 4 shows the peak temperatures for explosion fractions of 0.4 and 0.8. At high explosion fractions, the temperature on the inside surface quickly surpasses 2000 K, the softening point (that is, "melting" point) of quartz glass[3].

Dynamic Stress Analysis

Stress distributions of the envelope were calculated using NIKE3D, an implicit 3-D finite-element structural mechanics code. At an explosion fraction of 0.19, hoop stress profiles were plotted (**Fig. 5**).

Radial stress profiles were also plotted (**Fig. 6**). The radial stress due to thermal and pressure loading is in compression and is about one order of magnitude less than the hoop stress. The hoop stress peaks in compression on the inside surface with a magnitude of 23 MPa at 0.7 ms. This high compressive stress is due to thermal loading.

On the inside surface of the envelope, the pressure load is a significant effect only for the first 0.1 ms of the pulse, pushing the inside surface briefly into tension. After the first millimeter past the inside surface, pressure is the dominate load throughout the pulse, causing maximum tensile stresses of about 8 MPa.

NIKE3D quasi-static analysis was used to calculate the envelope stresses. Because the pulse length is so short, two other numerical methods were used to validate the quasi-static assumption. The same analysis was run using the NIKE3D Newmark integration option, which takes inertial terms into account. The problem was also analyzed with the explicit code DYNA3D, for a fully dynamic analysis. A detailed comparison of the stress profiles showed

Figure 3. Temperature profiles under normal operation conditions ($f_x = 0.19$).

Figure 4. Temperature profiles on inside surface (where peak temperature occurs) for several explosion fractions.
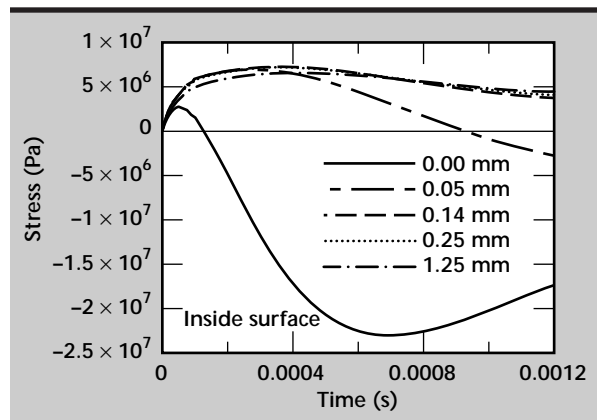
Figure 5. Profiles of hoop stress under normal operating conditions ($f_x = 0.19$).
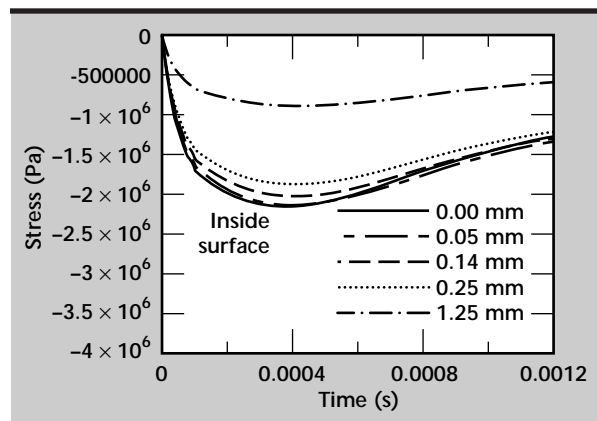
Figure 6. Radial stress profiles under normal operating conditions ($f_x = 0.19$).

close agreement between three methods, validating the quasi-static assumption. Run time for the NIKE3D quasi-static method was by far the lowest, so that method was used for the remainder of the analysis.

Studying the envelope stresses under normal operating conditions does not answer questions about one-shot failure of flashlamps. In fact, these stresses are well below the tensile yield strength of quartz glass of 50 MPa and the compressive yield strength of 1100 MPa[3]. But even at high explosion fractions, the dynamic stresses do not begin to approach the yield strength of quartz glass. **Figure 7** is a plot of hoop stress profiles at $f_x = 0.8$.
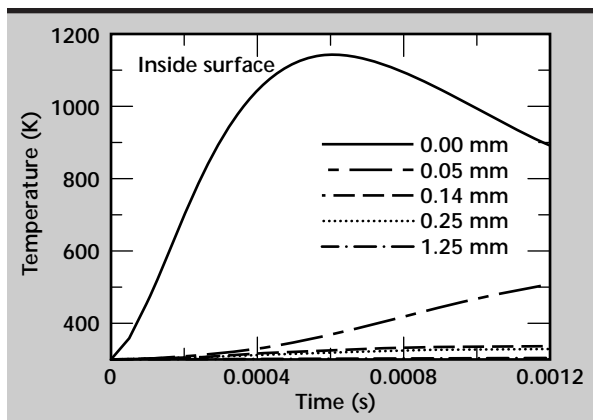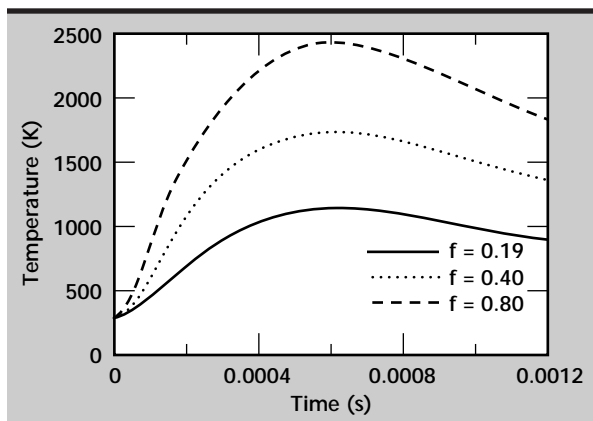
So why do flashlamps fail at high explosion fractions?

This evidence suggests that the most likely mechanism involves phase change due to thermal loading. At high explosion fractions, the temperature on the inside surface of the envelope exceeds the melting point of quartz glass. This suggests a possible scenario for failure. During the pulse, the inside surface is under high compressive stress. The inside surface also melts during the shot, deforming and relieving this compressive stress. Within the first millisecond after the pulse, the inner surface solidifies, locking in its current deformed state and creating

high tensile stress. This tensile stress is of the same order of magnitude as the original compressive stress during the pulse, which exceeds the tensile strength of quartz glass, causing failure.

This suggested scenario would imply that during one-shot failure, the flashlamp actually explodes within the millisecond *after* the pulse rather than during the pulse, which would not be easily detectable.

Future Work

Though the initial theory for one-shot flashlamp failure involving cerium doping was not borne out by the analysis, a very likely mechanism has been suggested by these results. Further data on the softening curve of quartz glass at high temperatures should be incorporated into the model. The analysis should be run out to longer times until the envelope has cooled to well below softening temperature. The tensile stresses frozen into the quartz may then be above failure. This work has also identified that a NIKE3D quasi-static analysis gives the required accuracy, and that a multi-band coupled radiation/conduction/thermal stress analysis is not necessary. This greatly simplifies the analysis, so that a larger range of flashlamp models may be tested to verify the hypothesis.

References

1. Havstad, M. A., and C. W. Dingus (1995), *Participating Media Radiant Transport at High Temperature,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-95).

2. Erlandson, A. C., F. R. Holdener, and A. C. Platt (1992), *Calculations of Dynamic Stresses in the Envelopes of Pulsed Xe Flashlamps,* Lawrence Livermore National Laboratory, Livermore Calif., (UCRL-JC-110021).

3. Chemical Rubber Company (1976), *Handbook of Chemistry and Physics, 57th Ed.,* Cleveland, Ohio.

Figure 7. Stress profiles at a high explosion fraction ($f_x = 0.8$).

Simulation of Turbine Fragment Containment and Mitigation

Gregory J. Kay and Edward Zywicz
New Technologies Engineering Division
Mechanical Engineering

Researchers at Lawrence Livermore National Laboratory (LLNL) and the Propulsion Research Engineering Division of the Boeing Commercial Airplane Group have been working to improve understanding of the dynamics of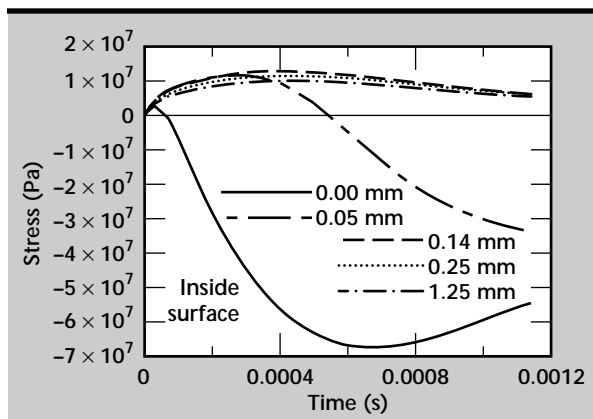 jet engine blade failure (fan blade-off events). The work focuses on simulating a blade-off event and the subsequent blade fragment containment. Favorable comparisons between observed fan blade fragment containment and simulation results were achieved in this study. Clarifications of important fan blade-off issues were made as a result of this modeling. Many improvements to DYNA3D were implemented simultaneously.

Introduction

The latest commercial gas turbine engine designs incorporate advanced material technologies that can lead to fan blade failures through either manufacturing imperfections or in-service deterioration. A fan blade failure can lead to high-velocity fragment generation which has the potential for further engine degradation or even airframe damage. The FAA currently requires that commercial aircraft engines be capable of containing the fan blade fragments. LLNL and the Propulsion Research Engineering Division of Boeing Commercial Airplane Group have established a working partnership to further understand and predict the dynamics of these blade-off events. Thus far the work has focused on simulating a blade-off experiment in a generic high-bypass commercial jet engine. In this event, a single blade was purposely failed at the root. The primary blade failure precipitated the failure of an adjacent fan blade but left the remaining blades essentially intact.

Progress

Previous DYNA3D simulations of the blade-off experiment overestimated the consequences of a single blade removal and predicted the loss of a substantial portion of the fan blade assembly. Areas identified as being responsible for the simulation deficiencies included the accuracy and stability of the DYNA3D constitutive relationships in spinning bodies; the initialization of rotating body force loads;

the engine structural responses to dynamically-induced large unbalanced loads; and the response of the fan blade containment cases to blade release and subsequent debris collisions. The first three areas are addressed in this report. The last area, while treated in this study, will be addressed in a project pending between the FAA, LLNL, Boeing, Pratt-Whitney, and Allied Signal.

During the first half of FY-96, deficiencies in the original engine mesh, obtained from an industry source, were identified and enhancements to DYNA3D were made. The DYNA3D modifications included improvements to the tied shell-edge slide surfaces during rigid-body spinning. Automatic contact was expanded to include start/stop time specifications, multiple automatic contact definitions, deletion of failed shell elements and nodes, and specification by materials or domains. Other DYNA3D enhancements included multiple definitions by materials for the applications of translational and rotational body force loads, and velocity damping in the automatic contact Lagrangian penetration option. A DYNA3D feature that writes a stress initialization file after completing a dynamic relaxation phase (to reach an initial equilibrium state) was also added. This feature permits the user to modify problem parameters such as load curves and contact definitions between the dynamic relaxation and transient analysis phases.

More recent tasks to clarify the fan blade-off modeling issues and improve the simulation agreement with the observed blade-off responses are briefly described below.

1) DYNA3D's ability to accurately and stably simulate spinning bodies was examined. A simplified shaft and blade assembly, with the same radial dimensions and material properties as the more detailed engine mesh was considered. The eight-blade model was spun for ten revolutions. **Figure 1** shows the constant effective stress at a point in one of the rotating blades as a function of angular blade displacement. The stress deviates only slightly from its theoretical constant value.

2) Initialization of body forces due to angular rotation and the impulsive effect of a blade release on the subsequent unbalanced shaft motion were studied. These effects were found to be of second order when compared to either the interaction of the blade fragments and the engine, or the unbalanced shaft load due to a blade loss. The stress state achieved in the DYNA3D analysis as a result of the steady state operation of the engine is shown in **Fig. 2**. Note, as desired, the rotational body forces produce a circumferentially uniform state of effective stress throughout the blade assembly. The initial radial effective stress state shown in **Fig. 2** is consistent with the state of stress expected in the curved fan blade assembly. Similarly, the radial stresses in the fan blade roots produce internal forces which are in agreement with the expected rotational body forces.

3) A simple oscillator mesh (a rotating mass at the end of a truss) was checked for the effects of partial body force initialization and subsequent rotational stability in the transient analysis. This provided information on possible variations in the DYNA3D transient analysis solutions under conditions where the initialized rotational body forces are not compatible with the prescribed initial velocity.

4) The engine mesh was checked for correct blade masses as well as for the initial clearances between the aluminum containment case and the fan blade assembly.

5) The stability of the DYNA3D solution was examined for variance with the code's chosen time step. This was motivated by the aspect ratio of the shaft shell elements, which in some cases had thicknesses in excess of 1.5 times the largest element length. Results of this check indicated that the current time step algorithm provides an appropriate time step size.

6) The preload due to the Kevlar winding on the aluminum case was simulated with thermal preloading.

7) The engine shaft bearing's rigidities were checked to ensure that excessive radial translations were not occurring.

8) The initial blade release (via the DYNA3D tie-breaking shell slideline option) was checked for stability during rotational motion.

9) Specific contact surfaces were used as a check on the accuracy of the current automatic contact algorithms. Previous fan blade-off simulations used the DYNA3D automatic contact option to model the interaction of the blade fragments, the following blades, and the containment cases. The use of explicit contact surface definitions proved to be among the most productive improvements in the blade-off simulation study.

Results of the calculation with the pre-defined contact surfaces can be seen in **Fig. 3**, which shows the engine approximately one third of a
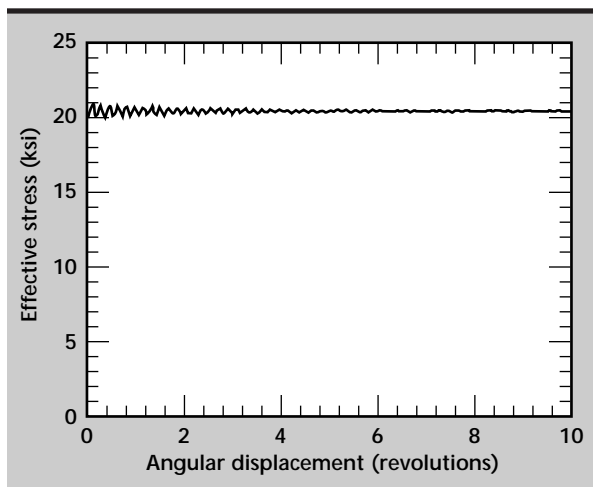

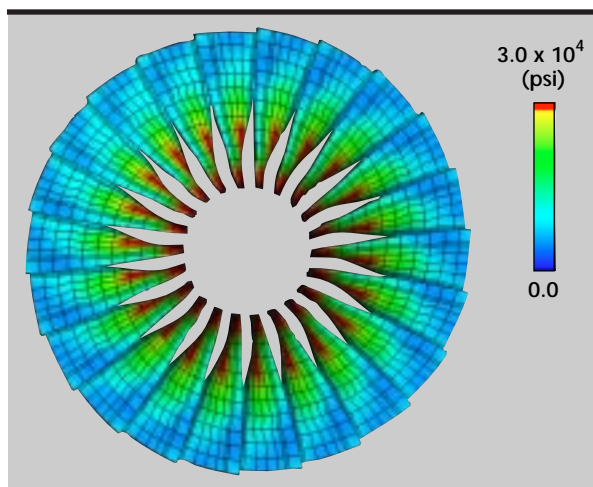Figure 1. *Effective stress in a rotating blade with body force loading.*

Figure 2. *Initial distribution of effective stresses in rotating blades. DYNA3D's stress initialization accurately captures the circumferentially uniform stress state arising from the initial steady engine operation.*
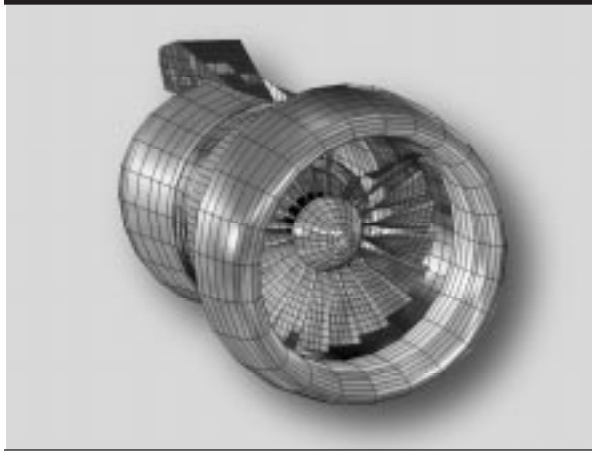
Figure 3. Fan blade-off simulation with DYNA3D.

revolution after the initial blade release. The red regions in the blades indicate yielding due to the interactions between the blades, the containment cases, and the released blade fragments. The mesh consisted of 108 materials, 11,814 nodes, 2095 beam elements and 10,923 shell elements. Initial blade release and secondary blade failure sequences were numerically invoked with tied nodal sets using experimentally obtained activation times.

Future Work

This project has set the foundation for the FY-97 FAA project by demonstrating that the basic tools are available for high-confidence blade-off simulations involving commercial aircraft and engine debris. However, enhancements to the current automatic contact algorithms are necessary if fan blade-off simulations are to be effectively performed with DYNA3D. In particular, the shell element algorithms should be made more robust to ensure adequate contact representations. Work on a next generation of material failure algorithms is necessary to further advance current blade-off simulation technology.

Currently, fan blade fragmentation is achieved by the release of tied nodal sets at times specified by test results, and the aluminum containment case failure is based on a simple failure criterion (effective plastic strain). These capabilities should be improved to better represent the physics in a blade-off event. The consequences of off-balance engine loads on the overall aircraft structure, including flight stability, and the ramifications of uncontained engine debris on the engine and critical airframe components are also of interest to the FAA and the aircraft industry. They should be further examined as well.

Raymond P. Mariella, Jr., Thrust Area Leader

The Microtechnology thrust area conducts activities for the Lawrence Livermore National Laboratory (LLNL) missions of global security, global ecology, and bioscience. Our activities are associated with devices, instruments, or systems which require microfabricated components, including electronics, photonics, microstructures, and microactuators. Our staff of more than 50 people have advanced degrees and expertise in electronics, engineering, mechanical engineering, chemistry, physics, biology, and chemical engineering.

Our longer-term direction is to invent and develop the technology to move into the areas of opto-electronic integrated circuits and into integrated microinstruments. In everything we do, we work through close collaborations with the various LLNL programs, selecting those areas for our attention that offer the greatest leverage. For example, our largest collaboration is with the photonics effort within the Physics Department, and in conjunction with this, our researchers received a patent and an R&D 100 award for a cross-talk-free semiconductor optical amplifier (SOA). This solved the industry's outstanding problem regarding SOA's.

All of the microtechnology work revolves around our Micro Technology Center. Over the last year, with a total budget of over $10M, including approximately $0.8 M of technology-base projects, we have successfully performed collaborations with industry, and we have had an instrumentation project for the polymerase chain reaction for DNA amplification, funded by DARPA, achieve international success with a demonstration of real-time detection of product.

Our first-of-its-kind project with Silicon Video Corporation (our microfabricated components were in the critical path of success) has achieved national recognition in its successful development of the technology to fabricate field-emitter-based flat-panel displays. Our principal investigator on that project received the Federal Laboratory Consortium Award.

As a critical part of the invention of seed technologies, we continue to perform LDRD and technology-enhancement projects. Our work in FY-96 included the fabrication of specially-doped silicon films, both wet and dry etching and bonding of substrates, including silicon-on-insulator and large substrates, and biological and medical instrumentation and devices.

In the area of microstructures and microactuators, our personnel have fabricated: (1) silicon-based microaccelerometers for kilo-g and for micro-g accelerations; (2) circular-cross-section capillary channels in silicon wafers (produced via mirror-image lithography and wafer/wafer bonding) which have cleanly separated a mixture of industrial gases; (3) sub-mm adjustable-position bimorphic actuators that function in air or in water; (4) thin-film shape-memory-alloy actuators that have brought our researchers considerable attention for fabricating the world's first self-powered microgripper on a guide wire; (5) thin fresnel lenses; (6) clear ultra-low-stress silicon nitride films, 100 nm thick with 1-cm^2 area; and many other microstructures.

In interdisciplinary work, we also have had patent claims allowed on our new physical configuration to collect perpendicular light scatter in flow cytometry, which increases S/N by more than an order of magnitude and increases the intrinsic accuracy of the measurement by a factor of 3. We also have demonstrated the world's first real-time detection of polymerase-chain-reaction (PCR) product in a portable instrument. In October of 1996, we took our flow cytometer and PCR instruments to the Joint Field Trials in Dugway, UT, for the open competition.

We have received external support for miniature micromachined valves; integrated photonics; flat-panel display fabrication; portable PCR instrumentation; low-temperature optical modulation; and electrophoresis.

Microtechnology

3

3. Microtechnology

Advanced Plasma Etch Processes for High-Aspect-Ratio, Sub-Micron-Feature-Size Applications

Jeffrey D. Morse
Engineering Research Division
Electronics Engineering

The objective of this effort was to develop and establish plasma etch processes in the newly acquired Lam Transform Coupled Plasma (TCP) Etch Reactor. Our research supports micro-electro-mechanical systems (MEMs), sensors, and programmatic activities at Lawrence Livermore National Laboratory (LLNL) that require high-aspect-ratio, sub-micron, or critical geometry structures to be fabricated in thin film or bulk materials.

Introduction

Plasma etching with advanced tools such as the Lam TCP 9400 reactor is a high-density plasma system that allows for independent control over the plasma efficiency (top electrode power) and the directionality of the ions (bottom electrode bias power).

The result is that extremely small feature sizes (<1000 Å) and highly anisotropic structures (>5:1) can be fabricated, with extremely high etch rates achievable through operation of the plasma at relatively low pressure regimes (1 to 10 mTorr) in comparison to other etch tools. Technologies that exploit such structures range from MEMs devices, flat panel displays, sensors, and high-efficiency gratings for specialized optics.

For MEMs applications, the feature sizes are larger (1 to 5 μm), yet the requirements of aspect ratio (>10:1), anisotropy, and selectivity of etching one material with respect to another (Si vs SiO_2, for example) are critical. It is desirable to have etch processes for these applications which require only a single photoresist mask or, at a minimum, a single layer of oxide or nitride as a mask, with etch selectivity of silicon better than 10:1 to form the appropriate structures for defined applications.

Progress

Several experimental matrices using the Lam TCP reactor have been investigated to better understand the effects of variations in the operating parameters (power, pressure, gas mixture) on etch performance. These matrices determined the effects of varying the operating parameters on the etch rates of different dielectric materials, including thermal and plasma-enhanced chemical vapor deposited (PECVD) SiO_2, silicon nitride, and polycarbonate.

The JMP Statistical Analysis Program is used to analyze the experimental data. The resulting prediction profiles for these materials are illustrated in **Fig. 1** for variations in TCP (top) and bias (bottom) power, pressure, and percentage of CF_4 in the gas mixture. These profiles become the basis for process designs, and subsequent statistical process control, which uses specific performance parameters that are input to the model.

As an example, these results could be used to design a process in which one of the materials is a mask for the other, or one in which one or more of these materials has to be etched in a single etch step, with minimal undercutting of the films. Depending on the specific response of the etch step, a desirability function can be implemented in the statistical model, thus providing the optimal process conditions for the desired response.

Another experimental matrix directly investigated the etch rate and selectivity of etching one material over another, in this case the Si/SiO_2 system. This particular material system has importance since SiO_2 can be thermally grown from Si. One approach is to develop a chlorine-based chemistry, which chemically forms $SiCl_4$ on the Si surface as the volatile etch product. Thus the etching of Si is chemically assisted

with Cl$_2$ plasma gases, and the volatile etch product readily dissociates from the surface when impacted by low energy incident ions or neutrals.

For SiO$_2$, no volatile etch product is formed at the surface, therefore etching only occurs when ions of sufficient kinetic energy impact the surface. Proper control of the substrate bias by the bottom electrode power, hence ion energy within the plasma, allows the oxide etching to be minimized, thereby increasing the selectivity of the etch rates. The trade-off with selectivity is anisotropy.

For more chemically-assisted etch processes, the tendency is to increase the lateral etch rates as well. Chlorine-based chemistries have thus far resulted in Si etch rates as high as 3000 Å/min, with excellent selectivity over SiO$_2$, greater than 50:1. Due to the corrosive nature of chlorine chemistry, however, experimental matrices were limited in scope, and the effort focused on the use of fluorine chemistry.

Fluorine-based gas chemistries have similar issues, except that volatile etch species are formed on SiO$_2$ as well, although those forming on SiO$_2$ require more energy to dissociate from the surface. Therefore, with careful process selection and control, selectivity of etching Si over SiO$_2$ can be optimized.

Another consideration in the use of fluorine-based gas chemistries is the optimization of polymer forming radicals in the plasma, such as CF$_2$, which readily deposits onto surfaces being etched, thereby reducing the etch rate. Thus, to increase etch rates, gases providing higher ratios of free F$_2$, such as C$_3$F$_8$ are typically used.

Figure 1. Prediction profiles for variations in pressure, TCP (top) and bias (bottom) power, and percentage of CF$_4$ in the gas mixture.
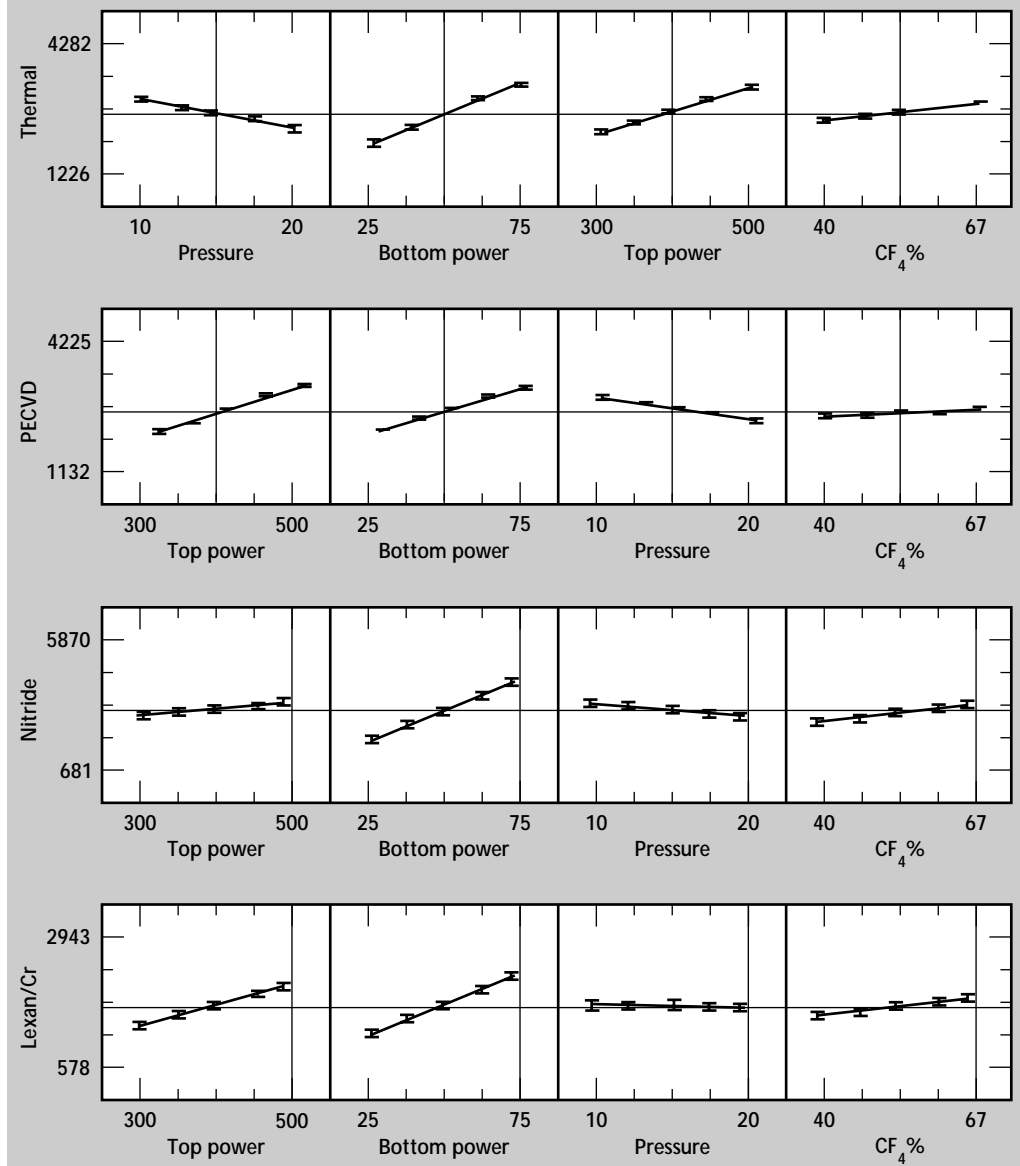
Other approaches use SF_6 chemistry. Passivating layers during etching can be taken advantage of as well, to inhibit lateral etching, in this manner achieving excellent anisotropy, such as with the addition of O_2 or H_2 to the chemistries.

Considering a SF_6-based chemistry, the resulting effects of varying the bottom electrode power are shown in **Fig. 2**. The other operating parameters are detailed in the figure as well. While the etch rates decrease for both materials as the power is decreased, the percentage decrease of the Si etch rate is less, thus the overall selectivity increases. For this example, the selectivity of etching Si over SiO_2 is >10 selectivities. Further improvements in increasing this selectivity are anticipated by broadening out the range of parameters to be studied.

Figure 3 illustrates the dependence of silicon etch rate on percentage of O_2 in the gas mixture for two sets of operating conditions. While the results are quite diverse, it is believed that the one operating condition creates significant polymer-forming radicals in the plasma, which limit the etch rate of Si. With the addition of O_2 to the gas mixture, the polymer-forming radicals are scavenged by the O_2, and the Si etch rate is enhanced significantly.

Figure 4 illustrates the etch profile as a function of process pressure. In effect, several parameters can be varied to optimize the processes for specific integration approaches to achieve the desired selectivity, etch rate, and anisotropy. **Figure 4** further illustrates that deep trench etching of Si is possible. In this case, 20-μm-deep Si trenches have been formed at a nominal etch rate of 1.8 μm/min.

Figure 2. Dependence of etch rate and selectivity for Si/SiO$_2$ system for bottom electrode power for operating parameters indicated above.
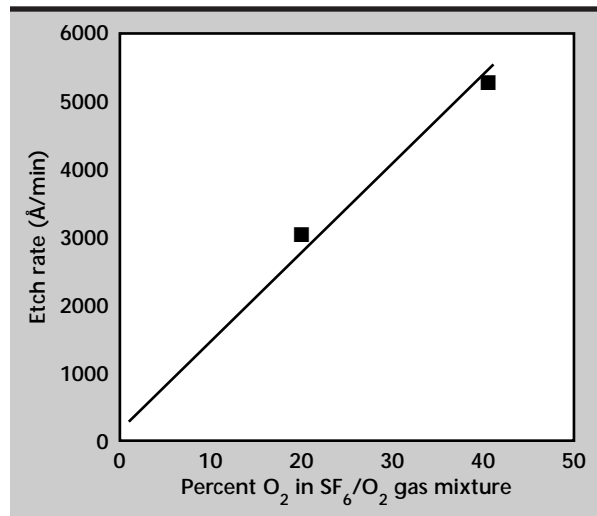
Figure 3. Dependence of Si etch rate on percentage of 0$_2$ in gas mixture.
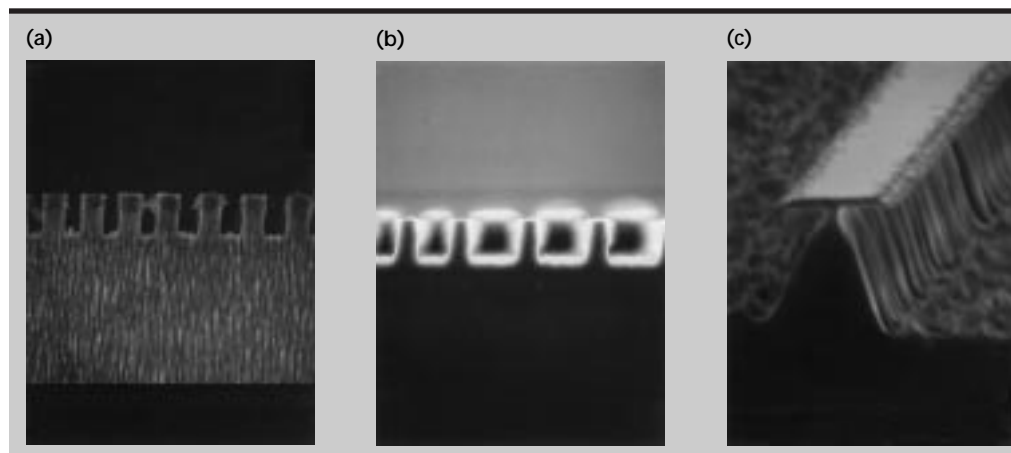
Figure 4. Etch profile as a function of process pressure: (a) 10 mTorr; (b) 20 mTorr; and (c) 60 mTorr.

PECVD-oxide etching was studied using the CHF_3/CF_4 chemistry with the objective of achieving deep sub-micron (<0.2 µm), anisotropic, high-aspect-ratio trenches. The mask used was a thin layer of Cr with ~1000 Å diameter vias patterned in it using an advanced lithography tool at LLNL. For the design of experiment (DOE) the following input variables were considered in the model;

10% <CF_4 (sccm)/($CHF_3 + CF_4$) (sccm) <40% (total flow is 200 sccm)

200 <TCP power (W) <500

30 <bias power (W) <120

10 <pressure (mTorr) <60

Once again, JMP statistical process control software was used for this DOE, and a Fractional Factorial resolution IV design was constructed. This type of design can estimate the main effects for each input variable, and three of the six interactions. The confounding scheme for the design is presented below;

Fractional Factorial Structure
Factor Confounding Rules

$$CF_4\% * Press(mTorr) = TCP(W) * Bias(W)$$
$$CF_4\% * TCP(W) = Press(mTorr) * Bias(W)$$
$$CF_4\% * Bias(W) = Press(mTorr) * TCP(W)$$

The responses measured were the oxide etch rate and the degree of sidewall slope, that is, the tangent of the angle with respect to normal. Thus, a value of zero represents a perfectly straight sidewall.

Oxide Etch Rate Analysis

Etch rate increased with both TCP and bias power, and decreased with increasing pressure. $CF_4\%$ was not an important parameter for determining etch rate. The highest observable oxide etch rate was 5000 Å/min. Using the F-test capability in the software to select significant model parameters, a very good model fit of the etch rate was obtained with just three parameters: TCP power, TCP*Press, and TCP*Bias cross-terms. The model for the etch rate, then, is;

Figure 5. Interaction profile for PECVD oxide etch rate for parameters using LAM TCP reactor. Profiles were determined using JMP statistical analysis software.
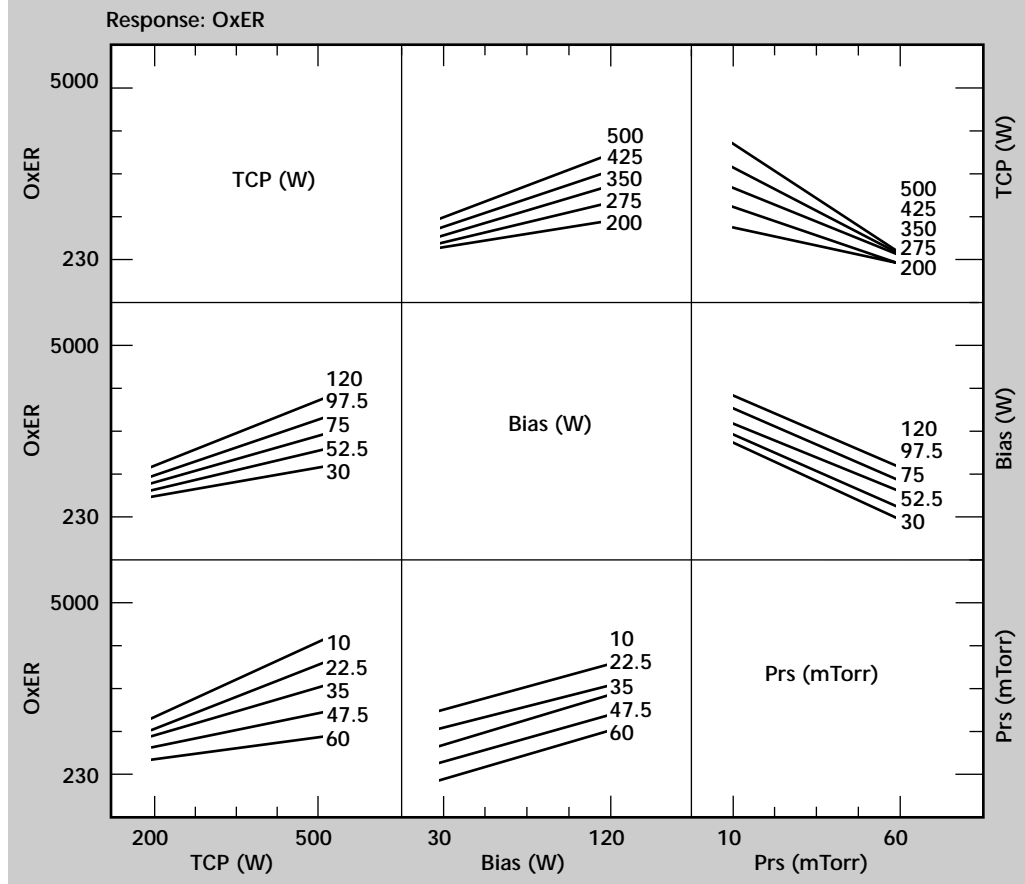
$$ER\left(\mathring{A}/\min\right) = 315$$
$$+ 0.0439 * TCP(W) * Bias(W) - 0.113$$
$$* \mathrm{Pr}\,ess(mTorr) * TCP(W) + 5.249 * TCP(W)$$

For the entire range of parameters explored, the etching is dominated by physical, ion-assisted processes leading to anisotropic etching profiles. There are two distinct operation regimes as a function of pressure. At higher pressures, the etch rates are significantly lower. In this regime, the effect of TCP power on etch rates is minimal, but there is some increase with increased bias power. These effects are most likely caused by the decreased substrate-impact ion energy at 60 mTorr pressure as a result of decreased ion mean-free-path.

At lower pressures, increasing both TCP power and bias power will increase the etch rate. At all TCP power levels, increasing the bias power and decreasing the pressure will increase the etch rates. The interaction profiles are illustrated in **Fig. 5** for the oxide etch rates.

Trench Profile Analysis

The measurements of the trench profile from scanning electron micrographs (SEM) was somewhat complicated due to polymer build-up on the sidewalls. The most anisotropic, vertical profiles were obtained at lower pressure, higher bias power, and higher CF_4 concentrations. For the range of input variables studied, two terms were most significant in predicting the slope: pressure, and interaction of pressure with CF_3%.

At lower pressures, CF_4 concentration has virtually no effect on trench profiles. At high pressures, where ion bombardment is not as effective, increased CF_4 concentration reduces the sidewall polymer formation, thus improving the sidewall profile. TCP power had no significant effect on the final trench slope.

Figure 6 illustrates the interaction profiles for the sidewall slope. **Figure 7** is a SEM that illustrates one of the trenches etched with the Lam TCP reactor. In this example, a pattern was formed in a resist film on top of a thin Cr layer.
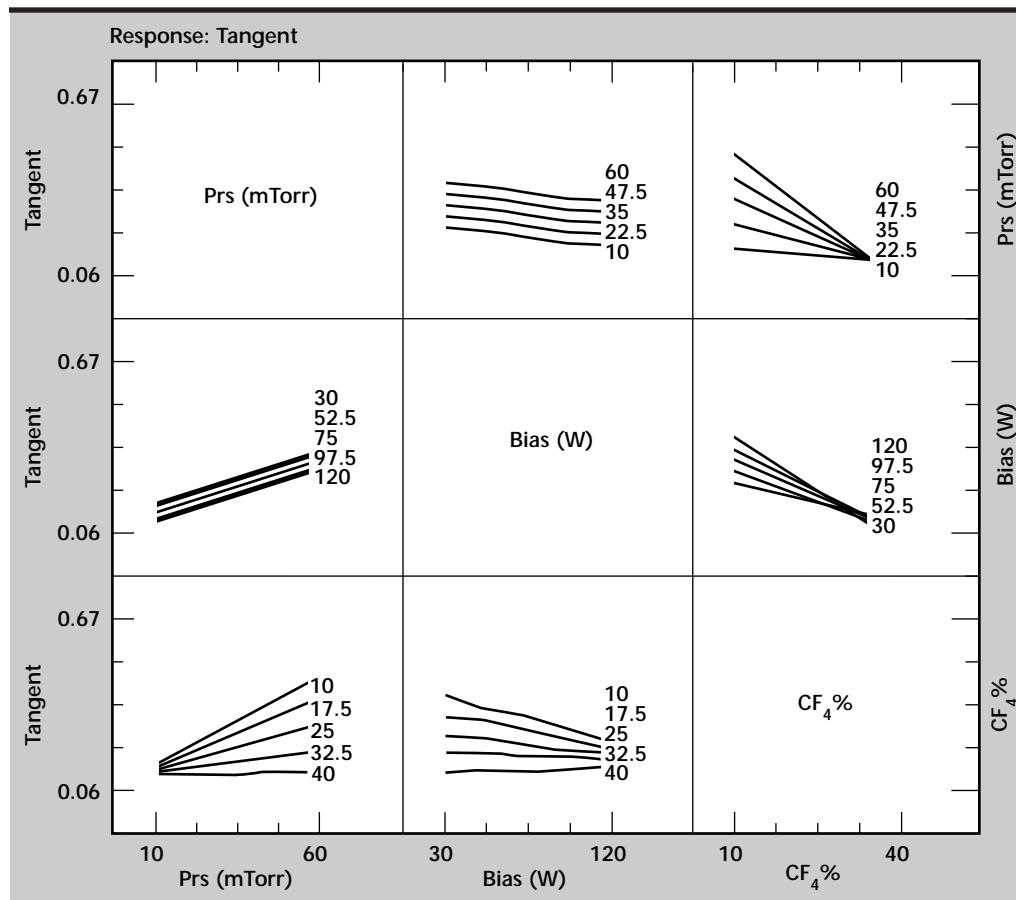
Figure 6. Interaction profile for oxide trench sidewall slope determined using JMP statistical analysis software.
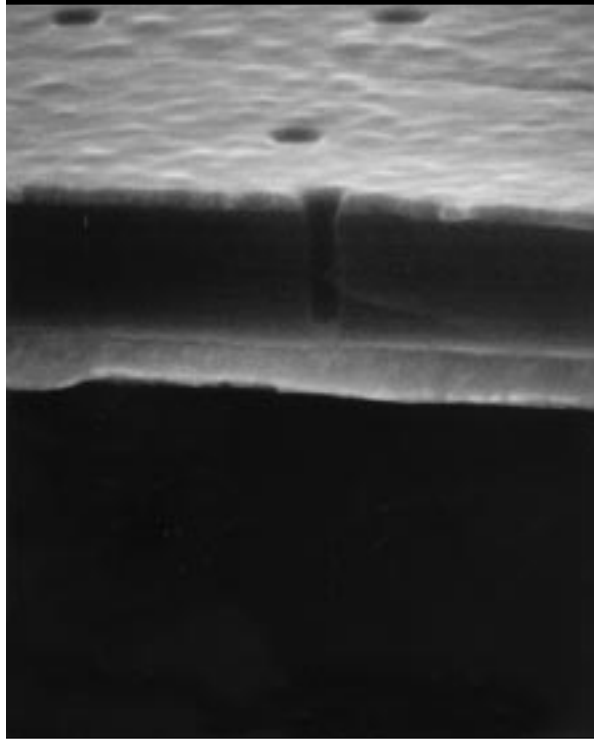
Figure 7. High-aspect-ratio (>5:1), deep sub-micron via (<0.1 μm) etched with LAM TCP. Top layer is Cr metal, and the via is etched through 5000-Å SiO_2.

Using this pattern as a template, a via having a diameter of 1000 Å is etched in the Cr layer. Subsequently, a CHF_3/CF_4 chemistry is used to etch a via through the underlying SiO_2 layer which is 5000 Å thick, to an etch stop. From the figure it can be seen that the features are preserved, and the sidewalls are very straight. This represents state-of-the-art performance in terms of feature size and aspect ratio. Further, it is estimated that without the use of the etch stop, >2-μm vias could readily be etched, thus achieving aspect ratios of better than 20:1.

Conclusion

We have made significant progress towards optimizing plasma etch processes for a variety of applications. The present effort has demonstrated 1) moderately high aspect ratios (>5:1); 2) high etch rates (1.8 μm/min for Si); 3) moderate selectivity (>10:1 for the Si/SiO_2 system); and 4) etching of a variety of materials, including Si, SiO_2, Si_3N_4, Cr, Mo, polycarbonate, photoresist, quartz, HfO_2, and other rare earth oxides.

These etch capabilities are routinely finding new applications within the programs where other approaches fail to achieve the desired results.

ntegration of PCR Amplification and Capillary Electrophoresis in a DNA Analysis Device

M. Allen Northrup, Dean R. Hadley, Stacy Lehew, and William J. Benett
Engineering Research Division
Electronics Engineering

Microfabricated silicon polymerase chain reaction (PCR) reactors and glass capillary electrophoresis (CE) chips have been successfully coupled to form an integrated DNA analysis system. To demonstrate the functionality of this system, a 15-min PCR amplification of a β-globin target cloned in M13 was immediately followed by high-speed CE chip separation in under 120 s, providing a rapid PCR-CE analysis in under 20 min. A rapid assay for genomic *Salmonella* DNA was performed in under 45 min, demonstrating that challenging amplifications of diagnostically interesting targets can also be performed. This work establishes the feasibility of performing high-speed DNA analyses in microfabricated integrated fluidic systems.

Introduction

Our objective in this project is to integrate at least two micromachined MEMS devices into a working system for DNA analysis. The past few years have seen the development of single, stand-alone micromachined devices. These need to be coupled to other MEMS-based devices to show that they have added value over existing "macro" technologies. Valves and reaction chambers, for example, do not stand alone but are part of integrated systems that also need or could benefit from miniaturization.

In the current project, we have focused on developing a fluidic interface to control reagent flow between a micromachined reaction chamber for the PCR and microchannel CE devices for product analysis. The valving/interface system will be applicable to means of controlling flow between two MEMS devices other than reaction chambers and CE channels.

We have taken a three-tiered approach to integration: 1) to simply combine a silicon-based reaction chamber[1,2] with a micromachined glass CE channel, in collaboration with the University of California at Berkeley, (UCB), to perform combined micro-PCR-CE experiments; 2) to develop and test a unique microvalve as a possible PCR-CE interface valve; and 3) to explore the development of unique silicon processing to make an inherent/passive valving mechanism.

Progress

Successes were achieved in all three areas, including the acceptance of the results of the integrated micro-PCR-CE system for publication.[3]

In particular, the direct interfacing of the PCR chamber to the CE channels worked quite well, and a large amount of real analytical data was obtained.
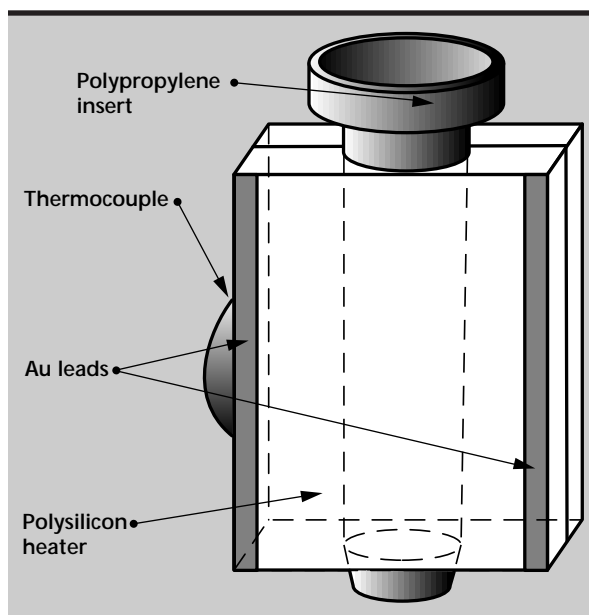
Figure 1. PCR reaction chamber designed to interface to electrophoresis chip.
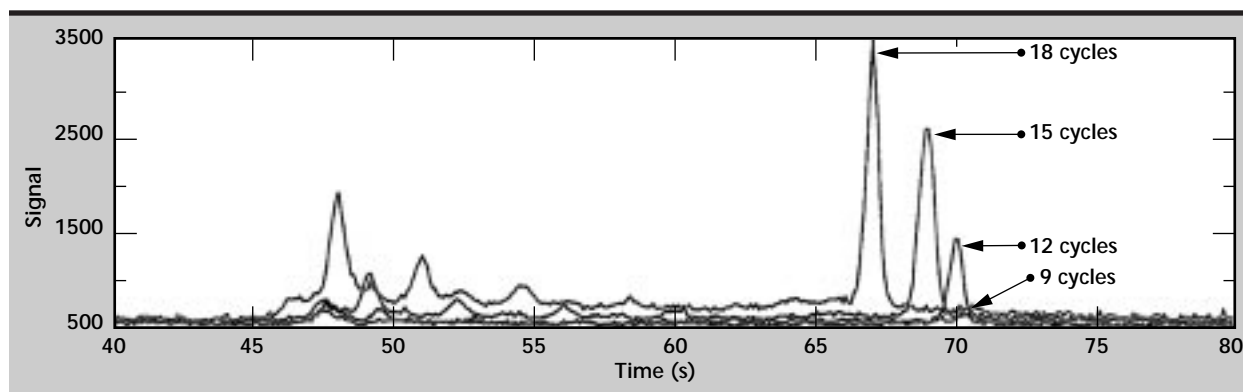
Figure 2. Real-time micro-PCR with microelectrophoretic detection of a PCR product (268 bp, β-globin genetic marker). The product can be seen on the right, the build-up of non-specific PCR product on the left. (The traces were shifted slightly to the left for ease of viewing.)

The microelectrophoresis system had already been developed at UCB[4,5]. **Figure 1** is a schematic of the integrated PCR-CE microdevice.[3]

The separation channels are 100 μm wide, 8 μm deep, and 46 mm long, with a 40-mm distance between the injection channel and the end. The injection channels were 50 μm wide, 8 μm deep, and 12 mm long, with a distance of 2 mm from the reservoir to the separation channel.

The etched glass plate was thermally bonded to the top glass substrate, which had access holes drilled to align with the etched reservoirs. These holes were 1.5 mm in diameter. The silicon PCR chambers with extruded plastic liners were glued in place over the access holes, allowing a direct connection between the reaction chamber liners and CE injection channels. Detection of the electrophoretically-separated PCR products was performed by laser-induced, confocal microscope detection.

Integrated PCR amplification and CE on a β-globin genetic marker (268 bp) was performed in less than 20 min (30 PCR cycles, 83-s electrophoretic separation and detection). This would normally take about 2 h on a commercial PCR instrument with gel electrophoretic detection.

As important as the speed in the integrated approach is its characteristic "hands-off" assay, in that there was no direct sample handling between the PCR amplification and detection.

This is the first time that these two techniques have been combined on a miniature scale, attaining advantages in both speed and integration. We were also able to detect *Salmonella* bacteria with reagents provided by the U.S. Department of Agriculture in the integrated system, and were able to perform real-time product analysis during PCR.

The real-time assay was performed by thermal cycling for a set number of cycles, and performing a 5-s injection step followed by electrophoretic-separation product detection. Results of such an experiment are shown in **Fig. 2**. In that case, the product was detected at 9, 12, 15, and 18 cycles, showing the exponential increase in product concentration. Subsequent cycle analyses showed the leveling-off of product production past 21 cycles, commensurate with predicted reagent limiting factors known to PCR.

It can be seen from the figure that there is an increase in smaller, non-specific product past the twelfth cycle. This is another important feature of an integrated, real-time system: the ability to perform studies on the efficiency of the PCR process itself. In the above case, for example, there was no advantage (in fact, there was a disadvantage) to performing more than 12 PCR cycles. This means that the entire analysis could have been performed in under 10 min.

Future Work

We have been able to demonstrate that improved performance can be attained by integration of two-micromachined devices in a bio-analytical system. The next steps are to continue to improve the integrated system, and add further functionality, such as miniaturized detection components. It will also be important to explore manufacturability and reusability of the systems, to perform real-world analyses on samples of interest, and to package the system into a portable instrument.

References

1. Northrup, M. A., M. T. Ching, R. M. White, and R. T. Watson (1993), "DNA Amplification in a Microfabricated Reaction Chamber," *Transducers '93, Seventh International Conference on Solid State Sensors and Actuators, Yokohama, Japan*, IEEE, New York, N.Y., (ISBN 4-9900247-2-9), pp. 924–927.

2. Northrup, M. A., R. F. Hills, P. Landre, S. Lehew, D. Hadley, and R. Watson (1995), "A MEMS-based DNA Analysis System," *Transducers '95, Eighth International Conference on Solid State Sensors and Actuators, Stockholm, Sweden*, IEEE, New York, N.Y., pp. 764–767.

3. Woolley, A. T., D. Hadley, P. Landre, A. J. deMello, R. A. Mathies, and M. A. Northrup (1996), "Functional Integration of PCR Amplification and Capillary Electrophoresis in a Microfabricated DNA Analysis Device," *Anal. Chem.*, **68**(23), pp. 4081-4086.

4. Woolley, A. T., and R. A. Mathies (1994), "Ultra-high-speed DNA Sequencing Using Microfabricated Capillary Array Electrophoresis Chips," *Proc. Natl. Acad. Sci., USA*, **91**, pp. 11348–11352.

5. Woolley, A. T., and R. A. Mathies (1995), "Ultra-high-speed DNA Fragment Separations Using Microfabricated Capillary Electrophoresis Chips," *Anal. Chem.*, **67**, pp. 3676–3680.

Microactuators for Optical Interferometry

Abraham P. Lee, Charles F. McConaghy, and Peter A. Krulevitch
Engineering Research Division
Electronics Engineering

William J. Benett
Defense Technologies Engineering Division
Mechanical Engineering

Gary E. Sommargren
Advanced Microtechnology Program

We are developing a phase-shifting diffraction interferometer as a metrology tool for extreme ultraviolet lithography.

Introduction

Extreme ultraviolet lithography (EUVL) combined with reflective optics is the primary method to achieve critical dimensions in IC processing below 0.25 μm, the current state-of-the-art line width. By 2007 AD, when 0.1-μm lithography shifts into high gear, the multi-billion dollar IC industry will require vendors to manufacture lithography tools for EUVL. This requires non-conventional EUV metrology that can measure precisely to below 1 nm. There is currently a multi-million dollar CRADA in EUVL between Lawrence Livermore National Laboratory (LLNL) and the IC industry.

At LLNL we are developing a phase shifting diffraction interferometer (PSDI)[1] as a metrology tool for EUV optics (**Fig. 1**). This interferometry system is so attractive that vendors want it immediately for conventional optics. Eventually, the success of the EUVL project will be judged not only by its ability to characterize high-precision optics, but also by the ease with which the technology can be transferred to the manufacturers.
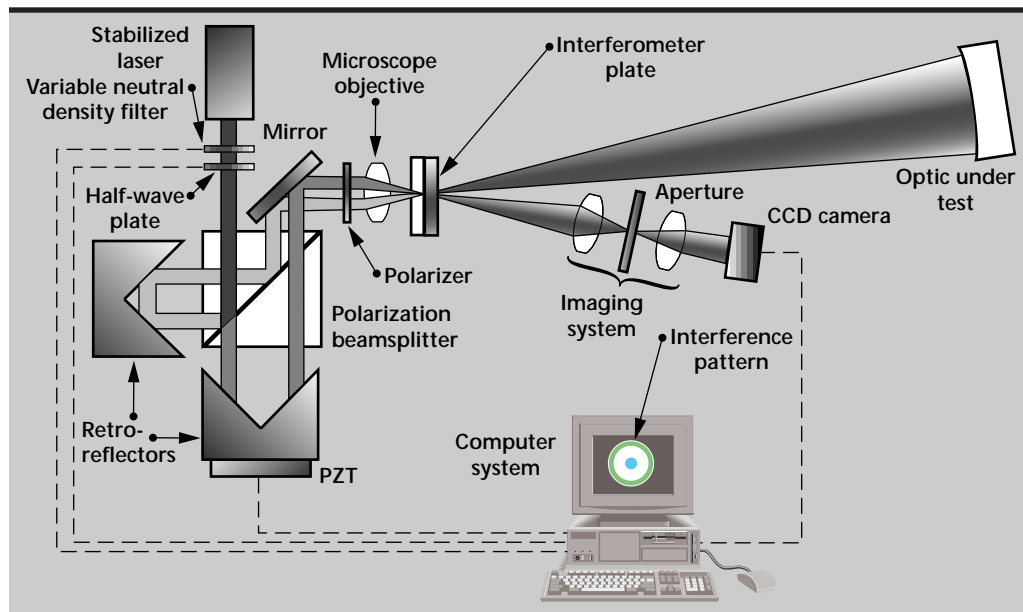
Figure 1. Phase-shifting diffraction interferometer system.

Progress

Given the scale of most specialized vendors, companies will not be able to hire a Ph.D.-level expert to maintain such a system. An ideal system would be a black box that they can buy to measure their manufactured optics without much maintenance. It would be extremely difficult for the existing system to be user-friendly with its 30-lb., 3-PZT actuator component, which is mounted vertically.

Our project is to replace this PZT phase-shifter with a miniaturized micromirror. We believe this is a major building block to a black box, and has potential of attracting funding from the industry as well as from the government. The realization of the project proposed here will greatly benefit the optical metrology efforts at LLNL and create new opportunities for industrial collaboration.

Microelectromechanical systems (MEMS) enable the integration of actuation, sensing, control and signal processing onto a single silicon chip. Microactuators that have been developed in the last two years at LLNL's Microtechnology Center include polyimide thermal bimorphs, smart polymers, shape memory alloys, and hydraulic and magnetic actuators. At LLNL we also have experience in designing, fabricating, and testing electrostatic microactuators. Our microvibromotors [2,3] are being applied to optical modules at the University of California, Berkeley, in collaboration with HP.

The resolution achieved by linear microvibromotors is 0.25 µm. Angular microvibromotors have rotational speeds up to 60,000 rpm. These microactuators are fabricated by adapting IC technology and can be batch-fabricated to reduce manufacturing costs. At LLNL, MEMS and optics are two major core competencies. MEMS components for optical measurements and calibrations have generated successes in many x-ray and laser experiments. Our early successes in Fresnel zone plates, x-ray windows, and gratings, have focused on passive microstructure components, where no moving elements were involved, and most were *ad hoc* projects with no long term plans to advance this promising new field.

By combining existing LLNL optics and micro-fabrication expertise with our recent development of microactuator technologies, a potential major impact area for LLNL can be seeded.

Future Work

Besides the EUVL project in LLNL's Lasers Program, many other programs can benefit from the success of MEMS optics work. Examples are:

(1) Fourier Transfer for infrared (FTIR) detection for spectroscopy/chemical and biochemical analysis. A precision position-controlled micromirror enables the filtering of wavelengths by the Fabry-Perot interfering principle. Tunable cavity lasers can be developed using the same concept.

(2) An optical phase array. A 2-D array of these micromirror phase shifting interferometers can set up a platform to steer and focus optical beams.

(3) Others: MEMS combined with photonics is a very hot research and development field. Microscanners, microshutters, microdynamic focusing mirrors, micromirrors for *in situ* coupling of optical fibers, flat panel displays, and movable diffraction gratings, are all components that have had initial success. Many more applications are being developed. The next generation of photonics will most likely depend on the maturing of this technology.

References

1. Sommargren, G. E. (1996), "Diffraction Methods Raise Interferometer Accuracy," *Laser Focus World*, August, p. 61.

2. Lee, A. P. and A. P. Pisano (1992), "Polysilicon Angular Microvibromotors," *Journal of Microelectromechanical Systems*, Vol. **1**, (2), pp. 70–76.

3. Lee, A. P., D. J. Nikkel, and A. P. Pisano (1993), "Polysilicon Linear Microvibromotors," *Proceedings of the 7th Int. Conf. on Solid State Sensors and Actuators, Transducers '93, Yokohama, Japan,* June, pp. 46–49.

Thin Silicon Windows

Dino R. Ciarlo
Engineering Research Division
Electronics Engineering

Charles E. Hunt
Department of Electrical Engineering and Computer Science
University of California at Davis

We have explored the use of carbon as an etch-stop material for fabricating thin silicon windows. We have produced one of the largest thin silicon windows ever made, 0.4 µm thick, with an outside frame dimension of 24 mm × 24 mm.

Introduction

Thin silicon windows find many applications in scientific experiments at Lawrence Livermore National Laboratory (LLNL). They are used as supports for x-ray interferometers, in radiation detectors, as debris shields, as supports for Moiré grids, as exit windows for electron sources, in x-ray microscopes, and in other apparatus where one requires a thin transmissive window for the radiation of interest.

The preferred method to fabricate such windows is to diffuse approximately 1% boron (atoms) into the surface of a silicon wafer and then chemically etch away the undiffused silicon from the other surface. This is the so called "p-stop" etch method, which takes advantage of the reduced etch rate of silicon when it is doped with a heavy concentration of boron. Germanium is usually added to compensate for the tensile stress caused by the boron. This material is referred to as SiGeB. In our research we have explored the use of carbon as an etch-stop material for window fabrication, since it has a much higher etch selectivity. This new material is referred to as SiGeC.

Progress

The silicon wafers used in this research were (100) orientation, with a thickness of 530 µm and a diameter of 100 mm. A carbon/germanium-doped silicon film was grown by epitaxy on the surface of these wafers by Lawrence Semiconductor Research Laboratories (LSRL) in Tempe, Arizona. LSRL is interested in this material for a number of applications, including its use as an etch stop in the fabrication of silicon-on-insulator (SOI) material for high-speed, high-voltage microcircuits. They plan to study this material at the University of California at Davis.

The thickness of the epitaxial layer in our wafers (which eventually determines the window thickness) ranged from 0.1 µm to 1 µm. The carbon concentration was varied from 1 to 7% (atoms). Various amounts of germanium (18 to 40%) were added to the carbon-doped layers to compensate for the stress caused by the carbon.

Once the wafers were received at LLNL, they were coated with a 1000-Å thick film of silicon nitride by low pressure chemical vapor deposition (LPCVD). This process is carried out at 800 °C and does not cause appreciable diffusion of the carbon out of the epitaxial layer. The silicon nitride is used as an etch mask to etch the thin window. Photolithography is used to pattern the wafer on the surface opposite the epitaxial layer, and then the nitride is etched in a Freon-14-plus-oxygen plasma.

Several silicon etches and temperatures were tried, mostly consisting of various concentrations of potassium hydroxide (KOH) which was saturated with isopropyl alcohol (IPA). **Figure 1** is a photograph of a large-area (12-mm × 12-mm) silicon window fabricated using this method. The window thickness is 0.4 µm and the outside dimension of the

thick silicon frame is 24 mm × 24 mm. This is a very large window for this thickness. Such large sizes are often required in experiments, such as in x-ray interferometers for laser plasma diagnostics.

Some of the wafers we received were used to determine the etch rate of this new material. For this purpose, we used a chrome-gold etch mask. Silicon nitride could not be used for these experiments since the plasma process used to etch the nitride also slightly etched the underlying silicon. The etch selectivity for this carbon-doped material was found to be much higher than the etch selectivity of the boron-doped material. For instance, when using 44 % (wt) KOH saturated with IPA, at 60 °C,
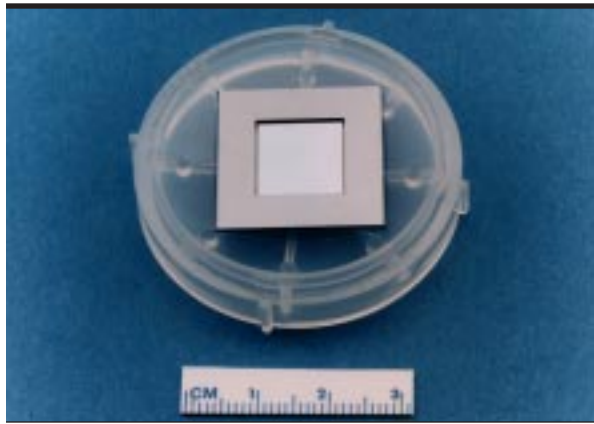
Figure 1. Large-area (12-mm × 12-mm) silicon window, 0.4 μm thick. This window was fabricated using the SiGeC etch-stop material.

the etch rate of the SiGeC material is approximately 1000 times slower than undoped silicon. On the other hand, the etch rate of the SiGeB material is only 100 times slower. This increased selectivity with the SiGeC material makes window fabrication much easier, since the material will tolerate a much longer over-etch.

The ultimate burst strength of the SiGeC windows was compared to the ultimate burst strength of the SiGeB windows for the same window sizes. For these experiments we used windows that were 0.5 μm thick and had a surface area of 12 mm × 12 mm. The SiGeC windows failed at 1 psig; the SiGeB windows failed at 3 psig.

Future Work

Our SiGeC material appears to be a superior thin window material, primarily because of the increased etch selectivity. Some remaining work needs to be done on the optimum carbon and germanium concentration for this application. Also, we need to establish the thinnest window that can be fabricated for a given area.

In the above experiments, we found that 0.4 μm was about the thinnest window that would survive the fabrication process for the large area of 12 mm × 12 mm.

Finally, there needs to be a theory developed which explains the etch-stop behavior of this new material.

Eutectic Bonding and Fusion Bonding

Steve Swierkowski
Engineering Research Division
Electronics Engineering

The first goal of this project, the eutectic bonding with gold/tin alloy (Au/Sn), has been shown to produce strong bonds in Si, and these bonds are shown to be excellent for cryogenic applications. Initial experiments in bonding GaAs to Si were not successful, but this can be attributed to contamination of the braze interface. The key idea of this effort should still work since no fatal flaws in the plan have been seen. More effort is needed.

The second goal of this project, the fusion bonding of glass, has been demonstrated with excellent results. Glass substrates up to 58 cm in length have been 100% void-free bonded in Lawrence Livermore National Laboratory's Microtechnology Center (MTC) with some newly developed apparatus and procedures, resulting in an invention disclosure.

Introduction

The two goals of this project encompass large area bonding of substrates. The initial project focus was to develop a Au/Sn eutectic bond between GaAs and Si substrates, in the range of 2 cm^2, that will withstand cryogenic temperatures (77 K). This first goal would make possible much larger IR imaging array chips for advanced surveillance. The scope of the project was then enlarged to include a completely different, larger area substrate bonding need, the fusion bonding of borosilicate types of glass with areas as large as 400 cm^2. This second goal makes feasible completely new microstructures that are now possible only with the recent completion of large area lithography technology within the MTC, that was used to first pattern the substrates.

Progress

Eutectic Bonding

The bond between substrates with different temperature expansion coefficients can be subjected to a great deal of stress for sufficient sizes and temperature changes. Our goal is to bond thick GaAs to wafer thin Si. Since GaAs shrinks more than Si, when this combination is cooled, the Si will be subjected to compressive strain. If the GaAs is thick enough, it will remain approximately flat. The composite can then be used as a carrier to mount focal plane arrays, with similar temperature coefficients to GaAs, onto the Si with currently developed indium bump-bond technology. To use this scheme, a hard, strong eutectic metal bond is needed between the substrates that will withstand the stresses generated by cooling down to cryogenic temperatures such as 77 K. A Au/Sn eutectic alloy has been successfully used in small areas of Si and III-V devices, and is a promising candidate.

Initial testing of evaporated films with separate deposition in layers from separate Au and Sn sources showed that the predicted melting points were not obtained and the surface morphology was questionable. Next, a composite Au/Sn alloy sputtering-source target was obtained, and this source produced films that seemed usable for further testing.

We deposited eutectic Au/Sn sputtered films on Si and GaAs wafers and performed bonding experiments. First tests done on coated Si wafer parts showed that the wetting was good and that the alloy could be heated to past its melting point, without evidence of non-wetting or globule formation. Strips of alloy about a millimeter wide on Si were subjected to adhesive tape pull tests and also to temperature stress tests with no evidence of peeling or bond failure. The most extreme test was

to plunge the bare eutectic strip/substrate into liquid nitrogen. For this temperature shock test, the alloy peeled back from shrinkage to form an arch-like shape, shown in **Fig. 1** upside down. The once-bonded side is up, showing that the alloy peeled back from the Si, but was 100% covered with Si that was pulled out of the host wafer. It shows up as the grooved material at the top of the arch. The alloy is the thin bright white line, visible only at the base of the arch on the side edge. There was no evidence anywhere of alloy separation at the Si interface.

The next bonding tests were done with a Si/eutectic-eutectic/Si wafer pair. A 2-cm^2 wafer pair of this type was cycled several times to 77 K with about 1 min cycle times, and no evidence of delamination was observed. Before and after cycling tests, samples were subjected to the so-called "razor blade" test. A scalpel was used to try to wedge apart the two wafers. The edges of the wafers are quite rounded, so it is possible to apply a great deal of separation force to the wafers. Ultimately, the Si was observed to fracture at the beginning of the alloy, but no separation of the alloy bond was observed, as illustrated in **Fig. 2**. The rounded edge of the top wafer piece runs diagonally across the lower wafer piece. On the left

end of the top wafer piece, is the shattered Si region where the end of the scalpel was eventually pushed out sideways with a great deal of force. It went through the Si, but without delaminating the alloy bond region.

Initial tests with eutectic on GaAs thin wafers were promising, in that the alloy seemed to thoroughly wet, and possibly alloy into the GaAs. Again, no peeling or globule formation was observed. Subsequent tests of thin wafer bonding GaAs to Si were not successful. The brazing jig appeared to have been contaminated chemically, and an oxide-like film visually appeared on both wafers. This seemed to prevent the two eutectic alloy coated surfaces from wetting each other thoroughly. More parts remain to be tested and this should be done because of the extremely promising positive results obtained with Si.

Glass Fusion Bonding

The goal of this work was to establish apparatus and procedures to thoroughly bond large (7.5 cm x 58 cm) borosilicate glass plates that are densely patterned with microchannel capillary arrays for DNA sequencing. The capillaries are 20 to 150 μm in cross-section and 48 cm long.

Figure 1. Si on a ribbon arch of eutectic alloy. The Si is the grooved material. The alloy is the thin white line on the bottom of the arch and visible on the side of it.

Figure 2. Two wafers of Si eutectically bonded. After several cycles to 77 K, the pieces were subjected to the razor blade wedge test. The metal wedge shattered the upper wafer piece on the left end, but no eutectic bond failure is seen.

Recently developed large area lithography technology in the MTC was used to pattern and etch the glass substrates. A special and unusual bonding jig and procedure was developed to align and hold the two large glass substrates together in a manner compatible with the high temperature furnace (500 to 700 °C). The apparatus and procedure, which are integral to the success, are described in invention disclosure IL-10039 (S. Swierkowski, 9/29/96). Alignment can be less than 20 μm on these huge plates. (20 μm out of 48 cm is 40 ppm.)

The bonding experiments have shown 100% full fusion bonding with no voids. The original glass interface is undetectable, even under cleaving experiments and SEM observation. The interface region appears as continuous homogeneous glass. The distortion of μm-sized channels appears to be less than a few percent. The channels have been pressure-tested and dye-tested with no delamination or bonding voids evident. Initial tests of DNA electrophoresis sequencing experiments in the prototype capillary plates have produced very good results.

A cleaved section of a bonded plate pair is shown in **Fig. 3**. The SEM view is oblique. The large upside-down U-shaped channel (appears dark) in the top plate, for sample injection, is matched to the shallow trough-shaped channel in the bottom plate and the cross-section is a cleaved interface. There is no evidence anywhere of interface delamination. In fact, there is no observed evidence of an interface except for the pre-etched channels. Another cleaved cross-section from the same bonded plates, but in a region where only the small channel exists, is shown in **Fig. 4**. This nearly end-on view shows the channel shape to be ideal for separation experiments. It has no sharp acute-angled concave corners and is quite smooth. Again, there was no evidence of delamination along the original interface along all of the 101 etched channels.
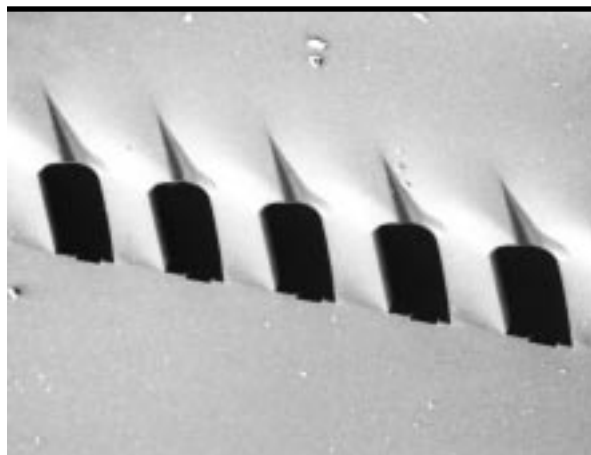
Figure 3. An SEM oblique view of fusion-bonded glass with large grooves bonded against small capillaries. The view is of a cleaved section of the composite and shows no evidence of delamination.
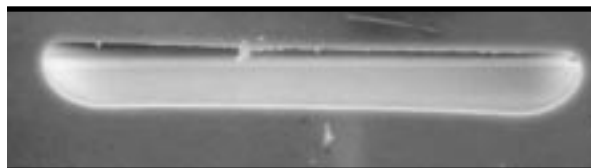
Figure 4. The end-on view of a cleaved section of the same fusion-bonded glass of Fig. 3, but in a different region, where only the small capillary is present. Its dimensions are 19 μm by 153 μm, or about the size of a human hair. The original glass interface was along the plane of the top of the capillary, but there is no trace of it with the full fusion of the glasses.

Acknowledgment

Solid-source MBE-Grown GaAs/AlGaAs Ridge-waveguide Semiconductor Optical Amplifiers

Jeffrey D. Walker, Sol P. DiJaili, William Goward, Holly E. Petersen, and Gregory A. Cooper
Engineering Research Division
Electronics Engineering

Frank G. Patterson and Robert J. Deri
Physics and Space Technology Directorate

We have established the growth of phosphide-based semiconductors at Lawrence Livermore National Laboratory (LLNL) by solid-source molecular beam epitaxy (MBE), and developed a robust etch-stop technology for applications to ridge-waveguide optoelectronic devices. We have met the technical objectives of this project and have applied it in several areas. In this report we present first the general operating parameters for growing phosphide thin films with a solid-source valved cracker that is now installed and operating successfully. Second, we demonstrate high-quality InGaP thin films that are lattice-matched to the GaAs substrate with less than 0.1% error. Third, we demonstrate wet etch-stop chemistries for selectively etching AlGaAs with respect to InGaP, with a selectivity of over 200 times. Fourth, we demonstrate 0.8 μm-wavelength ridge-waveguide semiconductor laser diodes (SLDs) and semiconductor optical amplifiers (SOAs) that use this technology to simplify processing and improve uniformity. The amplifiers are the first in the world to use an InGaP etch-stop technology to simplify processing. They have a chip gain of 32 dB, which is comparable to the best SOA results in the literature.

Introduction

LLNL has a track record in producing high quality arsenide-based III-V photonic devices such as the cross-talk-free SOA.[1] However, much of the world's data traffic now flows over fiber-optic systems that are built around phosphide-based III-V devices. In this "information age" phosphide-based devices are thus of great interest and our previous inability to grow these materials was identified as one of the weak links in LLNL's overall strategy for III-V photonics research for the next century. With this in mind, one of the major goals of this project was to establish the growth of phosphide-based III-V devices. While we are not the first in the world to grow phosphides by MBE, we are among the first to use a new type of safe phosphorus source known as a solid-source valved cracker. Our project makes a contribution in this area.

In addition to establishing growth of phosphide materials, this project sought to apply the new capability to a real device.[1] The device project that was targeted was to improve our existing 0.8 μm-wavelength ridge-waveguide GaAs/AlGaAs SOA process by introducing an InGaP etch-stop layer, as seen in **Fig. 1**. The etch-stop layer simplifies the SOA fabrication process, and increases yield, uniformity, and reproducibility of these devices.[2]
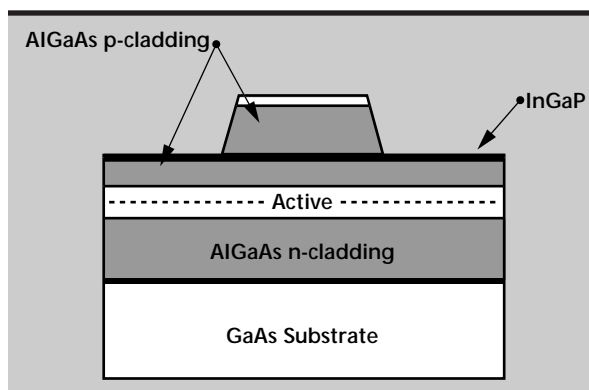
Figure 1. *Schematic of ridge-waveguide SOA device with InGaP etch stop.*

Progress

In the first phase of this project, a phosphorus source was added to our MBE system to give us the ability to grow phosphide thin films by all solid-source MBE. A new type of "safe" phosphorus source called a valved cracker was used to eliminate the need to work with phosphene gas (which is extremely toxic) or solid white phosphorus (pyrophoric, and also extremely toxic). These valved crackers are loaded with red phosphorus, which is safe to handle. Once the source is loaded and under vacuum, a small amount of red phosphorus is converted to white phosphorus, which is then used for the growth. Following installation of the phosphorus source and initial characterization, recipes were determined for the growth of InP-on-InP wafers and InGaP-on-GaAs wafers.
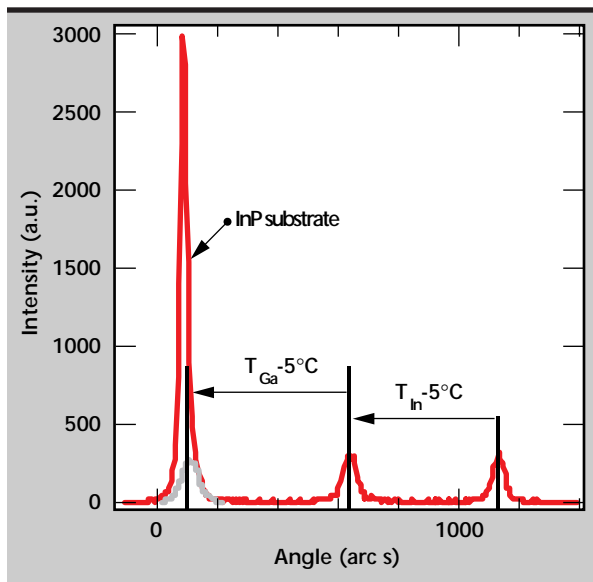
A major issue for the growth of InGaP is matching the lattice constant of the InGaP to that of the GaAs substrate. This requires very careful control of the In-to-Ga ratio in the material. This can be controlled by adjusting the In and Ga source temperatures. **Figure 2** shows double crystal x-ray diffraction results from a MBE wafer grown to calibrate lattice mismatch. For this wafer, three 0.4-µm $In_{0.49}Ga_{0.51}P$ layers were grown on the same substrate with different In and Ga source temperatures. The temperatures for the first layer were T_{In} = 885 °C and T_{Ga} = 875 °C; the In temperature was raised 5 °C for the second layer, and the Ga temperature was lowered 5 °C for the third layer.

The x-ray data shows a large peak from the substrate, and three smaller peaks from the InGaP layers, one of which overlaps with the substrate peak. From this wafer, it is determined that the lattice matching condition is for T_{In} = 890 °C and T_{Ga} = 870 °C, and that the lattice constant shifts –100 arc s/°C with In cell temperature and +105 arc s/°C with Ga cell temperature. Note that 100 arc s corresponds to 0.04% lattice mismatch. For high quality material, the lattice mismatch must be kept below 0.1%, and mismatches below this are consistently achieved with our MBE system.

Once we established the ability to grow InGaP lattice-matched to GaAs substrates, the next step was to develop a wet etch chemistry for selectively etching AlGaAs with respect to InGaP. For this

Figure 2. Double crystal x-ray diffraction results from a calibration sample consisting of three 0.4-µm InGaP layers grown on GaAs with different In and Ga source temperatures.
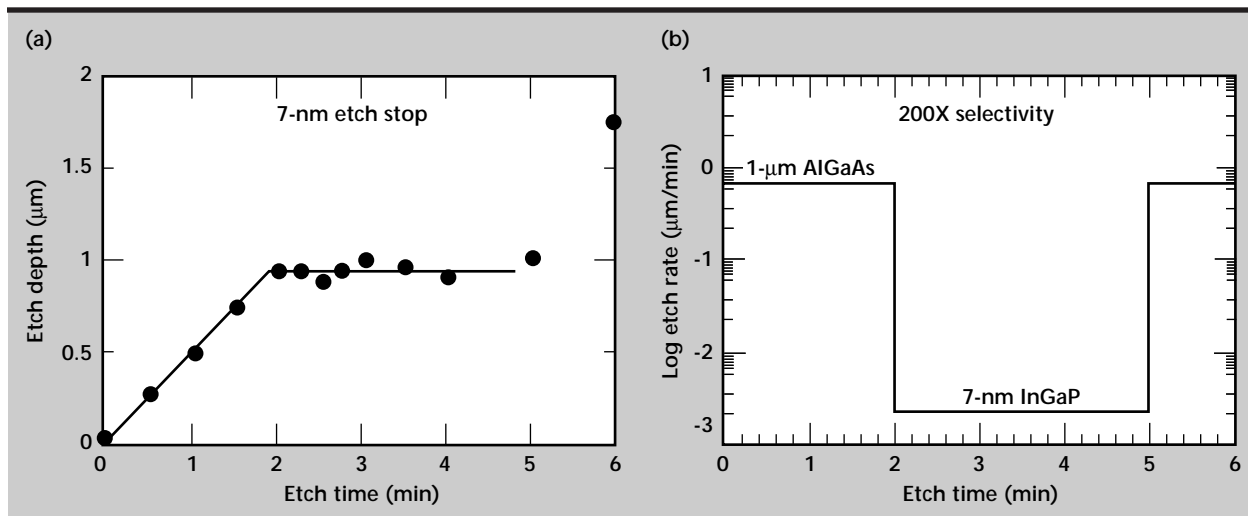
Figure 3. Selective etch results for removing AlGaAs down to a 7-nm-thick InGaP etch-stop layer with 1:8:100 H_2SO_4:H_2O_2:H_2O at room temperature. (a) Etch depth vs time. The 7-nm-thick layer holds off the etchant for 3 min; (b) etch rate vs time, showing etch selectivity greater than 200 times for AlGaAs over InGaP.

purpose, a test wafer was grown that incorporated an InGaP etch-stop layer in the middle of 2 μm $Al_{0.6}Ga_{0.4}As$. The thickness of the InGaP layer was chosen as 7 nm, or about 25 atomic layers. The selective etch used was 1:8:100 H_2SO_4:H_2O_2:H_2O, which etches AlGaAs at a rate of 0.5 μm/min at room temperature. To determine the selectivity of this etch, the test wafer was masked with stripes of photoresist, diced into 2 × 5 mm pieces, and etched for different times in the solution. The photoresist was then stripped from these wafers, and the etch depth was determined using a profilometer.

The results for the etch depth as a function of etch time are shown in **Fig. 3a**. From this figure, it is seen that the 7-nm-thick InGaP etch stop held off the etchant for about 3 min. Thus, stopping the etch any time between 2 and 5 min would yield an etch depth of exactly 1.0 μm and a smooth surface. In practicality, it was found that over-etching by too long affected the width of the ridge-waveguides, and over-etching of about 15 s provided better device results than over-etching by several minutes.

Figure 3b shows the etch rate as a function of time for the test wafer. An etch selectivity of over 200 times for removing AlGaAs with respect to InGaP was achieved.

Following development of the InGaP growth and selective etch recipes, this technology was used to fabricate ridge-waveguide semiconductor lasers and optical amplifiers. The same MBE wafer design is used for both lasers and amplifiers, and they differ only in the angle of the ridge waveguide with respect to the cleaved facets of the device. For lasers the ridges are perpendicular to the cleaved facets, which then act as the laser mirrors, and for amplifiers the ridges were tilted by 7° from the perpendicular to reduce the fraction of reflected optical power to below 10^4.

The MBE-grown wafers consisted of a n-type GaAs substrate, a 1.0 μm n-type GaAs buffer layer, a 1.0 μm n-type $Al_{0.6}Ga_{0.4}As$ lower cladding, a 0.4 μm graded-index single quantum well active region, a 1.0 μm p-type $Al_{0.6}Ga_{0.4}As$ upper cladding, and a 0.1 μm p+ GaAs contact layer (**Fig. 1**). A 7-nm-thick $In_{0.49}Ga_{0.51}P$ etch-stop layer was inserted in the upper AlGaAs cladding 0.2 μm above the active region. The structures were grown at 835 °C thermocouple reading with an arsenic over-pressure of 4.5×10^{-5} Torr, except for the InGaP layer which was grown at 480 °C with a phosphorus over-pressure of 3×10^{-5} Torr. The arsenic and phosphorus sources were both EPI valved crackers, and both cracking sections were operated at 920 °C.

Although our primary interest is in ridge-waveguide SOAs, it is easier to characterize the quality of the MBE wafer by fabricating broad area laser diodes, and the quality of the device processing by fabricating ridge-waveguide laser diodes. The quality of a wafer is determined by the threshold current density for broad area lasers fabricated from the
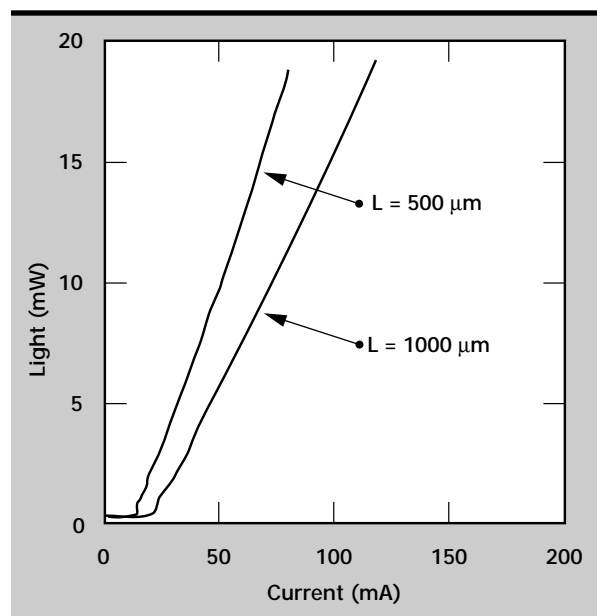
Figure 4. Pulsed (1 μs/1 kHz) light vs current for 500-μm and 1000-μm-long ridge-waveguide 830-nm laser diodes fabricated using the etch-stop technology. Threshold current is 11 mA for the 500-μm-long-device, and 17 mA for the 1000-μm-long device.
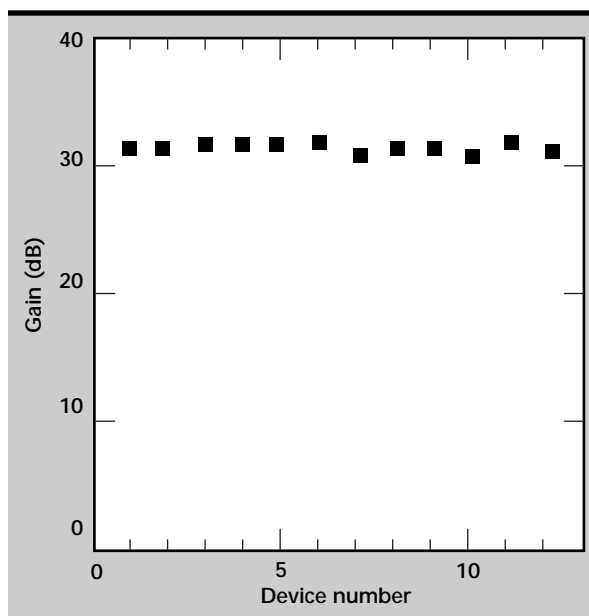
Figure 5. Pulsed (1 μs/1 kHz) chip gain at 100 mA drive current as a function of device number in a 12-element SOA array.

material. The wafers grown for this project showed minimum broad area threshold current density of 159 A/cm^2 for 1-mm long devices, which is an extremely competitive number for unstrained GaAs quantum well lasers such as this. Results from ridge-waveguide laser diodes are shown in **Fig. 4**. The threshold current was 11 mA for a 500-μm long laser, and 17 mA for a 1000-μm long laser. The internal quantum efficiency was measured to be about 95%.

These numbers are comparable to the best results to date for unstrained lasers from our MBE system. This is important because it proves that the addition of phosphorus to the MBE growth chamber has not had an adverse effect on the quality of our AlGaAs/GaAs laser material.

The results of 830-nm ridge-waveguide SOAs are shown in **Fig. 5**. The SOA gain is shown as a function of device number for a 12-element SOA array. The gain was measured as 31.6+/-0.35 dB for these devices (optical gain of about 1500 times). The uniformity of gain across the array is partially attributed to the uniformity of etch depth obtained through use of the InGaP etch-stop layer. These devices are the first SOAs fabricated using an etch-stop technology, and the gain is believed to be comparable to the best results in the literature for SOAs in general.

Future Work

Although this was a one-year project, the phosphorus growth technology developed here will be used in many future applications. It is already being used for the SOAs discussed above, for substrate removal etch-stops, and for 1.55 μm SOAs. Our future plans are to use this technology for active/passive integrated optics and for vertical-cavity laser diodes.

References

1. Walker, J. D., F. G. Patterson, S. P. DiJaili, and R. J. Deri (1996), "A gain-clamped, crosstalk free, vertical cavity lasing semiconductor optical amplifier for WDM applications," *Integrated Photonics Research Conference, Boston, Mass*, **IWD1**-**1**, April 29–May 2.

2. Hobson, W. S., Y. K. Chen, and M. C. Wu (1992), "InGaAs/AlGaAs ridge waveguide lasers utilizing an InGaP etch-stop layer," *Semiconductor Science and Technology,* **7**, p. 1427.

Large Area Lithography

Steve Swierkowski
Engineering Research Division
Electronics Engineering

A general purpose large-area lithographic apparatus and process have been established, enabling the Lawrence Livermore National Laboratory's Microtechnology Center (MTC) to make unique programmatic parts. Etching and bonding have been completed for the first programmatic components from this technology: electrophoresis capillary array plates for use in the BBRP human genome program. This technology should also be suited to several substrate types commonly developed in the MTC that are not amenable to spin coating, even if they are not large. This would include substrates that are uneven or stepped, non-flat or curved, or thick and heavy. It is also ideal for double sided coating.

Introduction

The goals of this project were to design, fabricate, and test apparatus and processing procedures that will enable the MTC to make high resolution (5 to 10 μm) features on very large (18 × 61 cm; [7 × 24 in.]) substrates. The apparatus includes: 1) a photoresist dip coating machine for large substrates; 2) a UV flood exposure system for scanning substrate illumination with UV for contact printing the fine mask features; and 3) bake-out apparatus, developing trays and filters for recycling the large volume (6 L) of photoresist.

The procedures include 1) qualifying a special vendor, 2) obtaining high-resolution photo masks of flat panel display size (20 × 24 in.), and 3) producing high-resolution, full-length capillaries in high-density arrays etched in glass substrates, initially 3 × 18 in. in size, as needed by BBRP program.

Progress

Photoresist Dip Coating

Our first major effort in large area lithography was to design and build an apparatus for coating large substrates with specialized photoresist. The literature and vendors were reviewed and dip coating was selected as the method over other possibilities that include roller coating, spray coating, dry film transfer, meniscus coating, and spin coating. A large apparatus, about 2 m tall and about 1 m wide, was designed and built to hold the substrate and to lower and raise it slowly in a tank of photoresist.

It is essential that the speed is slow and uniform, and that vibration be at an absolute minimum.

The apparatus shown in the background of **Fig. 1** has produced uniformly coated plates as large as 7.5 × 60 cm. The pull rate is about 2.5 cm/min and the photoresist is about 6 μm thick. These values are right on the manufacturers design values. The thickness varies less than a micron; the lateral

Figure 1. Large-area (45 × 55-cm) master photomask, held in the foreground. In the background is the 2-m-tall photoresist dip coating puller apparatus.

spread in optical interference fringes is over several centimeters. The apparatus has easily interchanged photoresist tanks, so different width and depth tanks can be used to match substrate size and thereby minimize photoresist usage.

Photomasks

Producing photomasks for very large substrates is a highly specialized business, mainly focused on the flat panel display industry. By relaxing the normal line width specifications a little and concentrating the other mask specifications on minimizing the number of shorts or opens on a multi-design grating mask, it was possible to obtain a very high quality 45 × 55 cm mask with dense arrays of gratings, including tapered structures, with line widths varying from 20 to 100 μm, comparable spacings, and lengths of 48 cm.

There were no shorts or opens on this plate, which is remarkably high quality. This master photomask plate is shown in the foreground of **Fig. 1**. The five-design plate costs under $5K, comparable to the cost per design for other high resolution masks used for 7.5 cm wafers. This plate was scribed and cleaved into the five separate plate designs within the MTC, with a specialized technique to prevent scratching or contaminating the separate masks with scribe fragments. Adhesive tape and clean room vacuum were used.
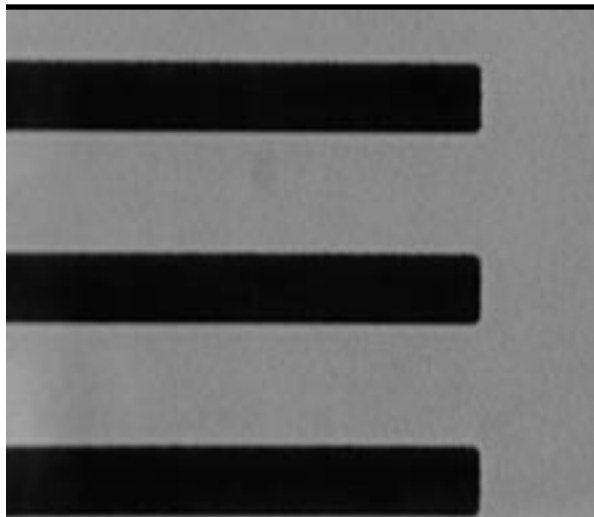
Figure 2. Typical printed features, subsequently etched in Cr/Au, shown from a portion of a 7.5 × 58-cm glass substrate. The line widths are 107 μm on a 300 μm pitch, and the lines are 48 cm long. The resist is positive and etched regions are shown as dark lines.

Mask Alignment

Grating sub-masks 7.5 × 55 cm long were used with manual alignment onto dip coated borosilicate float glass substrates in the contact printing mode. The manual alignment is done with a low power magnifier. Hand adjustment and alignment within 20 μm is easily obtained. The contact is maintained by simple spring steel binder clips around the perimeter. By clamping from one end to the other, it is easy to get very close contact, as evidenced by optical interference fringes and also by the final quality of the printed lines.

Contact Printing

A second major apparatus was constructed to expose the mask/substrate clamped pair. No convenient deep UV sources were available for this large substrate size. An existing MTC deep UV flood source has a field of exposure of about 10 cm. An apparatus was constructed to slowly drag the mask/substrate pair under a slot of light from the existing unit. Adequate exposure is obtained, even for this very thick resist, in about 10 min. This is conveniently done by swinging the current light source head out over the end of its table and mounting the exposure puller apparatus temporarily onto pre-set pins on the end of the table. This exposure apparatus is about 1.5 m long and would stick out into the workspace too much when it is not being used.

Lithographic Results

Initial tests on quarter-sized (to save costs) substrates showed excellent photoresist coatings. Exposure and development were optimized on these smaller substrates with the aid of a special 7.5 cm very high resolution test mask. The development test mask is a top quality e-beam-generated set of gratings and exposure test patterns with line widths from 10 to 80 μm. Even with the 6-μm-thick resist that is specialized for dip coating, it was possible to print the 10-μm test lines with a narrow range of exposure and development. For line widths of 20 μm and larger, it is very easy to print high quality patterns.

A more serious test for the project was to print the full length designs. This was done by printing onto a Cr/Au-coated 7.5 × 58-cm substrate with an actual BBRP program capillary array design. This

design had 101 channels of 100 μm line width with an initial pitch of 580 μm and a tapered middle section, followed by an output section pitch of 300 μm. From the entire process, a few defects were visible on the printed pattern that were caused by particulates, most likely in the dip coating part of the process. Only about three printed lines out of 101 were flawed in width by process defects, but none of the lines were shorted or open.

The printing quality onto the gold film is excellent over the whole plate, as shown in **Fig. 2**.

Acknowledgment

Phase-Shift Lithography

Dino R. Ciarlo
Engineering Research Division
Electronics Engineering

Don R. Kania
Advanced Microtechnology Program
Laser Directorate

We have used phase-shift photomasks to enhance the lithographic capability of steppers.

Introduction

Programs at Lawrence Livermore National Laboratory (LLNL) often find the need for sub-micron lithography. As an example, the Advanced Microtechnology Program (AMP) is working on giant magnetoresistance devices for advance-read heads for the magnetic disk industry. For this application, 0.25- to 0.5-μm lithography is required. In addition, sub-micron pinholes are required for point diffraction interferometers, and sub-micron resolution targets are needed for x-ray tomography experiments. Unfortunately, the state-of-the-art steppers used by industry to do this are very expensive: a 0.7-μm stepper costs $1.5 M; a 0.35-μm stepper costs $5 M. Also, there are not many companies willing to do sub-micron lithography as a service. Consequently, there is motivation to develop techniques to enhance the lithographic capability of available equipment. This can be done with phase-shift photomasks, as we demonstrated in this research project.

Progress

A number of phase-shift techniques have been explored by the IC industry. The one we used is referred to as edge phase-shift masks, and is illustrated in **Fig. 1**. With this technique, there is no opaque chromium on the photomask, as is normal. Instead, a phase-shifter region is etched into the surface of the photomask so that the electric field vector in the illuminating UV is shifted by 180°. To do this, the etch depth, d, has to be equal to:

$$d = 0.5\lambda/(\eta - 1)$$

where λ is the wavelength of the exposing radiation (i line = 0.365 μm) and η is the index of refraction (1.46 for quartz). Using these values, an etch depth of 0.3832 μm is needed for a quartz phase-shift mask. The intensity to expose the photoresist at the wafer is proportional to the square of the electric field intensity. Since the magnitude of the electric

i-line radiation 365 nm

Phase shift photomask

Electric field strength (E) at the mask

Electric field strength (E) at the wafer

Intensity at the wafer ($I = E^2$)
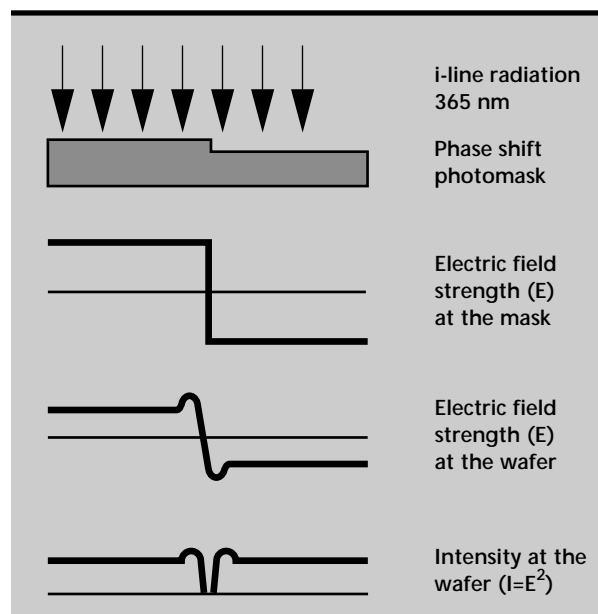
Figure 1. Illustration of edge phase-shift lithography process.

(a)

(b)

0.3 μm post
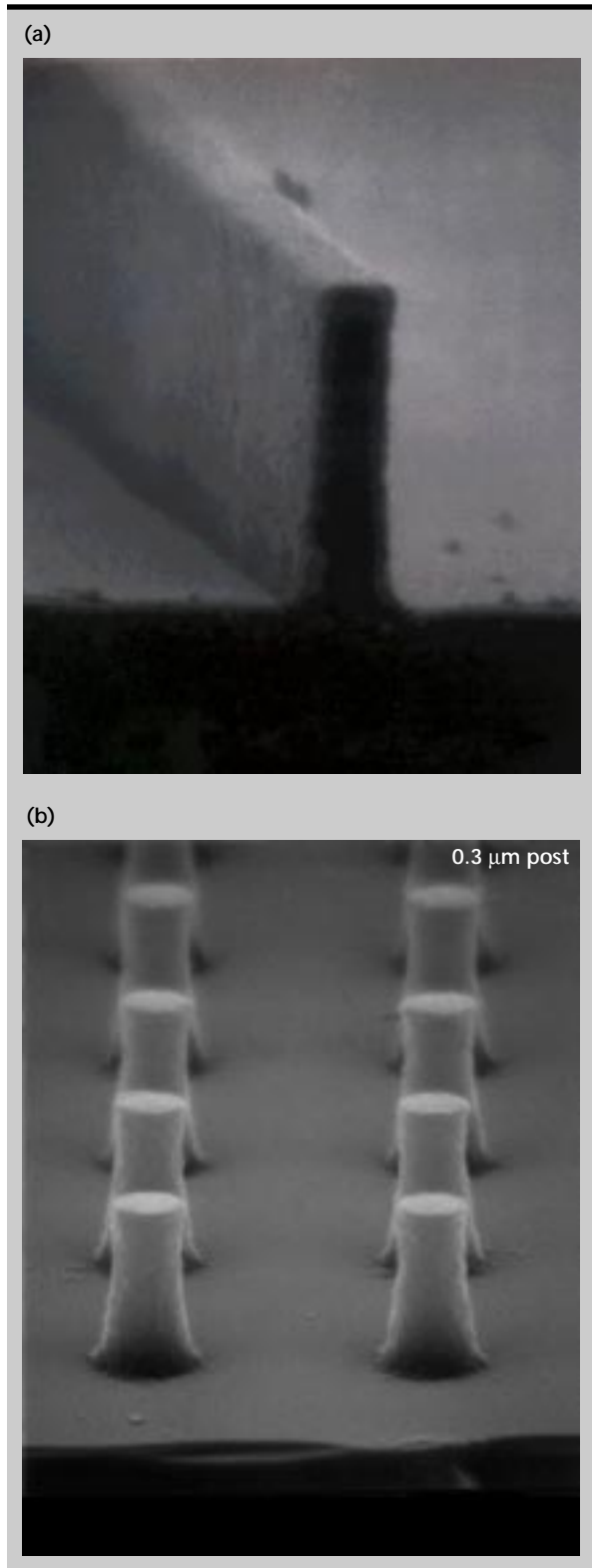
Figure 2. (a) Photoresist line, 0.2 μm wide and 1 μm thick, printed with an edge phase-shifted mask; (b) photoresist posts, 0.3 μm diameter, 1.0 μm thick, printed with a phase-shifted mask.

field vector is forced to pass through zero, the resulting intensity profile is narrower than would be obtained with a conventional chrome photomask. The photomasks we used were fabricated by Du Pont Photomasks for Ultratech's stepper. These photomasks contained a wide range of feature sizes and were fairly expensive, $7.7 K each.

Our experiments were performed at Ultratech, in San Jose using their SATURN 1X stepper, model 2550i, which uses 0.365-μm radiation, and has a numerical aperture of 0.365 and a minimum resolution of 0.7 μm. Our goal was to lower the minimum resolution of this machine by using edge phase-shift masks. We used 150-mm silicon wafers for the experiments. They were coated with resist, exposed and developed at Ultratech. The wafers were then brought back to LLNL where they were sectioned and inspected using our SEM. Many experiments were performed in an effort to optimize the process. The most important variables turned out to be the exposure intensity and the focus. Some modeling, with the PROLITH 2D program was used to guide the experiments.

Figure 2 shows some results of our work. In **Fig. 2a** we see an SEM of a 0.20-μm line of resist which is full thickness, 1 μm. This line was printed from a mask feature consisting of a phase-shifted region adjacent to a non-phase-shifted region. In **Fig. 2b** we see an SEM of a column, with a diameter of 0.3 μm and a thickness of 1 μm. In both cases, the resolution of the stepper is clearly improved. The column in **Fig. 2b** was printed using a 0.75 by 0.75-μm phase-shifted area on the photomask. Some experiments were performed doing a double exposure of a phased-shifted edge. The second exposure was rotated at 90° to the first. These experiments resulted in columns which were full thickness and had a diameter of 0.2 μm.

Future Work

The above results clearly show the advantage of using phase-shift lithography to improve the performance of a projection printer. The features printed consist mostly of straight lines and circular columns. Some experimenters require more complicated features with corners and/or elliptical shapes. These shapes are more difficult to accomplish using the phase-shift approach.

The features printed so far were on bare silicon surfaces. We need to explore this technology on other surfaces, such as those coated with metal and/or containing topography.

Thermally Robust Optical Semiconductor Devices Using AlGaInAs Grown by Molecular Beam Epitaxy

Sol P. DiJaili and Jeffrey D. Walker
Engineering Research Division
Electronics Engineering

Frank G. Patterson and Robert J. Deri
Condensed Matter Physics Division
Physics and Space Technology Directorate

We have proposed and demonstrated a thermally robust semiconductor optical amplifier (SOA) at 1.5 μm wavelength. The novel contributions are the use of thermally robust gain medium at 1.5 μm wavelength and the use of a highly thermally conductive dielectric for the fabrication of the SOA. The devices possessed high gain, over 30 dB, and thermally robust behavior in pulsed operation from 60 to 80 °C.

Introduction

We have addressed the problem of the thermal sensitivity of optoelectronic devices at 1.5 μm wavelength. The reasons for the thermal sensitivity are that the usual semiconductor material used for the active gain region, InGaAsP, suffers from poor electron confinement and Auger recombination, a strongly temperature-dependent non-radiative recombination mechanism. The poor electron confinement is due to small conduction band off-sets of the active region with respect to the surrounding cladding.

Our solution is twofold. First, we use a thermally robust gain material, AlGaInAs/InGaAs, that has shown excellent thermal characteristics for laser diodes at 1.5 μm wavelength. Second, we use a highly thermally conductive dielectric material near the active region that both serves as a waveguide and conducts heat away from the active region where most of the heat is generated. This demonstration of making a high-gain thermally robust SOA using AlGaInAs/InGaAs is a first, to our knowledge.

Indeed, making a SOA thermally robust is more difficult than the demonstration of thermally robust laser diode since the carrier density level is typically a factor of three higher than in a good laser diode design. Thus, Auger recombination is almost a factor of 30 higher, since it depends on the cube of the carrier density. Ultimately, we hope to operate these SOAs without thermoelectric coolers.

Thermoelectric coolers operate with currents on the order of a few amps and often are the reliability bottleneck in an optoelectronic module.

Briefly, we will describe a SOA. The basic function of any optical amplifier is to make the light intensity brighter. In the case of a high performance SOA, the light is made brighter by about one thousand times, or 30 dB. This number is reduced by input and output fiber coupling losses which can be around 6 dB total. **Figure 1** shows a schematic of a ridge waveguide SOA. The light is confined in two dimensions by an optical waveguide. The p-type and n-type semiconductor regions have a lower index than the intrinsic region, thus forming an optical waveguide in the vertical direction. In the horizontal direction, a waveguide is formed by the etching of the ridge. Outside the ridge, the effective index for
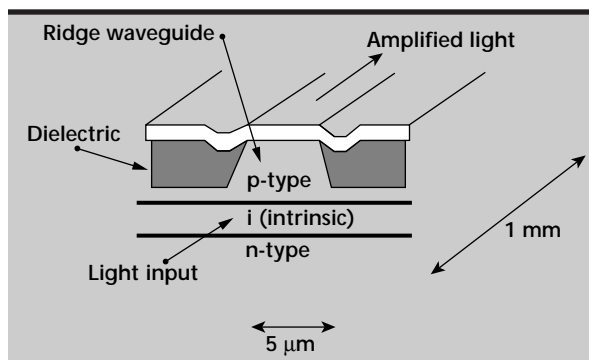
Figure 1. Basic diagram of a semiconductor optical amplifier.

the optical mode is lower than in the center since the index for the dielectric is made lower than the p-type semiconductor. The ridge height is critically adjusted to keep the waveguide single mode in the horizontal direction. The ridge width is typically about 5 μm.

The light beam propagates down the ridge waveguide, approximately 1000 μm in length. Gain is provided by stimulated emission that occurs in the intrinsic or active region. By injecting current through the device, holes are injected into the intrinsic region from the p-type region and electrons, at a higher energy, are injected in from the n-type region. The electrons and holes recombine and give off the energy in the form of light. If a light beam is present, then the release of energy will be "stimulated" and the net effect is a coherent addition of energy to the original light beam. Thus, as the beam propagates down the ridge waveguide it gets brighter and brighter. Ultimately, saturation of the gain due to the signal and spontaneous emission sets an upper limit on the amount of gain that can be obtained.

Optical amplifiers are used to make up for losses in the transmission of an optical signal. The losses can arise from serial transmission of a light beam in an optical fiber. The amplifier is used in a fiber optic link as a booster, regenerator, or preamplifier between the transmitter and receiver. Also, as optical networks become more prevalent, losses can occur from the distribution of the signal to many users or by splitting losses in a switched network. Optical amplifiers can be used to make up for splitting losses as well. For semiconductor optical amplifiers, the advantages over other types of optical amplifiers are in cost, size, and the fact that they can be switched at a rate more than six orders of magnitude faster than the doped fiber amplifiers. The main problem preventing their widespread use has been the problem of cross-talk between different wavelength channels and gain recovery in the time domain. These problems arise from the dynamic gain variations that occur in the SOA. Recently, we have demonstrated a solution to this problem that we believe will enable the widespread deployment of SOAs.[1]

Progress

Figure 2 shows the band diagram for the active regions and surrounding p- and n-type material. The carriers (electrons and holes) are confined to the active region by the increase in the band gap energies between the active region and the surrounding

n- and p-type regions. As the temperature is increased, the resulting electron energy distribution causes the higher energy electrons to surmount the band gap of the n- or p-type regions and leak out of the active region. The effective mass of the electrons is much lighter than that of the holes, and the loss of carriers due to thermionic emission out of the active regions tends to be dominated by the electrons. The quantum wells provide improved carrier confinement and make the stimulated emission of photons more efficient.

In the case of the AlGaInAs/InGaAs, the difference in energies in the conduction band between the bottom of the quantum well and the conduction band in the n- or p-type regions is around 500 meV. This energy difference compares quite favorably to the energy differences in the conventional material used at 1.5 μm wavelength. For InGaAsP active regions and InP n- and p-type regions, the energy difference is around 300 meV.

Also, in the case of the AlGaInAs/InGaAs material system, strained InGaAs quantum wells are used. In an energy versus momentum diagram, this strain causes the light and heavy hole bands to split in energy. The result is that conservation of energy and momentum required for Auger recombination to occur is significantly reduced. Auger recombination is a non-radiative recombination mechanism between the electrons in the conduction band and holes in the valence band, and thus suppresses efficient stimulated emission.

The molecular beam epitaxy (MBE) machine was used to grow the thermally robust material. To grow high quality AlGaInAs films on InP substrates, high quality epitaxial InP needs to be grown. To accomplish this, a solid source phosphide cracker was prepared and installed into the MBE. The film stacks that were grown consisted of a n-type InP buffer, a n-type AlGaInAs region, a graded AlGaInAs region to a lower energy band gap, an intrinsic
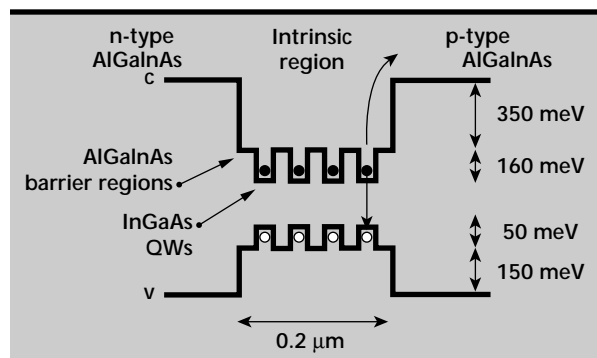
Figure 2. Energy band diagram for the AlGaInAs/InGaAs multi-quantum well heterostructural material.

region consisting of AlGaInAs barriers and InGaAs quantum wells, a graded AlGaInAs region to a higher energy band gap, a p-type AlGaInAs region, a p-type InP buffer region, and a p-type InGaAs cap layer. All the layers were lattice-matched to the lattice constant of InP, except for the InGaAs, where an approximate 0.7% tensile strain was introduced.

An x-ray diffractometer was set up and the lattice match conditions were measured. Several wafers were grown and the MBE growth conditions were adjusted to achieve the necessary lattice match conditions. Broad area lasers were processed from the resulting laser material and the laser thresholds were measured. A laser threshold current density of 400 A/cm^2 was achieved. For MBE grown AlGaInAs/InGaAs, this performance number is the best by a significant margin as compared to currently published work.

The laser wafer material was then processed into ridge waveguides using a self-aligned process developed at Lawrence Livermore National Laboratory (LLNL). The ridges were tilted at 7° with respect to the (01$\underline{1}$) crystal plane. The etch process was developed so that smooth and controlled etches are achieved on InP. A self-aligned and highly thermally conductive dielectric was deposited using a magnetron sputtering machine. The device is self-aligned since the electrical contact opening in the dielectric and the top of the ridge waveguide are intrinsically aligned in the process. This dielectric provides a lower index of refraction so that a waveguide is achieved in the lateral direction. Also, it provides for a highly thermally conductive material that has a significantly higher thermal conductivity than that of regrown InP. The devices are cleaved along the crystal planes. The tilt in the ridge with respect to the crystal facet is used to suppress the coupling of the back reflection into the guided mode. The back reflections into the guided mode should be reduced to about 10^{-4}; otherwise undue spectral gain ripple or lasing can occur. The process is routinely performed and good yields are achieved. **Figure 3** shows a SEM of the completed SOA device.

A plot of the amplified spontaneous emission (ASE) versus wavelength from the device is shown in **Fig. 4**. The output was taken with a multi-mode fiber coupled to the output of the SOA. The signal was measured using an optical spectrum analyzer with a 10 nm resolution bandwidth. The center wavelength is at 1535 nm and the full width at half-maximum (FWHM) is approximately 60 nm, representing about 8 THz of optical bandwidth. The operating current for the device was 300 mA.

To facilitate rapid measurements of the gain of the device, we used a power meter to measure the ASE power output of the device. The formula for ASE power coupled into the guided mode of the SOA is given by $P_{ASE} = 2 h\nu \Delta\nu (G - 1) \chi$, where G is the gain of the SOA, hν is the photon energy at 1.5 µm wavelength, $\Delta\nu$ represents the optical bandwidth of the ASE spectra, and χ represents the amount of inversion in the gain medium and ranges from 1 to 3 dB.
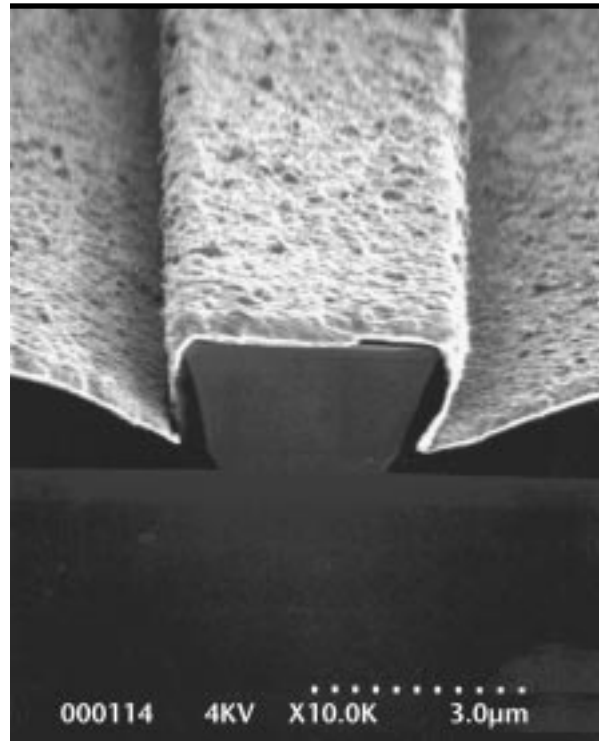
Figure 3. SEM photo of the completed thermally robust SOA using AlGaInAs on InP substrates.
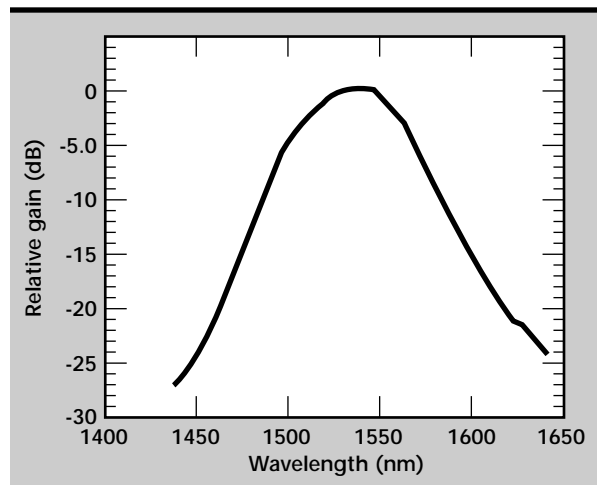
Figure 4. Amplified spontaneous emission spectrum for the 1.5 µm SOA device.
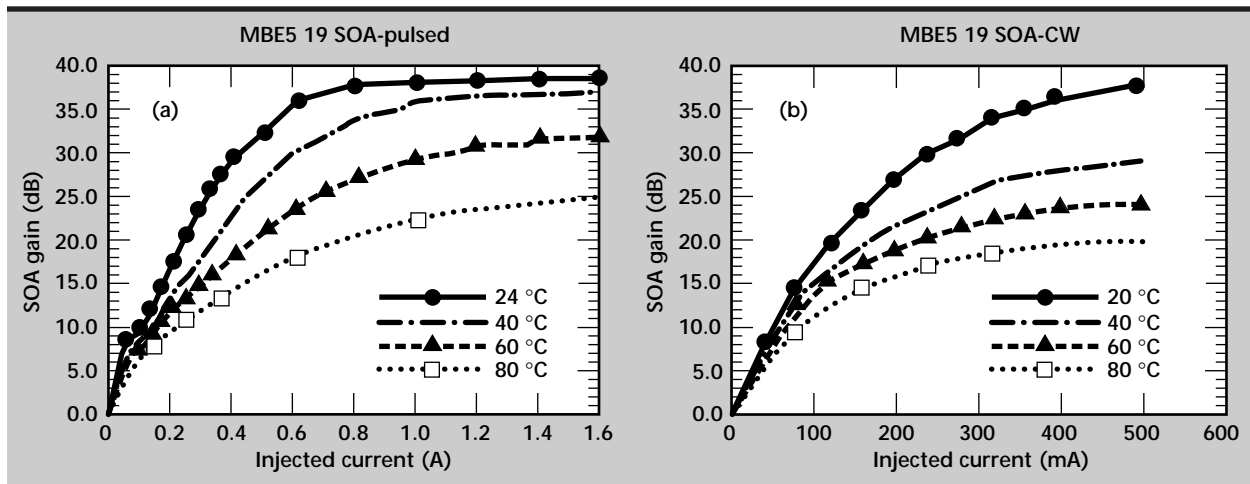
Figure 5. Gain versus current at various temperatures for the 1.5 μm SOA for (a) pulsed and (b) CW operation.

Thus, by measuring the ASE power output, we can obtain a rapid measurement of the gain present. **Figures 5a** and **5b** show the pulsed and CW gain of the SOA device as a function of applied current at various heat sink temperatures. The current pulses were 1 μs in duration at a 1 kHz repetition rate. At room temperature, the SOA device showed a maximum gain of over 35 dB pulsed and 28 dB CW. At 60 °C, the device showed 32 dB gain pulsed, and 17 dB CW. At 80 °C the pulsed gain is still over 20 dB. The pulsed data with temperature being varied indicate the performance of the active region alone. The drop in gain due to CW operation indicates how well the heat is being extracted from the gain region.

Unfortunately, the current consumption is quite high. We are taking steps to reduce the current by redesigning the optical mode so that a narrower ridge width and adequate ridge height to reduce current leakage is achieved while still maintaining the effectiveness of the anti-reflection tilted facets. We are confident that we can achieve a low current consumption device and have the CW gain versus current data approach the pulsed data to within several dB. It should be pointed out that in spite of the high current consumption, the devices, as they stand, can operate to 60 °C and still have an adequate gain.

Future Work

Next year we look forward to improving the device results and then achieving polarization-independent performance, which will allow us to start field testing the thermally robust high performance SOA devices at 1.5 μm wavelength in real world optical data links.

Acknowledgments

The authors would like to acknowledge the diligent efforts of B. Goward, H. Petersen, and P. Stefans.

References

1. Walker, J. D., F. G. Patterson, S. P. Dijaili, and R. J. Deri (1966), "A gain-clamped crosstalk free vertical cavity lasing semiconductor optical amplifier for WDM applications," *Integrated Photonics Research Conf., (I WD1), Boston, Mass.*, April 29–May 2.

Porous Silicon Formation and Characterization

Norman F. Raley, Conrad M. Yu, and M. Allen Northrup
Engineering Research Division
Electronics Engineering

We have formed various porous silicon films using electrochemical and chemical etching techniques and have measured basic properties of the films, as well as adsorption/desorption and gas transport characteristics. We have also developed fabrication procedures for porous silicon membrane structures on silicon and pyrex.

Introduction

Porous silicon has attracted significant interest recently, primarily due to its light emission characteristic. However, the material can have many other applications for microelectromechanical system (MEMS) devices, including use as gas sensing or fluidic elements, which have been explored in this study.

Progress

Equipment

The pore diameter of porous silicon is critical to the resultant film properties. In this study we have considered pore size diameter varying from about 10 nm, 100 nm and up to the 1 µm range.

The smallest pore material was formed using an existing conventional electrochemical etch apparatus with single-cell anodization.[1] Current is applied to a platinum wire immersed in a HF solution in a teflon beaker which flows to a (100) silicon substrate mounted with an O-ring seal opposite the wire. The largest pore material was formed using wet etching of (110) silicon in a heated KOH solution.

Intermediate pore material is possible using either technique. For electrochemical formation, a modification of the etch apparatus was made to provide backside illumination without heating.[2] This was accomplished using a small tungsten lamp with an infrared filter and cutting away the substrate backside contact plate except at the wafer edges. Packing of the bottom of the teflon beaker was necessary to prevent light penetration into the etch solution. A simple silicon cover was placed on top of the beaker to prevent background light input.

Film Formation

Most of the films made were electrochemically formed without backside illumination using a HF etch solution on p+-type substrates. This procedure is expected to give approximately circular cross-section pores at about 10 nm diameter. The heavily doped substrates were chosen as a compromise between high porosity and the mechanical integrity of the films.[3] The formation rate was about 1.8 µm/min, with no apparent area or voltage dependence. Reproducibility of the etch rate was verified by consistent etch weight loss within 5%, with repeatable etch solution desorption, also measured by weight. Film thicknesses from about 50 to 150 µm over a 4.35-cm-diameter area were made. The thicker films were used for adsorption measurements on a weight balance.

The films formed using chemical etch were made with a KOH etch solution on p-type substrates. The pore pattern was an alternating array of 5-µm rectangular slots aligned parallel to the (110) wafer flat. Film thicknesses up to 60 µm were made.

Only initial attempts were made to form circular pores of intermediate to large size using the electrochemical etch with backside illumination. Wafers with n-type dopant and appropriate resistivity were acquired to form 0.2 to 2.0 µm diameter pores. These experiments are on-going. Resist patterning for formation of circular holes and rectangular slots of intermediate size are not trivial and are under review.

Film Basic Properties

The porosity of the small diameter pore material was measured at 32%, using weight measurements before and after formation and removal. The etch rate was extremely fast in KOH wet etchant at room temperature, at a value of 25 μm/min, about 1250 times faster than that of crystalline silicon. However, the etch rate was reduced using CF_4-O_2 dry etch to a value of 3.6 μm/min, only 18 times faster than that of crystalline silicon. The difference is presumably due to the more limited chemical etch component of the dry etch, with corresponding reduced attack on a smaller value of surface area. Film cleaning in standard RCA-1 solution gave some oxidation, but no measurable weight change. The surface of the porous silicon was hard and did not scratch during typical process steps.

The porosity of the large diameter material was about 70% by SEM observation. Etch rate in wet KOH or dry CF_4-O_2 is virtually identical to that of crystalline silicon. Film cleaning in RCA-1 solution had no effect.

Film Adsorption/Desorption

The adsorption/desorption characteristic is important for gas sensing applications. Adsorption measurements were made at room temperature by exposure of thick porous silicon films, made with small pore diameter, to the vapor pressure of a particular liquid under sealed conditions, followed by weight increase measurements on a balance. Desorption was generally measured at room temperature also, by weight loss. Typical liquids used were water and ethanol.
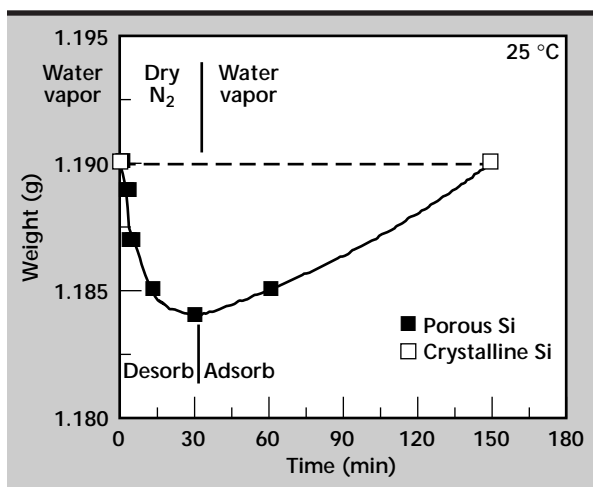
Figure 1. Adsorption/desorption characteristic for small-pore size porous silicon film of 150 μm thickness and 4.35 cm diameter.

Initial data indicates that the calculated amount of adsorption is about 10 to 40% of film saturation. The adsorbed capacity of the porous silicon film for water was about 32 mg/cm^3 for saturation, whereas that for ethanol was at least 150 mg/cm^3, since saturation was not believed to be obtained in the latter case. The adsorption time could be as low as 30 to 60 min, also typical for desorption.

A measured adsorption/desorption characteristic is shown in **Fig. 1**. The actual measured adsorption flux for water was about 0.3 mg/cm^2-h, whereas that for ethanol was higher, at 0.9 mg/cm^2-h. The desorption flux for water was about 1.1 mg/cm^2-h, whereas that for ethanol was again higher, at 1.7 mg/cm^2-h. The difference between the water and ethanol rates can be most likely explained by their corresponding difference in molecular weights. Note that the desorption for water was observed to increase significantly upon heating at 100 °C to give 10.1 mg/cm^2-h flux.

We hope to decrease the adsorption/desorption times to about 1 to 2 min for exposure at atmosphere, as opposed to liquid vapor pressure at room temperature, which may be more practical for a useful application. Initial reproducibility tests indicate that desorption is repeatable for ethanol, except for a base weight shift observed after multiple vapor exposures. This shift may be due to residues from the ethanol vapor and not from a change in the adsorption characteristic of the material.

Film Gas Transport

The time for gas transport through a porous silicon membrane is important for flow-through sensing applications to keep the sensing time realistic. Lower pressure transport across a membrane made with small pore material was first measured using liquid ethanol as a test vapor source. The membrane was exposed to vapor on one side of the membrane, and an additional porous silicon layer on silicon substrate on the other side of the membrane was used to monitor vapor transport by adsorption. The monitor wafer was removed after exposure and its weight increase measured on a balance. The transport time was determined by this method to be about 1 to 2 min for a 50-μm-thick membrane, with linear increase of adsorbed amount in the monitor for further exposure time.

Higher pressure transport measurements closer to atmosphere were then addressed. A test set-up was devised and assembled to permit pressurization of one side of a porous silicon membrane with nitrogen through a shut-off valve. The valve is then closed and

the pressure drop monitored with a pressure gauge. The minimum gas transport rate possible is about 10^{-4} cm^3/min/psi, as limited by the leak rate of the set-up. The flow rate through a 2-mm-square membrane made with small pore material was measured to be linear with pressure and equal to about 10^{-2} cm^3/min at 10 psi.

Film Fluid Transport

The time and amount of fluid transport are important for fluid flow applications. A test set-up was devised to permit filling of a flow line with liquid on one side of a membrane. The fluid is then pressurized and the amount of fluid transport is monitored by optical microscope observation on the other side of the membrane. Heating of the membrane to see its effect is done by exposure to a hot air gun or by use of an integral heater element adjacent to the porous silicon membrane. Measurements on actual membranes are planned.

Flow Rate Calculations

The gas or fluid flow rate though a porous silicon membrane can be simply estimated using the equation for flow velocity in the pores, appropriate for an incompressible fluid.[4] The flow rate is then equal to the product of pore velocity, pore area and the number of pores. Assuming a pore diameter of 10 nm, a pore surface-to-volume ratio of 2.3×10^6/cm, and a membrane geometry of 2 mm^2 with 45% conversion and 50 μm thickness, the pore area is 7.8×10^{13} cm^2, and the number of pores is 1.3×10^{10}.

For air gas, the viscosity is 1.85×10^{-5} N-s/m^2, which gives a velocity of 2.3×10^{-2} cm/s for a pressure of 10 psi, or an air flow rate of 1.4×10^{-2} cc/min. Note that the calculated flow rate is quite close to that measured at 10^{-2} cc/min.

This result helps to justify the use of the equation of velocity for an incompressible fluid for a compressible gas.

For water fluid, the viscosity is higher, at 0.9×10^{-3} N-s/m^2, which gives a velocity of 4.7×10^{-4} cm/s at 10 psi or a water flow rate of 2.9×10^{-4} cc/min. For applications where about 1 nL volume is required, the pressure activation time is then 0.2 s, which is quite reasonable.

Microstructures: Membranes on Silicon and Pyrex

Flow-through gas sensing or fluid flow requires that the sense gas or fluid be drawn through a porous silicon element. A likely implementation of this element is a membrane on either silicon or pyrex substrates. The membrane would ideally be made from small diameter pore material, and be relatively thick to enhance adsorption sensitivity for gas sensing applications, for example.

A special fabrication procedure was developed to make the membranes on silicon, as shown in **Fig. 2a**. Membranes with small pore material were demonstrated, with a 5 mm × 5 mm area as well as a 2 mm × 2 mm area, both with 50 μm thickness. The achievable conversion from crystalline to porous silicon ranged from 25 to 75%. No areas of optical transmission were visible under microscope view. Membranes with large pore material were also demonstrated.

Some bow of the membrane with small pores was observed after formation. The center deflection of the membrane ranged from about 60 μm for the 5 mm area geometry to as little as 4 μm for the 2 mm area. The bow is presumed due to compressive stress in the membrane, caused by native oxide growth as the oxide thickness is comparable to the dimension of the silicon fibers for the small diameter pore material.[5] No bow was observed for membranes with large pores.
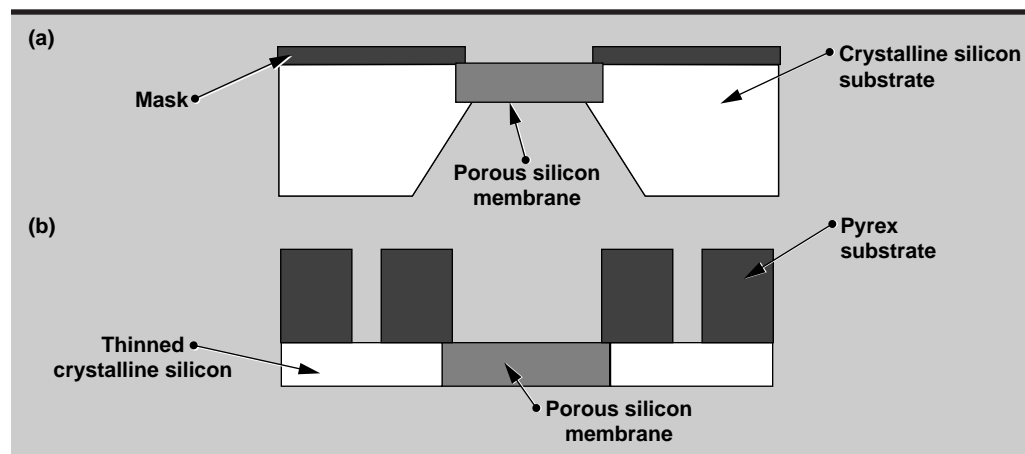
Figure 2. Porous silicon membrane structures fabricated on various substrates: (a) silicon and (b) pyrex. Electrode backside contact to silicon required for electrochemical etching is made directly in (a) and through holes in the pyrex in (b).

A special fabrication procedure was also devised to make membranes on pyrex, as shown in **Fig. 2b**. Tests of all fabrication steps indicate that membrane formation of large-diameter pore material using a KOH etch solution over a 5-mm diameter area at 50 μm thickness is realistic. More experiments to form the small-pore size material on pyrex substrates using electrochemical etching are needed.

Summary

Usable formation procedures for both small- and large-pore size porous silicon have been demonstrated. Significant adsorption has been measured on small-diameter pore material. Vapor transport at low pressure and gas transport at high pressure have been achieved as well. Test apparatus has been devised for high pressure gas and fluid transport. Fabrication procedures for porous silicon membranes on both silicon and pyrex substrates have been developed.

Future Work

More experiments are needed to complete characterization of basic properties for small-size porous membranes with small- as well as large-diameter pores. Further measurements on specific film properties, such as atmospheric pressure gas and fluid transport, are planned. Longer term efforts to produce intermediate pore size material, which is more difficult to make, are on-going.

Acknowledgments

References

1. Jung, K. H., S. Shih, and D. L. Kwong (1993), "Developments in Porous Si," *J. Electrochem. Soc.*, **140**(10), p. 3046.

2. Ottow, S., V. Lehmann, and H. Foll (1996), "Processing of Three-Dimensional Microstructures Using Macroporous n-type Silicon," *J. Electrochem. Soc.*, **143**(1), p. 385.

3. Herino, R., G. Bomchil, K. Barla, and C. Bertrand (1987), "Porosity and Pore Size Distributions of Porous Silicon Layers," *J. Electrochem. Soc.*, **134**(8), p. 1994.

4. Fung, Y. C. (1969), *Continuum Mechanics,* Prentice-Hall, Englewood Cliffs, N.J., p. 230.

5. Bai, G., K. H. Kim, and M.-A. Nicolet (1990), "Strain in porous Si formed on a Si (100) substrate," *Appl. Phys. Lett.,* **57**(21), p. 2247.

Kenneth L. Blaedel, Thrust Area Leader

The mission of the Manufacturing Technology thrust area at Lawrence Livermore National Laboratory (LLNL) is to have an adequate base of manufacturing technology, not necessarily resident, to conduct future business.

Our specific goals continue to be (1) to develop an understanding of fundamental fabrication processes; (2) to construct general purpose process models that will have wide applicability; (3) to document our findings and models in journals; (4) to transfer technology to LLNL programs, industry, and colleagues; and (5) to develop continuing relationships with the industrial and academic communities to advance our collective understanding of fabrication processes.

The environment in which the thrust area operates includes Laboratory Directed Research and Development (LDRD), Cooperative Research and Development Agreements (CRADAs) with outside industry, and development work that is performed and paid for within an LLNL program. In general, LDRD incorporates the higher risk research associated with manufacturing technology, while the CRADAs embody the lower-risk reduction-to-practice associated with industry. The ideas developed in the thrust area fall somewhere in between.

Our niche continues to be bringing our strengths in precision engineering to bear on problems in manufacturing, particularly trying to increase the precision-to-cost ratio of some of the processes and devices that we think important to our future. The motivation for this is that the demand for the utmost in precision is now accompanied by a demand for less expensive precision. CRADAs have presented this demand over the past few years. Recently, as support for CRADAs has waned, work for LLNL programs has waxed, with surprisingly similar demands. Thus, application of manufacturing technology, the lower-risk end of the spectrum, has shifted from CRADA partners to LLNL programs.

Manufacturing Technology

4
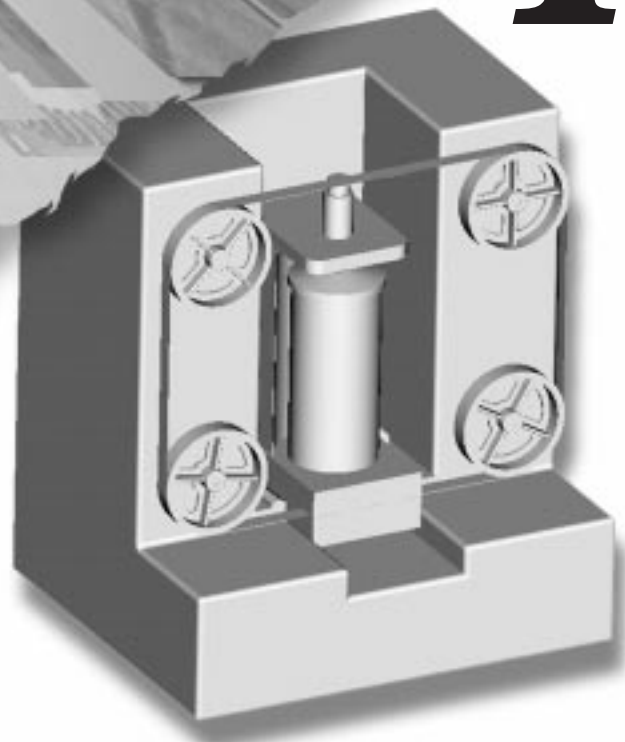
4. Manufacturing Technology

Design of a Precision Saw for Manufacturing

Jeffrey L. Klingmann
Manufacturing and Materials Engineering Division
Mechanical Engineering

A precision saw, similar in concept to a common bandsaw, for slicing crystallographic materials is investigated. The saw, which uses a hydrodynamic bearing for improved cut precision, would allow a significant reduction in the number of secondary finishing operations in the production of, for instance, KDP crystals for the Lawrence Livermore National Laboratory (LLNL) Laser Program, and silicon wafers for the semiconductor industry.

Introduction

The precision slicing of crystalline materials is a technology that presents an opportunity to reduce the cost of fabricating KDP crystals for the National Ignition Facility (NIF). If the precision of the first processing stage of the KDP boule can be improved significantly, costs will be reduced through elimination of subsequent fabrication steps for the approximately 800 required crystals.

A similar opportunity exists in the semiconductor industry, which is currently increasing the diameter of the typical silicon wafer from 200 to 300 mm, and planning a further increase to 400 mm. Drivers in this process are not only precision, but also reduced kerf width and high throughput. We have instituted a project to develop a precision saw to meet the needs of these two applications.

Progress

The three focus areas of the Precision Saw Project are blade development, prototype testing and conceptual machine design. During FY-96, progress was made in blade development and in the design features of a production machine. Testing of prototypes is discussed in the Future Work section below.

Blade Development

The hydrodynamic bandsaw blade design can be separated into the hydrodynamic lubrication analysis of the bearing and the cutting features on the leading edge of the blade. In hydrodynamic design, we have continued work started in FY-95 with the modeling of the fixed-incline bearing, and extended this effort to include a parallel-step bearing design that will be much easier to fabricate.

Since existing designs use only tension to maintain blade straightness in the cut, the tensioned blade or wire provides a useful comparison for the hydrodynamic design.

Using 2-D hydrodynamic models,[1] we calculated the restorative force as a function of disturbance offset (**Fig. 1**) for the two bearings and the tensioned blade. The slopes of these curves represent the stiffness of the blade in the cut. The hydrodynamic bearing in this calculation has 12-μm-tall features and a nominal 50-μm gap, while the tension on the blade or wire is assumed to be 70,000 N.

Since the models used are only 2-D and do not include the effects of leakage to the sides of the wedge (to the front and back of the blade), the actual stiffness of the hydrodynamic bearing would
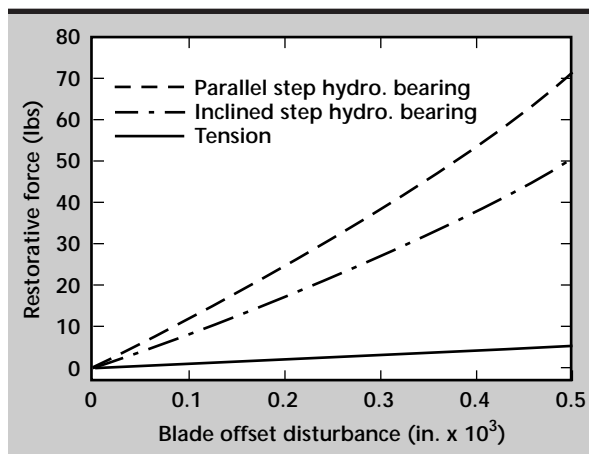
Figure 1. Comparison of restorative forces on the blade.

be somewhat less than indicated in **Fig. 1**. However, the high stiffness of the proposed blade design, compared to just a tensioned blade, should minimize the effect of disturbances that would cause poor cut precision. Beyond these two cases, computational fluid dynamics (CFD) methods will have to be used to examine 3-D effects and to consider alternative bearing designs such as the "herringbone."

The other aspect of the bandsaw blade is the design of the cutting edge and how this feature is fabricated. While the cutting parameters such as diamond size, surface speed, and feed-rate are important, just as critical is a reliable method of fabricating the cutting edge on a metal band with the bearing features previously etched on its surfaces. It is desirable that the width of the cutting edge be small to keep the kerf small, and that the diamonds on the edge be durable for reasonable productivity and cost-effectiveness. Also, the diamond coating should be uniform and symmetric on the blade to minimize cutting errors.

Three methods of producing the blade cutting edge were investigated. The first is the industry standard which is to electroplate nickel with diamond onto the edge of the blade. The limitations to this process are that typically large diamonds are used in a fairly thick layer, resulting in a larger kerf. Additionally, the diamonds are held physically into the nickel coating and are therefore easily detached from the blade.

The second method uses a process developed at LLNL to apply a single layer of diamonds to a substrate with a much stronger chemical bond. This method has been tested for other applications and produces a very thin, durable abrasive coating due to the chemical bond to the substrate. Investigations into the applicability of this method are on-going.
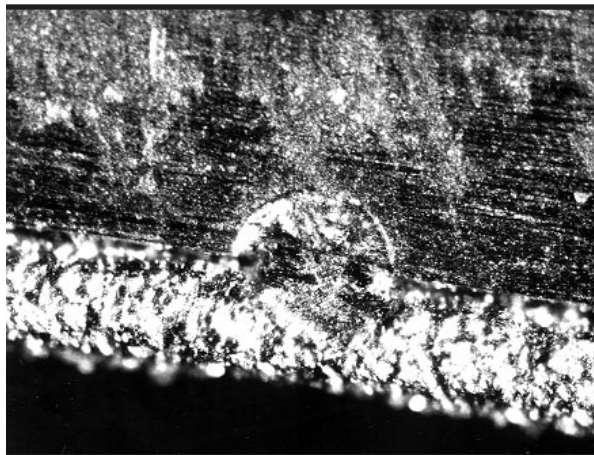
Figure 2. Diamond-impregnated wire brazed to a stainless steel substrate (100×).

Finally, the third method is to physically bond a fine diamond-impregnated wire to the leading edge of the blade. Typical wire saws for the silicon wafer saw industry use a loose abrasive process. However, for other applications, Laser Technology Inc. produces a fixed abrasive wire. **Figure 2** shows a 250-µm wire with 15-µm diamond abrasive that was laser-brazed to the leading edge of a 100-µm stainless steel band. Cutting and durability tests are continuing on this sample, but it seems to be a viable option for a prototype saw blade to test, for instance, the effect of diamond grit size on performance. The chemically-bonded process is preferred for production blades.

Conceptual Machine Design

Attempting the design of a production precision saw at this early stage of the project forces designers and engineers to consider the entire problem of slicing crystalline materials, not allowing a narrow focus on only a few aspects. Additionally, a conceptual machine design has use in describing various aspects of the project to industrial contacts. The following list of machine design requirements and features have been identified.

1) Boule and wafer fixturing. Due to the potential for large forces being generated by the hydrodynamic blade bearing, the wafer must be substantially supported to prevent fracture. If this fixturing for the boule and wafer can be a structurally stiff loop that also contains a guide for the blade entering and exiting the boule, the wafer thickness variation and flatness can be improved. This fixturing must also allow the boule to be oriented so that cuts are made in the proper relation to the crystallographic plane.

2) Axis configuration. Traditional bandsaws have the blade wheels rotating about horizontal axes and the part translating on a linear horizontal axis. Our current thinking maintains this approach but uses four wheels and slices on the blade span between the bottom two wheels; hence the symmetry axis of the boule is now vertical, and parts may be removed from the bottom of the saw.

3) Immersed cutting. To maintain a constant water supply to the hydrodynamic blade bearing, it is preferable to immerse the cutting region in a bath of cutting fluid. Of course, a consideration is the frictional loss due to the viscous fluid action on the rotating wheels.

4) Productivity. For a silicon slicing application, the saw must have high throughput to match competitive technologies. The productivity of

the current machine design will likely come from its improved performance, obviating the need for secondary finishing steps. The current industry solution for 300-mm silicon wafers is the diamond slurry wire saw which appears to be capable of producing 300 to 3000 wafers/day (12 to 120 wafers/h). It appears likely that a precision saw can be designed to cut multiple (5 to 10) boules simultaneously, so that a throughput of 60 wafers/h is possible.

5) Cutting force support. A current unknown is how much force is generated in the in-feed direction during cutting. This force must be resisted to keep the blade on the wheels. The desired solution is to use actively controlled magnets and a magnetic blade material to maintain blade position.

A simplified view of the proposed machine design incorporating these ideas is shown in **Fig. 3**.

Future Work

The most critical task to pursue in the future is to test the hydrodynamic and cutting characteristics of the blade. The current plan begins with blade testing in a guide with the hydrodynamic features in replaceable guides and mechanical wheel bearings. The next steps would be to incorporate the desired features into the blade with a photo-etching process and to design and fabricate hydrostatic bearings for the blade wheels. Finally, a cutting blade would be tested on a machine with a linear axis to translate the boule.
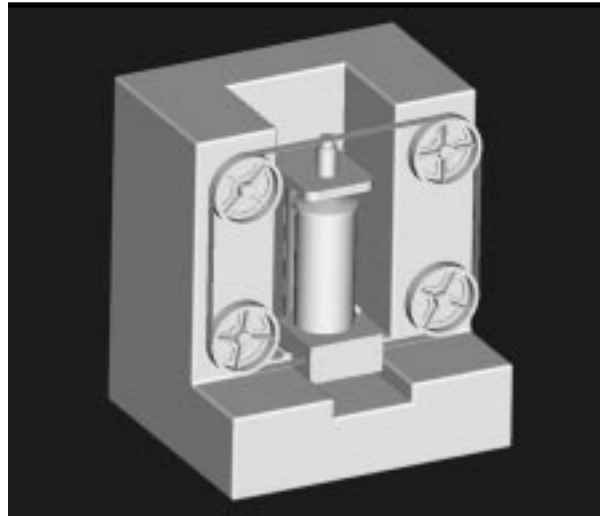
Figure 3. Simplified view of a conceptual production precision saw.

Modeling of the hydrodynamic features should continue with a CFD code to perform detailed design on both simple bearing configurations, such as the parallel and inclined step, and more advanced concepts such as a herringbone design. A conceptual machine design should also be developed to examine critical aspects of a production machine so that necessary concepts can be tested on a prototype machine.

References

1. Hamrock, B. J. (1994), *Fluid Film Lubrication*, McGraw Hill Co., New York, N.Y.

Deposition of Boron Nitride Films via PVD

David M. Sanders and Steven Falabella
Material Fabrication Division
Mechanical Engineering

Daniel M. Makowiecki
Materials Science and Technology Division
Chemistry and Materials Science

A neutral beam-assisted deposition system that has demonstrated the potential of producing high-quality films at reduced stress was assembled and evaluated in a chamber in the Vacuum Processes Laboratory of Lawrence Livermore Laboratory (LLNL). Transparent boron nitride (BN) coatings were produced, with some varying crystal structure, as evidenced by absorption differences in FTIR scans of films produced with varying amounts of neutral bombardment. However, no conclusive evidence of the cubic phase was noted. The films produced were very adherent to silicon substrates and exhibited low residual stress, indicating that this process may be able, with additional optimization, to produce technologically useful cubic BN films.

Introduction

One of the strengths of LLNL in the precision manufacturing field is its diamond turning capability. Rotationally symmetric workpieces can be machined to tolerances of just a few microinches. However, one of the severe limitations of this technology is that only a select few non-ferrous metals (including copper, gold, aluminum, and nickel) can be machined with surface roughness less than 5 nm with a single-crystal diamond tool. The high quality workpiece surface results from the very sharp initial edge of the single-crystal diamond, and the resistance of the diamond edge to wear against these metals. In contrast, carbide-forming metals such as steel, titanium, and beryllium cannot be similarly machined, because they cause excessive chemically-induced wear of the diamond.

There is great demand for precision turning of workpieces made from carbide-forming metals. Within LLNL, for example, precision components of steel and beryllium are of great interest. Within the commercial sector, the steel bearing process would benefit economically if bearings could be turned instead of ground. Within the Department of Defense, there are many current and future components which would benefit from high precision turning, such as mirrors machined directly from silicon.

Boron-based coatings are also of interest for other applications, including electronics, optical filters, wear-resistance, and high-hardness applications.

Our goal for this project is to turn high-precision, very smooth surfaces on carbide-forming materials, particularly steel, through the use of a hard-coated tool. This goal is distinguished from many other hard-coated tool efforts that are directed not at the finishing stage, but rather at the intermediate or roughing stage of manufacture. In this finishing domain, a coating thickness of perhaps 1 μm is considered a thick coating, and a tool edge radius of 1 μm is considered dull. The reward for success is that the large investment we already have in the infrastructure to perform diamond turning could be brought to bear on other desirable materials with a small incremental cost of a coated tool.

The most successful method, ion-beam assisted deposition, has a problem with stress delamination limiting the thickness of films. We used two methods to produce BN films. The first, sputtering of a boron target in a nitrogen atmosphere is a continuation of last year's work, and the second, e-beam evaporation of boron while simultaneously bombarding the substrate with an energetic, neutral nitrogen beam, is documented in this report. We will refer to the latter approach as neutral-beam assisted deposition (NBAD). The sputtering work was done in

LLNL's Chemistry and Materials Science Department, while the NBAD work was done in its Vacuum Processes Laboratory.

Progress

A NBAD system that has demonstrated the potential of producing high-quality films at reduced stress[1] was assembled and evaluated in a chamber in the Vacuum Processes Laboratory at LLNL. An Ion Tech 3-cm-diameter ion source and custom neutralizer cone was used to focus the neutral nitrogen onto the substrates held on a radiatively-heated stage. A Balzers e-beam system was installed in the chamber to evaporate the boron. These two systems were assembled for the first time in this chamber. Since it took much effort to get to a point of reliable operation there was insufficient time to optimize the deposition process.

Transparent BN coatings were produced, with some varying crystal structure as evidenced by absorption differences in FTIR scans of films produced with varying amounts of neutral bombardment. However, no conclusive evidence of the cubic phase was noted. The films produced were very adherent to silicon substrates and exhibited low residual stress, indicating that this process may be able, with additional optimization, to produce technologically useful cubic BN films.

Description of Process

The formation of cubic BN (c:BN) has been shown to require three conditions: a source of boron, a source of energetic nitrogen, and a substrate temperature in the range of 400 to 800 °C. A photograph of the system is shown in **Fig. 1**. A Balzers e-beam evaporator was used as the source of boron. The evaporation rate was monitored and controlled by an Inficon IC6000 controller and quartz crystal monitor placed near the substrate. The evaporation rate was varied in the 0.5 to 2.0 Å/s range for this series of experiments.

The energetic nitrogen was supplied as a neutral beam. This beam was produced by directing an ion beam onto the inside of a cone. An Ion Tech 3-cm, gridded ion source was used, with hollow cathode

Figure 1: Photograph of apparatus used for NBAD of BN films.
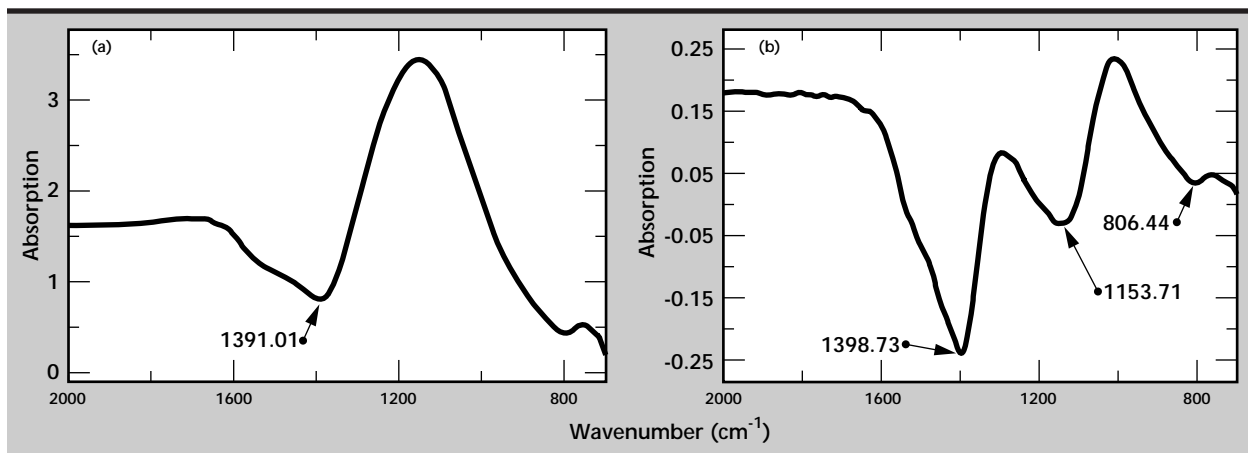
Figure 2. (a) FTIR scan of typical BN film produced by NBAD system. The absorption near 1390 cm^{-1} is indicative of the hexagonal phase. (b) FTIR scan of an area that shows absorption near 1150 cm^{-1}, which is evidence of sp^3 bonding.

plasma source using a 50/50 nitrogen/argon gas mix. Inside the outer cone, there was a smaller, inverted cone to block all line of sight to the substrate. The cone provides surface neutralization of the ion beam, while the particle energy is reduced only slightly due to the glancing angles the ion beam makes with the cones. The ion energy was varied from 300 to 550 eV and ion currents from 25 to 50 mA over the course of these experiments.

The 2-in. silicon wafers used as substrates were held at the exit of the neutralizer cone, in the evaporation plume of the boron source, on a heated stage. The temperature of the stage was varied from 400 to 550 °C for this series of experiments. The stage has the capability of biasing with either rf or dc, but this was not used for this work, as the goal was to use neutral nitrogen bombardment of the substrate.

Results

As noted above, transparent BN films were produced using the above equipment over a fairly wide range of conditions, but in the limited time available, conditions that produce the cubic phase were not found. The crystal structure of the films was determined using the absorption characteristics measured using FTIR spectrometry in the 700 to 2000 cm^{-1} range. A representative scan is shown in **Fig. 2a**. Most of the films demonstrated a large, broad absorption at 1390 cm^{-1} and smaller peak near 800 cm^{-1}, indicative of sp^2 (hexagonal) bonded BN. As shown in **Fig. 2b**, one sample exhibited absorption at 1150 cm^{-1}, which is close to the cubic phase signature of 1060 cm^{-1}, but we feel that the shift is too great to be explained by intrinsic stress, and may be due to a wurtzite phase.[2] Although it is sp^3 bonded, this phase is metastable and not the hard, cubic phase useful in tool bit coatings.

Conclusions

Although not demonstrated using NBAD, there were indications that the system could produce c:BN films of good adherence and low enough stress to make coatings thick enough for tool bit use.

Acknowledgments

We wish to thank R. Sanborn of Analytical Sciences Division for producing the FTIR scans of our films, and S. Bryan of MMED for assembling and troubleshooting the Balzers e-beam system.

References

1. Lu, M, A. Bousetta, R. Sukach, A. Bensaoula, K. Walters, K. Eiperssmith, and A. Schultz (1994), "Growth Of Cubic Boron Nitride On Si(100) By Neutralized Nitrogen Ion Bombardment," *Applied Physics Letters*, Vol. **64**, (12), pp. 1514–1516.

2. Mirkarimi, P. B., K. F. McCarty, and D. L. Medlin, "Critical Review of Advances in Cubic Boron Nitride," to be published in *Science & Technology of Hard Coatings*, Sundgren and Barnett, Editors.

Manufacturing and Coating by Kinetic Energy Metallization

T. S. Chow
Defense Technologies Engineering Division
Mechanical Engineering

We have performed systematic laboratory experiments and modeling analyses to demonstrate that (1) it is feasible to form thin and thick deposits by kinetic energy metallization (KEM) using various metal powders onto several types of substrates; (2) very low as-deposited porosity can be achieved by using fine (<10 μm) to ultra fine (<3 μm) powders; (3) KEM deposition rates are favorable for thick buildup for practical applications; and (4) the thick deposits will require post-KEM processes for improvement in both density and tensile properties. In addition, computer modeling predicts that the magnitude of the impact velocity is an important factor and that high deformations are produced in both impacting particles and target materials after impact. Features in the micrographs are consistent with these predictions.

Introduction

KEM is a low-bulk temperature coating process not requiring the introduction of thermal energy. Bonding energy is induced by high-speed impact of fine metal particles on a substrate. KEM is different from the conventional thermal spray coating processes in which the metal particles are heated to the liquid state and sprayed by a moderate- to high-speed carrier gas at the substrate to produce coatings. The processes are high-temperature operations. KEM is therefore very attractive for near net shape products.

ITI (Las Cruces, NM) was the first company in the U. S. to produce thin coating deposition samples through KEM technology and has applied for the patent rights. A group of Russian researchers have published similar work[1] called "cold gas-dynamic deposition."

Our objectives are to demonstrate the technical feasibility of the KEM process through a parametric experimental study in a controlled environment and to establish better understanding of the particle-to-substrate interaction mechanisms. The coating and forming experiments based on a mutually agreed upon test matrix were performed by ITI under a contract with Lawrence Livermore National Laboratory (LLNL).

ITI produced a set of thin coatings, on the order of 10's of mils, from which process parameters were selected to generate three bulk (thick) samples, Ti, Cu, and Zn, for further testing and analyses at LLNL.

Preliminary finite element analyses of the impact process were carried out at LLNL. The goal of the analytical studies is to determine the effects of particle velocity, size, material, and substrate material.
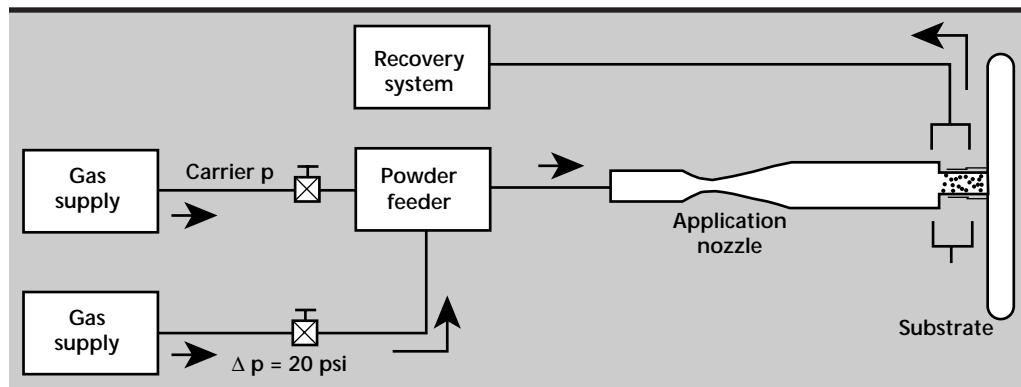
Figure 1. Schematic representation of the KEM system.

Progress

KEM Test System and Process Description

A schematic of the KEM system is shown in **Fig. 1**. The system consists of a supersonic nozzle for driving μm-sized particles to speeds in excess of 300 m/s and a recovery hood for the over-sprayed powder. The gas source used two K-bottles of commercial grade He, regulated to nozzle inlet pressures of 500 psig or less. The fluidizing gas pressure driving the feeder was regulated to a Δp of 20

Table 1. KEM test matrix based on six parameters.

Parameters	Lower bound	Upper bound	Increment	Bulk samples
1 Substrate hardness	Al	Cu or Inconel-600		Al
2 Powder type: reactivity melting point hardness		Al Cu Ni Ti Zn		Cu Ti Zn
3 Powder size	<3 μm	<45 μm	3 sizes	<10 μm, 45 (Ti)
4 Gas type	N	He	2 gases	He
5 Nozzle inlet pressure	200 psig	450 psig	3 settings	200 to 300
6 Deposition duration	5 s	60 s	3 values	up to 60 s

Table 2. Summary of all runs using Al substrate.

Powder	Avg. deposition rate ($\mu m/cm^2$-s)	He gas (psig)	Porosity	Hardness RH, 15T	Base metal RH
Al (<10 μm)	13 ± 7	200	<5%	71 ± 2	77 ± 2
Al (<10 μm)	6 ± 5	450	<5%	72 ± 2	
Al (20 μm), sph	99 ± 20	200	5 to 10%	61 ± 4	
Al (20 μm), sph	384 ± 20	450	5 to 10%	64 ± 3	
Cu (<10 μm)	8 ± 3	100	<5%	77 ± 3	76 ± 1
Cu (<10 μm)	84 ± 17	200	5.6 ± 4.5	85 ± 2	
Cu (<10 μm)	122 ± 23	450	<5%	89 ± 2	
Cu (<45 μm)	20 ± 1	200	> 10%	76 ± 2	
Cu (<45 μm)	59 ± 6	450	> 10%	78 ± 3	
Ni (<3 μm)	14 ± 5	200	–	88 ± 2	87 ± 1
Ni (<3 μm)	23 ± 12	450	–	87 ± 2	(Inconel–600)
Ni (<10 μm)	13 ± 4	200	5 to 10%	88 ± 4	
Ni (<10 μm)	44 ± 7	450	–	92 ± 1	
Ti (<45 μm)	29 ± 7	200	18 ± 4.5	78 ± 3	80 ± 1
Ti (<45 μm)	42 ± 13	450	> 10%	83 ± 6	
Zn (<10 μm)	82 ± 8 to 160 ± 44	200 to 450	<5%	63 ± 4	
Zn (10 to 25 μm)	15 ± 6 to 20 ± 3	200 to 450	<5%	58 ± 4	

to 30 psig above the process line. Nitrogen gas was periodically tested throughout the parametric study to determine if the KEM depositions were feasible with a lower expansion velocity gas.

The test matrix consists of the following parameters with a wide range of values: powder materials, powder sizes (commercial grades), carrier gases, pressures, and substrate materials based on hardness and deposition thickness (thin film and bulk samples, as summarized in **Table 1**).

Results

The detailed experimental data and results are given in ITI's final report.[2]

Table 2 summarizes the results for deposit on aluminum substrates, showing the effects of powder size, shape and gas pressure on the deposition rate, porosity, and hardness. The large uncertainties associated with the deposition rates were attributed to powder flow rate variations that were highly dependent on the powder properties and the feeder behavior. Generally, higher inlet pressures and smaller particles tended to increase the deposition rates. A few selected KEM depositions of various powders were made on substrates of Cu, Al, and Ni.

The micrographic characterization of all the samples listed in **Table 2** is included in Reference 1.

Micrographic comparison of the particle size effects are shown in **Fig. 2**, for Cu (<10 μm) and Cu (<45 μm) on Al. The bulk and interfacial porosities were quite low (<5%) for the <10-μm Cu, while the <45-μm Cu exhibited moderate porosity (10 to 15%). The micrograph of a Cu-on-Cu (**Fig. 3**) showed a much lower level of interfacial porosity, as compared to a Cu-on-Al deposit under otherwise identical conditions, suggesting that similar metals form stronger bonding than dissimilar ones.

ITI generated the bulk (thick) samples (Cu, Ti, and Zn) on (1mm) Al (**Table 3**) using the parameters derived from the test matrix. All KEM bulk samples were fairly rough on the deposition side because of manually operated nozzle translation. The Cu and Ti bulk samples exhibited significant residual stress that resulted in bending of the samples. The Zn sample showed much less residual stress.

The post-KEM processes included annealing of the Cu sample, and annealing followed by HIP-ing of the Ti sample. These treatments greatly improved the mechanical properties of the deposits. **Table 4** is a summary of the analysis and testing results.

Annealing of KEM-deposited Cu had moderate reduction in porosity, (from 6% to 4%), and a slight decrease in hardness. Annealing and HIP-ing of the KEM-deposited Ti reduced the porosity from 17% to 0, and greatly increased the hardness.

Figure 2. Micrographs of Cu on Al, <10 μm (a) and <45 μm (b).

Figure 3. Micrograph of Cu on Cu.

Table 3. Bulk sample deposition parameters.

Bulk sample	Cu	Ti	Zn
Dimension	8.0 cm × 1.6 cm × 0.6 cm	8.0 cm × 1.7 cm × 0.5 cm	8.0 cm × 1.8 cm × 0.3 cm
Powder size	<10 μm	<45 μm	<10 μm
He pressure	300 psig	250 psig	300 psig
Height of curvature	3.0 mm	2.5 mm	0.5 mm
Radius of curvature	34	32	203

A tensile specimen (**Fig. 4**) machined from an annealed Cu sample was tested, with the result plotted in **Fig. 5**. It can be seen that, with 38 ksi ultimate stress at 0.75% elongation, the specimen was brittle but relatively strong.

Elemental analysis of the Cu sample shows that the bulk oxygen content was in line with the level of commercial grade Cu powders. The KEM process failed to purge out the oxide coating naturally occurring on the Cu particle surfaces.

Theoretical Modeling Analysis

There are three major aspects to the KEM phenomenon: (1) the gas dynamics of the supersonic, gas/solid jet impinging on the target that may or may not modify the shock waves in front of the target; (2) the particle impact interaction processes at the surface; and (3) multiple impacts on (thick) KEM samples.

We have concentrated our analytical efforts[3,4,5] on the last two problems, that is, the interaction of the particle with the solid target upon impact at high velocities, particularly the time dependent deformation characteristics of both the particles and the surface on impact. The axi-symmetric 2-D CALE code that can calculate deformation and moving stress waves in the materials at high strain rates was used.

We have made calculations for various particle-impact scenarios. Here we discuss several sample cases: Cu particle impact on an Al target at various velocities (200 m/s, 500 m/s, and 1,000 m/s); Cu particle impact on a Cu target at 1 km/s; and two Cu

Table 4. Bulk samples test summary.

	Treatment[1]	Rockwell hardness[2]	Density	Tensile property	Machine-ability	Oxygen content
Cu	As-deposited	87 (15T)	94%	—	poor	0.209% ± 0.057— comparable to general MSDS spec. for Cu powders
	Annealed	80 (15T)	96%	38 ksi 0.7% (ultimate)	good	
Ti	As-deposited	82 (15T)	82%	—	poor	—
	Annealed	32 (C)	93%	—	poor-fair	
	HIP'ed	38 (C)	100%	—	fair	
Zn	As-deposited	74 (15T)	96%	—	fair	—

[1] Annealing condition: 2/3 melting temperature for 3 h; HIP condition: 15 ksi at 1000 °C for 3 h.

[2] [15T] is the Rockwell superficial hardness: max. scale = 92.5, equivalent to 19.9 on [C] scale.
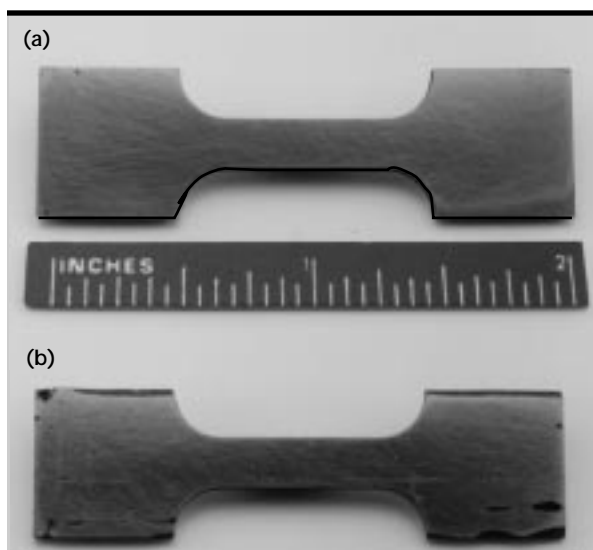

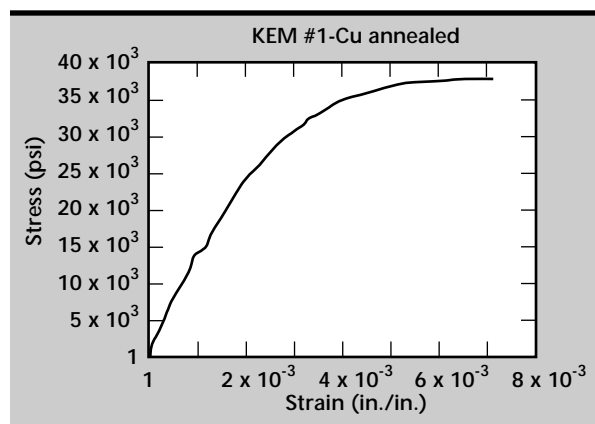Figure 4. Annealed KEM-deposited Cu tensile specimen.

Figure 5. Stress/Strain of annealed KEM-deposited Cu.

particles impinging on an Al target at 1 km/s. The model assumes a moving particle (a cylindrical disk), about to strike a target at t = 0. Due to axi-symmetry, only the top half of the field is needed in the simulation.

Case A: Effects of Particle Impact Velocity on Particle/Target Interaction. The deformation patterns of the Cu and Al materials at t = 0.36 μs after impact, when the impact velocity was 200 m/s, resulted in only a small amount of deformation. When the impact velocity was increased to 500 m/s, greater deformation took place. The stress waves in the Al travel first as compressive and then relieving tensile, so that the dilation of the material brings about the bulging crater-like protrusions around the impacting Cu particle.

When the impact velocity is raised to 1,000 m/s, which is a realistic velocity considering the high Mach number and velocity for the He carrier gas, our results show marked penetration of the particle as well as fragmentation of the materials. These analytical predictions show indentations and detach-ment of the target materials that resemble those found in the ITI test samples.

Case B: Cu Particle Impact on Cu Target. We consider this case to determine if there are any differences in impact response of like materials. The results at t = 0.24 μs after impact for the defor-mations in both the particle and the target show that these deformations are less than those for Cu on Al at t = 0.24 μs after impact at 1,000 m/s, signifying that Al deforms more readily than Cu when struck by a Cu particle.

Case C: Two Cu Particles on a Cu Target. The initial configuration at t = 0 is that both particles are traveling toward the target at 1,000 m/s. At t = 0.22 μs after impact, the first particle has already struck the target and has proceeded to inter-act with the target, while the second particle is still moving toward the target. The interaction shortly after the impact of the second particle shows that, at t = 0.24 μs, the penetration of the particles into the target at the centerline is now even greater and, in addition, the relieving tension waves cause the peripheral region of the target to protrude even more.

The velocity vectors show that the particles are still moving into the target, which will cause more deformation in both the target and the impacting particles at subsequent times.

Summary

This study has demonstrated that it is feasible to use the KEM process to deposit a wide range of metal powders to form both thin and thick coatings on various substrates. The work also has validated KEM as an efficient method for spraying fine reactive and pyrophoric powders such as Ti, in an inert gaseous atmosphere without using high-vacuum deposition technology. Commercially available Ti powder was used.

Post KEM processes such as annealing and/or HIP-ing will greatly improve the mechanical properties of the KEM-produced free-standing components. For refurbishment or thin coating applications the post-processes may not be as important.

The effect of the measured residual oxide on the KEM material property is not well understood. But it does suggest that powder conditioning and/or *in-situ* generation of powder as an integral part of the KEM process should be studied.

We have done preliminary modeling analyses of the KEM phenomenon. The results indicate that impacts by solid particles at high velocities could cause deformations on both the incident particles and the target, and that the deformation in the latter material is qualitatively similar to that seen in the test samples.

Future Work

Our plans are to investigate the KEM process through experiments and modeling analyses to ulti-mately produce the optimum process conditions. The proposed tasks are:

1) conducting particle/target interaction experi-ments using a well-instrumented set-up, such as measuring the particle velocity and responses of both it and the target following the impact;
2) setting up a U-deposit facility and conducting U-6Nb experiments at LLNL to demonstrate the feasibility of the technology as applied to the U-alloy powder;
3) studying the effects of powder synthesis and conditioning on deposition; and
4) performing process modeling analyses.

Many important aspects of this complex metal-deposition phenomenon remain to be addressed, such as the interaction of the particles with each other and with the target in fully 3-D impact situa-tions; the bonding process at the surface; and the behavior of the supersonic 2-phase (gas/solid) flow in front of the target and its effect on the shock wave and, especially, on the velocity distributions of the incident particles.

Acknowledgment

References

1. Alkhimov, A. P., V. F. Kosarev, and A. N. Papin (1990), "A Method of Cold Gas-Dynamic Deposition," *Soviet Physics Doklady*, **36**(12), pp. 1047–1049.

2. Tapphorn, R., and H. Gabel (1996), *Final Report—KEM Process Feasibility Study*, ITI, Las Cruces, New M.

3. Chow, T. S. (1996), "Manufacturing and Coating by Kinetic Energy Metallization (KEM)," *Progress Report for Manufacturing Thrust Area*, Lawrence Livermore National Laboratory, Livermore, Calif., March.

4. Chow T. S., and S. W. Kang (1996), *Preliminary Evaluation of KEM for Fabrication*, Lawrence Livermore National Laboratory, Livermore, Calif.

5. Tapphorn, R., and H. Gabel (1996), "Kinetic Energy Metallization Process Feasibility Study—Final Report," *Proceedings for the 12th General Meeting of JOWOG31, Livermore, Calif.,* ITI, Las Cruces, New M.

Magnet Design and Applications

Thomas M. Vercelli
Applied Research Engineering Division
Mechanical Engineering

The primary goal of this project is to demonstrate and provide closure for the magnet design process used in the physics/accelerator community. The magnet design closure process compares the measured field to the design model for a magnet designed by the Accelerator Technology Engineering Group (ATEG) for the Stanford Linear Accelerator (SLAC). Revision of modeling techniques will be based on this data comparison to more accurately predict future magnet designs. A second goal is to produce a conceptual design of both a specific and generic magnetic field mapping device. A conceptual design for a specific magnet mapper is demonstrated by the application of an existing coordinate measuring machine at Lawrence Livermore National Laboratory (LLNL) to map a variety of standard dipole and quadrupole magnets. A conceptual design of a generic magnet mapper applicable to large scale magnets is demonstrated for the Muon Magnet of the PHENIX Detector project at the RHIC accelerator at Brookhaven National Laboratory (BNL).

Introduction

High-energy physics experiments frequently require large-aperture spectrometer magnets as well as high-quality bending and focusing magnets. Their magnetic fields vary widely in uniformity and symmetry and must be predicted very accurately. The sophisticated 2-D and 3-D computer simulations often must be verified by direct magnetic field measurements. Fully automatic positioning and magnetic field measuring equipment is required.

At present, there is no facility in the United States that has a full-fledged operating magnetic measurement laboratory. Without magnetic measuring capabilities, projects requiring magnets are relying on computer modeling techniques to accurately design the magnet to desired specifications. To date, no one has verified the accuracy of these models.

Our project looked at validation of the computational model used to design quadrupole focusing magnets for the B-Factory project at SLAC. The magnets were designed by the ATEG in ARED. Magnetic field measurement data was then compared to the predicted model by ATEG staff. Differences in the predicted fields versus the measured fields were assessed for computer code modeling accuracy.

This project also looked at a conceptual approach for mapping large-scale magnets by designing a scheme to map the PHENIX muon magnets. The approach can be applied to other large-scale magnets, such as the Compact Muon Solenoid (CMS) at CERN and the Main Injector Neutrino Oscillation Search (MINOS) at Fermi National Laboratory. A conceptual design for mapping bending and focusing magnets was done by adapting a coordinate measuring machine at LLNL to accurately position a probe arm and probe assembly inside a magnet bore.

Progress

Two modified PEP-I insertion quadrupole magnets will be used on the high-energy ring (HER) in the interaction region (IR) of the PEP-II asymmetric B-Factory at SLAC, with the name QF5 quadrupoles. The insertion quadrupoles were designed to provide an extremely uniform (DB $\leq 10^{-4}$) and pure ($B_n/B_2 \leq 10^{-4}$) quadrupole field.

Two-dimensional field calculations were performed using POISSON, a code widely used and accepted within the high-energy physics/accelerator community. Because the magnet geometry required a model with very small elements in a large cross-section, the code was challenged by the large size of the element count as well as by the necessity to have the

dimensions of air/iron interfaces on the order of the element dimensions. Magnetic measurements were performed at LLNL using a 161-mm diameter rotating "search" coil.

The measured multipole errors (normalized harmonics) were the same order of magnitude as the accuracy limitations of POISSON. Yet, a comparison between the calculated and measured major systematic multipoles (n = 6,10) indicates good agreement. The comparison indicates that POISSON is an excellent tool for magnet design and that the non-conventional use of single elements in air/iron interfaces was successful for this application.

A conceptual design of the PHENIX mapper was created, along with a parts list and cost estimate of fabrication, assembly, and testing. This design concept uses a combination of laser interferometry and CCD cameras as the tracking and positioning measurement system. Linear probe position (z-axis) is accomplished by a linear slide mounted on a boom. A cart carrying the Hall probes is attached to the linear slide. The boom rotates at one end and is attached to a wall crawler at the opposite end to achieve an x-y position (**Fig. 1**). The entire system is controlled by LABVIEW software.

This approach can be applied to other large-scale magnets such as the CMS. The probe-positioning system is specific to the PHENIX magnet volume, but the positioning concept, tracking, and control system can be generically applied to other magnets.

A conceptual design for using a coordinate measuring machine (CMM) as a mapping tool was created. This design uses the "DEA" CMM as the positioning device for the magnetic field measurement probes. A cantilevered probe arm assembly is attached to the CMM positioning head. The magnet to be mapped is placed on top of the granite table of the CMM. The CMM is programmed to position the field measuring probe inside the magnet bore. A data acquisition system (DAS) records the magnetic field.

The DAS is integrated with the CMM control system, so programmed movements of the measuring probe are coordinated with the field mapping. The CMM measures a reference target on the magnet to locate the center and axial location of the magnet bore. Probe position can then be programmed for very accurate probe placement during mapping (**Fig. 2**).
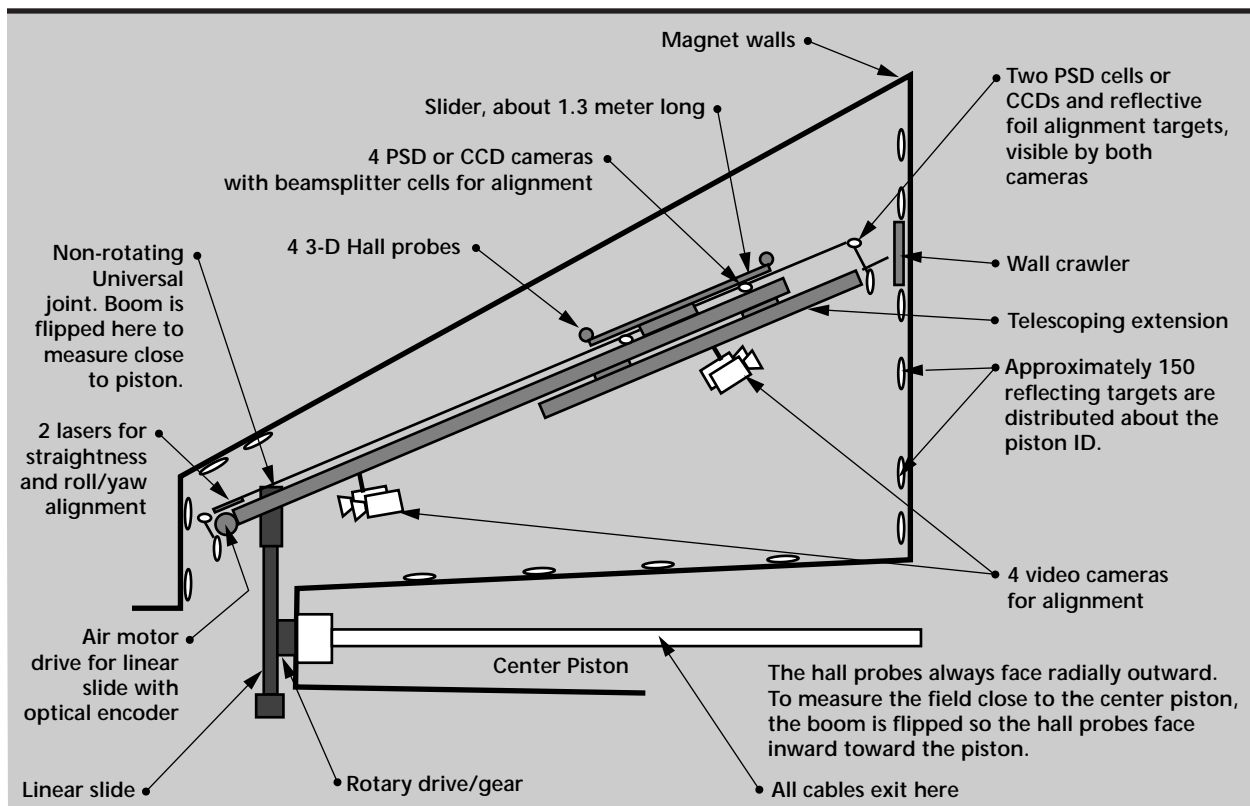
Figure 1. Conceptual design for PHENIX mapper.
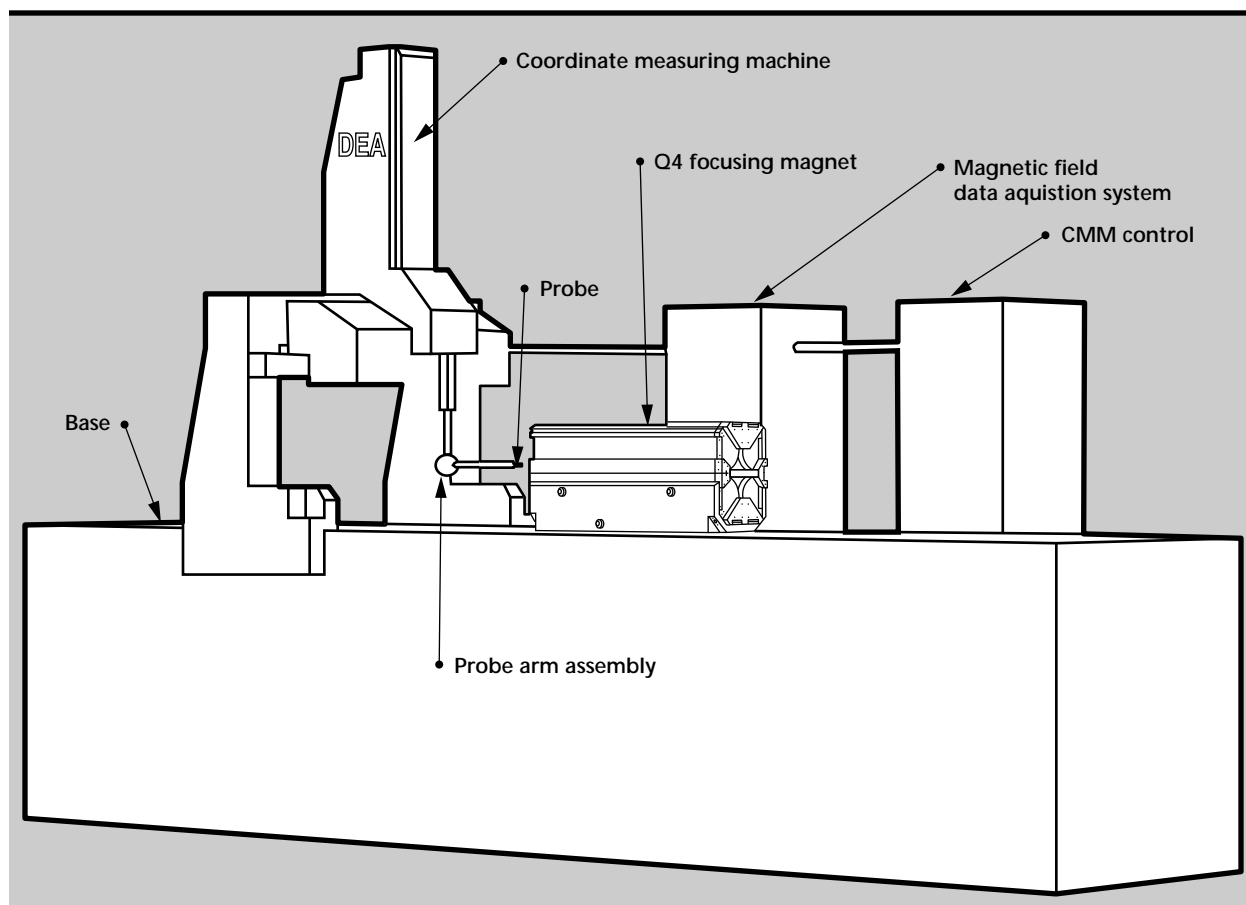
Figure 2. Conceptual design for coordinate measuring machine integrated with data acquisition system.

Future Work

Modeling validation has been done for only one magnet design. This is certainly not fully representative of the different magnet designs used. Further magnetic field data comparison should be done on a variety of magnets for increased confidence in the design code accuracy. Further code development is also necessary to calculate 3-D end effects of chamfers and shims, features currently modified empirically in the field.

Use of the CMM as a mapping tool will require a more detailed design and development effort in the areas of probe arm design; data acquisition and positioning integration; characterization of the probe arm assembly displacement errors due to magnetic field and thermal effects; and fabrication and testing of the system. The completed product will then have a direct application to map magnets such as the HIF and AHF magnets.

Our specific mission can continue to evolve and broaden in scope. LLNL possesses core competencies and expertise that can be directly applied to the design, development, testing, manufacture, and implementation of magnetic field mapping systems. This thrust area can lead to the establishment of a magnetics engineering center at LLNL to support many programs as well as the high-energy physics/accelerator community.

Acknowledgment

The author thanks M. Kendall and J. Swan.

Donald R. Lesuer, Thrust Area Leader

During FY-96, work within the Materials Science and Engineering thrust area was focused on material modeling. Our motivation for this work is to develop the capability to study the structural response of materials as well as materials processing. These capabilities have been applied to a broad range of problems, in support of many programs at Lawrence Livermore National Laboratory.

Recent examples of structural response problems studied are material fracture (including interface failure); damage in laser optics; the response of weapons components (including HE); and the failure of composite materials. For materials processing, typical problems studied include metal forming, laser processing, casting and heat treating.

To improve our ability to model material behavior, much of our work involves developing new material models and failure models, and applying the codes to new problems. Most investigations involve experimental studies to gather basic information on material response and to validate codes or material models. Projects are inherently multi-disciplinary, involving several investigators with expertise in materials and mechanics.

In addition to material modeling, the thrust area is also actively involved in superplastic materials and composite materials. Our technology base in these areas is being maintained or, in some cases, expanded through on-going programmatic work.

Superplastic materials are crystalline solids that can be deformed in tension to very large strains with low flow stress. We are currently working on several projects involving superplastic aluminum alloys and superplastic (net shape) forming with institutes in the Former Soviet Union. Our projects have industrial partners in the U.S. and we are developing these technologies for commercial applications.

Work in composite materials has involved metal and polymer matrix composites as well as laminated metal composites. Much of this work involves the use of these materials for lightweight applications.

Our thrust area studies for FY-96 are described in the following five articles: (1) Strength and Fracture Toughness of Material Interfaces; (2) Damage Evolution in Fiber Composite Materials; (3) Flashlamp Envelope Optical Properties and Failure Analysis; (4) Synthesis and Processing of Nanocrystalline Hydroxapatite; and (5) Room Temperature Creep Compliance of Bulk Kel-F.

Materials Science and Engineering

5

5. Materials Science and Engineering

Strength and Fracture Toughness of Material Interfaces

Robert A. Riddle and David S. Hiromoto
Manufacturing and Materials Engineering Division
Mechanical Engineering

In this project, we have specified desired material properties and created material combinations which in a component will have the required density, stiffness, strength, fracture toughness, manufacturability and cost for a particular application. We are concerned with the effect of processing parameters on the interface strength and interface fracture toughness between ceramic materials and ductile metals. This has lent insight into the macroscopic strength and fracture toughness properties of a metal matrix composite (MMC). We have measured the strength and fracture toughness of aluminum/alumina interfaces on specimens produced by vacuum liquid phase bonding methods, and compared the results to specimens prepared by ultrahigh vacuum diffusion bonding in previous years. We have also developed J Integral methods for quantifying the energy release rates for crack growth at ductile material/ceramic interfaces. As postulated from the test results, the energy-release rate as voids propagate as cracks and then blunt, leads to fairly unique failure behavior.

Introduction

The ductility of the metal phase plays an important role in increasing the fracture toughness of an MMC, while the ceramic particles increase the stiffness and strength. The interface under study here is between aluminum and alumina, both single crystal (sapphire) and polycrystalline. This study is intended to increase our understanding of commercial MMC materials, particularly those that consist of an aluminum alloy reinforced with Al_2O_3.

MMCs, when properly designed and fabricated, show superior properties, such as increased specific stiffness, reduced coefficient of thermal expansion, improved creep and wear parameters, and reduced fatigue damage at elevated temperatures. To take advantage of these improved properties in component designs, these properties must be both demonstrable and predictable in terms of their fabrication and use history. The production of MMC parts must also be cost-effective. Hence, fundamental understanding of the composition/structure and property/performance relations is required for these composites. A key ingredient in this understanding is the strength and toughness of the interface between the constituent materials.

Progress

We have used the microextensometer system developed in FY-95 to measure the elastic/plastic response of the ductile metal layer in the alumina/aluminum bend bar, tensile, and fracture toughness specimens. In this test system, two cylindrical rods are attached near the ductile layer interface. As the specimens are loaded in the test frame, the distance between the rods increases and a laser measures the increased spacing. The output of the load cell, the cross-head displacement, and the extension of the rod spacing is stored digitally in a computer system. This microextensometer system allows us to measure the constitutive response of thin layers of bonded materials loaded in tension.

We have updated a J Integral post-processor to calculate the energy-release rate available for crack advance at the metal/ceramic interface. This post-processor works in conjunction with a tie-break slide-line to simulate progressive failure at the interface.

Two key areas for the finite element modeling of the failure processes in the MMC include the appropriate failure criteria at the interface, and a straight-forward experimental method to measure the failure parameters. The strength of an interface is

measured based on an average stress with no defect present, while the fracture toughness relates to the stress field at a defect, as a function of the defect or crack length. In the modeling of failure at interfaces, a unified failure criterion is sought which combines the stress and strain field at the interface with a critical deformation parameter to describe decohesion of cracked and uncracked interfaces.

Experimental Results

Polycrystalline alumina materials were joined via aluminum foil in a vacuum brazing process. The alumina parts were Coors Ceramics AD-995, which is nominally 99.5% Al_2O_3. The finish of the alumina

Figure 1.   Part assembly for vacuum braze joining the alumina blocks and the aluminum foil.

bonding surfaces was polished. This best possible finish had a surface roughness that was measured at 5.1 μm peak-to-valley over 5.08 mm. The surfaces were cleaned with acetone, and then rinsed and wiped dry.

The aluminum foil was 99.997% pure aluminum, obtained from Alfa AESAR Puratronic. The foils were 0.05 mm thick.

The assembly of the alumina parts for the four point-bend specimens, with the aluminum foil as the braze medium, into a bonding fixture is shown in **Fig. 1**. The parts and the fixture were placed in a closed chamber, which was evacuated to 1.0×10^{-6} Torr, then heated to 1200 °C. The temperature was increased to this maximum over the period of 1 h.

In general, the lower temperatures produced poorer bonds. The smaller bend and tensile samples also were better bonded than the larger compact tension samples, suggesting an effect of surface roughness and foil thickness on the bond strength. Some tensile samples were notched to constrain failure to the bond region.

Because of the very small gauge length on the specimens, it was necessary to attach the cylindrical rods on the edge of the specimens, and use the displacement between the rods, as measured by a laser system, to characterize the load versus displacement behavior of the interface braze material.

Of the four bend specimens tested, two failed in the braze joint at an average load of 2544 N. Using the bend beam formula for stress, we calculate a maximum tensile stress of 345 MPa. The specimens
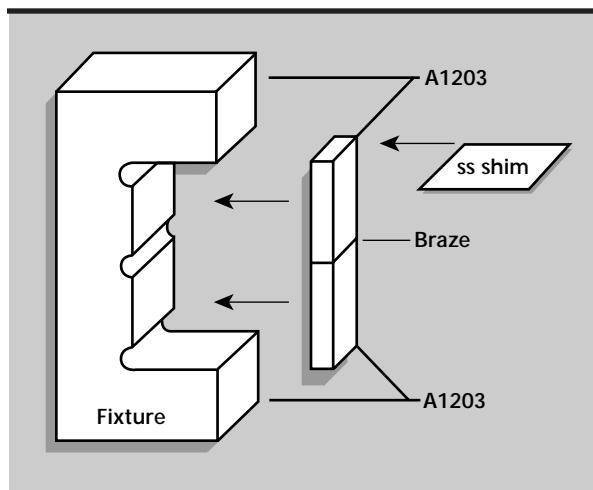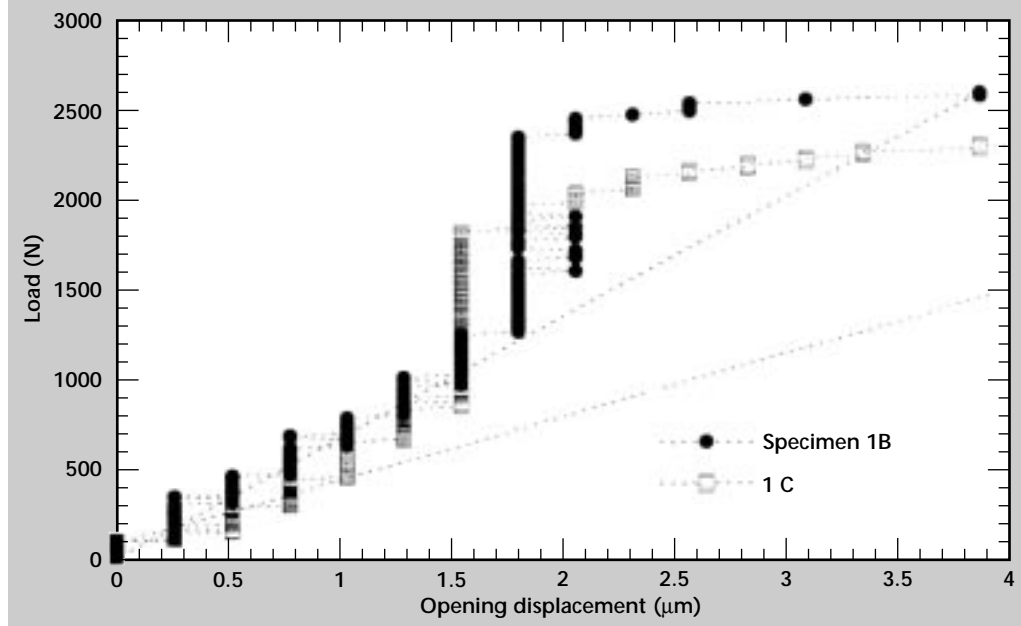

Figure 2.   Load vs opening displacement for bend specimens.

that did not fail at the braze failed at an average 7% higher load. These numbers compare favorably with the typical flexural strength listed in the Coors company sales literature as 379 MPa. The bonds were successful in the sense that they delivered a strength very nearly equal to a typical strength of the substrate material.

These bonds strengths are somewhat larger than values obtained under the previous conditions of ultrahigh vacuum diffusion bonding.[1]

Figure 2 shows the displacement characteristics for the two bend specimens that failed in the bond. The significant non-linear deformation prior to final fracture suggests that damage in the form of voids or cracks is present at fairly low loads.[2] The fact that the load versus displacement curves stiffen in the mid-load region is a distinct characteristic of the specimen failure, and will be investigated further in the analysis section.

Analytical Results

Finite element calculations were performed to derive material failure properties from the bend specimen tests. **Table 1** shows the stress/strain relation for the alumina and aluminum materials.

A mesh of 2632 elements was created to represent the geometry of the alumina/aluminum interface. To get more definition of plastic deformation in the aluminum bond layer, eight elements were used to model its thickness in the region of crack initiation. One eighth the thickness of the bond layer is the smallest geometric feature of the finite element mesh, and controlled the creation of the rest of the mesh, based on limitations of element aspect ratios.

The alumina regions of the bend specimen were modeled to the extent where the inside edge of the cylindrical rods in the microextensometer displacement system would be present in the bend specimens. The right hand or lower side of the mesh was constrained to have zero normal displacement. The left hand or upper side of the mesh had a linearly varying pressure boundary condition, based on the stress distribution predicted in the simple beam theory calculations of stress in the bend specimens. The displacements on the upper edge of the mesh from the finite element calculations are meant to be directly comparable to the displacements measured at the cylindrical rods.

A tie-break slideline was introduced at the interface between the lower edge of the aluminum bond layer and the alumina. The operating features of this slideline are such that until a specified, critical effective plastic strain is reached in the elements

along the slideline, the nodes across the slideline remain tied, representing a bonded region. When the critical effective plastic strain is reached, the slideline opens, the nodal constraints are released, and void growth or crack growth is then simulated.

The crack growth and/or voids at the interface affect the displacements at the boundary edges as the stress is applied, and correspond to the measured loads and displacement, or compliance, as measured in the microextensometer system. A change in specimen compliance is the only measure of defect growth in these experiments.

By doing finite element calculations on several failure scenarios for the bondline, including starting with no crack and no slideline, with a crack, with no crack and a slideline, and with a debond and a slideline, it became clear that the failure scenario with an initial debond, or void, and a tie-break slideline best matched the data produced in the actual tests. **Figure 3** shows the results of a finite element analysis, shown in the open square symbols, of a mesh with an initial debond of 83 µm at the aluminum/alumina interface, centered at 250 µm from the tensile edge of the bend specimen. The agreement is well within the tolerances of the displacement measuring system, and our knowledge of the material properties, such as the modulus of the alumina.

The initial debond grew into a lenticular-shaped void as observed,[2] and then propagated along the tie-break slideline, with the critical effective plastic strain to failure in the aluminum material being 0.10.

Table 1. Constitutive properties of aluminum and alumina.

Aluminum	
E	68.942×10^9 Pa
υ	0.33

Stress (Pa)	Strain
3.50×10^6	0.0
1.22×10^7	0.002
1.65×10^7	0.005
2.40×10^7	0.0150
3.44×10^7	0.045
4.85×10^7	0.125
6.85×10^7	0.35
9.74×10^7	1.0

Alumina	
E	345.0×10^9 Pa
υ	0.21

Figure 3. Stress vs opening displacement for bend specimen calculations.
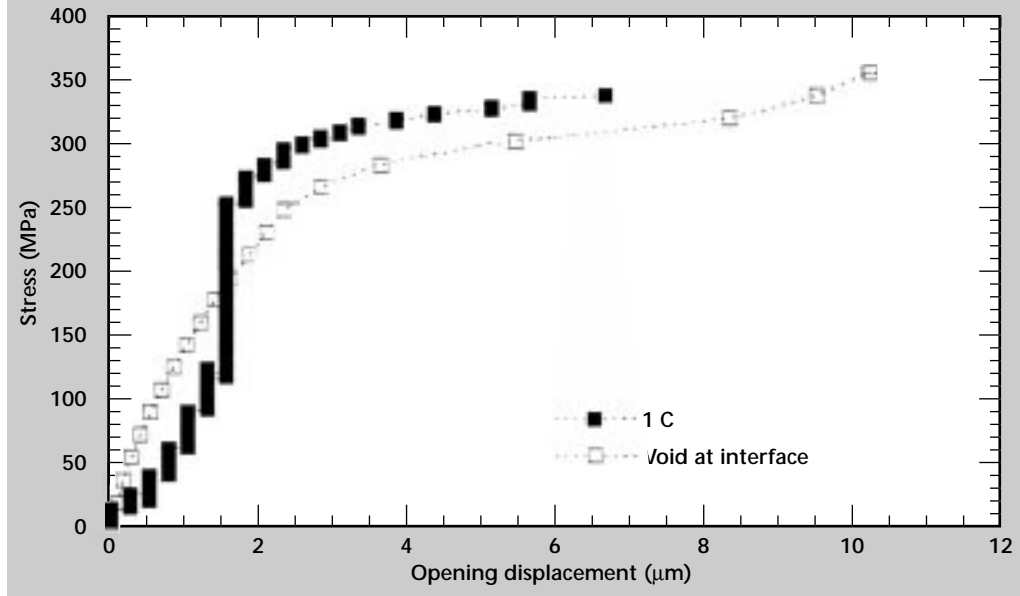
J Integral calculations for the energy-release rate at the first increment of crack growth yielded 430 Nm/m^2 as the fracture toughness of the interface.

Conclusion

The experimental techniques developed in these efforts to investigate the failure at metal/ceramic interfaces have produced new and interesting results regarding the growth of voids as a precursor to crack growth in the failure at the interface, and the implications on the compliance of specimens with initial debonds as failure proceeds. The development of the microextensometer system was essential in measuring specimen compliances on such a small scale.

Future Work

The analytical techniques, including the development of a more powerful J Integral post-processor, and its correlation with a tie-break slideline in predicting failure at interfaces, should allow application to a wide range of other problems where failure prediction is required.

The analytical and experimental techniques developed in the completion of this project may have a major beneficial effect on several on-going technology transfer projects, particularly efforts with GM on whisker reinforcement and with ALCOA and Pratt and Whitney on metal laminates. These techniques should also find use in predicting the reliability of interfaces between both metal and ceramics and metal and metals, of importance in the ADaPT program which is to ensure the safety and reliability of the nuclear weapons stockpile into the next century.

Acknowledgments

References

1. King, W. E., G. H. Campbell, W. L. Wien, and S. L. Stoner (1993), *Strength of Al and Al-Mg/Alumina Bonds Prepared Using Ultrahigh Vacuum Diffusion Bonding*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-93), **6**, pp. 19–22.

2. King, W. E., G. H. Campbell, D. L. Haupt, J. H. Kinney, R. A. Riddle, and W. L. Wien (1995), "X-Ray Tomographic Microscopy Investigation of the Ductile Rupture of an Aluminum Foil Bonded Between Sapphire Blocks," *Scripta Metallurgica et Materialia*, Vol. **33** (12), pp. 1941–1946.

Damage Evolution in Fiber Composite Materials

Steve J. DeTeresa, Scott E. Groves, Dennis C. Freeman, Patrick J. Harwood, Roberto J. Sanchez, and Marvin A. Zocher
Manufacturing and Materials Engineering Division
Mechanical Engineering

Edward Zywicz
New Technologies Engineering Division
Mechanical Engineering

The evolution of matrix microcracks and their effects on the elastic and viscoelastic properties of laminated fiber composites was investigated. Proposed models satisfactorily predict the degradation in tensile modulus of cross-plied laminates, but simple ply discounting is a viable alternative. Shear moduli reductions are less than predicted by the models and ply discounting is clearly too drastic. Short-term creep measurements before and after microcracking showed that tensile behavior is elastic, but there is time-dependent behavior in shear which is enhanced by microcracking. A new elastic bend test is being developed in an attempt to determine, for the first time, a fracture energy for fiber tensile failure.

Introduction

One of the strengths of fiber-reinforced structural materials is an ability to sustain relatively high levels of damage without catastrophic failure. The heterogeneous and anisotropic microstructure of these materials contribute to microcracking mechanisms which do not significantly reduce static load-bearing capability of properly designed structures. The accumulated damage can, however, have a significant effect on the thermal elastic and viscoelastic response of the material. For stiffness-critical applications such as dimensionally stable support structures and primary aerospace components, the degradation in these properties can have serious impact on performance.

Additionally, in any modeling scheme that seeks to follow the progression of damage and the concomitant redistribution of stresses, there is a need to know the relationship between constitutive properties and state of damage.

A significant body of theoretical work addresses the problem of predicting the effects of matrix microcracking on the elastic response of fiber composites. This type of damage is shown in **Fig. 1** and is generated in laminated structures by both mechanical and thermal loads. The various models

for predicting the *in-situ* elastic properties of a cracked lamina include, but are not limited to, shear-lag,[1–3] minimum complimentary strain energy[4,5] and internal state variable approaches.[6]

Experimental validation has been restricted mostly to the degradation in tensile modulus and Poisson's ratio in coupon specimens. The one study where the effect of matrix cracking on shear modulus was determined used a cantilever beam technique.[3] Because of the non-uniform stress state in the cantilever beam, analysis of the experiment is not straightforward. A preferred method to measure shear properties is torsion of a thin-walled tube, which is the only method that yields pure shear.

We have used a cylindrical specimen and test technique developed previously to simultaneously measure the effects of microcracking on both the tensile and shear moduli of laminates.

Recent work funded under one of our TTI CRADA projects has led to the development of a finite element model for a matrix-cracked viscoelastic laminate.[7] Similar modeling efforts are planned for the plastic-bonded high explosives and organics in our surviving stockpile. To date there have been no studies to determine experimentally the effects of matrix cracking on the viscoelastic response of composites. As part of our effort to

validate material models, we have investigated the magnitude of enhanced creep in both tension and shear using rectangular coupon and tube specimens.

While matrix cracking is the most prevalent form of damage in fiber composites, damage due to fiber direction stresses can also occur. We have studied the onset of damage in collimated and misaligned fibers under tension.

Progress

Elastic Behavior of Matrix-Cracked Laminates

Thin-walled, 2-in. diameter cross-ply tubes of AS4 carbon fiber and S2 glass fiber with an epoxy resin system (Fiberite 949) were produced with the transverse 90° layer from carbon fiber and the longitudinal 0° layer from either carbon or glass. The 90° layer is cracked during tensile loading in the 0° direction, and moduli in tension and in-plane shear are measured at each crack density.

The loads for these tests are applied as shown in **Fig. 1**. Two lamination patterns were used. A $0^i/90^c_3$ laminate where i = c (carbon) or g (glass) allowed cracks in the exposed 90° layer to be monitored directly during the tests using a Microvision MV 2100 video system. The 0° layers of glass are more compliant than the carbon and were selected to allow higher crack densities to develop. Strains were monitored using a biaxial extensometer, designed and constructed by Freeman, which measures axial deformation and rotation simultaneously. Surface-mounted strain gages were ineffective once cracks formed underneath them. Laminates having

a $0^i/90^c_3/0^i$ orientation were cracked and measured for moduli, then subsequently sectioned to determine the crack density. Strain gages were used for these specimens. All tests were conducted under tension load and torque control on a MTS biaxial hydraulic test machine.

The evolution of matrix cracks has been modeled using different damage progression schemes. Data collected here were compared to predictions based on a shear lag analysis with a strength criterion[2] and a strain energy release rate approach.[5] Each requires a material property (strength or fracture energy) to predict the damage state as a function of the applied stress. Using typical numbers for epoxy composite systems (9 ksi transverse tensile strength and 350 J/m^2 fracture energy) the fits to data collected for the $0^g/90^c_3$ specimens are shown in **Fig. 2**. Reasonable agreement with both models was obtained for the glass/carbon hybrid system, but not the all-carbon system. The same failure parameters were used for each case since the cracked layer in both laminates is the 90^c_3.

The elastic tensile moduli of $0^c/90^c_3$ and $0^g/90^c_3$ tubes as a function of crack density are compared with shear-lag (Tan-Nuismer) and complimentary energy (Hashin) models in **Fig. 3**. The Hashin model is a lower bound estimate for modulus as evident in the figure. Despite the scattered nature of the test results, it appears that both predictions capture the reduction in tensile modulus with matrix cracking. The data for the glass/carbon hybrid tubes tend toward and below the Hashin prediction. Similar results have been seen for other laminates having exposed 90° layers.[5]
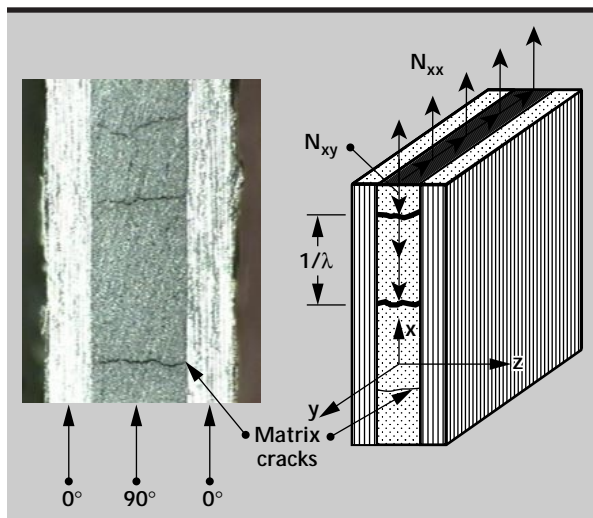

Figure 1. Microcracks in cross-ply laminated fiber composites and coordinates for tensile, N_{xx}, and shear, N_{xy}, loading of damaged materials.
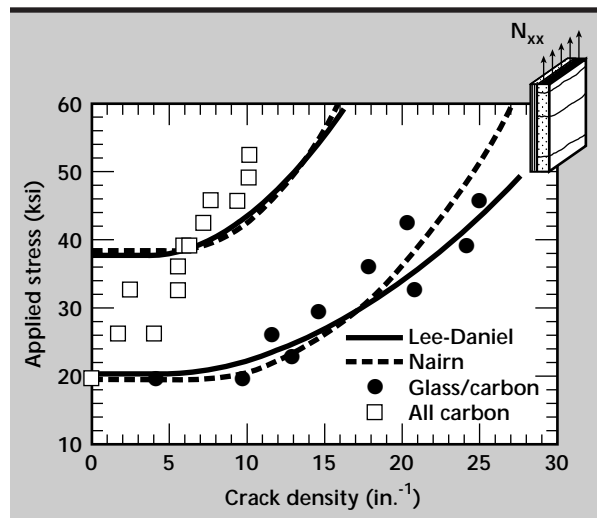

Figure 2. Comparison of microcrack evolution with model predictions for $0^i/90^c_3$ tubes.

A possible explanation is that microcracks in these layers form in a more random fashion than those in sandwiched 90° plies, and our observations confirm this. The discrepancy is that the model assumes equal spacing, but the differences are not significant.

In general, the modulus results are in agreement with both model predictions and previous experimental results. Also in the plots shown in **Fig. 3** are the limiting values of modulus reduction obtained by simply "discounting" the properties of the cracked ply. Simple damage progression models use complete discounting (that is, they assign vanishingly small properties) of the damaged ply at the onset of cracking to determine the new stress state. With only moderate levels of damage, which can be seen in **Fig. 2** to be achieved fairly quickly, the tensile modulus approaches this limit. Therefore, the simple ply discount scheme for tensile modulus appears to be a reasonable approximation.

A different result is obtained for the shear moduli of the damaged laminates. **Figure 4** shows the shear modulus reduction for the $0^i/90^c_3$ tubes. Similar results were obtained for the other laminates. These are first time results for the effect of matrix cracking on the shear modulus measured in a pure shear mode. In all cases the predicted loss in shear stiffness is greater than measured. Moreover, the ply discount of shear modulus is a gross overestimate of the actual modulus reduction, despite the fact that some high crack densities were generated. In this case it is clear that a damage relationship is needed to adequately model the shear modulus reduction.

Viscoelastic Behavior of Matrix-Cracked Laminates

Several attempts were made to determine the short-term creep response of $0/90_2/0$ laminates in tension and at various temperatures. Despite having near-saturation crack densities exceeding 50/in., no significant difference between the creep of pristine and damaged samples could be found. This loading is fiber-dominated and creep is extremely limited even at temperatures approaching the glass transition temperature of the matrix. We concluded that fiber-dominated behavior is essentially elastic even after significant damage has been accumulated.

Under longitudinal shear, there is significant creep even at room temperature. Because these tests were conducted on the thin-walled tubes in a servo-hydraulic machine under torque control, we were limited to approximately one hour test periods. At longer times, machine drift from a constant torque became a problem.
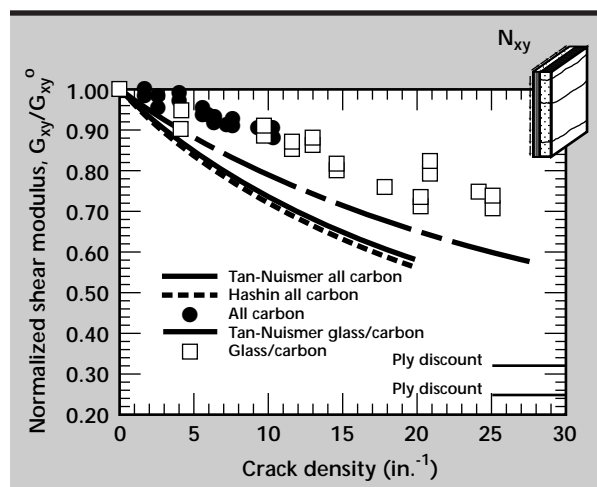

Figure 3. Tensile modulus degradation due to microcracking in (a) $0^c/90^c_3$ and (b) $0^g/90^c_3$ tubes.

Figure 4. Shear modulus degradation due to microcracking in $0^i/90^c_3$ tubes.

An example of the creep curves obtained for a single $0^g/90^c_3/0^g$ tube before and after generating nearly 50 cracks/in. is shown in **Fig. 5**. In addition to the obvious increase in the elastic component of the compliance, there is also a significant increase in the creep rate in the damaged tube. Qualitatively similar results were obtained for the $0^c/90^c_3/0^c$ tubes, but the magnitude of the creep rate increase was much less. This is undoubtedly due to the lower crack densities achieved in these tubes. We are currently analyzing these results using the viscoelastic model developed earlier.[7]

Fiber Direction Damage Modes

During the course of study of the viscoelastic response of $0/90_2/0$ laminates, we discovered that a processing defect in thermoplastic matrix materials severely limited the tensile strength (and, therefore, the attainable crack density) of these composites. The materials are processed above their melt temperatures to achieve adequate consolidation. During cool-down, the difference in thermal expansion between the fiber and tooling can put the fibers in compression. Any fiber compression above the melt temperature cannot be sustained by the composite because the matrix, which supports the fiber against buckling, is a viscoelastic liquid.

The transfer of load into the composite is via a shear lag mechanism, which quickly dissipates through the thickness of the composite. The result is an in-plane microbuckling which occurs periodically near the surface of the specimen. An example of these microbuckles is shown in **Fig. 6**, where the severe fiber misalignment is evident.

Tensile tests were performed on $0/90_2/0$ specimens having microbuckles in the 0° plies. Failure occurred at about half the expected tensile strength. The failure mode was changed from the normal longitudinal splitting observed for defect-free materials to a predominantly transverse failure surface. Examination of specimens loaded just below ultimate stress levels, such as that shown in **Fig. 6**, revealed that the tensile cracking was initiated at the region of highest misalignment in the microbuckles. Furthermore, the reduction in tensile strength was found to be roughly equivalent to the width of the microbuckle defect at the failure plane. This last observation is consistent with the notch insensitivity of the tough thermoplastic composites.

Our first attempts to generate stable fiber tensile fracture in aligned composites were with hybrid composites of two types of carbon fibers. The fibers were identical except for a nearly 2:1 disparity in tensile strength. It was thought that the failure of the weaker fiber layer would be stabilized by surrounding plies of the high-strength material. However, in all tests, the failure of the weak plies was sudden and complete. Release of the stored elastic strain energy at failure resulted in total delamination between the plies.

To reduce the amount of energy available to propagate cracks once fibers fail, we devised an elastic bend test using thin strips of unidirectional composite. Early results are encouraging in that

Figure 5. Effect of microcracking on short-term shear creep response of $0^g/90^c_3/0^g$ tubes.
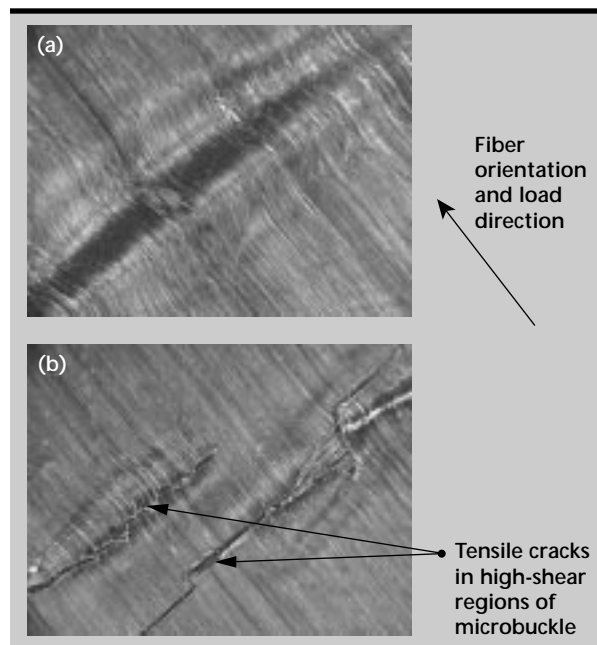
Figure 6. In-plane microbuckle defects (a) and localized tensile fracture (b) in thermoplastic $0^c/90^c_2/0^c$ tubes.
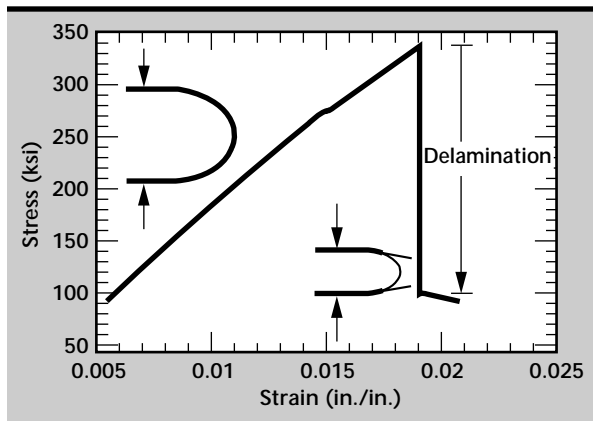
Figure 7. Mechanical response of a thin unidirectional carbon fiber/epoxy strip in the "U-bend" test.

fiber failure results in a degree of delamination which can be measured and therefore potentially analyzed to calculate a fiber fracture energy. Typical results for these "U-bend" tests are shown in **Fig. 7**. Data are being compared to analysis using a new composite delamination element developed for the NIKE and DYNA structural codes.

Future Work

The viscoelastic response of damaged composites will be compared with results obtained using the finite element model for matrix-cracked viscoelastic laminates. Further developments in modeling the time-dependent response of other composite systems, such as high explosives, in a pristine and damaged state are needed. The new bend test will be used in conjunction with numerical analysis to determine if a fiber fracture energy can finally be determined.

References

1. Nuismer, R. J., and S. C. Tan (1988), "Constitutive Relations of a Cracked Composite Lamina," *J. Comp. Mater.,* **22**, pp. 306–321.

2. Lee, J.-W., and I. M. Daniel (1990), "Progressive Transverse Cracking of Crossply Laminates," *J. Comp. Mater.,* pp. 1225–1243.

3. Tsai, C.-L., and I. M. Daniel (1991), "The Behavior of Cracked Cross-Ply Composite Laminates Under Simple Shear Loading," *Comp. Eng.,* **1**(1), pp. 3–11.

4. Hashin, Z. (1985), "Analysis of Cracked Laminates: A Variational Approach," *Mech. Mater.,* **4**, pp. 121–136.

5. Nairn, J. A. (1989), "The Strain Energy Release Rate of Composite Microcracking: A Variational Approach," *J. Comp. Mater.,* **23**, pp. 1106–1129.

6. Lee, J.-W., D. H. Allen, and C. E. Harris (1989), "Internal State Variable Approach for Predicting Stiffness Reductions in Fibrous Laminated Composites with Matrix Cracks," *J. Comp. Mater.,* **23**, pp. 1273–1291.

7. Zocher, M. A., D. H. Allen, and S. E. Groves (1996), "Stress Analysis of a Matrix-Cracked Viscoelastic Laminate," *Int. J. Solid. Struct.* (in press).

Flashlamp Envelope Optical Properties and Failure Analysis

Mark A. Havstad
Laser Science Engineering Division
Mechanical Engineering

We have measured the high-temperature transmittance of cerium-doped quartz. Measurements were made in the ultraviolet and visible spectral ranges (0.25 to 0.7 μm) for sample temperatures from 26 to 1000 °C. Cerium-doped quartz is of interest because it is the envelope material for the 10,000 flashlamps planned for the National Ignition Facility (NIF) at Lawrence Livermore National Laboratory (LLNL). Our measurements of the increase in absorption of the doped quartz quantify the decrease in performance of the lamps as they heat during each firing cycle. They also make it possible to calculate the transient thermal stress in an envelope due to temperature-dependent absorption through the material thickness. Our data has been used to show that it is unlikely that temperature-dependent absorption plays a role in the catastrophic flashlamp failure mode sometimes observed early in the lifetime of a lamp.

Introduction

The great majority of optical elements and materials operate below 100 °C. In many applications excursions in excess of a few degrees from room temperature are avoided. Similarly, the sophisticated and precise instruments evolved by the optics industry are rarely capable of treating samples at elevated temperatures. However, the transmittance of the glass envelopes of flashlamps, commonly used to provide the optical pumping in a wide range of laser systems, plays an important part in the system performance, lifetime, and operating conditions for material temperature ranging from ambient to 800 °C.

The cerium-doped quartz envelopes of the flashlamps planned for LLNL's NIF are a key component because they transmit both pump band energy (0.4 to 1.0 μm) for the laser amplifiers and thermal energy (beyond 1.0 μm) and UV energy (below .4 μm). The thermal energy heats the laser slab and causes a wait of several hours before the slab is sufficiently cool and uniform in temperature to fire again (adequate beam quality and no risk of fracture of the laser slab due to thermal stress are concerns). The UV energy damages materials with each firing and reduces the lifetime of components.

Not only is flashlamp transmission important to NIF operation, absorption of flashlamp radiation in its own envelope is also a concern because heating causes more absorption, reduced lamp output, and transient thermal stress. During the early portion of the 350-μs firing period the room-temperature and low-temperature transmittance of the envelopes is operative, but by the final portions the transmittance at 800 °C causes heating on the inner skin of the envelope and constitutes a positive feedback loop where absorption causes heating, which causes more absorption. Since the absorption is not uniformly distributed over the thickness of the envelope, one expects a resultant transient thermal stress that could cause lamp failure (explosion of the envelope).

In this report we present our measurement methods and the temperature-dependence of the transmittance of cerium-doped quartz from 0.25 to 0.75 μm for temperatures from 26 to 1000 °C. Last year we presented visible and near infrared transmittance measurements for this material, but with a very coarse spectral resolution and only at 400, 600, 800 and 1000 °C. This year we revised our optical system to reach the UV range and to obtain much finer resolution, both in temperture and wavelength. Our measurements of last year indicated that there was little temperature-dependent absorption from the far end of the visible to 1 μm, so we have limited ourselves to 0.25 to 0.75 μm in the present study.

Progress

The optical system (shown in **Fig. 1**) was designed to be used over a broad spectral range, 0.2 to 9 μm. The carbon filament lamp (power usage 24 A at 12 V DC), imaged along the vacuum furnace axis by the spherical mirror on a kinematic base is useful in the visible and IR. The arc lamp (200-W mercury/xenon bulb), imaged by a UV-grade fused-silica lens, provides UV, visible and near IR output. For visible and UV measurements the carbon lamp is off and the spherical mirror is removed from the kinematic base. This arrangement allows both light sources to be imaged down the furnace axis and, once aligned, to be fixed in position. Apertures along this optic axis limit the beam to a diameter of 0.6 mm.

The furnace includes a 102-mm long by 64-mm diameter heating element (0.64-mm thick carbon composite) which dissipates 600 W of DC power at 21 A to give sample temperatures of 1000 °C. Samples are held in graphite rings that are supported by four alumina rods that span the axial length of the hot zone and are themselves supported by the niobium thermal shields at either end of the furnace. A clear aperture of 12.7 mm diameter is bored through the thermal shields for the transmitted beam, but larger beams could be accommodated; the graphite rings holding the sample allow a clear aperture through the sample as large as 40 mm.

Five layers of thermal shielding are used at the furnace ends and around the furnace circumference. These last shields isolate the hot zone from a water cooled jacket that is fitted between the shields and the vacuum vessel walls.

The furnace is pumped by a 60 l/s turbomolecular pump backed by a mechanical roughing pump. The windows for optic access are calcium fluoride disks mounted in conflat vacuum flanges by a commercially proprietary silver brazing process. The window on the entrance side is mounted in a 33.8-mm O.D. flange and has a clear aperture of 11.8 mm and, on the exit side, a 69.8-mm flange with a clear aperture of 32.4 mm.

The optical system uses two beam paths: a sample path where the beam passes through the furnace and the sample, and a reference path where the beam is diverted by a turning mirror and routed to the detector by a path which remains at essentially room temperature for the duration of the measurements. The reference path (~210 cm long) is equal in length to the sample path (to within about 1 cm). The two turning mirrors in each path were fabricated on the same coating run. The focusing lens, filter and detectors common to the two beam paths are relatively isolated from the furnace thermal emission.

During initial optical alignment through the furnace, the output of the two light sources is imaged on the far wall of the laboratory so that the
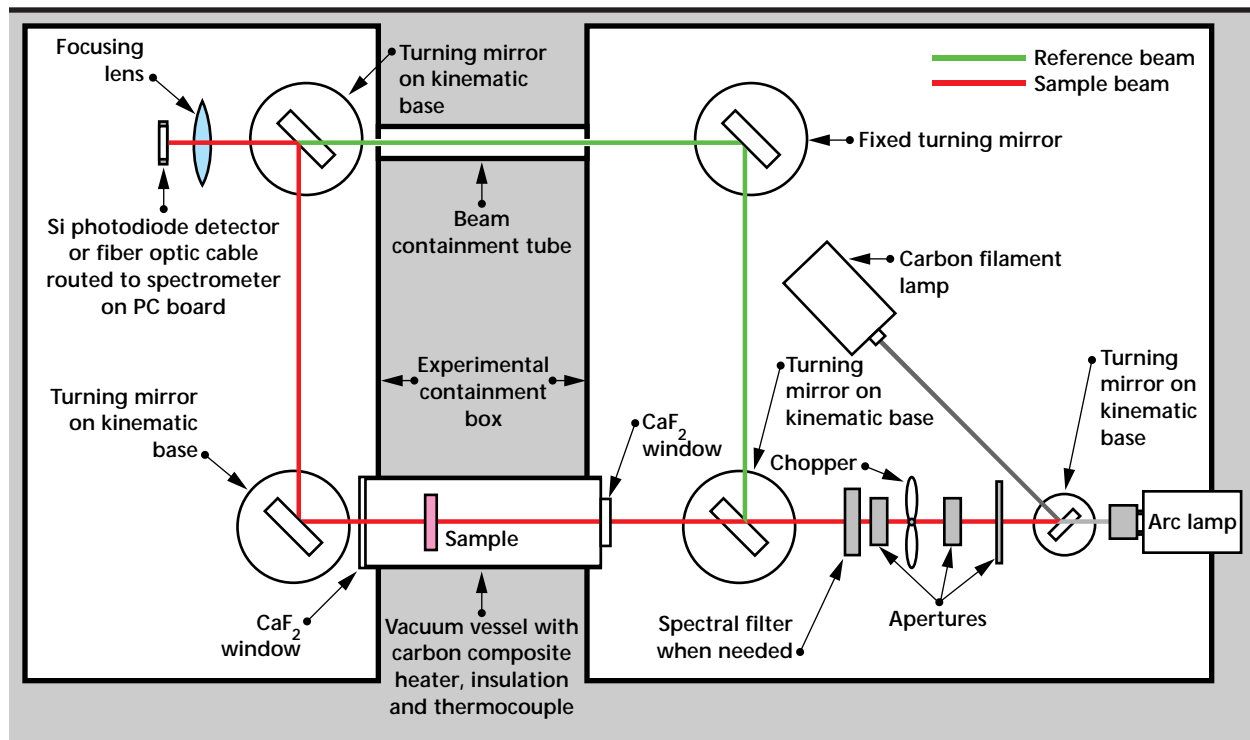

Figure 1. Schematic of optical system.

light incident on the sample is nearly collimated. Since all the turning mirrors are planar, the thermal emission from the furnace is not imaged at the detector and its intensity declines as the square of the distance from the furnace to the focusing lens.

The CaF_2 focusing lens images either of the light sources on the detection optics. For spectrally coarse measurements a bandpass filter is placed between the the lens and a silicon photodiode or an InSb/HgCdTe sandwich detector. For this system, the chopper, a lock-in amplifier and a diode pre-amplifier make up a phase-sensitive detection system. For more spectrally-resolved measurements (0.25 to 0.7 μm) the transmitted beam is imaged into an optical fiber. The fiber extends to a commercially supplied computer board. At the board back plane are a lens, a miniature grating and a 1024-element CCD detector array. The spectral calibration (pixel number to wavelength) is provided with the board and can be checked with any light source with sharp spectral lines.

The detector response for the sample and beam paths (minus the system dark signal) can be represented by:

$$V_{sam} - V_{dar} = B_s \cdot \tau_{W1} \cdot \tau_{sam} \cdot \tau_{W2} \cdot \rho_{T1} \cdot \rho_{T2} \cdot \tau_L \cdot \tau_f \cdot R_D$$

$$V_{ref} - V_{dar} = B_{s'} \cdot \rho_{T3} \cdot \rho_{T4} \cdot \tau_{L'} \cdot \tau_{f'} \cdot R_{D'}$$

where ρ is the turning mirror reflectivity, τ is transmittance, R_D is detector responsivity and B_s is lamp output through the chopper and apertures. Subscripts L, f, T1, T2, T3 and T4 refer to the focusing lens, bandpass filter and turning mirrors. Subscripts W1 and W2 refer to the entrance and exit windows on the furnace. The use of a prime symbol on some of the subscripts indicates a second order deviation between the primed and unprimed values. For example, the transmittance of the focusing lens is nominally the same for either beam path, but small differences in beam size or incident angle for example, result in slight deviations.

At any given wavelength and sample temperature, the ratio of transmittance at temperature to the transmittance at room temperature is computed from the above two relations:

$$\frac{\tau(temp)}{\tau(amb)} = \frac{\left(\left|\frac{B_s \cdot \tau_{W1} \cdot \tau_{W2} \cdot \rho_{T1} \cdot \rho_{T2} \cdot \tau_L \cdot \tau_f \cdot R_D}{B_{s'} \cdot \rho_{T3} \cdot \rho_{T4} \cdot \tau_{L'} \cdot \tau_{f'} \cdot R_{D'}}\right|\right) temp}{\left(\left|\frac{B_s \cdot \tau_{W1} \cdot \tau_{sam} \cdot \tau_{W2} \cdot \rho_{T1} \cdot \rho_{T2} \cdot \tau_L \cdot \tau_f \cdot R_D}{B_{s'} \cdot \rho_{T3} \cdot \rho_{T4} \cdot \tau_{L'} \cdot \tau_{f'} \cdot R_{D'}}\right|\right) amb}.$$

Thus the apparatus is used to make a relative transmittance measurement, τ_{hot}/τ_{amb}. A Carey 5 spectrophotometer, was used to make a precision absolute transmittance measurement at room temperature.

The temperature-dependent transmittance of a cerium-doped quartz sample provided from LLNL's NIF development work is shown in **Fig. 2**. Absorption in the UV and short wavelength portion of the visible increases with temperature. The short wavelength cut-off in transmission shifts to longer wavelengths steadily at a rate of approximately 3 nm per 100 °C.

The increase in absorption in the UV is desirable from the viewpoint of reducing damaging UV passed from the lamp to the remainder of the laser system, and is probably inconsequential to lamp failure mechanisms. It would take a much larger effect with temperature to account for lamp heating leading to explosion. The shift in the absorption edge to longer wavelengths is unwanted because it reduces the pump band output of the flashlamp, but the shift noted here is modest and most of it confined to wavelengths less than those designated as pump band, 0.4 to 1.0 μm.
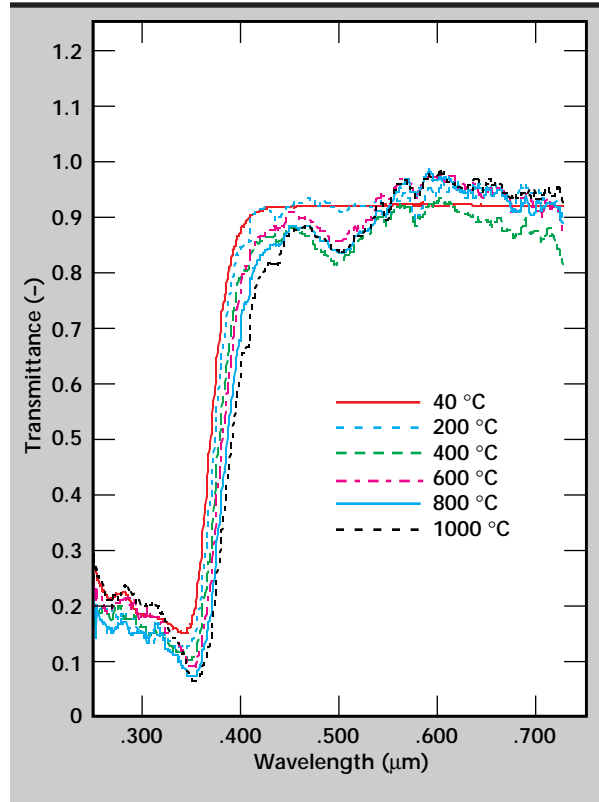

Figure 2. Transmittance of cerium-doped quartz in the ultraviolet and visible regions.

A second unwanted but also modest effect is the absorption at .500 μm. This feature as well as the decline in transmission at .400 μm can be observed in last year's work (with coarse spectral resolution), but here they are well quantified. The data shown indicates that there are no absorption phenomena of significance beyond .550 μm (as was indicated in previous work also), but quantitatively the indication of higher transmittance at elevated temperature than at ambient is probably incorrect.

The optical constants of cerium-doped quartz, which determine the material transmittance, can be formulated several ways; for example, by the index of refraction and the extinction coefficient, k, or the index and the absorption coefficient, $\mu = 4\pi k / \lambda$.

Computation of the absorption coefficient (which contributes to the transient thermal stress in the envelope of each flash) from a single transmittance measurement requires an assumption concerning the relationship between the two optical constants. Here we assume that in the short wavelength portion of the spectrum, up to .36 μm, a collection of 8 Lorentz oscillators adequately characterizes the relationship between the optical constants:

$$\varepsilon^c = \left(\varepsilon^a\right)_s + \sum_{j=1}^{n} \frac{\left(\nu_j\right)^2 \cdot \Delta\varepsilon_j}{\left(\nu_j\right)^2 - \nu^2 + z \cdot \nu \cdot \gamma_j}$$

The complex permittivity is ε^c (yet another pair constituting the optical constants), ν is the

Table 1. Oscillator parameters calculated from least squares fit.

	Temp. (°C)					
	26	**200**	**400**	**600**	**800**	**1000**
Oscillator number			**Mode strength (dimensionless)**			
1	0.138	0.467	0.2	0.116	0.182	0.394
2	0.618	0.796	1.616	2.107	2.514	3.404
3	0.7	0.8	0.8	0.7	0.8	0.7
4	0.3	0.4	0.4	0.3	0.4	0.3
5	8	8.7	8.7	8	9	7
6	1.5	1.9	1.9	1.8	2	1.9
7	0.3	0.4	0.4	0.3	0.5	0.3
8	7.02	7.25	7.25	7.02	7.5	6
		Linewidth (damping)÷oscillator frequency (dimensionless)				
1	0.027	0.071	0.05	0.039	0.031	0.055
2	0.043	0.054	0.064	0.058	0.063	0.065
3	0.2	0.2	0.2	0.2	0.2	0.2
4	0.15	0.15	0.15	0.15	0.15	0.15
5	0.3	0.3	0.3	0.3	0.3	0.3
6	0.12	0.12	0.12	0.12	0.12	0.12
7	0.04	0.04	0.04	0.04	0.04	0.04
8	0.319	0.319	0.319	0.319	0.319	0.319
		Oscillator mode (location of resonance) ν (nm)				
1	338	338	338	338	338	338
2	347	349	350	351	353	355
3	280	280	280	280	280	280
4	313	313	313	313	313	313
5	335	335	335	335	335	335
6	263	263	263	263	263	263
7	252	252	252	252	252	252
8	300	300	300	300	300	300

wavenumber, $\Delta\varepsilon_j$ is the oscillator strength (or plasma frequency), ν_j is the oscillator resonant frequency and γ_j is the damping constant. The static dielectric constant is $(\varepsilon^c)_s$. All quantities except the wavenumber are functions of temperature.

This equation satisfies the Kramers-Kronig relations, and thus provides a physically consistent relationship between the two optical constants.[1] The choice of modes (oscillator resonant frequencies) is somewhat arbitrary, particularly in the shortest wavelength ranges, but the fit as shown here gives a close match to the data and yields a real index of refraction that is consistent with previously published data for quartz at room temperature.[2] For the 0.25- to 0.36-μm range the oscillator parameters obtained from least squares fitting of the measurements are given in **Table 1**.

For wavelengths beyond 0.36 μm the absorption edge is well fit by the "Urbach rule"[3]:

$$\mu(\nu,T) = \mu_o \cdot \exp\left[-b \cdot \left(E_o - E(\nu)\right)\right]$$

$$\tau(\nu,T) = (1-\rho)^2 \cdot \exp\left(-\mu(\nu,T) \cdot t\right) \quad .$$

The absorption coefficient at ambient is μ_o, photon energy is $E = h\nu$, E_o is the bandgap, and b is the steepness parameter. The part thickness is t and ρ is the reflectivity (actually also a function of wavenumber and temperature, but in the absorption edge region and to longer wavelengths within the visible, the dependence of ρ on either ν or T is insignificant). Various explanations for this dependence of the absorption edge on wavelength have been given, including thermal fluctuation of the bandgap and electric fields arising from defects, phonons, and impurities.[4,5]

The transmittance data for ambient has been fit to the relations above (for wavelengths from 0.36 to 0.45 μm) with four parameters (μ_o, E_o, b, and ρ) varied until the sum of the squared error was minimized. The data at elevated temperature was fit similarly, but with ρ fixed and just the other three variables minimized. The results of these fits are given in **Table 2**.

Of the three quantities (μ_o, E_o, and ρ_o) varied to obtain each elevated temperature fit, μ_o is approximately constant and the other two are well fit to linear functions (**Fig. 3**). The shift of E_o with temperature, a shift of the Urbach tail to longer
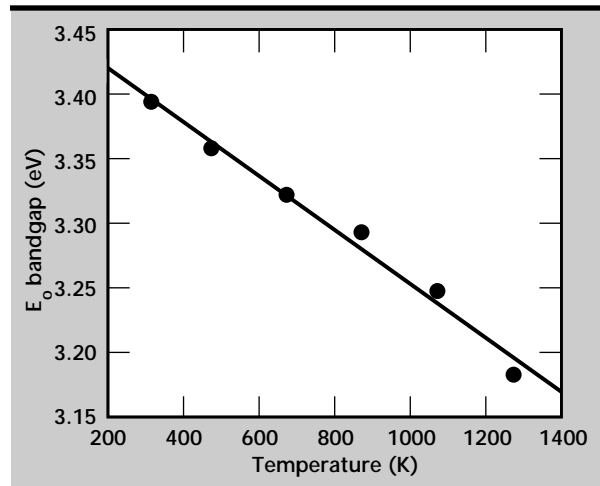
Figure 3. Shift in bandgap energy fit to linear function.

Table 2. Transmittance data.

Temperature (°C)	μ_o (1/cm)	b (1/eV)	E_o (eV)
26	307.9	8.508	3.394
200	308.6	8.131	3.358
400	319	6.767	3.322
600	312.4	7.088	3.293
800	302.4	6.323	3.248
1000	260.6	5.801	3.183
Average	301.8	linear	linear

Reflectance = 0.039

	y = mx + b	
	m	b
E_0	-2.08×10^{-4}	3.4619
b	-2.73×10^{-3}	9.234
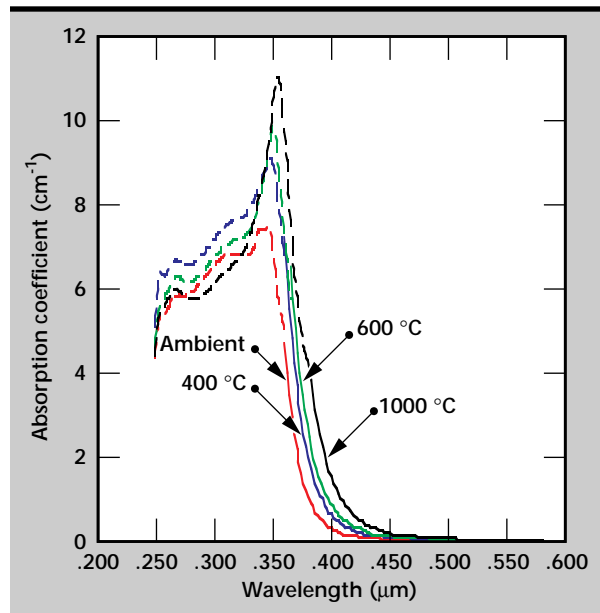
Figure 4. Absorption coefficient of cerium-doped quartz.

wavelengths, is common in insulators and semiconductors. As temperature increases the short and long wavelength limits of transparency tend to shift inwards, shrinking the region of transmission.

The oscillator and Urbach models allow a compact specification of the transmission data or the absorption coefficient for use in computer modelling of radiant transport and transient thermal stress in the flashlamp envelope. The absorption coefficient computed from the parameters in **Tables 1** and **2** is plotted in **Fig. 4** for four temperatures. The models do not attempt to match the absorption noted in **Fig. 2** at 0.50 μm; a more complex model for the absorption would have to be developed to match that region as well.

This data has been provided to computer models which have estimated the spectral and temperature-dependent absorption through the thickness of a simulated flashlamp envelope. Thermal stress calculations were then used to estimate transient thermal stress in the cerium-doped quartz.[6]

References

1. Thomas, M. E. (1991), "Temperature-Dependence of the Complex Index of Refraction," *Handbook of Optical constants of Solids II*, Academic Press, New York, p. 177.

2. Myers, V. H., A. Ono, and D. P. DeWitt (1986), "A Method for Measuring Optical Properties of Semitransparent Materials at High Temperatures," *AIAA Journal* **24**, pp. 321–326.

3. Lynch, D. W. (1985), "Interband Absorption - Mechanisms and Interpretation," *Handbook of Optical Constants of Solids*, Academic Press, New York, p. 189.

4. Skettrup, T. (1978), "Urbach's Rule Derived from Thermal Fluctuations in the Bandgap Energy," *Phys. Rev.* B **18** p. 2622.

5. Dow, J. P., and D. Redfield (1972), "Toward a Unified Theory of Urbach's Rule and Exponential Absorption Edges," *Phys. Rev.* B **5**, p. 594.

6. Maltby, J., V. C. Garcia, and B. T. Kornblum (1996), "Radiation-Induced Thermal Stresses in High-Power Flashlamps," *Engineering Research, Development and Technology,* Lawrence Livermore National Laboratory, Livermore, Calif., UCRL-53868-96.

Synthesis and Processing of Nanocrystalline Hydroxyapatite

T. G. Nieh
Materials Science and Technology
Chemistry

Donald R. Lesuer
Engineering Sciences
Mechanical Engineering

We have developed a technique to construct bulk hydroxyapatite (HA) with different cellular structures. The technique involves the initial synthesis of nanocrystalline HA powder from an aqueous solution using water-soluble compounds, followed by spray-drying into agglomerated granule. Depending upon the processing parameters, the morphology of the spray-dried granules can be varied. The granules were further cold-pressed and sintered into bulks at elevated temperatures. Since the starting HA powders were extremely fine, a relatively low activation energy (21.3 kJ/mol) for sintering was obtained. From the sintering study, both porous and dense structures were produced by varying powder morphology and sintering parameters.

Introduction

Calcium phosphate apatite (CPA) is one of the most viable implantable materials because of its biocompatibility. The major subphase of the bone mineral consists of submicroscopic crystals of an apatite of calcium and phosphate, resembling HA in its crystal structure [$Ca_{10}(PO_4)_6(OH)_2$]. The HA is nanocrystalline with grains ranging in size from 10 to 50 nm. These nano-sized grains are hierarchically assembled into connective hard tissue, that is, bone-skeleton.

Currently, there are three types of bone grafting: autograft, allograft, and synthetic (that is, coralline blocks). Recent research has been performed to develop composite materials such as HA/bioglass® ceramic composites with ZrO_2, Al_2O_3 and SiO_2, and HA/bioglass® polymer composites.[1] However, cellular structures with macro- (> 100 µm) and micro-interconnected pores made directly from HA and with sufficient fracture strength have not been successfully produced. This research is aimed at constructing various HA cellular solids (bone-like) from nanocrystals.

Progress

The chemical precipitation method used to synthesize HA particles is described by the following chemical reaction:

$$10 Ca(OH)_2 + 6 H_3 PO_4 \xrightarrow[\text{Room Temp.}]{NH_4 OH}$$
$$Ca_{10}(PO_4)_6(OH)_2 + 18 H_2 O \qquad (1)$$
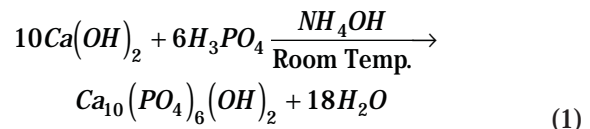
The precipitated HA reveals a needle-shaped morphology (~10 to 20 nm in width and 60 to 90 nm in length), similar to that of the apatite particles existing in natural bones. These needle-shaped nanocrystalline particles were further suspended and sprayed, resulting in large-sized powders with different shapes.

The sprayed powders were cold-pressed into green bodies (density = 1.42 ± 0.15 g/cm^3, approximately 44% of the theoretical density) at a uniaxial pressure of 187 MPa. The green bodies were sintered for 4 h in air at temperatures varying from 700 to 1300 °C. Experimental results are listed in **Table 1**. As indicated, the density value increases with sintering temperature and reaches a nearly fully dense value (99%) at the sintering temperature of 1300 °C. The micro-hardness values, H_v, ranging from 208 to 483 MPa, were also found to increase with sintering temperature. The grain structure of the 99%-dense sample is shown in **Fig. 1**, which revealed equilibrated, hexagonal-shaped grains with sharp apexes with a mean grain size of about 4 µm. X-ray analysis indicates that the sample also contains a slight amount of whitlockite (TCP), which

is somewhat expected, since apatite is essentially a line compound and technically difficult to obtain in pure form.

The density of sintered samples as a function of temperature is plotted in **Fig. 2**. Within the temperature range of the present study, density increases monotonically with sintering temperature. Data from the literature[2,3] are also included in **Fig. 2** for comparison. Our density values are noted to be higher than those reported previously.[2,3] This is because the starting HA powders used in the present study are ultrafine (nanocrystalline). Therefore, there exists an exceedingly large amount of surface energy acting as an additional thermodynamic driving force for sintering. In addition, enhanced diffusion processes (that is, short-circuit diffusion) are expected to operate, as discussed later. Both factors tend to accelerate the sintering process.

The shrinkage of a non-transforming solid during sintering, assuming a non-isothermal process, can be described by **Eq. 2**[4]

$$S\% = k\exp(-Q/RT)t^{1+n} \qquad (2)$$

where t is the sintering time, Q is the activation energy, T is the absolute temperature, R (= 8.314 J/mol) is the gas constant, k is a constant, and n is a constant which is about 0.3 to 0.4 for oxide ceramics. From **Eq. 2**, the activation energy for sintering can be computed. It is calculated to be 21.3 kJ/mol (5 kcal/mol), which is significantly lower than those reported previously (~200 kJ/mol). The relatively low activation energy agreed with the early suggestion that short-circuit diffusion processes (for example, surface and interface diffusion) dominate during the sintering of nanocrystalline HA.

Recent clinical tests indicated that the optimum pore size for an implant with a cellular structure ought to be greater than 100 μm.[5,6] At this size, soft body tissues are able to exhibit a good ingrowth while the implant still retains its mechanical integrity. In the present research, several efforts were made to fabricate HA with a cellular structure.

Table 1. *Physical properties of sintered HA at various temperatures.*

Temperature (°C)	Relative density (ρ %)	Shrinkage (S%)	Micro-hardness (Hv)
1300	99	55	483
1200	88	50	—
1180	84.5	48	—
1100	84	47.6	303
1000	71	38	263
900	64	31	208
700	44	0	—
Green body	44	—	—

Figure 1. *The grain structure of the 99%-dense sample, revealing equilibrated, hexagonal-shaped grains with sharp apexes, with a mean grain size of about 4 μm.*
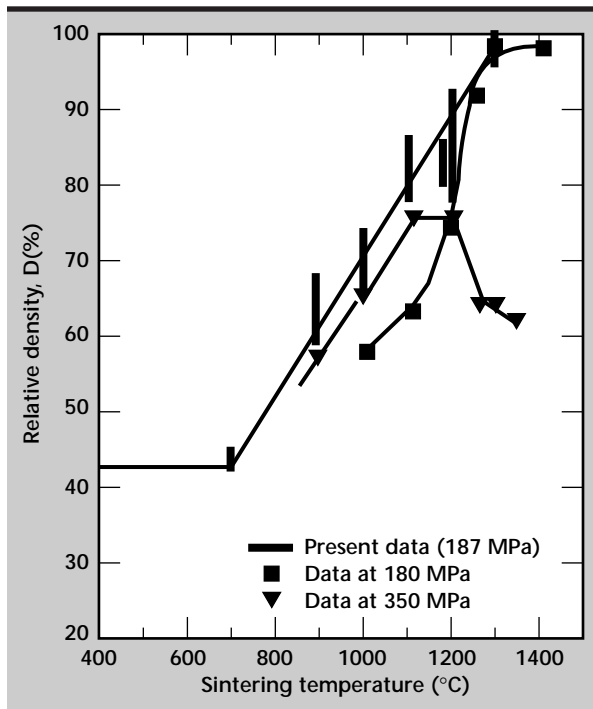
Figure 2. Relative density of sintered HA as a function of temperature.

Figure 3. SEM surface morphology of the composite (HA plus polymer beads) sintered at 1100 °C for 4 h.

Based upon the sintering results, experiments were carried out by adding polymer beads with a diameter of approximately 100 μm to HA in the green state. The polymer beads were subsequently pyrolyzed during high-temperature sintering to stimulate the formation of interconnected macro-pores in HA skeleton. The sintering temperatures of 1000 and 1100 °C were selected.

Figure 3 is the SEM surface morphology of the composite sample (HA plus polymer beads) sintered at 1100 °C for 4 h. The sample has a density of 55% with a mean pore size of about 20 μm, and the pore distribution is relatively uniform. SEM stereo projection indicates that the pores are interconnected. The 3-D interconnectivity of these pores was further confirmed with x-tomographic microscopy.

Future Work

This research will not be continued in FY-97.

References

1. Wilson, J., L. L. Hench, and D. Greenspan (1995), *Bioceramics, Volume 8*, Pergamon/Elsivier Science Ltd., Oxford.

2. Goto, T., N. Wakamatsu, H. Kamemizu, M. Iijima, Y. Doi, and Y. Moriwaki (1991), "Sintering Mechanism of Hydroxyapatite by Addition of Lithium Phosphate," *J. Mater. Sci.—Medicine,* **2**, pp. 149–152.

3. Wang, P. E., and T. K. Chaki (1993), "Sintering Behavior and Mechanical Properties of Hydroxyapatite and Dicalcium Phosphate," *J. Mater. Sci.—Medicine*, **4**, pp. 150–158.

4. Kingery, W. D., H. K. Bowen, and D. R. Uhlmann (1976), *Introduction to Ceramics*, John Wiley & Sons, Inc., New York, N.Y.

5. Klawitter, J. J., and S. F. Hulbert (1971), "Application of Porous Ceramics for the Attachment of Load Bearing Orthopedic Application," *J. Biomed. Mater. Res. Symp.*, **2**, pp. 161–169.

6. Hench, L., and J. Wilson (1993), *An Introduction to Bioceramics*, World Scientific Publishing Co. Ltd., Singapore.

Room Temperature Creep Compliance of Bulk Kel-F

Marvin A. Zocher, Steve J. DeTeresa, and Scott E. Groves
Manufacturing and Materials Engineering Division
Mechanical Engineering

Experiments have been conducted to determine the room temperature tensile creep compliance of bulk Kel-F 800. Tests of linearity have been included as well. Specimens were subjected to loads resulting in nominal homogeneous stress states of 138, 310, and 490 kPa. Under all three loading conditions, the material was observed to be linear. The degree of crystallinity in the specimens tested was low. They were essentially amorphous.

Introduction

Kel-F 800 is a copolymer formed from the monomers chlorotrifluoroethylene (CTFE) and vinylidine fluoride (VF). The nominal composition consists of a 3:1 mole ratio of CTFE to VF, with the monomers arranged in random alternation. Kel-F 800 also contains a small amount (< 2 wt%) of perfluorodecanoic acid, which is added as an emulsifier. Originally developed by the M. W. Kellog Company, it is now manufactured by 3M where it is referred to as FK-800.

In bulk form, the copolymer is a thermoplastic and has a T_g of about 31 °C. Kel-F 800 is used by Lawrence Livermore National Laboratory (LLNL) as the binder for the explosive 1,3,5-triamino-2,4,6-trinitrobenzene (TATB) in the plastic bonded explosive (PBX) LX-17. LX-17 is 92.5% (by weight) TATB and 7.5% Kel-F 800. The Los Alamos National Laboratory (LANL) PBX-9502 is similar to LX-17, consisting of 95% TATB and 5% Kel-F 800.

The thermo-mechanical behavior of Kel-F 800 is not well characterized at the present time. The literature contains data on tensile properties determined from quasi-static testing, and on shear properties determined from dynamic mechanical analysis (DMA).[1–4] Cady and Caley[1] performed tensile tests on dog-bone specimens cut from molded sheets having a uniform thickness of 0.05 cm. These tests were conducted in an Instron loading machine with cross-head speeds of 0.127, 1.27, and 25.4 cm/min. All tests were conducted at room temperature (RT). Tensile yield was observed to increase linearly with degree of crystallinity (DOC) while ultimate elongation decreased linearly with DOC. Ultimate strength was found to be

largely independent of DOC. The DMA tests were conducted in the forced torsional pendulum mode at 0.2 Hz over a temperature range of –160 to 110 °C.

Grotheer[2] performed DMA tests in the forced torsional pendulum mode at 1 Hz over a temperature range of –75 to 100 °C. Grotheer's specimens were cut from sheets molded at 80 °C (Cady and Caley did not address the processing of the sheets from which specimens were taken).

Hoffman, Matthews, and Pruneda[3] performed DMA and DSC on specimens molded under moderate pressure at 120 °C (the processing temperature for molding LX-17). The DMA was conducted over a temperature range of –100 to 150 °C using five different frequencies.

DMA has been conducted by Herold[4] as well. Additional data on Kel-F 800 may be found in the LLNL Explosives Handbook[5] and in a survey, currently under way, by Overturf and LeMay.[6]

The objective of the current program was to add to our understanding of the thermo-mechanical behavior of Kel-F 800, particularly with respect to time-dependent behavior. In keeping with this, experiments have been conducted for the purpose of determining the RT creep compliance of bulk Kel-F 800. Tests of linearity have been included in the experimental program.

Progress

Specimen Fabrication

As alluded to in the previous section, DOC can have a strong influence on the thermo-mechanical response of Kel-F 800. The DOC is in turn

strongly influenced by thermal history. Consequently, it is the opinion of the authors that any data presented on Kel-F 800 that is not accompanied by processing history and an assessment of DOC is of little value.

Kel-F 800 comes from the manufacturer in a form that is similar in appearance to common rock salt, but with a subtle yellowish tint. This raw material was molded into sheets, roughly 28 cm in diameter and 2 mm thick. This was accomplished with heat and pressure. The method used for processing the sheets is as follows:

Step 1. 150 g of Kel-F 800 was heated at 120 °C for 1 h in a crude mold (a hose clamp lined with Kapton film) measuring approximately 10 cm in diameter, with a height of approximately 1.3 cm. After an air cool, the Kel-F was removed from the mold as a loosely consolidated disk, looking similar to a rice cake. Thirty-two such disks were formed, one for each sheet to be processed.

Step 2. Processing of the disk was accomplished in a 667 kN PHI 4-post hydraulic press with platens roughly 61 cm × 76 cm. This platen area was large enough for four sheets to be formed at a time. The processing region of the PHI was fitted with a furnace. Sheets of Kapton were used to prevent the Kel-F 800 from adhering to the platens.

a) The oven was preheated to 120 °C per thermocouple located at the center of the platens.

b) Once the temperature was stabilized at 120 °C, the samples were loaded to 106.76 kN (24,000 lb$_f$) in compression at a loading rate of 53.38 kN (12,000 lb$_f$)/min. The samples were held at 106.67 kN (24,000 lb$_f$) for 3 min, then unloaded to about 445 N (100 lb$_f$) and held at that load for 1 min. The unloading rate was the same as the loading rate, that is, 53.38 kN/min.

c) The load/unload cycle described above was repeated three times (**Fig. 1**). The processed sheets were then allowed to air cool.

Uniaxial tension specimens (**Fig. 2**) were cut from these sheets using a "cookie-cutter" type of stamp (MMED Stamp No. 18816) which is based on ASTM D412. Each of the 32 sheets was large enough to provide either five or six tensile specimens (most provided six). Approximately 12 specimens were used to establish test procedures and 20 were used in the current study of RT creep compliance. This leaves several specimens available for follow-on work.
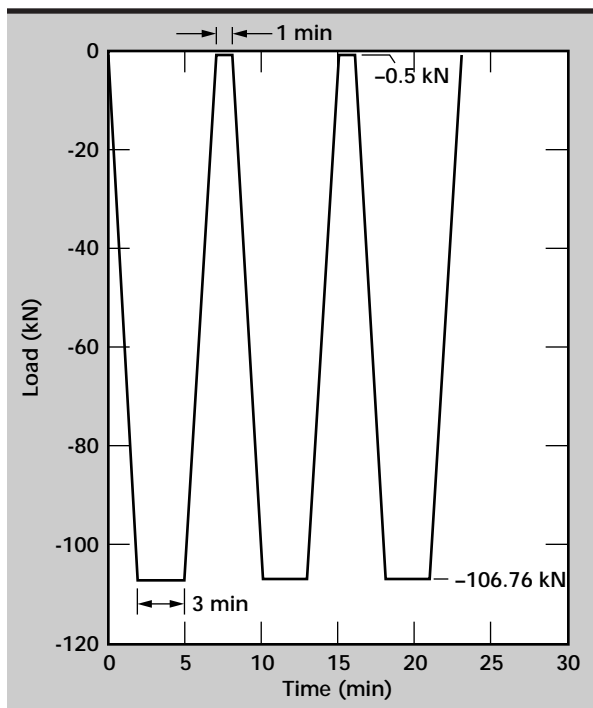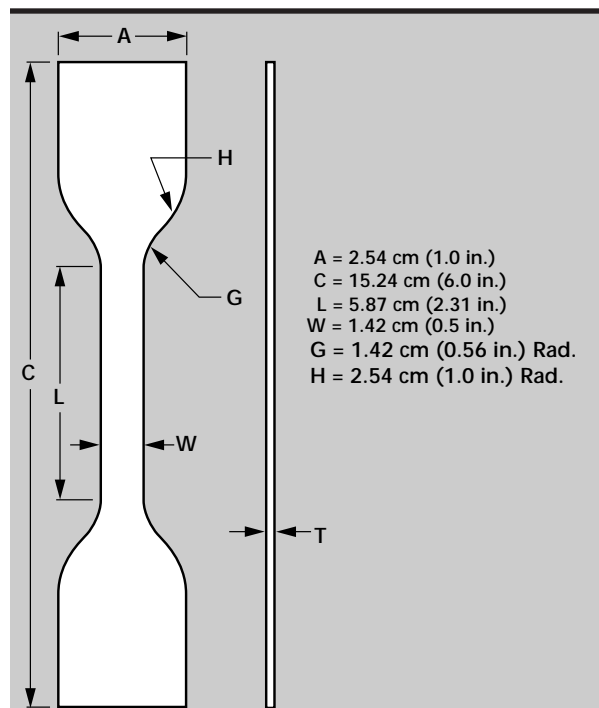

Figure 1. Pressing sequence.

A = 2.54 cm (1.0 in.)
C = 15.24 cm (6.0 in.)
L = 5.87 cm (2.31 in.)
W = 1.42 cm (0.5 in.)
G = 1.42 cm (0.56 in.) Rad.
H = 2.54 cm (1.0 in.) Rad.

Figure 2. Specimen geometry.

Test Procedures

Both creep and creep-and-recovery tests were conducted at three different load levels. These tests were conducted in a room maintained at 21.5 °C (± 0.5) and 50% relative humidity. Creep tests were conducted to determine creep compliance, and creep-and-recovery tests were conducted to test linearity. A viscoelastic material is linear if and only if its response to stimulus satisfies the mathematical principles of homogeneity and superposition. These principles may be stated as:

Homogeneity:

$$\sigma_{ij}\{c\varepsilon_{kl}\} = c\sigma_{ij}\{\varepsilon_{kl}\} \qquad (c = \text{constant}) \qquad (1)$$

Superposition:

$$\sigma_{ij}\{\varepsilon_{kl}^a + \varepsilon_{kl}^b\} = \sigma_{ij}\{\varepsilon_{kl}^a\} + \sigma_{ij}\{\varepsilon_{kl}^b\} \qquad (2)$$

where ε_{kl}^a and ε_{kl}^b may be either the same or different input histories and the braces are used to denote history-dependence. It is noted that satisfaction of superposition guarantees satisfaction of homogeneity, but that the converse statement is not true.[7] An important consequence of this fact is that a creep-and-recovery test provides all that is needed to assess material linearity.

Creep tests were conducted on 12 specimens (**Table 1**). Five were loaded with 3.56 N, four with 8.01 N, and three with 12.45 N. These load levels are nominal (actual values varied somewhat from specimen to specimen due to factors such as variations in grips and extensometer spring forces). This resulted in nominal stresses within the test section of about 138, 310, and 490 kPa, respectively. Test duration ranged from a low of 70 to a high of 220 h.

The test procedure was as follows.

Table 1: Test matrix.

	Test 3	Test 4	Test 5	Test 6	Test 7
Station 1					
Specimen	4003-1	4004-1	4005-1	4006-1	4007-1
Creep load (N)	3.56	3.56	8.01	8.01	3.56
Thickness (cm)	0.1996	0.2017	0.1999	0.2108	0.2057
Width (cm)	1.2522	1.2649	1.2675	1.2548	1.2573
Creep stress (kPa)	142.4	139.5	316.1	302.8	137.7
Station 2					
Specimen	4003-2	4004-2	4005-2	4006-2	4007-2
Load (N)	8.01	8.01	8.01	8.01	8.01
Thickness (cm)	0.2022	0.1999	0.2019	0.2121	0.2070
Width (cm)	1.2675	1.2649	1.2624	1.2598	1.2598
Stress (kPa)	312.5	316.8	314.3	299.8	307.2
Station 3					
Specimen	4003-3	4004-3	4005-3	4006-3	4007-3
Load (N)	12.45	12.45	12.45	12.45	12.45
Thickness (cm)	0.2002	0.1991	0.2007	0.2090	0.2080
Width (cm)	1.2573	1.2522	1.2548	1.2598	1.2548
Stress (kPa)	494.6	499.4	494.4	472.8	477.0
Station 4					
Specimen	4003-4	4004-4	4005-4	4006-4	4007-4
Load (N)	3.56	3.56	3.56	3.56	3.56
Thickness (cm)	0.2007	0.1999	0.1976	0.2062	0.2080
Width (cm)	1.2598	1.2598	1.2548	1.2624	1.2649
Stress (kPa)	140.8	141.4	143.6	136.8	135.3

Step 1. Aluminum grips, with surfaces specially prepared to prevent slippage, were attached to one end of the specimen with threaded fasteners torqued to 2.03 J.

Step 2. The specimen was hung by the other end in a spring-loaded mechanical grip, and an extensometer was attached. A spring was attached to the extensometer to balance it so as to minimize the force imposed on the specimen by the extensometer.

Step 3. The specimen was left hanging overnight (generally about 16 h) with only the weight of the aluminum grips applied.

Step 4. The creep load was applied. This was accomplished by releasing a set screw which allowed the load (weights on a pan) to free-fall a short distance, thereby applying the load almost instantaneously.

As seen in **Fig. 3**, which shows the complete test set-up, a load cell is positioned at the top of the load train. Beneath the load cell is the mechanical grip, specimen (with extensometer attached), aluminum grip with slot to catch a ball attached to the weight
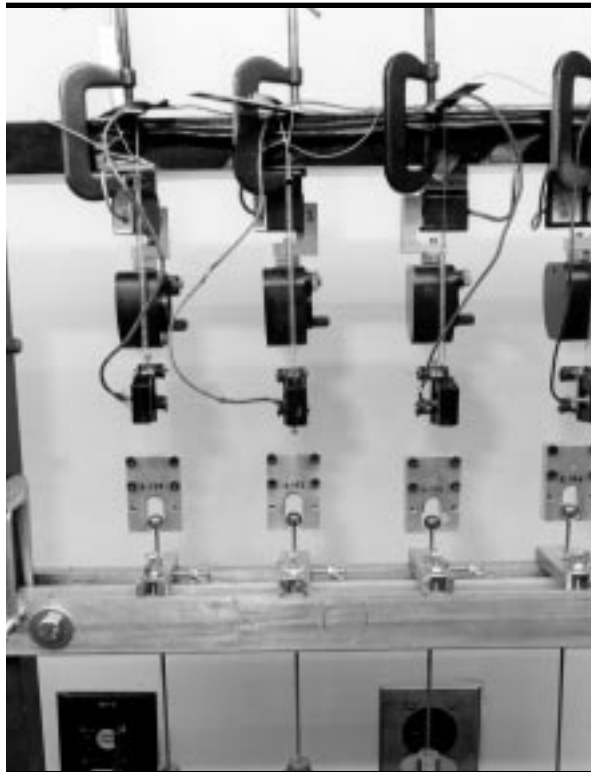

Figure 3. Test set-up.

pan when the pan is released, and finally, the weight pan. The specimens are difficult to see in **Fig. 3** because they are almost as clear as glass.

The load cells used were manufactured by Instron and were rated to 222.41 N (comparable load cells manufactured by Interface were used in Test 7 only). The extensometers, which were manufactured by Instron, had a 2.54-cm gage length and a +50% displacement range (except on station 4, where the displacement range was +10%). Four specimens were tested at a time.

The measured creep compliance, based on an average of the twelve test results, is presented in **Fig. 4**.

Ideally, a creep test involves the instantaneous application of a given load, P, at some time t_0, to a previously unloaded specimen. The state of stress for $t \geq t_0$ is ideally spatially homogeneous and temporally constant. The creep tests conducted here deviate from the ideal in two respects: (1) the loading was not instantaneous, and (2) the specimens were not fully unloaded for $t < t_0$. The first deviation mentioned above is of little concern since the load was applied rapidly enough that the event was essentially instantaneous (especially for tests that ran several hours). The second deviation exists in large part as a consequence of the difficulty incurred in trying to apply the load instantaneously and, at the same time, obtain accurate measurement of strain.

Using spring-loaded mechanical grips at the bottom of the specimens was not practical since one could not attach them to the specimen and then release them in such a way as to approach a step input loading. Adding to the difficulty, one could not manually attach and release these mechanical grips without producing erroneous strain measurements (from jostling the specimen) at the instant of loading.
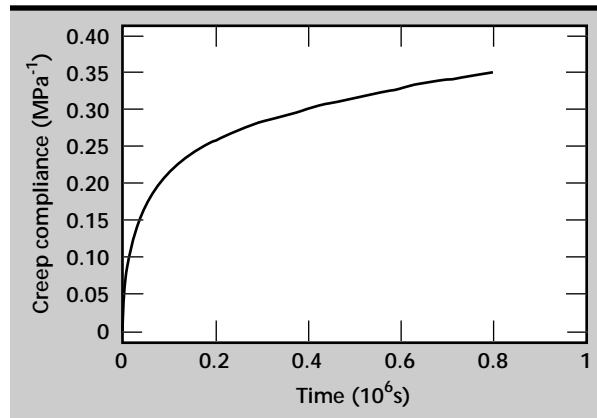

Figure 4. Creep compliance.

Hydraulic grips are encumbered by the same difficulties. The alternative chosen here, was to pre-load the specimen with the weight of the aluminum grips. This pre-loading took out any bending that may have been present due to the forces imposed by the extensometers and, at the same time, enabled the application of a load input that looks very much like a step function. By loading the specimens in this manner, we failed to satisfy the no-load for $t<t_0$ objective. We sought to minimize this deviation from the ideal by making the grips light (typically 0.88 N or less).

It is the opinion of the authors that the approach taken is justified on two counts: (1) fading memory serves to minimize the effects of the pre-loading; and (2) using a mechanical or hydraulic grip presents difficulties, as already mentioned.

Creep-and-recovery tests were conducted on eight specimens. Three were loaded (during the creep portion of the test) with 3.56 N, three with 8.01 N, and two with 12.45 N, nominally. The specimens were unloaded (except for the aluminum grips) during the recovery portion. Test procedures were identical to those described above for the creep tests, except that the load was removed after a given amount of time (generally about 50 h) to allow for recovery. This was accomplished by rapidly raising the weight pan a small amount and then tightening a set screw to hold it in place. The creep-and-recovery tests revealed the material to be linear under all three loading conditions.

DOC was measured by modulated DSC on selected specimens both before and after loading. All samples were found to be essentially amorphous, possessing 2% or less crystallinity.

Future Work

One obvious area for future work is testing Kel-F 800 at temperatures other than RT. It is possible that such testing will reveal the presence of a time/temperature superposition principle. This is all the more likely since the results of the current work revealed the material to be linear.

Another area for future work is to test at higher loads to determine the onset of nonlinearity.

A third topic for future study is to quantify the effect of DOC on creep compliance by conditioning specimens in such a way that creep compliance can be determined as a function of DOC.

A fourth area for future study is aging. It would be interesting to apply the methods of Struik[8] to determine the effect of physical aging on creep compliance.

A fifth area for future work is the determination of the effect of specimen thickness. It may be that the properties of thin films of Kel-F 800 are not the same as the bulk properties.

Acknowledgments

References

1. Cady, W. E., and L. E. Caley (1977), *Properties of Kel-F 800,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-52301).

2. Grotheer, E. W. (1988), *Characterization of PBX Binders*, Allied Signal, Kansas City, Mo., (BDX-613-3849).

3. Hoffman, D. M., F. M. Matthews, and C. O. Pruneda (1989), "Dynamic Mechanical and Thermal Analysis of Crystallinity Development in Kel-F 800 and TATB/Kel-F 800 Plastic Bonded Explosives. Part I Kel-F 800," *Thermochemica Acta*, Vol. **156**, (2), pp. 365–372.

4. Herold, R. D. (1990), *PBX Binders FY90-2,* Allied Signal, Kansas City, Mo, (KCP-613-4362).

5. Dobratz, B. M., and P. C. Crawford (1985), *LLNL Explosives Handbook, Change 2,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-52997).

6. Overturf III, G. E. and J. D. Lemay, (Draft), *Survey of Kel-F Literature for W87 LEP*, Lawrence Livermore National Laboratory, Livermore, Calif.

7. Schapery, R. A. (1974), "Viscoelastic Behavior and Analysis of Composite Materials," *Mechanics of Composite Materials*, G. P. Sendeckyj, Ed., Academic Press, New York, N.Y., pp. 85–168.

8. Struik, L. C. E. (1978), *Physical Aging in Amorphous Polymers and Other Materials*, Elsevier Scientific Publishing, Amsterdam, Netherlands.

Mark A. Newton, Thrust Area Leader

The Power Conversion Technologies thrust area identifies and sponsors development activities that enhance the capabilities of engineering at Lawrence Livermore National Laboratory (LLNL) in the area of solid-state power electronics.

Our primary objective is to be a resource to existing and emerging LLNL programs that require advanced solid-state power electronic technologies. Our focus is on developing and integrating technologies that will significantly impact the capability, size, cost, and reliability of future power electronic systems.

During FY-96, we concentrated our research efforts on the areas of (1) Micropower Impulse Radar (MIR); (2) novel solid-state opening switches; (3) advanced modulator technology for accelerators; (4) compact accelerators; and (5) compact pulse generators.

Power Conversion Technologies

6

6. Power Conversion Technologies

Advanced Modulator Technology for Heavy Ion Recirculators

Roy L. Hanks, Hugh C. Kirbie, and Mark A. Newton
Laser Engineering Division
Electronics Engineering

At Lawrence Livermore National Laboratory (LLNL) we are building a small-scale recirculator to demonstrate the feasibility of accelerating ions in a closed circular path. This report describes progress made in the development of a new solid-state induction modulator cell that is a key element in this effort.

Introduction

Recirculating induction accelerators are being investigated as potential low-cost drivers for inertial fusion energy.[1,2] A recirculator is a circular induction accelerator where beams of heavy ions are accelerated and deflected in a closed path, as illustrated in **Fig. 1**. Unlike linear machines, the acceleration sequence reuses each induction accelerating cell many times to reduce cost.

The Heavy Ion Fusion (HIF) project at LLNL is constructing a small recirculator that will demonstrate concept feasibility. This small-scale recirculator is a 4.5-m diameter ring that will accelerate
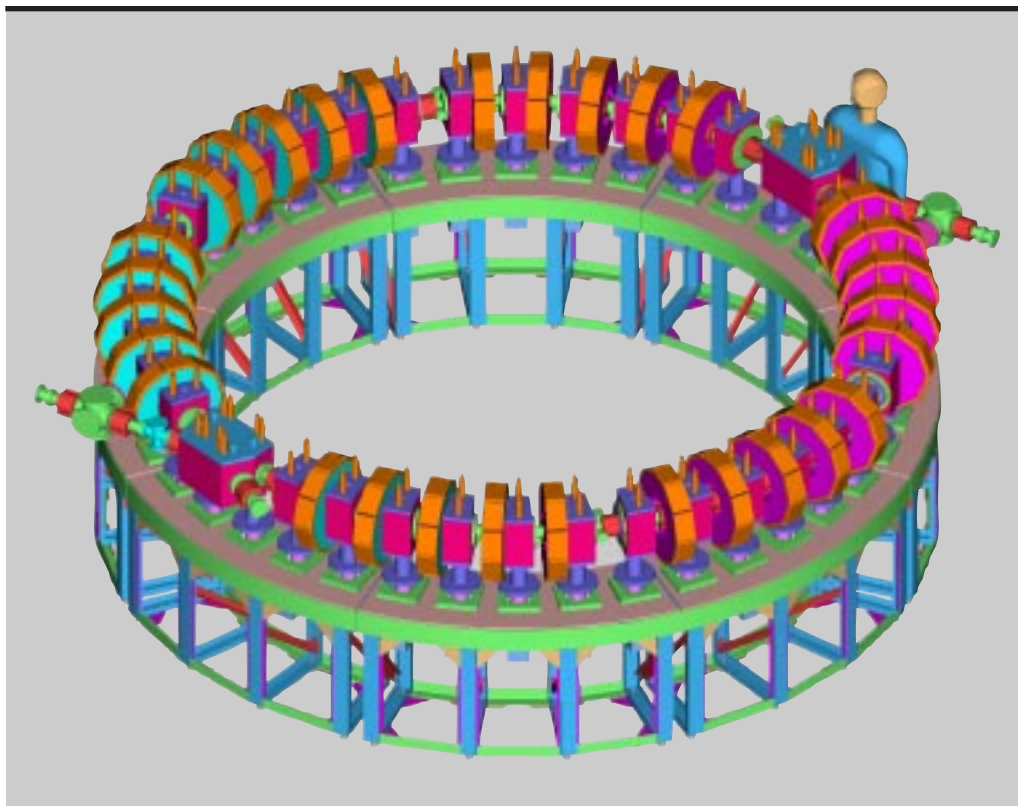

Figure 1.
Demonstration recirculator under construction at LLNL.

singly ionized potassium ions to 320 keV, after 15 laps past the 34 induction accelerating cells that make up the ring.

Electrical Requirements

A sequence of pulses must be generated by each of the 34 modulators to accelerate the beam on each lap around the recirculator. The time between
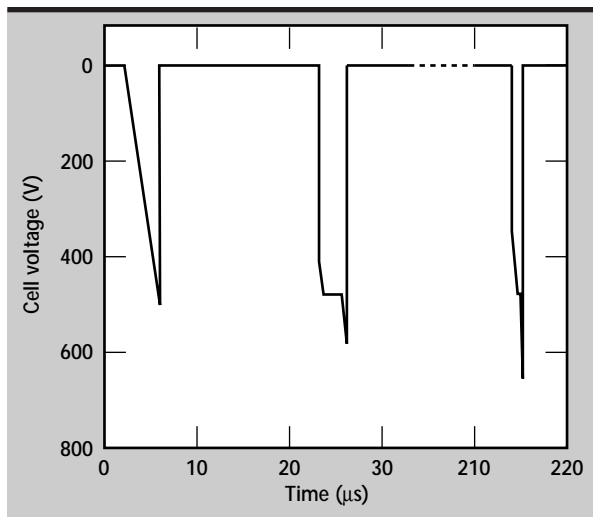

Figure 2. Simulation of a typical recirculator pulse schedule showing that each lap has a different pulse requirement. Notice a time-scale change after 30 µs, which eliminates 12 pulses from view.

pulses is equal to the time it takes for the beam to make one full circuit of the ring. As the beam energy increases, the lap time decreases from 25 µs to 10 µs, which corresponds to pulse repetition rates from 40 to 100 kHz.

Another facet of the recirculator experiments is the demonstration of longitudinal beam compression. As beam velocity increases and compression takes place, the modulator must generate pulses that reduce in width from 4 to 1 µs. To achieve compression of the ion beam, additional features are added to the modulator pulses. These features take the form of a variable risetime on the leading edge of the rectangular modulator pulses and a ramp that varies in amplitude and risetime appended to the trailing edge, as shown in **Fig. 2**. They have the effect of slowing down the head of the ion beam while speeding up the tail.

To meet these requirements, a system was proposed that uses two modulators within each induction cell. Their outputs are summed to achieve the desired voltage waveforms, as shown in **Fig. 3**. The first modulator produces the main 500-V rectangular pulse that features risetimes of 250 ns to 1 µs, pulse widths of 1 µs to 3.5 µs, and pulse top regulation of better than 1%. The second modulator generates a 30 V to 150 V ramp that modifies the trailing edge of the main pulse. This ramp varies in risetimes from 250 ns to 1 µs during the evolution of the pulse train.

Figure 3. Block diagram of induction cell.
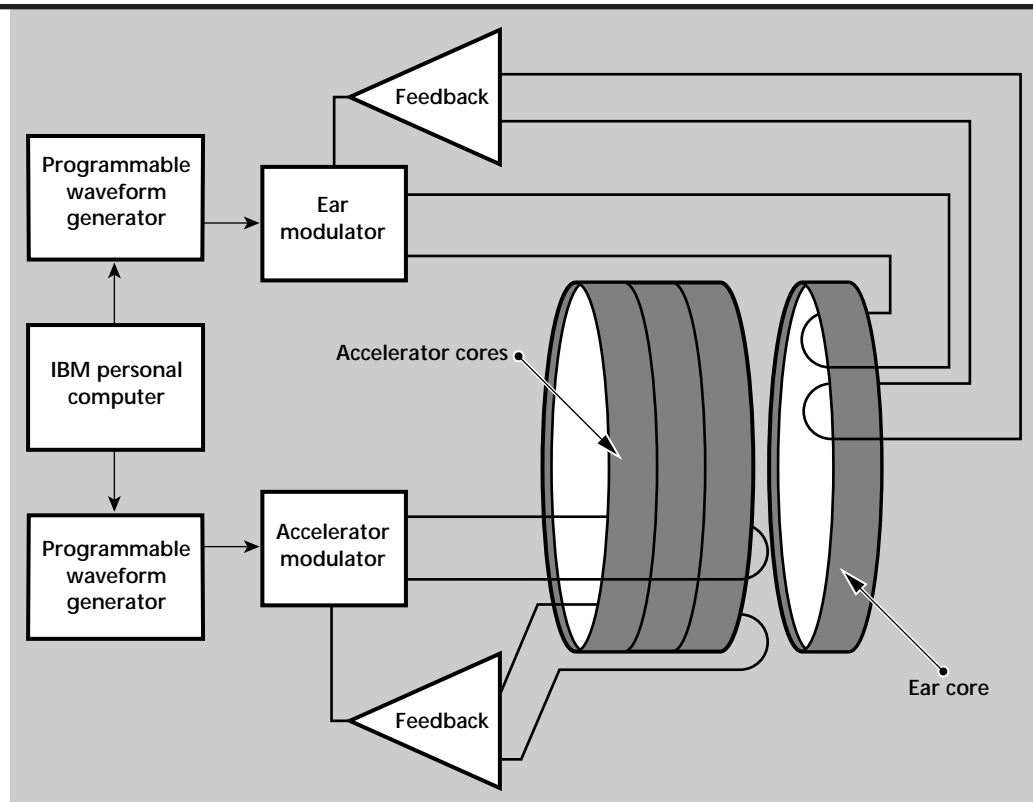
Progress

Initial evaluation of the main-pulse modulator was carried out with cores wound with Metglas 2605 CO. As a result of testing and research, it was determined that enhanced performance would be obtained with a new core of longitudinal field annealed Metglas 2605SC. This material promises to provide superior performance in terms of greater pulsed permeability and saturated flux density.

The ramp-generating modulator was first constructed using a single piece core of PE11B ferrite material. Early pulse tests indicated an unexpectedly low saturation flux density and permeability. We next wanted to evaluate the relative performance of CMD5005 ferrite material. Due to the large diameter of the toroidal core, a single piece unit was unavailable. The alternative was to purchase a core constructed of five epoxied

Figure 4. Low frequency B-H core mapping, revealing large differences in performance between PE11B ferrite (a) and CMD5005 segmented ferrite (b).

segments. To better understand their magnetic properties, both cores underwent low frequency B-H loop mapping. The test results, shown in **Fig. 4**, reveal substantial differences in the performance of the two ferrite materials.

From this data it was possible to determine the reason for the poor results encountered in the earlier pulsed PE11B core tests. Due to its "square" B-H characteristics combined with incomplete electrical resetting, the core was allowed to operate over only a limited portion of the available B-H loop.

In contrast, the CMD5005 ferrite core behaved much like a gapped core inductor. We believe this is a direct result of its segmented construction, leading to no measurable residual flux density. This has the advantage of requiring only passive electrical resetting. On the negative side the permeability is also 20 times less than that of PE11B. Soon we will select the appropriate ferrite material and modify the ramp modulator to accommodate the choice. This decision will take into account parameters such as core size, transient response, electronic circuit complexity, and modulator cost.

The modulator electronics have undergone several revisions to drive and control circuitry over the past year. Here are a few examples:

1) Apex WB05 was replaced by the Elantec EL2008. The originally selected high-current buffer was costly and scheduled to be discontinued in 1997. It was replaced by a higher bandwidth driver for one tenth the cost.

2) Analog Devices AD840N was replaced by the National LM7171. The video op-amp that provides voltage regulation was upgraded in
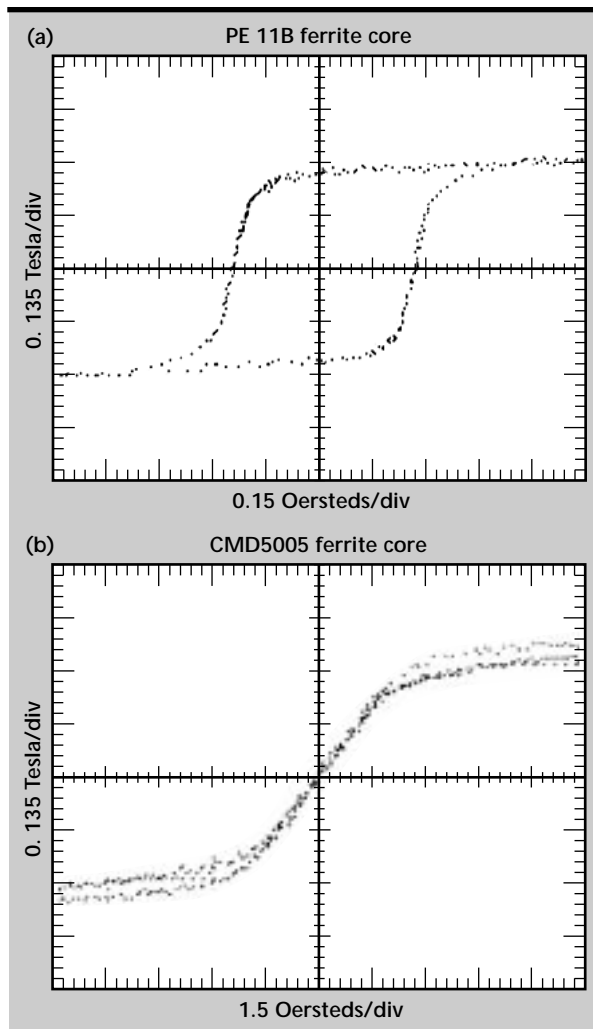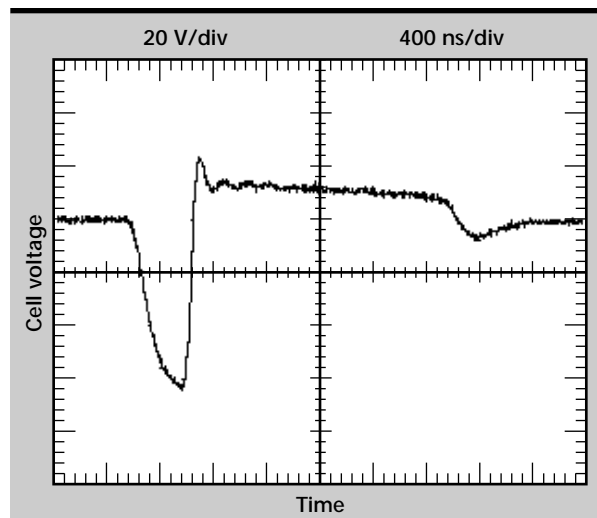

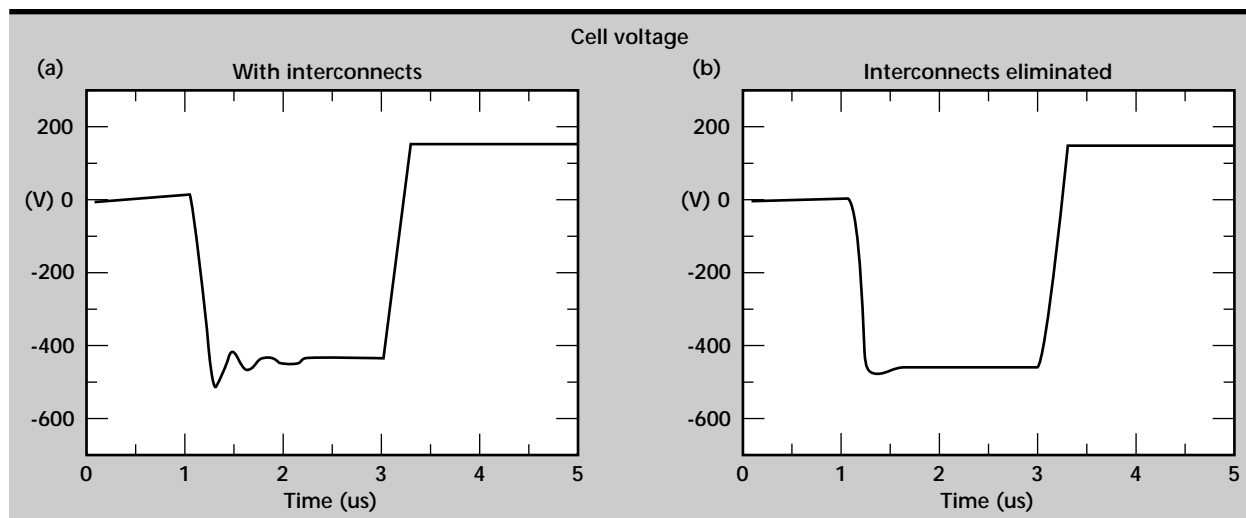Figure 5. Typical ramp voltage waveform.

Figure 6. MicroCap IV simulations, showing a significant reduction in overshoot if interconnect inductance can be eliminated.

order to take advantage of the latest generation of high-speed/high-current op-amps with improved phase margin.

3) Many configurations of the voltage feedback network have been evaluated to meet the requirements of the induction cell and also to achieve the best balance of transient and DC regulation. **Figure 5** shows a typical ramp voltage waveform using five voltage-regulated ramp modulators and a PE11B ferrite core.

4) The effect of stray inductance and capacitance on modulator performance was investigated. The primary stray circuit elements we considered have the form of package and interconnect inductance and stray capacitance that is attributed to the mechanical structure of the cell. Based on MicroCap IV simulations, a reduction in MOSFET gate inductance and interconnect lengths, shown in **Fig. 6**, could yield substantial performance improvements in the form of improved rise and falltime with less overshoot.

Future Work

As a result of the work that was conducted this past year two issues have come to light.

First, our present active reset circuit is not capable of determining if the core is initially at its positive or negative remanence state. This information is important to take advantage of the full core flux swings. The modulator circuitry will be revised such that the core will be initialized to the desired remanence state.

Second, our modulators are constructed using discrete component on printed circuit board technology. The next step is a collaboration with AlliedSignal on the design and construction of a hybrid microcircuit (HMC) that will increase packaging density and address the problem of stray and interconnect inductance. The first samples of these HMCs will be evaluated in FY-97. From these tests we hope to demonstrate the anticipated improvements in size and performance provided by this new technology.

References

1. Friedman, A., J. Barnard, M. Cable, D. Callahan, F. Deadrick, D. Grote, H. Kirbie, D. Longinotti, S. Lund, L. Nattrass, M. Nelson, M. Newton, C. Sangster, W. Sharp, T. Fessenden, D. Judd, and S. Yu (1995), "Progress Towards a Prototype Recirculating Induction Accelerator for Heavy-Ion Fusion," *Proceedings of the 1995 Particle Accelerator Conference, Dallas, Tex.,* (UCRL-JC-119538).

2. Newton, M., H. Kirbie, and R. Hanks (1995), *Development of Advanced Modulator Technology for Heavy Ion Recirculators*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-95).

Evaluation of a Compact High-Voltage Power Supply Concept

Robert L. Druce
Defense Sciences Engineering Division
Electronics Engineering

Randall E. Kamm and Roy L. Hanks
Laser Engineering Division
Electronics Engineering

We have evaluated a Russian-designed opening switch for use in a compact high-voltage power supply. The switch is for use in a charge-pump circuit which would eliminate the transformer while allowing the power supply to charge a capacitor to high voltage. The design parameters of the power supply were < 1.5 kV input voltage, 100 kV output voltage, 0.15 µF load capacitor, and < 1 s charging time. The diode switches tested proved to be too lossy for this application. Additional evaluation is not planned for this concept unless devices with characteristics better suited to this application become available.

Introduction

High-voltage power supplies typically achieve voltage multiplication by means of a transformer. The size and weight of such power supplies are ultimately limited by the amount of wire in the transformer windings and the core material. Increased output voltage increases the number of turns necessary, while increased output power increases the size of the wire and amount of core material necessary. As a result, high power and high voltage conspire to increase the size and weight of the transformer and the power supply. Increasing the operating frequency improves conditions somewhat, but winding capacitance and switch speeds limit the possible improvement.

The design of high-voltage power supplies using transformer technology is a mature discipline, precluding major improvements in efficiency or size. Charging a capacitive load provides an additional challenge to the overall efficiency, since transformer-coupled supplies are constant-voltage sources in series with an impedance. A new, transformerless approach is required to provide additional significant reduction in size and weight of these supplies.

Progress

Concept

One possible alternative is a charge-pump circuit such as that shown in **Fig. 1**. This circuit uses an opening switch with an inductive store to pump small charge packets into the capacitive load. Since this circuit functions primarily as a constant-current source, the efficiency can be improved over that of a constant-voltage source supply for capacitor charging. The main challenge with this circuit is finding an appropriate opening switch to drive the inductive store. This switch must be capable of operating at the desired output voltage at a high repetition rate and must have low losses.
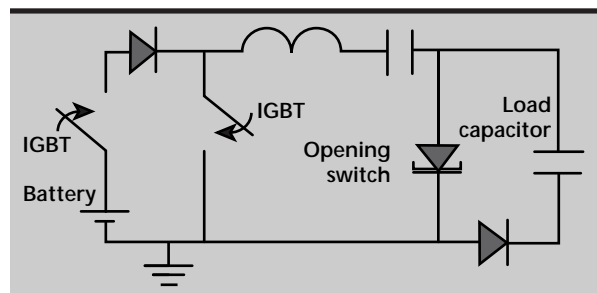

Figure 1. Simple charge-pump circuit using an opening switch for voltage multiplication.

Opening Switch

The likeliest candidate for an opening switch in this application is a solid-state device. The switch chosen for evaluation is a diode of Russian design. This device consists of a series/parallel stack of diode junctions with each junction tailored to open at a very high rate after being biased with a specific waveform. The diode stacks are rated at 100 kV opening voltage against 800 A current and 12 ns opening time. The stacks are about 10 cm long. The diodes must be driven with a ringing waveform to give the proper initial conditions for opening. These diodes are designed to be used as a pulse sharpening element. We planned to evaluate them in a somewhat different regime, giving rise to some risk.

Diode Evaluation Circuit

A schematic of the evaluation circuit is shown in **Fig. 2**. The circuit was fabricated using strip lines to give minimal parasitic circuit elements. The circuit parameters were varied over the range shown in **Table 1**. The maximum voltage was limited to a range similar to that attainable by commercial solid-state closing switches since any practical power supply would use these devices. Operation of the circuit with no opening switch diode installed agreed with calculated waveforms to within a few percent.
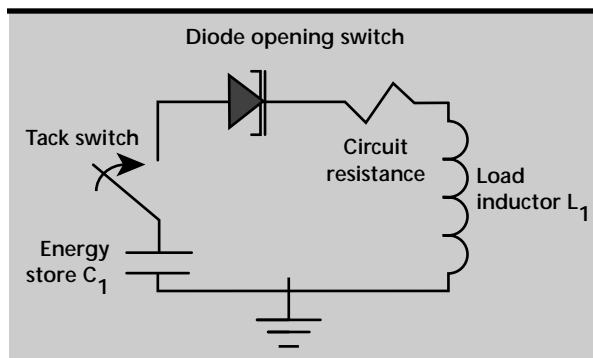
Figure 2. Schematic of the circuit used to evaluate the opening switch diodes. The switch is a solid-dielectric tack to give fast turn-on and low impedance over a wide range of voltage and current.

Table 1. List of circuit parameters and range of variation.

Circuit parameter	Range
Source Capacitance (C_1)	10 –100 nF
Circuit Inductance (L_1)	1–10 μH
Charge voltage	100 – 2,400 V

Diode Evaluation

Initial DC evaluation showed a forward bias gap voltage of $\approx$ 85 V and a forward resistance of about 2 Ω. The DC small-signal reverse resistance indicated $\approx$ 10 MΩ. The forward bias gap voltage was verified with the pulsed experiments but the forward impedance was not consistent and was possibly masked by other effects. No attempt was made to verify the reverse characteristics in the pulsed tests. The diodes showed a definite threshold current (and charge) for switching. **Figure 3** shows a typical inductor voltage and current waveforms for a switching event. The parameters for this event were as follows: V_{chg} = 500 V, L = 7.5 μH and C = 51 nF.

Note that the voltage of the inductor spikes up as the current falls when the diode switch opens. This event had a voltage gain of approximately 0.95 into an open circuit. The inductor voltage early in the waveform is approximately 200 V. This leaves about a 300 V drop across the opening switch diode. **Figure 4** is a plot of the same

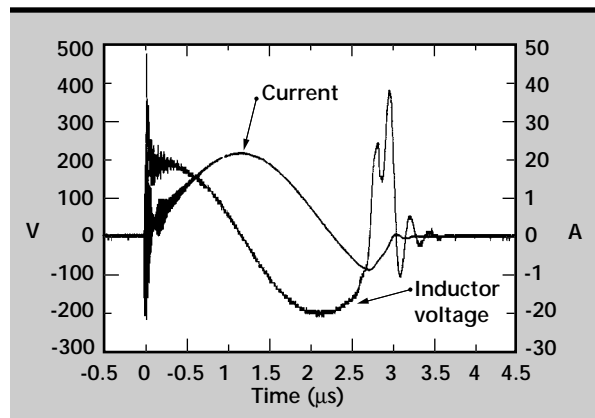*Figure 3. Plot of inductor voltage and total circuit current for a typical switching event. The parameters for this event were V_{chg} = 500 V, 7.5 μH inductor and 51 nF capacitor.*
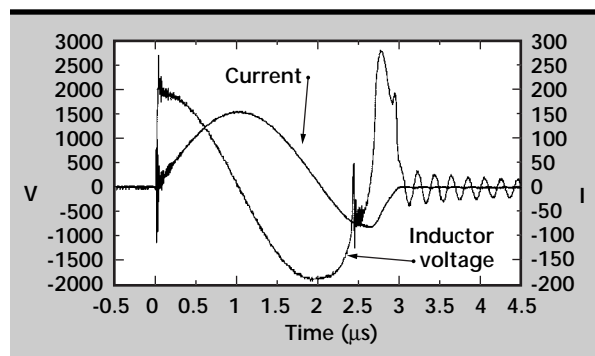
Figure 4. Plot of inductor voltage and circuit current. The charge voltage is 2400 V for this event. The circuit inductance is approximately 7.5 μH; the source capacitance is 51 nF.
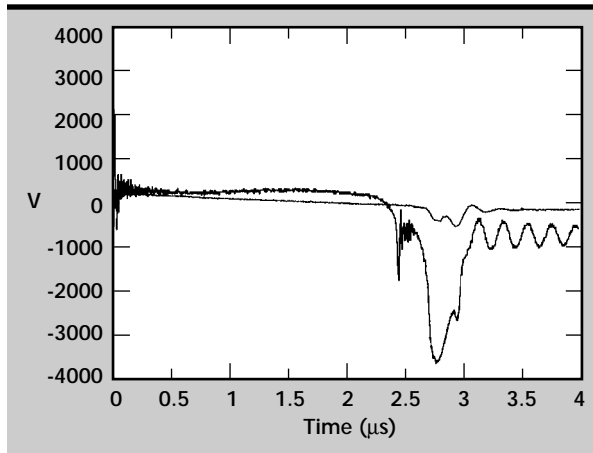
Figure 5. Comparison of the diode voltage for 500 V and 2400 V charge switching events. The waveform with the larger negative spike is the 2400 V charge waveform.

waveforms with a 2400 V charge and the same circuit parameters. The voltage gain is approximately 1.2. Note that the voltage gain is increasing with increasing voltage and current, indicating that the diodes will perform well at much higher voltage and current. **Figure 5** is a comparison of the voltage across the diode for the same two switching events, showing that the diode is

extremely lossy but that the losses are not linear with voltage or current. We have not been able to determine definitively the source of this loss.

These waveforms show typical behavior for the switch. The waveforms with differing capacitance and inductance in the circuit are similar. The gain is typically lower with less circuit inductance and small inductance showed a higher switching threshold.

Conclusion

Though the concept of a charge-pump high-voltage power supply is quite appealing, the new opening switch diode technology evaluated is not appropriate to produce the voltage gains needed in the parameter space required. These devices were evaluated in a regime that they were not specifically designed for, on the possibility that they would produce the necessary voltage gain. In light of the large gap between the required parameters and those demonstrated by the devices tested, we discontinued the investigations without extensive research into the nature of the losses in the devices.

The results of these investigations do not imply that the devices are not appropriate for pulse sharpening at higher voltage and current.

Millimeter-wave Microradar Development

Stephen G. Azevedo and Thomas E. McEwan
Laser Engineering Division
Electronics Engineering

John P. Warhus
Defense Sciences Engineering Division
Electronics Engineering

The objective of this project was to enhance the low-cost impulse radar systems made famous at the Lawrence Livermore National Laboratory (LLNL) with new ranges of frequency, resolution, and directionality. Moving into these ranges opens up many new areas of application while expanding our expertise in small low-power radar systems. Several areas were identified that are leading to projects in support of LLNL programs, as well as generating new outside funding. Small, low-cost radar systems enable many applications that use arrays of transmit/receive elements, and we are pursuing many such imaging system concepts.

Introduction

In a previous report,[1] we described many of the different types of Micropower Impulse Radar (MIR) systems that have been developed in support of numerous government and commercial systems in the last few years. The main ideas behind MIR, invented by T. McEwan at LLNL, are the generation and detection systems for extremely low-power ultra-wideband pulses in the gigahertz regime using low-cost components. These ideas, coupled with new antenna systems, timing and radio-frequency (RF) circuitry, computer interfaces, and signal processing, have been the catalysts for a new generation of compact radar systems. The systems generally fall into four sensor categories: 1) motion sensors; 2) distance sensors; 3) imaging sensors; and 4) communication devices. (For more information on these systems, and on the overall MIR Project, see past reports[2] or our world-wide web page at http://www-lasers.llnl.gov/lasers/idp/mir/mir.html.)

There are still many new directions that we plan to explore, to continue our leadership role in MIR. Rather than repeat the broad list of sensors and systems within the MIR scope, in this report we will describe our efforts in FY-96 to develop higher performance (yet still small and low-cost) radars that will expand the technology base and offer more opportunities. As an entirely new technology area, rather than a single circuit, MIR has the potential to address a rich set of applications for which there is not yet program, government, or commercial support. The modular systems and higher-frequency sensors developed in the last year will inspire numerous novel concepts that have high expected impact and return on investment. We plan to embark on areas that will enhance our engineering expertise and technology base, while providing new opportunities and capabilities for programs.

Progress

The objective of this project was to expand and advance our current capabilities in MIR technology along several directions. We discuss the key technical MIR developments in terms of the technology improvements, with applications for each.

Modular MIR Components

In an effort to produce functional radar modules that can have many features to aid system development, we have designed and developed a family of MIR boards, antennas, interconnections and software with standard interfaces that "plug and play"

together. **Figure 1** is a schematic diagram of MIR modular components that fall into four general categories of software, computer interfaces, timing and baseband processing circuits, and high-speed front ends. The interconnections between the categories are based on industry standard hardware or software components (for example, SMA or audio connectors, TTL voltages, and 50-Ω terminating resistances).

The focus of our effort has been on the timing and high-speed circuitry, where much of the family has already been implemented. Incorporated into these modules are many of the anticipated enhancements needed to develop future MIR hardware systems. For this reason, the modular MIR boards are generally larger than the first generation of boards, but their layout is such that subcircuits can be easily

reconfigured to build custom boards, or potentially application-specific integrated circuits (ASIC's), for particular applications. Both government and commercial projects are better served by the modular approach, because prototyping of radar systems becomes a simpler task and direct characterization of the individual components is more straight forward than with previous MIR systems.

Significant progress has been made in generating radar components that produce consistent and repeatable responses under most conditions. In contrast to earlier MIR prototypes, the current surface-mount board designs are very robust to shock, interference, and temperature changes. For example, a constant-fraction discriminator (CFD) has been added to the Dipstick and Rangefinder
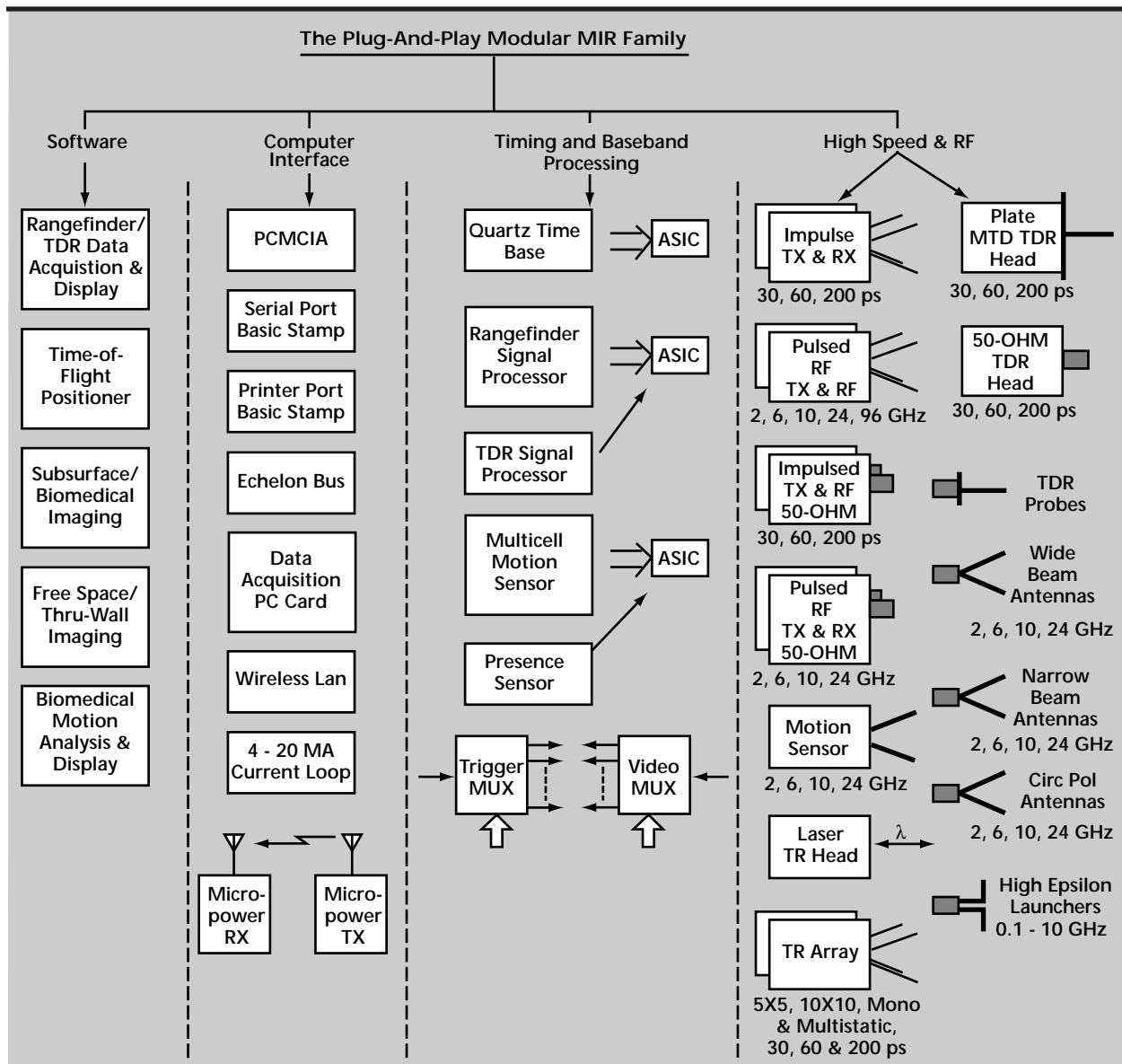

Figure 1. The MIR modular system.

systems that automatically adjusts a threshold detector to a temperature-compensated reference voltage. Then distance or range is measured by pulse-width modulation of the time between threshold crossings of MIR impulses.

For a 100-ps-wide pulse, the leading edge rise-time corresponds to about 1.5 cm in range. To hold range errors below 1 mm, we need a threshold detection accuracy of better than 6%, regardless of pulse amplitude. Compensation for voltage and thermal changes (about 1% fluctuation over outdoor temperatures) is automatic and flexible enough to operate under harsh conditions.

Out of the modular components described in **Fig. 1**, there are several complete radar systems that have been developed in the last year.

1. MIR Motion Sensor.[3] This is an enhanced version of the single-board motion sensor that can be easily reconfigured to match a specific need. Like the original, it is range-gated, low power (multi-year battery life), low cost, channel-less (multiple MIR units can operate in close proximity without RF interference), and nearly impossible to detect. A photograph of the board with simple quarter-wave antennas is shown in **Fig. 2**.

Only motion-modulated signals or changes from a baseline measurement are detected, thereby eliminating false triggers from stationary room "clutter." The motion pass-band can be changed by modifying the on-board filter components to match the application. An independent laboratory has verified that the MIR

motion sensor can satisfy FCC Part 15 regulations. Applications are in security and energy control systems, industrial safety, robotics, vibration sensing, and speech processing.

2. MIR Electronic Dipstick.[4] This is a two-board low-cost time-domain reflectometer (TDR) system that was designed to detect the height of fluid in a reservoir or container by measuring the pulse-echo time of an MIR pulse launched along a transmission line—a simple wire. The two modular boards are the quartz time base and TDR signal processor described in **Fig. 1**. Measurement of the fluid height is typically resolved to 0.1% of maximum range. There are many applications of the system in measuring fluid and material levels in industrial containers (tanks, vats, silos), hazardous materials, down-hole water levels, automotive tank monitoring, and in providing automatic fill control.

3. MIR Rangefinder.[5] This is a five-board complete impulse radar transceiver system with swept range-gate and ultra-wideband antennas. A photograph of the full modular Rangefinder is shown in **Fig. 3**. The five boards used are the quartz time base, Rangefinder signal processor, 60-ps impulse receiver, and two transmitter boards (60-ps impulse and 6.5-MHz pulsed oscillator boards). The receiver works equally well with both impulse and pulsed-oscillating transmitters. Waveform outputs of the two transmitters are shown in **Fig. 4**. It generates an equivalent-time A-scan (echo amplitude vs range, similar to a WW-II radar) with a typical
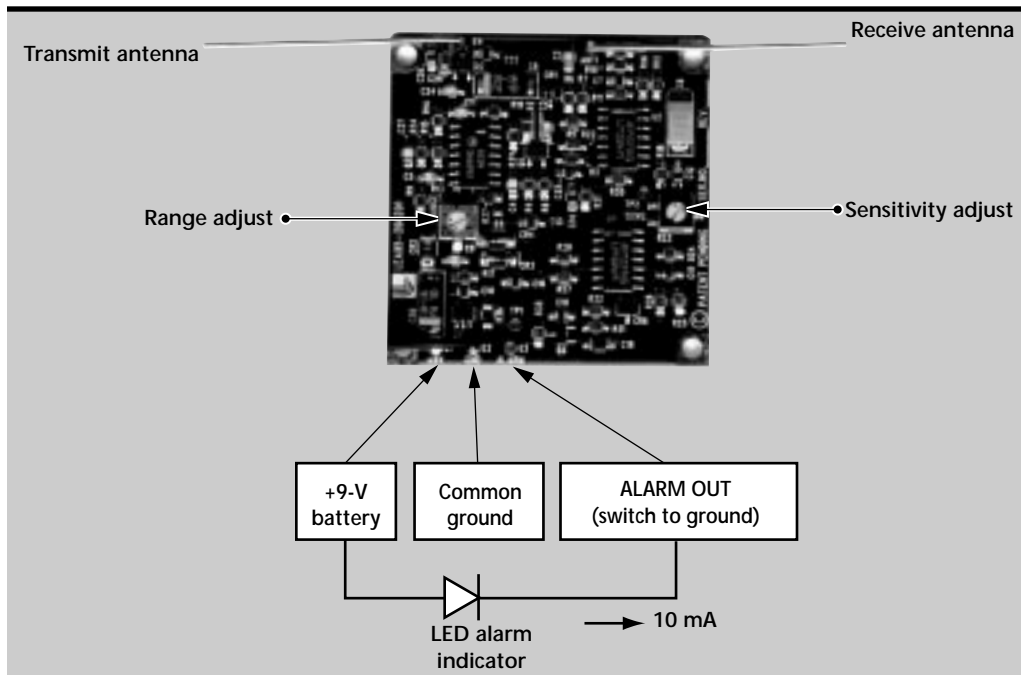
Figure 2. Photograph of the MIR Motion Sensor with simple interconnects.

range sweep of 10 cm to 3 m, and an incremental range resolution, as limited by noise, of 0.3 mm. It operates in spectral regions that readily penetrate walls, wood panels, and to an acceptable extent, concrete and human tissue.

The MIR Rangefinder is the most sophisticated of the dozens of MIR prototypes; it is the basis of all imaging applications and of many reimbursable projects. Uses of the Rangefinder include replacement of ultrasound rangefinders for fluid-level sensing (a dipstick without the stick), light-weight altimeters for unmanned airborne vehicles, localizing breathing motion behind walls, vehicle height sensing, and robotics control. When positioned over a highway lane, it can collect vehicle count, vehicle profile, and approximate speed data for traffic control.

Figure 3. Photograph of the Modular MIR Rangefinder with simple interconnects.
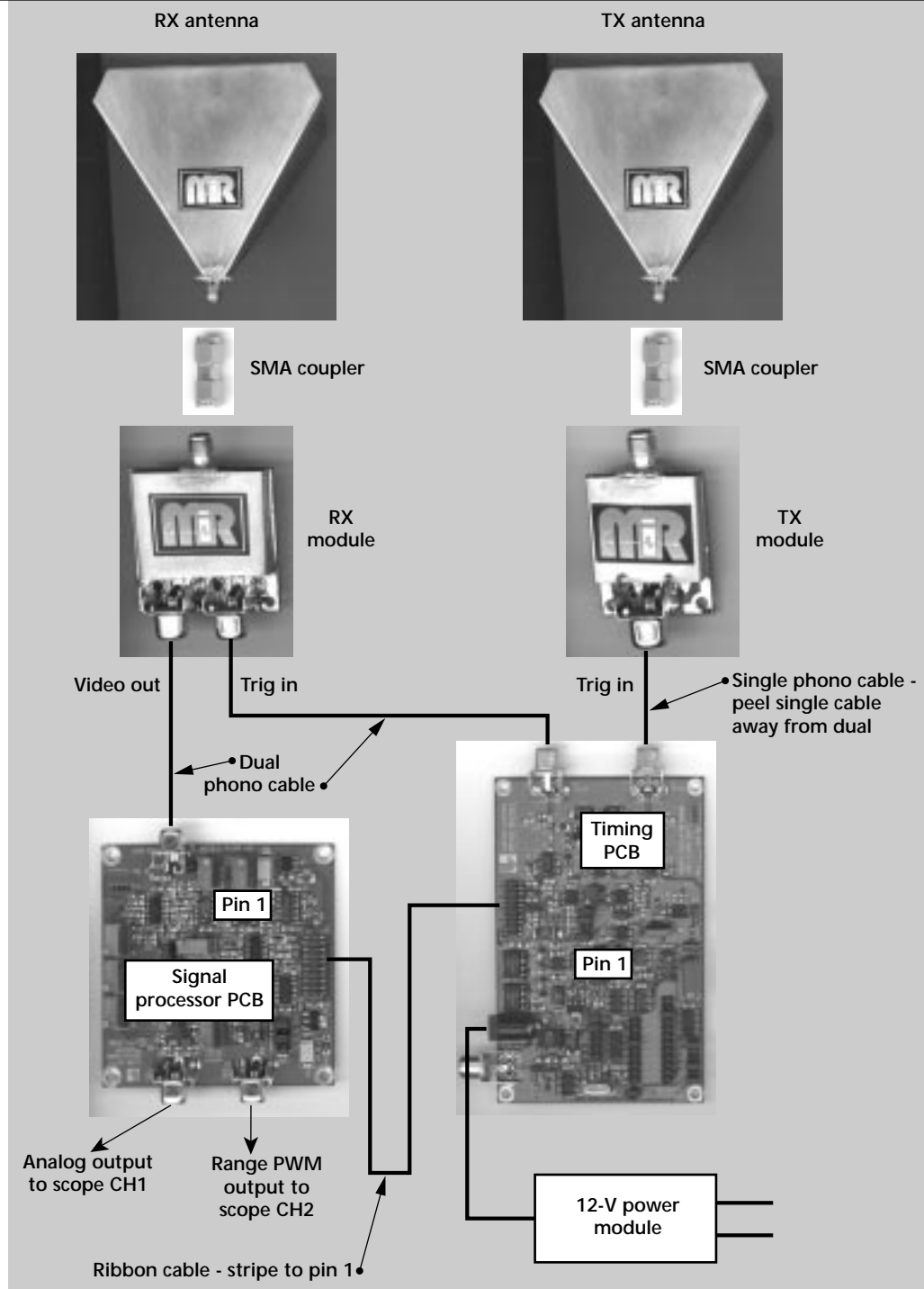
Rangefinder Extensions

As mentioned in the last section, the modular MIR architecture makes it possible to extend the Rangefinder to new levels of performance and capability. Where the original Rangefinder was a fully-integrated unit with fixed antennas, the modular design adds flexibility to our prototyping efforts. During the development of the modular design, we also added improvements to the standard MIR features (without substantially adding to the cost). For example, the CFD described above is now directly integrated into the Rangefinder design. Also included are more accurate time-base generation (0.1% over the maximum range using a crystal oscillator), high- and low-frequency cut-off filter controls, and wide flexibility in the start location and length of the scan. In this section, we describe some additional extensions to the Rangefinder made possible by the plug-and-play concept.

The modular MIR Rangefinder provides the perfect vehicle for ultra-wideband imaging applications, and we are currently working on several such projects through outside sponsorship. Multiple transmit/receive modules (yet only one time base and one Rangefinder signal processor) can be configured into arrays of radars and coupled to a computer to form either synthetic aperture or real-array imaging.[6] Radar return signals are digitized and stored in the computer. Reconstruction of cross-sectional images from B-scan or waterfall type data is performed by diffraction tomography software on the computer.[7] Images of the scene are displayed directly on the screen within 10 s (in 2-D).

We have demonstrated the use of this radar package for integration into an imaging array that is small, lightweight, low power, and inexpensive, relative to existing radars. Some of the imaging applications we are exploring are road-bed and bridge-deck inspection,[8] land mine and buried ordnance detection,[9,10] detection of underground utility lines, through-wall detection of people (for military, law enforcement, and search and rescue teams), and nondestructive evaluation (NDE) of concrete (civil structures, earthquake damage), wood (lumber evaluation, power pole rot), or, to a limited extent, living tissue (hematoma detection, kidney stones). Other materials are also possible candidates for material inspection, such as low density foams and composites.

The current arrays use monostatic imaging, but future versions will be capable of multistatic operation. The modular system makes it possible to separate the antennas to any distance and to perform both reflection and transmission experiments. In conjunction with the NDE Thrust Area, we are developing systems to explore these and other possibilities (see the article by J. Mast, in the NDE section of this report).

For the bridge-deck inspection project,[8] a high-speed radar (HSR) front-end system was developed that requires single pulse detection with no averaging so that the vehicle can travel at highway velocities and still detect subsurface flaws. While still small, the HSR has higher performance specifications (with associated higher cost and higher power) than its MIR counterpart. However, the HSR front-end hardware was made so that it attaches to the modular MIR antennas and back-end circuitry as another module to the tool set. This is another example of how all applications can now fit the common architecture.

Antennas

It will be necessary to integrate the electronics with the antennas in future versions of the imaging radar systems. This is needed to keep the
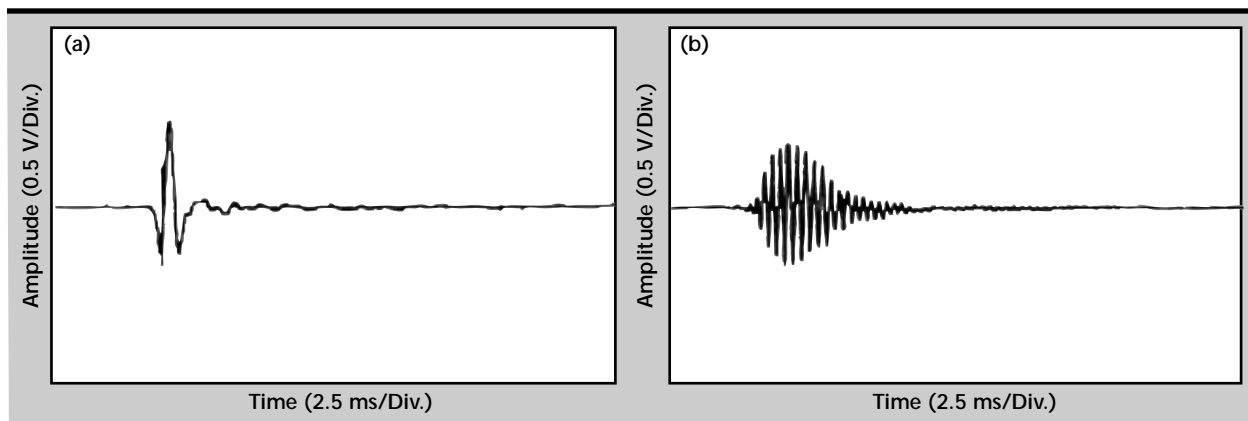

Figure 4. Video output of the MIR Rangefinder using the impulse transmitter (a) and the pulsed-oscillating transmitter (b). The transmitting and receiving antennas are pointed directly toward one another with about 20 cm stand-off.

electronics small, robust, and fast. Integrated electronics of this type will advance our capabilities in radar well beyond the current state of the art, and open numerous new areas of program development. Higher power and other directional imaging arrays may also be required for specific applications. A list of antenna constraints for the MIR Rangefinder may include:

1) s_{11} and s_{21} characteristics that are flat across a very broad band and exhibit smooth, linear phase s_{21} roll-off at the band edges;

2) group delay (dΦ/df) substantially less than the impulse width across the operating band, such as less than 50 ps across 1 to 10 GHz (equivalently, very clean step function response);

3) well-controlled, low sidelobes—no change in pulse shape vs angle;

4) low feedline-coupling into the antenna; and

5) low-cost, compact, rugged, and simple construction.

In conjunction with the Computational Electronics and Electromagnetics Thrust Area, a significant amount of effort has been directed toward stable, repeatable, and scaleable ultra-wide-band antenna designs. This work has been instrumental in improving the beam-width, bandwidth, impedance, launch point, size, shape, and cross-talk characteristics of the complete MIR system. Several antenna designs are currently being used and are pending patent consideration. The basic imaging antennas have a very broad beam width

and correspondingly low gain. They are suitable for synthetic aperture imaging where broad illumination is desirable. Narrower beam widths and higher gain can be obtained on a broadband basis with horns, reflectors, or dielectric lenses.

Frequency Extensions

Extending the frequency range of MIR from the maximum of about 4 GHz (microwave) into the millimeter-wave bands has been initiated in the last year. We have accomplished this with both impulse systems (wideband up to 12 GHz) and pulse-driven oscillators (at 20 GHz and 94 GHz). **Figure 5** shows the relationships of frequency range to resolution and penetration depth. By going to higher frequencies, wavelengths become shorter, antennas become smaller, and resolution improves to the millimeter range. Pulses at these higher frequencies can be launched more directionally and with lower sidebands. Penetration into materials is much less, but the resolution (on the order of 1/4 wavelength) is much greater. Also, for most applications the pulse will be launched in air, where losses are small. In all cases, we use the modular MIR timing and video processing on the back end while emphasizing research in the pulse driver circuitry.

An important advantage of moving into the millimeter-wave bands is that very high-gain antennas can be designed, which are small when compared to antennas with similar gain in the

Figure 5. Frequency ranges of current and proposed MIR subsystems. Notice that the resolution improves (gets smaller) as the frequency increases, but the penetration in most materials decreases at the same time.
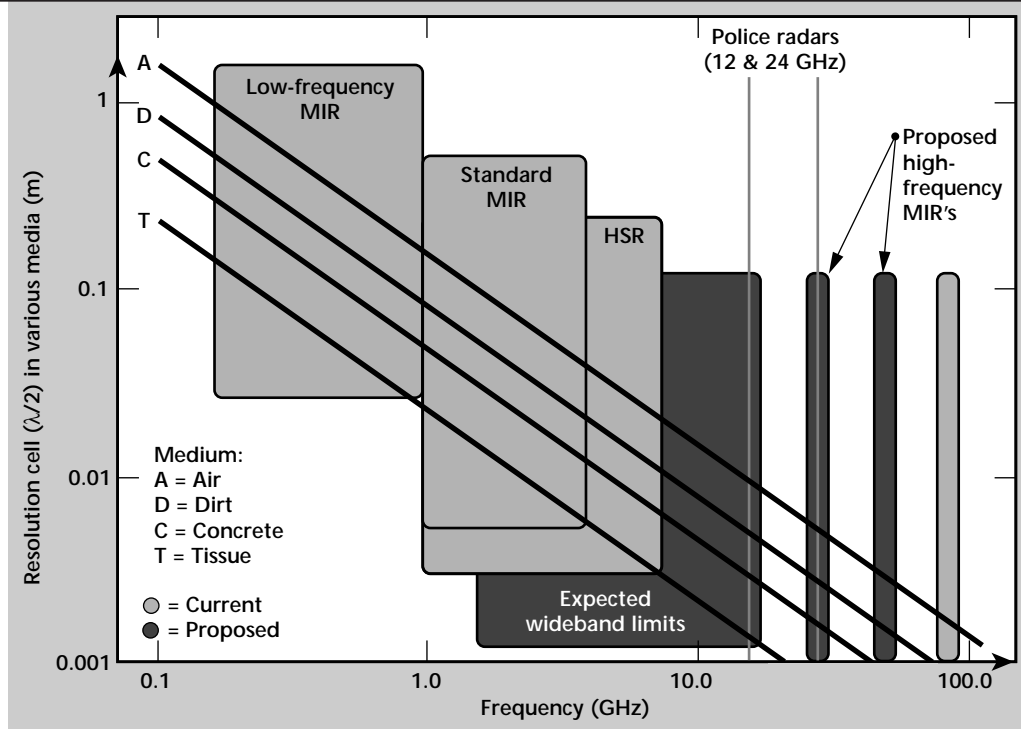
current MIR operating frequency band. Reasonably-sized millimeter-wave antennas, with improved gain and directivity can be produced to extend the range of MIR, and improve angular resolution and system portability. In addition, compact high-gain antennas will enhance the performance in some of the applications for which MIR sensors have already been proven. Examples of such millimeter-wave MIR applications include bullet/projectile tracking; high-resolution target acquisition for tanks and other weapon systems; robotic collision avoidance; airport ground traffic surveillance/ tracking; automotive systems; high-resolution personnel imaging; radar aids for the blind; and NDE of composite materials.

A prototype impulse driven 94-GHz radar has been assembled to gauge the operational capabilities of impulse-driven ultra-wideband millimeter-wave radars. The test system consists of a standard MIR backend timing and signal processing module that drives a high-speed IMPATT diode sampler. This diode oscillates for several cycles at 94 GHz and drives a high-frequency transmit antenna. Return signals are measured with a similar sampler on the receiver side, and again attached to the MIR signal processing module. A plot of the resulting waveform is shown in **Fig. 6**. Initial test data indicates a radar bandwidth of >5 GHz and transmit power >10 mW. Antennas used had a -3 db beam width of 10°. Tests indicate that a properly designed impulse driven 94-GHz radar should have a transmit power >25 mW and a bandwidth >15 GHz.

With additional development it may be possible to reach full waveguide bandwidth of 35 GHz (75 to 110 GHz). Radar repetition rates from kHz to >10 MHz are easily reached. Proper design of IMPATT diode matching networks for the input-driven impulse and RF output matching could allow IMPATT diodes to be used to create radars in the 1 to 40 GHz band with bandwidths on the order of 10 to 20 GHz. Further work needs to be done to better quantify the full capability of impulse-driven millimeter-wave IMPATT diodes.

Single-antenna Systems

The early MIR systems all had two antennas, one for transmit and one for receive. In this last year, we have performed work on single-antenna systems that use the same antenna for both functions. For example, a type of motion sensor we call the Field Disturbance Sensor (FDS) has been developed with 2- and 4-GHz pulsed-oscillators. The FDS is a range-gated homodyne motion sensor with similar

characteristics to the MIR, yet only a single antenna is needed, and that antenna can be a standard directional one rather than omni-directional.

The Rangefinder can also be used in a single-antenna configuration by means of another recently developed "directional sampler" circuit. To receive while still transmitting, a method is needed to cancel the transmit pulse at the sampler (receiver) input. In the directional sampler, the transmit pulse is applied to the top of a resistive bridge. The transmit pulse is divided equally by the bridge resistors and applied to a differential sampler comprised of a pair of charge-holding capacitors and diodes. The output of the differential sampler is applied to an amplifier where, properly tuned, the sampled transmit pulse is differenced to zero.

Future Work

We envision many additional refinements for the MIR systems of the future. Many of those are centered around insights gained from government and commercial interactions. Most applications also require some level of effort, such as a change in range/sensitivity/directionality, or size/-power/penetration, or pulse shape, or signal processing, to reduce it to practice. As proprietors of the MIR technology, we anticipate performing much of the "proof-of-principle" development work, while attracting private industry participation for mass production.

For FY-97, we plan to continue our progress toward higher frequencies and more modular systems. In this way, we can continue our efforts to characterize MIR systems and tailor them more
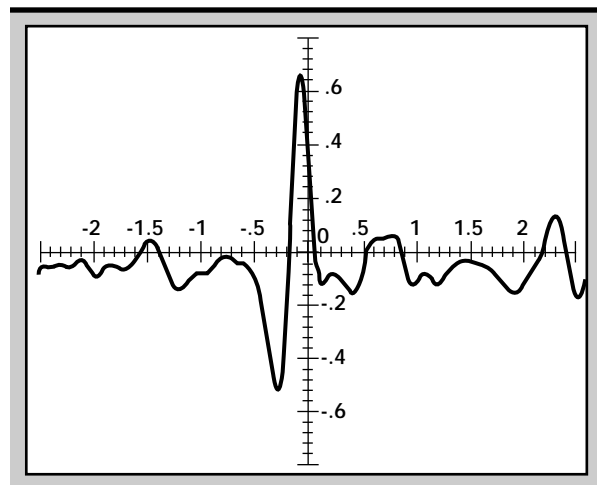

Figure 6. Pulsed 94-GHz received waveform from a metal plate at 1.32 m from the transmit/receive horns. The 94-GHz signal is sampled by the MIR back-end and shifted to the 1 to 4 GHz range.

easily to the applications at hand. There is still significant R&D needed to have reliable turn-key high-frequency MIR systems. For example, we have long considered the possibility of integrating much of the back-end MIR circuitry into a single silicon ASIC. This is still expected soon, but we will most likely carry this out in conjunction with industrial partners. While our licensees and partners are continuing to develop commercial applications of MIR, our internal efforts must stay focused on the longer-term systems issues that will take us to the next step.

Acknowledgments

References

1. Azevedo, S. G., T. E. McEwan, and J. P. Warhus (1996), "Microradar Development*," Engineering Research, Development and Technology: Thrust Area Report*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-53868-95), pp. 6–17.

2. Azevedo, S. G., and T. E. McEwan (1996), "Micropower Impulse Radar," *Science and Technology Review*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-52000-96-1/2).

3. McEwan, T. E. (1996), *MIR Motion Sensor User's Guide*, (Controlled Distribution), Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-124118).

4. McEwan, T. E. (1996), *MIR Electronic Dipstick User's Guide*, (Controlled Distribution), Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-124398).

5. McEwan, T. E. (1996), *MIR Rangefinder User's Guide,* (Controlled Distribution), Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-MA-125056).

6. Azevedo, S. G., D. T. Gavel, J. E. Mast, E. T. Rosenbury, and J. P. Warhus (1996), "Arrays of Micropower Impulse Radar (MIR) Sensors for Subsurface Detection," *Proceedings of the EUREL Conference on the Detection of Abandoned Land Mines, (IEE Conf.), Edinburgh, Scotland, United Kingdom*, Pub. No. **431**.

7. Mast, J. E., and E. M. Johansson (1994), "Three-dimensional ground penetrating radar imaging using multi-frequency diffraction tomography," *Advanced Microwave and Millimeter Wave Detectors,* (SPIE), Vol. **2275**, pp. 25–26.

8. Azevedo, S. G., J. E. Mast, S. D. Nelson, E. T. Rosenbury, H. E. Jones,T. E. McEwan, D. J. Mullenhoff, R. E. Hugenberger, R. D. Stever, J. P. Warhus, and M. G. Wieting (1996), "HERMES: A high-speed radar imaging system for inspection of bridge decks," *Nondestructive Evaluation Techniques for Aging Infrastructure and Manufacturing*, (SPIE), Vol. **2946**,(23) in press.

9. Gavel, D. T., J. E. Mast, J. Warhus, and S. G. Azevedo (1995), "An Impulse Radar Array for Detecting Landmines," *Proceedings of the Autonomous Vehicles in Mine Countermeasures Symposium, Monterey, Calif.*, pp. 6–112.

10. Azevedo, S. G., D. T. Gavel, J. E. Mast, and J. P. Warhus (1995), "Landmine Detection and Imaging using Micropower Impulse Radar (MIR)," *Proceedings of the Workshop on Anti-personnel Mine Detection and Removal, Lausanne, Switzerland*, pp. 48–51.

High-Performance Insulator Structures for Accelerator Applications

Stephen E. Sampayan and David O. Trimble
Laser Engineering Division
Electronics Engineering

George J. Caporaso and Yu-Jiuan Chen
Inertial Confinement Fusion Program
Laser Programs

Clifford L. Holmes
Applied Research Engineering
Mechanical Engineering

Robert D. Stoddard and Ted F. Wieskamp
Defense Technologies Engineering Division
Mechanical Engineering

M. L. Krogh and S. C. Davis
Allied Signal Corporation
Kansas City, Missouri

It is experimentally observed that insulators composed of finely spaced alternating layers of dielectric (<1 mm) and thin metal sheets have substantially greater vacuum surface flashover capability than insulators made from a single uniform substrate. A conclusive theory that fully explains this effect has yet to be presented. The increased breakdown electric field that these structures exhibit may result either separately or in combination from 1) minimized secondary electron emission avalanche (SEEA) growth; 2) shielding of the insulator from the effects of charging; or 3) a modification of the statistical nature of the breakdown process by separating the structure into N-1 additional sub-structures. We have previously performed measurements and reported on small- to moderate-sized insulator structures. In the previous work we showed these structures to sustain electric fields 1.5 to 4 times that of a similar conventional single substrate insulator. In addition, we previously reported on the capability of these structures under various pulsed conditions, in the presence of a cathode and electron beam, and under the influence of intense optical illumination. In this paper we describe our on-going studies investigating the degradation of the breakdown electric field resulting from alternate fabrication techniques, the effect of gas pressure, and the effect of the insulator-to-electrode interface gap spacing. Additionally, we report on initial testing that subjects the insulator to the effect of energetic radiation fields.

Introduction

A U.S. patent is pending for this work.

We have pursued the development of compact, high-current (>2 kA), high-gradient accelerator systems for various Department of Energy missions over the past several years. This work has mainly focused on a new high-gradient, prompt pulse (on the order of 10 to 50 ns) accelerator concept called the Dielectric Wall Accelerator (DWA).[1] The pulsed electric field in this accelerator is developed by a series of asymmetric Blumleins incorporated into the insulator structure (**Fig. 1**). When this structure is combined with the new high-gradient

vacuum insulator technology reported here, short-pulse high gradients of greater than 20 to 30 MV/m may be possible.[2,3]

The asymmetric Blumlein functions as follows. Two stacked pulse-forming lines, form the asymmetric Blumlein each with a different transit time and ideally, equal impedance. In the ideal configuration, these Blumleins consist of alternating layers of two dissimilar dielectric materials with permitivities, ε_r, which differ by a ratio of 9:1.

When the conductor in common with both lines is charged to potential, V_o, and shorted on the circumference of the accelerator structure, two reversed polarity wavefronts move at a velocity proportional to $\varepsilon_r^{-0.5}$ toward the beam tube.

For a fast pulse line length of time, t, and a slow pulse line length of time, 3t, an energy gain of $2V_o$ occurs across a single Blumlein structure into a matched beam load over the interval t to 3t.

The maximum gradient of this accelerator is defined by the dielectric strength of the wall dielectrics and the maximum pulsed surface breakdown electric field capability of the interior vacuum interface in the acceleration region. Most dielectric materials can support the required gradients; the vacuum insulator structures generally do not. To maximize these gradients, we have undertaken to improve the overall performance of vacuum insulators.

In addition to this particular accelerator application, we are pursuing other near-term applications. These include the Advanced Hydrotest Linear Induction Accelerator (AHF-LIA) proposed by Lawrence Livermore National Laboratory, and the Dual Axis Radiography Hydrotest (DARHT) Accelerator presently being built at Los Alamos National Laboratory. In these accelerators, high-performance insulators will be required to optimize accelerator gap design for long pulses (on the order of 2 μs) on the AHF-LIA system and also for multi-pulse options being considered for the DARHT system.

A high-gradient insulator consists of a series of very thin (<1 mm) stacked laminations interleaved with conductive planes. This insulator technology was originally conceived and disclosed by Gray in the early 1980's[4] and resulted from experimental observations that the threshold electric field for surface flashover increases with deceased insulator length (**Fig. 2**).[5,6] Further investigations showed substantial increases in the breakdown threshold of these insulator structures over conventional, single substrate insulators.[2] More recent data shows an increase of 1.5 to 4.0 times that over conventional insulator technology.[3] We have also explored the properties of these structures in the context of switching applications, investigating their behavior under high-fluence photon bombardment.[7]

A certain amount of understanding of the increased breakdown threshold of these structures can be realized from the basic model of surface flashover. The most simplified vacuum surface breakdown model suggests that electrons originating from the cathode-insulator junction are
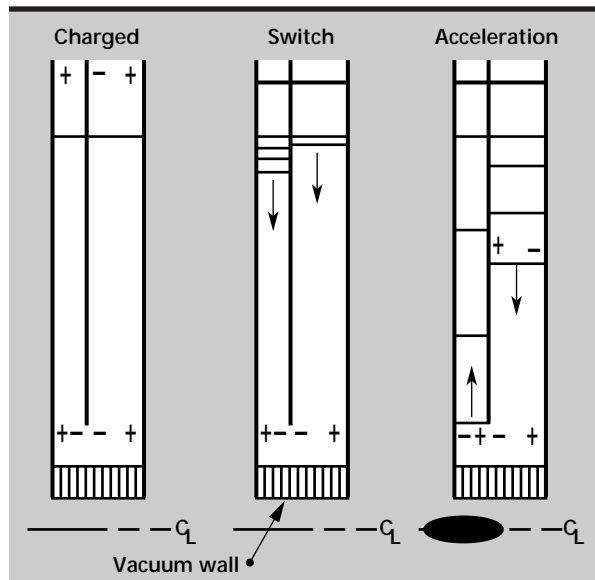

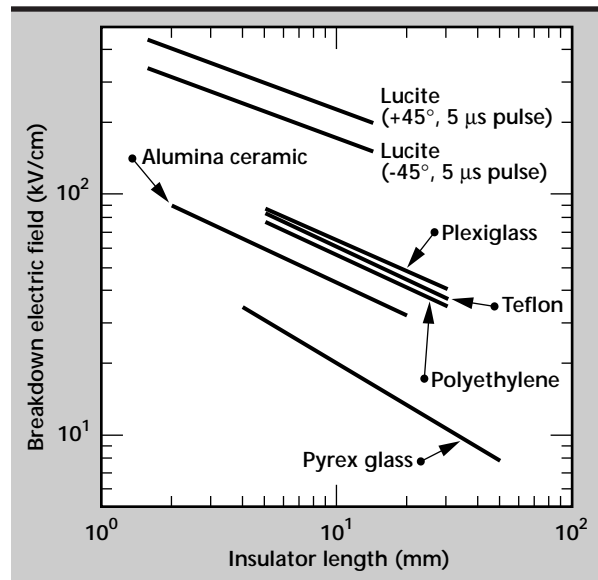*Figure 1. Principle of the asymmetric Blumlein in the Dielectric Wall Accelerator.*

Figure 2. The effect of length scaling on the surface breakdown electric field of vacuum insulators.

responsible for initiating the failure.[8] When these electrons are intercepted by the insulator, additional electrons, based on the secondary emission coefficient of the surface, are liberated.

This effect leaves a net positive charge on the insulator surface, attracting more electrons and leading to escalation of the effect or SEEA breakdown.

It has been shown that full evolution of the discharge occurs within 0.5 mm.[9] Thus, placing slightly protruding metallic structures at an equivalent interval is believed to interrupt the SEEA process and allow the insulator to achieve higher gradients before failure. Alternate modifications to this explanation include the effects of insulator shielding and equilibration of the induced surface charge. As a result, electron impact on the surface is modified. Or, alternately, by separation of the insulator into N-1 additional decoupled substructures, a local breakdown on the insulator cannot propagate to the remainder of the structure.

In this paper we describe our on-going work in which we have performed additional studies on the effect of various fabrication techniques, the effect of gas pressure, and the effect of the insulator-to-electrode interface spacing. Additionally, we report on initial testing which subjects the insulator to energetic radiation fields.

Progress

Small sample testing (approximately 2.5 cm diameter by 0.5 cm thick) was performed in a turbo-molecular pumped, stainless-steel chamber at approximately 10^{-6} Torr. High voltage was developed with a 10 J "mini-Marx." The Marx developed a pulsed voltage of approximately 1 to 10 μs (base-to-base) and up to 250 kV amplitude across the sample. Diagnostics consisted of an electric field sensor and a current viewing resistor. Failure of the insulator was determined by a prompt increase in Marx current and a prompt collapse in the voltage across the sample.

Several small sample insulators were fabricated by interleaving layers of 0.25-mm fused silica, formed by depositing gold on each planar insulator surface by a sputtering technique and then bonding the stacked layers by heating while applying pressure. Bond strength between the gold layer and substrate using this technique was measured to exceed 10 kpsi. To perform the breakdown experiments, the structure was slightly compressed between highly polished bare aluminum electrodes that establish the electric field for the tests.

To obtain a particular data set, the insulators were subjected to several low-voltage conditioning pulses. The voltage was then increased a small amount incrementally until breakdown occurred. Voltage was reduced for several shots and then incrementally increased again until a constant value was achieved. In these experiments, however, we generally observed that the insulators did not condition. Once a breakdown occurred at a particular field, reducing the voltage slightly and increasing it again did not cause an increase in breakdown field.

To produce a given data set we applied up to 150 to 200 shots to a given structure and attempted to determine if any damage to the structure occurred that significantly altered the breakdown characteristics. At these applied energies, we generally did not observe any degradation. These data were then reduced to reliability plots by determining the total number of successful shots over the total number of applied shots. In these data we define the electric field as the applied voltage divided by the total insulator length. We define reliability at a given electric field as the total number of successful shots over the total number of shots.

Using this method, we observed flashover of the small samples at approximately 175 kV/cm for the fused silica substrates (**Fig. 3**). The effect of pulse width from 1 to 10 μs on this breakdown threshold was well within the statistical nature of our data. The trend in conventional insulator technology (**Fig. 4**) for 0°-insulators indicates a breakdown threshold of approximately 50 kV/cm. Thus, with
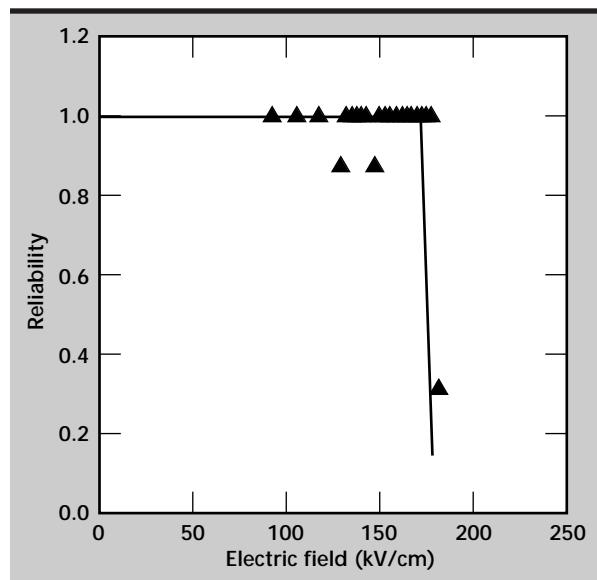
Figure 3. Pulsed surface breakdown reliability of a ground-fused silica high-gradient insulator.

these insulators there was a net increase in the performance over conventional technology by a factor of approximately 3.5.

To ensure concentricity on these first structures, a finish grinding operation was performed on the outside diameter. Since this process is a time-consuming second operation, an alternate fabrication means was pursued. To simplify fabrication, we attempted an ultrasonic machining process. Although it was possible to fabricate the part in a

single operation, the surface was left slightly rougher. Comparison of the breakdown characteristics of these samples showed significantly more scatter and on average a slightly decreased breakdown threshold of approximately 25% (**Fig. 5**).

The structures were also subjected to increased pressures to determine susceptibility to breakdown (**Fig. 6**). In these data, using the previously described procedure, a fixed reliability was established at the various pressures. All data was then

Figure 4. Pulsed surface breakdown electric field as a function of pulse width for single substrate, straight wall insulators.

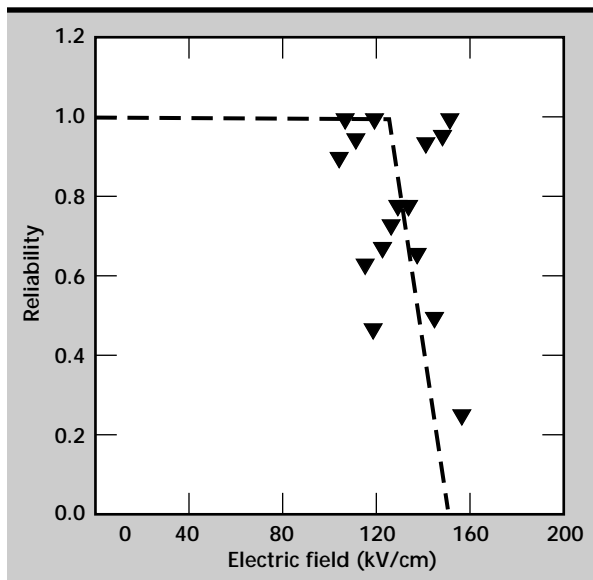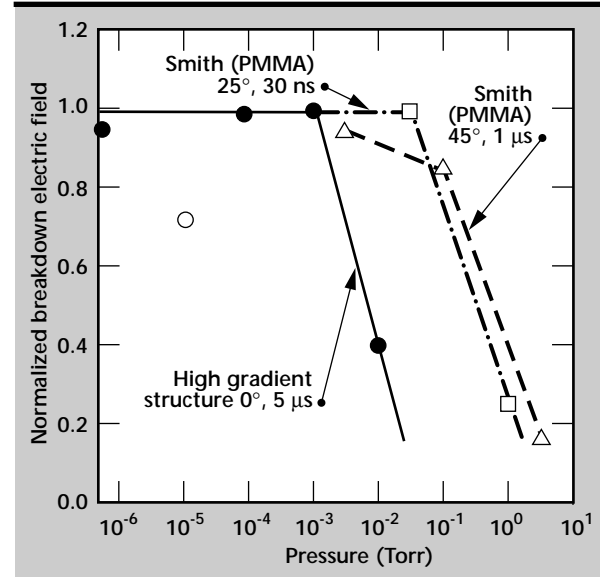Figure 6. Effect of gas pressure on the performance of high-gradient insulators.

Figure 5. Pulsed surface breakdown reliability of a fused silica high-gradient insulator fabricated using an ultrasonic fabrication technique.
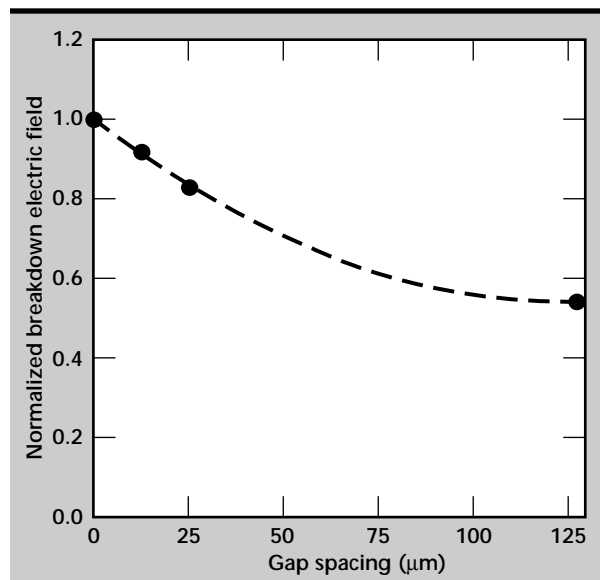

Figure 7. Effect of an increased electrode/insulator interface gap.

normalized to a mean breakdown electric field. Susceptibility to breakdown stays relatively constant up to about the 10^{-3} Torr range, at which point, the field at which breakdown occurs decreases rapidly. Also shown are data from previous work by Smith.[10] It appears that the new structures show a lower breakdown electric field threshold than that of the previous data.

Any insulator not in full contact with the electrode surface will show a higher susceptibility to breakdown and lower reliability at a given electric field. This effect results from the enhanced electric field that occurs between the insulator/electrode interface gap. To investigate this effect with these new structures, shims were placed between the cathode electrode and insulator, and the reliability at a given electric field were determined. This data, normalized to the configuration where the insulator was flush with the electrode, is shown **Fig. 7**. In these tests, we observed the reduction in the capability of the insulator to be strongly reduced from about 90% of full capability for a 12-μm interface gap to less than 60% for a 125-μm interface gap.

We have also begun testing these structures in the presence of an ion beam and various other radiation fields. In this test, we used a high-current pulsed-ion source. The ions are allowed to impinge on the structures near the cathode triple junction while a high potential is applied across the sample. Our observations to date are somewhat qualitative and indicate that the ion beam does not induce an immediate and prompt breakdown on impact. Rather, we observe only a somewhat reduced breakdown electric field capability resulting from direct ion impact on the insulator surface.

References

1. Caporaso, G. (1994), "Induction LINACS and Pulsed Power," Frontiers of Accelerator Technology, *Proc. 1994 Joint Topical Course, Maui, HI.*

2. Elizondo, J., and A. Rodriguez (1992), "Novel High Voltage Vacuum Surface Flashover Insulator Technology," *Proceedings of 1992 15th International Symposium on Discharges and Electrical Insulation in Vacuum, Vde-Verlag Gmbh, Berlin, Germany*, pp. 198–202.

3. Sampayan, S., G. Caporaso, Y. Chen, C. Holmes, E. Lauer, D. Trimble, B. Carder, J. Elizondo, M. Krogh, B. Rosenblum, C. Eichenberger, and J. Fockler, "High Gradient Insulator Technology for the Dielectric Wall Accelerator," (1995), *Proceedings of the 1995 Particle Accelerator Conference, (IEEE), New York, N.Y.*, pp.1269–1271.

4. Gray, E., (1984), private communication.

5. Milton, O. (1972), "Pulsed Flashover of Insulators in Vacuum," *IEEE Trans. Electr. Insul.,* Vol. **EI-7**, pp. 9–15.

6. Pillai, A. S., and R. Hackam (1982), "Surface Flashover of Solid Dielectric in Vacuum," *J. Appl. Phys.,* Vol. **53**(4), pp. 2983–2987.

7. Sampayan, S., G. Caporaso, M. Norton, D. Trimble, B. Carder, and J. Elizondo, "Optically Induced Surface Flashover Switching for the Dielectric Wall Accelerator," (1995) *Proceedings of the 1995 Particle Accelerator Conference, (IEEE), New York, N.Y.,* pp. 2123–2125.

8. Miller, H. C. (1988), "Surface Flashover of Insulators," G. E. Aerospace Report, (GEPP-TIS-1064-UC-13).

9. Glock, W., and S. Linke (1969), "Pulsed High-Voltage Flashover of Vacuum Dielectric Interfaces," Cornell University Report, Laboratory of Plasma Studies, Ithaca, N.Y., (No. LPS 24).

10. Smith, I. D. "Pulsed Breakdown of Insulator Surfaces in Poor Vacuum," unpublished AWRE report, Aldermaston, United Kingdom.

Compact Gas Switch Development

David A. Goerz, Michael J. Wilson, and Ronnie D. Speer
Defense Sciences Engineering Division
Electronics Engineering

Joseph P. Penland
Defense Technologies Engineering Division
Mechanical Engineering

We have developed a low-profile, high-voltage, spark-gap switch designed to be closely coupled with other components into an integrated high-energy pulsed-power source. We performed field modeling to determine the appropriate shape for the highly stressed insulator and electrodes, and employed special manufacturing techniques to mitigate the usual mechanisms that induce breakdown and failure in solid dielectrics. We have constructed and tested a prototype switch unit and achieved satisfactory operation at 100 kV levels. Preliminary tests to evaluate repetitive operation and lifetime have been encouraging.

Introduction

Spark-gap switches are commonly used in Marx generators and other high-energy pulsed-power sources. This type of switch consists of two metal electrodes, separated by a gap filled with pressurized gas contained within a dielectric housing. The switch operates when the electric field in the gap exceeds the breakdown level of the gas, or when charge carriers are introduced into the gap by some external means.

A Marx generator uses a fundamental circuit that charges many capacitors in parallel, then switches them into a series configuration to generate a high-voltage output equal to the sum of the voltages across each capacitor. In principle, Marx generators are simple, and their capability is largely determined by the basic capacitors and switches that make up each stage. In practice, the performance of Marx generators depends critically on physical layout and construction details.

In most applications, Marx generators are constructed with standard, commercially available components that are arranged in an orderly manner into compact assemblies. Since these components are made for general use, their packaging and terminal styles are usually less than optimal for achieving maximum performance and the compactness needed

for some applications. Significant improvements can be realized by modifying standard components or by manufacturing special components to fit into integrated packages.

We have developed a design for a repetitively-switched, Marx-type high-voltage generator based on custom components that can be closely coupled and integrated into an extremely compact assembly. This ultra-compact Marx (UCM) can be used in a variety of special applications requiring a compact high-voltage pulsed-power source. The concept relies on a low-profile, low-inductance, high-voltage, spark-gap switch with the following performance levels:

Hold-off voltage: 100 kV
Peak current: 30 kA
Repetition rate: 10 Hz
Charge transfer: 0.1 C
Inductance: 5 nH
Capacitance: 130 pF

Progress

We have evaluated many design options and performed field modeling to evaluate the electrical stresses and quantify field enhancements of different shaped assemblies. Through an iterative process, we have been able to identify the limiting features

and devise suitable design and construction methods to satisfy the basic requirements. **Figure 1a** is a sketch of the compact gas switch.

The shape of the metallic and dielectric parts is crucial to properly manage the electric fields and keep the stresses below the threshold for flashover or breakdown of a material. **Figures 1b** and **1c** show field modeling results where contours of equipotential lines are plotted. The metal electrodes and insulator surfaces are appropriately shaped to reduce electric field stresses in the weakest regions where dissimilar materials meet, and to spread the fields evenly throughout the dielectric materials, allowing them to operate closer to their intrinsic breakdown levels.

Figure 1b shows that for the chosen dimensions with an applied voltage of 100 kV, the electrical field is less than 30 kV/cm at the triple-point region where the metal, plastic, and gas all meet, and the field is less than 130 kV/cm along the envelope of plastic material containing the pressurized gas. These stress levels are below the thresholds reported by others as troublesome.[1,2]

The average field across the thinnest annular region of plastic is 500 kV/cm, whereas the highest field at the outermost enhanced region is 575 kV/cm.

The gap between the cathode and anode electrodes is sized according to the desired operating level and gas pressure for a particular gas species or mixture. A nominal gap of 2.0 mm will normally breakdown at 100 kV with SF_6 pressurized to 100 psia. Field modeling allowed us to correct the empirical relationship for gas breakdown of planar gaps by including the effect of field enhancement at the edges. As evident in **Fig. 1b**, the fields are enhanced along the surface of the electrodes, reaching 560 kV/cm. This is described as an enhancement factor (f*) of 1.120, given by the ratio of 560 kV/cm to 500 kV/cm.[3]

We constructed a special test fixture to evaluate various switches and demonstrate their performance over the desired range of operating parameters. **Figure 2a** is a photograph of the assembled unit next to a high-voltage probe used to monitor the charge voltage. The slotted metal cylinder is the outer current return path for performing current ringdown tests. The finned cylindrical structure below the test fixture is a high-energy current-viewing resistor used as a diagnostic. **Figure 2b** is a sketch of the cross-section of the test fixture showing a 1-cm-thick switch body nestled between a ring of ceramic capacitors on top and a metal spacer below, with channels for flowing gas through the switch electrodes.
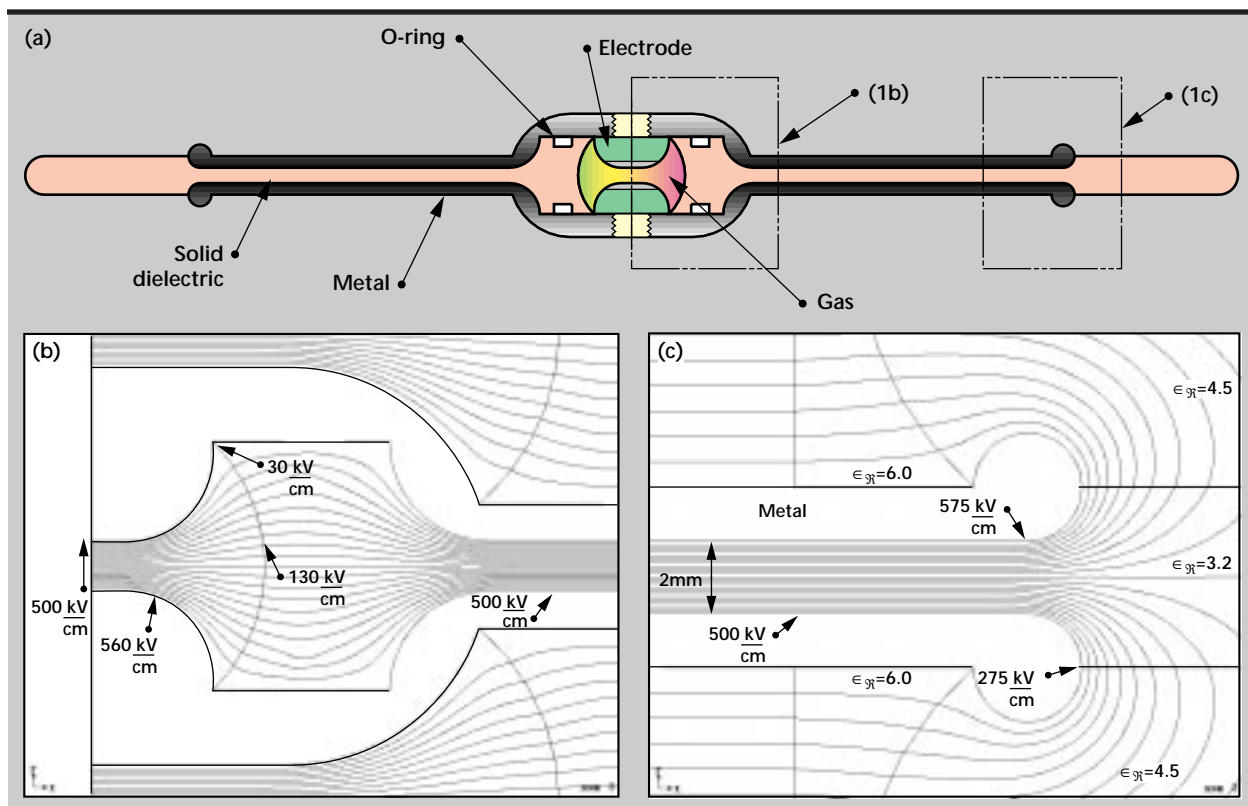

Figure 1. (a) Schematic of compact gas switch; (b) and (c) show field modeling results.
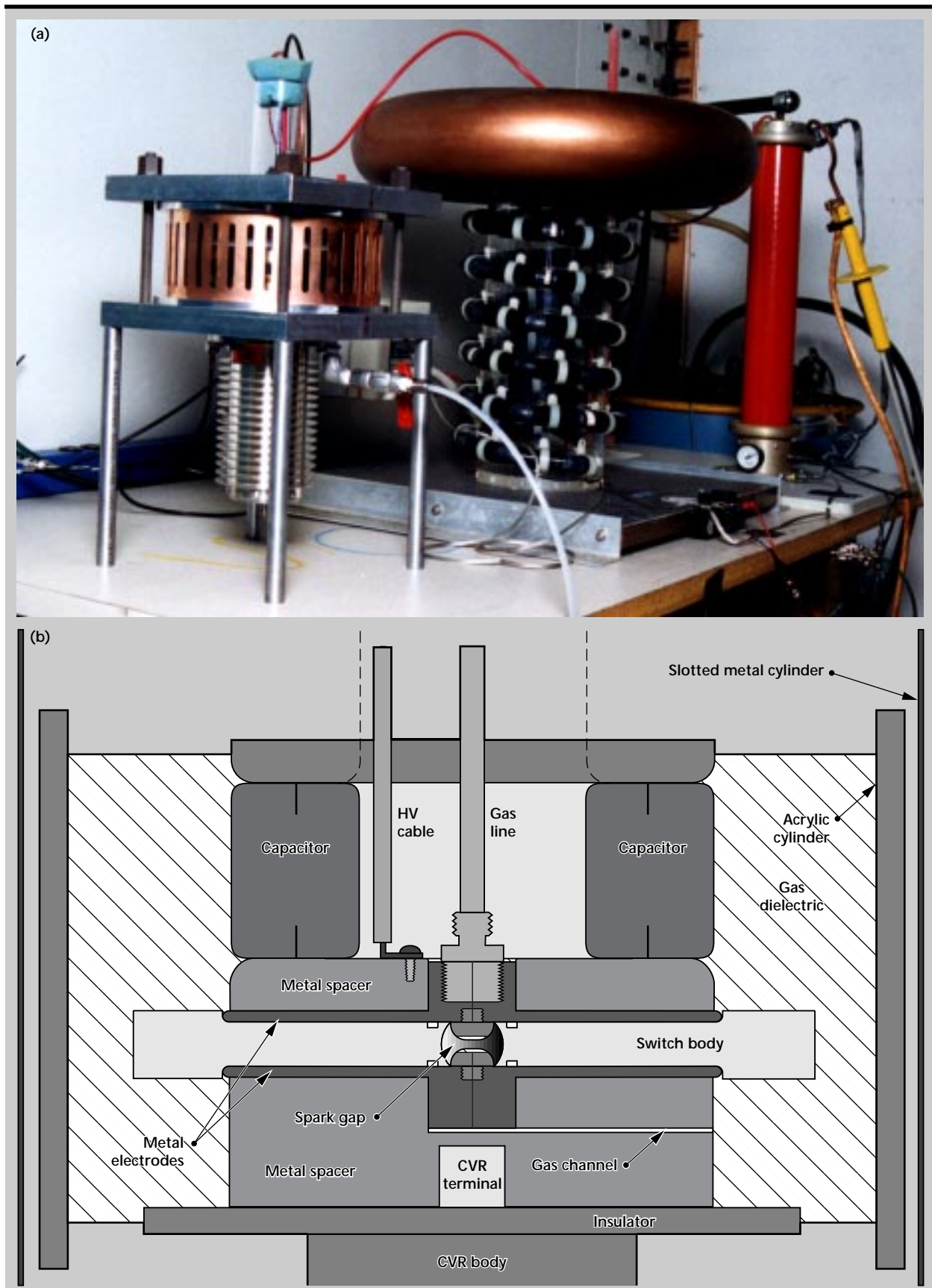
Figure 2. (a) Photograph of apparatus; (b) sketch of cross-section of test fixture.

Three different sets of electrodes were used to characterize the switch with nominal gap spacings of 2.0, 2.5, and 3.0 mm. **Figure 3** shows results from testing the static breakdown performance of the compact gas switch. The solid markers represent the breakdown measurements that were made at each pressure increment of 5 psi. The open markers represent the well-known relationship for breakdown in SF_6 gas[3] corrected for the field enhancements for these particular shaped electrodes. While there is some statistical variation in the breakdown levels, the general agreement with the empirical model indicates that the switch does indeed function as intended. Considerable testing confirmed that the electric field stress has been properly managed at the triple points and along the insulator surface, and that the switch performs as expected.

Further testing was done to exercise the compact gas switch at expected peak currents and charge-transfer levels to determine whether the insulator would survive repetitive high-energy pulses, or whether electrode erosion would become troublesome. **Figure 4** shows a typical voltage trace from a 4-s burst-mode operation of a 3-mm gap switch with SF_6 at 65 psia and the 100-kV power supply current limited by a 20-MΩ series resistor. **Figure 4** shows the current trace from the ringdown event. In this case the oscilloscope was set up to average the first 30 ringdown current waveforms.

This mode of operation was also used to determine the statistical variation of voltage breakdown levels to evaluate how well the switch would function in a Marx-type high-voltage generator. **Figure 5** shows a histogram of the number of switch breakdowns versus voltage level. For this 4-s burst the mean operating voltage was 88.8 kV and the standard deviation was 4.5 kV or 5.0%. This statistical variation is adequate for a Marx generator to function reliably with a reasonable amount of voltage coupling between stages to ensure successive switch operation.

To evaluate the expected lifetime of such a compact switch assembly, multiple 4-s bursts were taken, and the voltage and ringdown current waveforms were recorded for each shot. The switch still operated satisfactorily after more than 175 bursts, totaling more than 7000 individual shots. This series of tests was concluded when one of the capacitors

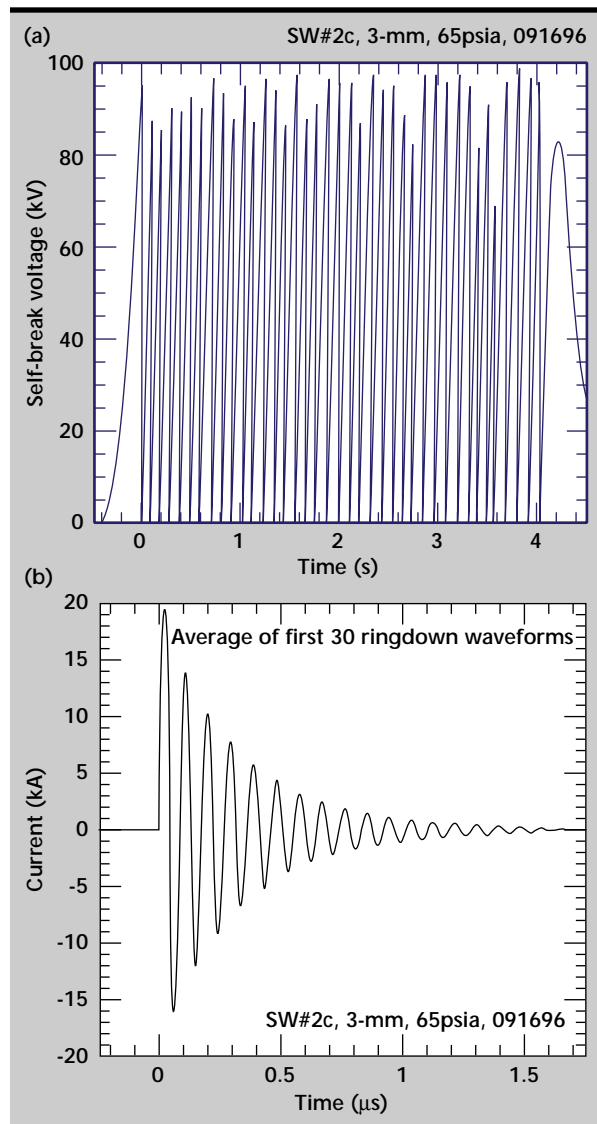Figure 3. Test results from static breakdown performance on compact gas switch.

Figure 4. (a) Typical voltage trace; (b) current trace from ringdown event.

failed. The switch was inspected and later reinstalled for further tests. **Figure 6a** shows a close-up photograph of the Lexan switch housing. No tracking or deterioration could be seen.

The anode and cathode button electrodes were weighed before and after the lifetime tests described above. The mass loss was 5.4 mg for each electrode. The total charge transferred through the switch during the period in which these electrodes were installed was 30.4 C. The specific mass loss amounted to 0.178 mg/C.

Several types of insulator materials were used to evaluate machining methods. Most of the electrical testing was done with a polycarbonate material (Lexan). **Figure 6b** shows a close-up photograph of one switch housing made from a large Lexan cylinder. Stress fractures were apparent after machining the inner switch cavity. This particular switch still operated satisfactorily for more than 1000 shots; however, the crazing became progressively worse. Switch housings were also manufactured using alumina-trihydrate (ATH) loaded epoxy and EPON-825 thermoset resin. ATH machined very easily, whereas some difficulty was experienced machining the EPON material.

Future Work

We plan to continue the development of the compact gas switch and demonstrate its performance in an ultra-compact, high-voltage Marx generator. More effort will go into evaluating higher-strength materials and developing advanced manufacturing methods for making intricately shaped parts.

We will conduct further tests on the different materials and assemblies to fully characterize their performance levels. We anticipate a need to perform 3-D analysis of electric field enhancements to determine optimal shapes of parts and interconnections for an integrated Marx package.
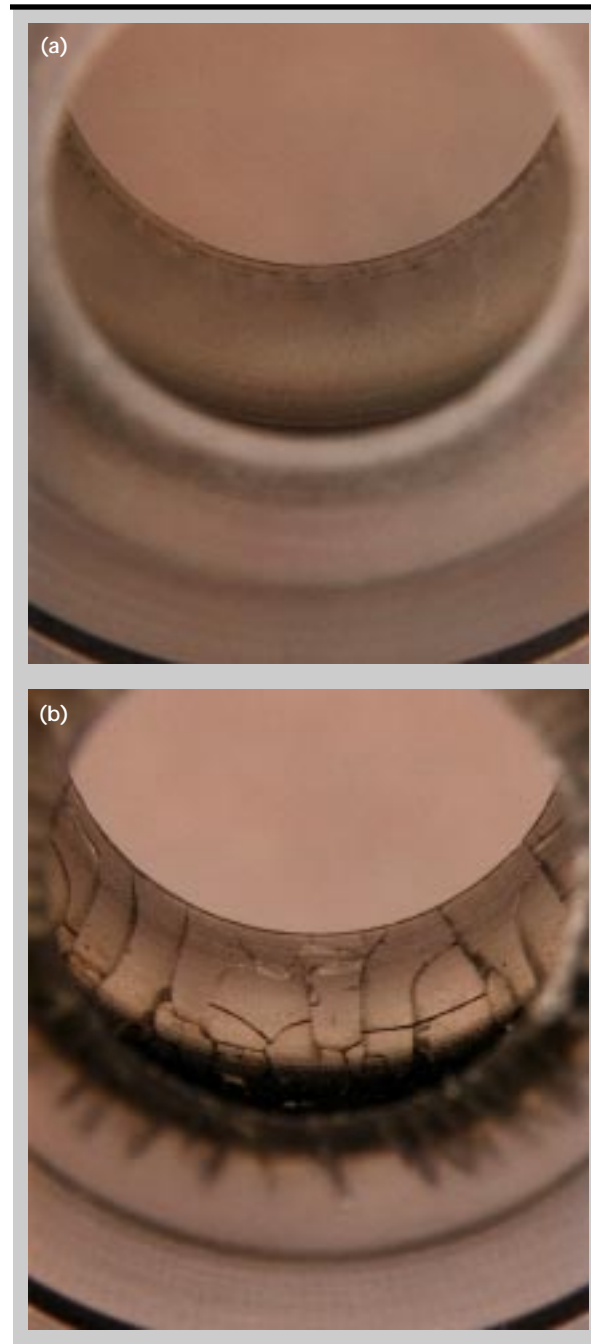
Figure 6. (a) Photograph of tested Lexan switch housing; (b) photograph of machined Lexan switch housing with stress fractures.

Figure 5. Histogram of switch breakdowns as a function of voltage level.

Acknowledgments

References

1. Laghari, J. R. (1981), "Surface Flashover of Spacers In Compressed Gas Insulated Systems," *IEEE Transactions on Electrical Insulation*, Vol. **EI-16** (5).

2. Laghari, J. R. (1985), "Spacer Flashover in Compressed Gases," *IEEE Transactions on Electrical Insulation*, Vol. **EI-20** (1).

3. Alston, L. L. (1968), "Breakdown Characteristics in Gases," High Voltage Technology, Oxford University Press, pp. 45–58.
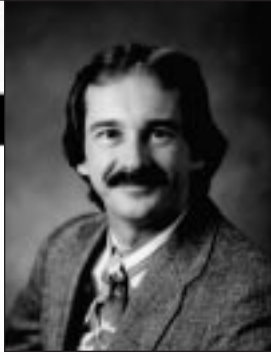
Harry E. Martz, Thrust Area Leader

The Nondestructive Evaluation (NDE) thrust area at Lawrence Livermore National Laboratory (LLNL) supports initiatives that advance inspection science and technology. Our goal is to provide cutting-edge technologies, that show promise for quantitative inspection and characterization tools two to three years in the future.

The NDE thrust area supports a multidisciplinary team, consisting of mechanical and electronics engineers, physicists, computer scientists, chemists, technicians, and radiographers. These team members include personnel that cross departments within LLNL, and some from academia and industry, within the United States and abroad. This collaboration brings together the necessary and diverse disciplines to provide the key scientific and technological advancements required to meet LLNL programmatic and industrial NDE challenges.

NDE provides materials characterization inspections of finished parts and complex objects, to find flaws and fabrication defects and to determine their physical and chemical characteristics. In addition, applying NDE throughout the life cycle of a part saves time and money and improves quality. NDE is being applied to new materials and to process design and development.

NDE also encompasses process monitoring and control sensors and the monitoring of in-service damage. Therefore, NDE is becoming a front-line technology that strongly impacts issues of certification, life prediction, and life extension.

To meet today's programmatic demands, it is important to increase collaboration among LLNL engineering thrust areas and departments. This year we have collaborated with the Power Conversion Technologies and the Information Engineering thrust areas. We have also worked with the Institute for Scientific Computing Research. Such collaborations enable us to stay at the leading edge of NDE technology, research and development.

The primary contributions of the NDE thrust area in FY-96 are described in the six reports that follow: (1) Advanced 3-D Imaging Technologies; (2) New Techniques in Laser Ultrasonic Testing; (3) Infrared Computed Tomography For Thermal NDE of Materials, Structures, Sources, and Processes; (4) Automated Defect Detection for Large Laser Optics; (5) Multistatic Micropower Impulse Radar Imaging for Nondestructive Evaluation; and (6) Multimodal NDE for AVLIS Pod Shielding Components.

Nondestructive Evaluation

7

7. Nondestructive Evaluation

Advanced 3-D Imaging Technologies

Erik M. Johansson
Defense Sciences Engineering Division
Electronics Engineering

Pierre-Louis Bossart
Manufacturing and Materials Engineering Division
Mechanical Engineering

We have developed a new software approach to the analysis and visualization of images and volumes. The efficiency of the visualization process is improved by letting the user combine small and reusable applications by means of a machine-independent interpreted language such as Tcl/Tk. These *hypertools* can communicate with each other over a network, which has a direct impact on the design of graphical interfaces. We have implemented a flexible gray-scale image widget that can handle large data sets, provides complete control of the color palette and allows for manual and semi-interactive segmentation. This visualization tool can be embedded in a data-flow image processing environment to assess the quality of acquisition, pre-processing and filtering of raw data. Our approach combines the simplicity of visual programming with the power of a high-level interpreted language. This report presents applications from biomedical imaging and industrial imaging as examples.

Introduction

Visualizing images or volumes helps extract qualitative information and quantitative measurements from raw data sets. Visualization software is thus becoming critical in virtually every domain of engineering. However, despite the number of packages available from commercial vendors and from the public domain, it is extremely difficult to find a package that fulfills the needs of a research laboratory dealing with large data sets. Indeed, all the packages are implemented based on the following scheme:

1) First, the file formats need to be decoded, and the raw data read in. The data can then be formatted or extracted. This step includes operations such as sub-sampling, interpolation or dimensional reduction, for example, when a slice is extracted from a 3-D volume.

2) Next, the data are normalized, typically to 8- or 16-bit integers, and displayed in a window after the colors have been allocated. Using pre-defined color look-up tables, stretching the histogram, reducing the image dynamic range and providing a colorbar help identify the relevant features of the data set.

3) Numerical values, extracted data, or regions of interest can then be written back into a file.

4) At any point, the user may interact with a graphical user interface (GUI) or issue commands to an interpreter, for example, to change the color look-up table or look at pixel values.

Practical experiments showed that none of the command-line oriented (such as VIEW, IDL, PVWave, or Matlab), dataflow-oriented (such as AVS or Explorer) or "self-contained" (such as Analyze or 3DVIEWNIX) visualization packages can handle large data sets whose sizes exceed the sizes of both the memory and the swap space. The user is thus compelled to manually extract smaller pieces of the data sets, which can be time-consuming and inefficient. Therefore, it is mandatory to tightly link the data extraction and data visualization steps, typically by reading and visualizing one slice at a time, instead of loading a 3-D volume in memory before visualizing its 2-D slices.

The efficiency of interactive visualization is also limited by the GUI design, which can almost never be customized by the user. Even when the source code is available, adding or removing features is very

difficult, if not impossible, since the GUIs are implemented as monoliths of hundreds of thousands of lines. However, the user needs to control the way the colors are allocated, for example, interactive thresholds, linear and non-linear color maps. Similarly, extracting profiles, histograms or non-rectangular regions of interest make quantitative measurements possible. Off-line data transforms or extractions are, in our opinion, too cumbersome in a research environment. In addition, unsupervised automatic segmentation performs poorly when the data are noisy. The alternative, that is, the use of manual and semi-automatic segmentation techniques, is limited by the lack of flexibility of most GUIs.

The considerations above led us to implement a new visualization tool geared to large data sets. Since we could not afford, nor did we have the experience required to write a self-contained application in X-motif, our approach was to divide the GUI into small reusable components by relying on Tcl/Tk, which is now used by thousands in every domain of graphical and engineering applications. Tcl/Tk provides simple ways to "glue" different modules together and can be extended easily, in contrast to other GUI builders.

However, the use of Tcl/Tk in the signal and image processing community is scarce. Two main reasons can explain this situation. First, efficient data management mechanisms and number-crunching capabilities in image processing are generally believed to require high-performance languages, in contrast to Tcl which only handles character strings. Next, the photo image widget (a Tcl image display widget) was not designed for interactive visualization, and its flexibility is very limited, mainly because image processing applications are not the main focus of Tcl/Tk developers.

In this report, we describe the implementation of a new gray-scale image widget. By focusing on data management and color allocation problems, we were able reach a level of performance which compares favorably with other image processing and visualization packages. A transparent overlay mechanism provides a link to manual and semi-interactive segmentation techniques. We will show how this widget was used to build VISU, a flexible visualization software package described below. VISU is made up of several stand-alone applications which communicate with each other over the network. The power and flexibility of these hyper-tools in image and volume visualization will be described with some examples. Applications from biomedical imaging and industrial imaging are presented as well.

Progress

The *Pict* Gray-scale Image Widget

As can be guessed from the name, the *pict* gray-scale image widget is based on the *photo* image widget. After a few experiments, it became clear that the *photo* widget was not appropriate for our application. First, this widget can only handle 8- or 24-bit color images, and it does not provide any mechanism to visualize floating-point images. Also, the colors are allocated statically and cannot be changed dynamically by the user. Thus, we decided to implement a new widget by focusing on data management and color allocation.

Data Management. Since we wanted to support byte, short, integer or float-point types, we modified the data structure to remove the fields related to color management and dithering. The raw data is allocated in a block of memory and can be accessed by using the data type information. The master structure also provides a pointer to a block of byte data, corresponding to the normalized raw data. The contrast can be increased or decreased by setting the dynamic range of the raw data. This feature proves most helpful when comparing two floating-point images whose dynamic range is different. The pixel values can be queried, and the resultant string contains the actual value, for example a floating-point value.

The Khoros and VIEW file formats are now supported which are local Lawrence Livermore National Laboratory (LLNL) formats. The GIF and PPM readers were modified to read only one color band. Readers/writers for the SUN Raster file format are provided for both the photo and the pict widget. Raw binary files can also be read from a file (or a channel if the code is linked against Tcl7.5) by specifying the dimensions, the data type and the number of bytes corresponding to the header. In all cases, only one slice is read at a time, which dramatically decreases the memory requirements. However, volume rotations and transpositions need to be done off-line. Support will be added in the near future for HDF, netCDF and ACR-NEMA file formats.

Profiling the Photo source code showed that memory management was fairly inefficient, since a lot of time is spent copying blocks between different addresses. The TkPictPutBlock and TkPictPutZoomedBlock routines were rewritten to make sure memory blocks are duplicated only when necessary.

Color Allocation. The human visual system cannot see more than 60 shades of gray, which makes color management for gray-scale images much more simple

than for color ones. Since we wanted to change the colors dynamically and use pre-defined color maps, we chose to display the images using an 8-bit pseudo-color display. This requirement is, in our opinion, fairly minimal. Besides, our experience proved that dithered images cannot be compared accurately.

To allow for fast array transformations, the colors are allocated from a contiguous set. For example, the palette can be inverted quickly by reversing the color indices. The images are displayed in false colors by choosing from a variety of pre-defined look-up tables. The histogram can be stretched or thresholded to produce a binary image. Furthermore, the colors can be allocated from shared, default or private color maps. Changing the colors of one shared color map will affect all the images that share it. This feature allows the user to visualize the same image displayed with different colors, or to compare the result of two different thresholds.

One additional benefit is that the color allocation can be used to display "semi-transparent" overlays. This is a feature that was found very useful in our interactive segmentation work. The user can, for example, "paint" on the image, draw polygons, Bezier or free-form curves, and yet guess the gray-level values. The overlays can be saved as a mask image. This feature enables the user to extract non-rectangular regions of interest. Alternatively, a mask image can be overlaid on top of the active image to check and verify the accuracy of an off-line segmentation, remove spurious pixels or compare two data sets. The user can interactively combine overlays with logical operations (such as or, xor, and) by changing the Graphic Context. The *pict* widget provides an interface to advanced image processing routines,[1,2] and the overlays can also be used in semi-interactive segmentation,[3] where a coarse initialization is specified. The segmented result can then be displayed.

VISU: A Volume Visualization Application

In this section, we describe how the *pict* widget is incorporated in VISU, a user-friendly volume visualization package. In addition to the low-level widget commands, we provide a set of default scripts and high-level commands which make the life of the average user easier. To remove any learning curve for LLNL users, the syntax of these commands reproduces that of VIEW, a general signal processing package used in our group. For example, the following commands will read the first slice of volume "hand", display the image, and then display the tenth slice:

```
% rdfile hand_0.sdt f
% disp f
% rdslice f 10
```

The simplicity of these commands enables the user to write his or her own set of macros. In addition to the command line, VISU also provides a graphical interface to most of the high-level commands. To reduce the GUI design, we chose to provide only one Control Panel and a one-to-one mapping between images and windows. At a given time, one image is considered "active". An image becomes active when the user clicks on it, and its window title is changed. The mouse location and pixel values displayed in the Control Panel correspond to those of the active window. Similarly, moving the slice scale will result in another slice being displayed. The dynamic range of the images can be typed in two entry boxes.

The Palette Menu lets the user change the colors with the mouse. Four scales can be moved to choose the low and high thresholds. The pixels whose values are between the low and high threshold appear white, and the rest appear black. As soon as the scale is changed, the look-up table is updated. Visualizing the actual values of the thresholds instead of normalized values proved very helpful. For example, in our CT applications, the user can see where the attenuation value exceeds a threshold. Pre-defined look-up tables can be loaded by clicking on one of the radio buttons. A color tool makes it possible to stretch the color map in a nonlinear way. The graphical interface allows the user to change the intensity and each of the RGB channels independently to create their own look-up table.

In the Overlays Menu, the active mask can be chosen and overlaid onto the active image. The user can choose how to combine the overlays by setting the overlay Graphic Context with the mouse. For example, the intersection of two binary masks can be seen and saved into a new image. We also provide a graphical interface to our segmentation routines.

The GUI can be customized within minutes without recompiling code. For example, displaying several images side-by-side could be done by packing them in the same canvas, instead of different windows. This flexibility enables us to design the best GUI and to take into account the requirement of a specific imaging application. In most cases, the user will be able to configure the GUI himself.

Releasing the source code on the Internet (ftp://redhook.llnl.gov/pub/visu) helped test VISU on various platforms and operating systems we did not have access to. The *configure* tool generates Makefiles automatically. While it is still dependent

on Xlib, VISU can be ported to Windows and Mac-OS without too much effort, to become a machine-independent visualization package. About 500 anonymous ftps were logged on our server. This figure may appear small, but most of the VISU users have very specific needs and would probably not use Tcl/Tk at all if this gray-scale image widget did not exist.

Hypertools

So far, the features of VISU we described are fairly standard, and can be found in other less flexible visualization packages. However, the Tk library makes it possible to use our widget to build visualization hypertools, defined by Ousterhout[4] as "stand-alone applications which can communicate with each other and be reused in ways not foreseen by their original designers". Indeed, it came as a surprise to us how easily the `send` command can be used by visualization tools based on the *pict* widget in peer-to-peer or master/slave relationships. To demonstrate the power of these hypertools, we discuss several examples.

Image Visualization. We mentioned several times in this report the importance of normalizing floating-point images in the same way. Typically, this can be done by querying the minimum and maximum value of each image, and by computing the absolute minimum and maximum of all the images being displayed. Setting the dynamic range has to be done each time an image is updated; for example, in volume visualization, the dynamic range is likely to be different for each slice. Moreover, visualization tools may run on different systems and display the images on the same screen. In summary, choosing a correct dynamic range is a tedious time-consuming task. However, the `send` command makes things simple, as the following script demonstrates:

```
proc set_range {} {
  # initializations
  set l [winfo interps]
  set min {}
  set max {}
  set visu_list {}

  # list visu applications
  foreach k $l {
    if { [string match "visu*" \
        [lindex $k 0] ] } {
      lappend visu_list $k
    }
  }
```

```
  # query  dynamic  range  of  \
      active images
  foreach k $visu_list {
    lappend min \
        [send $k {$curr_img getmin}]
    lappend max \
        [send $k {$curr_img getmax}]
  }

  # compute the global minimum and\
      maximum
  set fmin [lindex $min 0]
  set fmax [lindex $max 0]
  foreach k $min {
    if { $k<$fmin } {
      set fmin $k
    }
    if { $k>$fmax } {
      set fmax $k
    }
  }

  # set the new dynamic range
  foreach k $visu_list {
    send $k {$curr_img range} \
        $fmin $fmax
  }
}
```

In this script, the dynamic range is broadcast to all the VISU applications. Other possibilities include showing the pixel values at the same mouse location in different images, so as to compare images on a pixel-by-pixel basis. Using the VISU scripts, this option could be written as the following command:

```
foreach k $visu_list {
    send $k {set x $x}
    send $k {set y $y}
    send   $k   {set   pix   \
        [$curr_img get $x $y]}
}
```

Instead of sending values, it is possible to send commands, and, for example, to visualize the same slice in different data sets by broadcasting the command `rdslice $curr_img $curr_slice`. A direct application is the creation of simultaneous animations of volumes located in different file systems and accessed over the network.

Image Profiles and Histograms. Visualization of image profiles along an arbitrary line is an invaluable tool to evaluate the presence of artifacts in reconstructed images. For example, in x-ray CT, beam-hardening produces "cupping" artifacts that can easily be detected by extracting a profile.

Similarly, the accuracy of the reconstruction techniques can be assessed by visualizing the profiles of sharp edges. Typically, low-pass filtering and noise elimination smooth and spread the transitions.

Along the same lines, visualizing a histogram helps understand the statistics of an image and the distribution of its gray levels. In the case where the histogram is made up of several modes, the objects can typically be segmented out by applying different thresholdings to the data.

Although most image visualization packages provide 1-D signal viewers, our experience proved that they cannot be easily customized by the user and are not flexible enough.

In our applications, the 1-D signals are extracted by issuing a command to the *pict* image widget and sent to a BLT graph. Tcl/Tk handles only character strings and relies on the X protocol to communicate between applications. In addition, the profiles are generally made of less than 1000 samples, so that the overhead introduced by the communications is minimal. This feature allows the user to configure the graphical interface. The main benefit of this approach is that the user can dynamically create and extract new profiles or histograms of images viewed across the network, combine and compare them without having to save these 1-D signals in files, and transfer them by ftp.

Data-flow Environments. In our biomedical application, all the data pre-processing and segmentation is implemented in Khoros, a data-flow image processing environment. Khoros provides a network editor. The input data undergo a series of transformations each time a network node is executed by the scheduler. This data flow environment helps the user design an imaging application, as the output of each node can be visualized by connecting it to a visualization module. (Other packages such as AVS, Explorer, LabView use the same paradigm.) However, dynamically setting the range of all the images in a data flow environment is next to impossible. Similarly, writing a macro or a loop in a command-line interface sometimes makes more sense than editing a network of nodes.

These two problems were solved by implementing a Khoros module that executes a VISU script. This approach results in several important benefits. First, it compensates for some of the drawbacks of VISU, for example by providing simple ways to extract arbitrary slices. Basically, we rely on modules provided in the Khoros distribution, or write our own modules to filter and transform the raw data.

Next, the power of both command-line and data-flow environment can be combined. A command-line interface is provided by a remote controller derived from the `rmt` example of the Tk distribution, which can communicate either with a specific application or send messages to all of them. As a result, the user can rely on the power of visual programming, while being able to rely on a more traditional command-line interface. To our knowledge, this feature is not provided by any other visualization package.

Finally, the embedded visualization modules can communicate with each other. For example, the same profile could be extracted in different images and displayed in the same 1-D signal viewer. The same approach could be easily implemented in other data-flow environments such as AVS or Explorer.

Volume Visualization. As described before, VISU can display slices of a 3-D volume. Another way to visualize a volume is to extract iso-surfaces, either by using the 3-D Marching Cubes algorithm[5] or by reconstruction of a 3-D surface from 2-D contours.[6] Visualizing a surface helps understand the spatial distribution of objects in 3-space, in contrast to visualizing 2-D slices. Almost every visualization package provides these tools, but they are almost never combined or used simultaneously.

Several other approaches make use of Tcl/Tk in volume visualization. Schroeder, Lorensen and Martin recently released vtk (http://www.cs.rpi.edu:80/~martink), a visualization tool kit which can be used either by writing C++ programs, or through Tcl/Tk scripts. Similarly, Lacroute[7] released VolPack (http://www.graphics.stanford.edu/software/volpack/), a volume rendering library coupled with a Tcl/Tk-based graphical interface.

These two tool kits could be used in conjunction with VISU to build a volume visualization package by relying on the `send` command. This approach results in two main benefits.

First, the different visualization modules can be combined easily, even if they are not executed on the same machine. Typically, it will be possible to highlight a point of the surface and to see the intersecting slice. Conversely, viewing a new slice will automatically change the coordinates of this highlighted point. The color tools provided with VISU will also be used to choose iso-surface values, or the transfer functions in volume rendering. The same graphical interface can also be used to set the viewing parameters in surface and volume rendering. It is our belief that the combination of these visualization techniques will help identify more efficiently the relevant features of volumetric data sets.

Second, provided that the interface between hypertools does not change, the user does not need to know how the surface or volume rendering techniques are implemented. As a result, the best

visualization modules can be chosen by the user. For example, vtk supports an abstract rendering engine and can be slower than the tkSM widget (http://www.isr.umd.edu/~ihsu/tksm.html), which provides easy access to the OpenGL and Mesa libraries. In the case where the software modules were not linked against Tcl/Tk, they can still be used and executed from the Tcl/Tk environment.

Applications

We present two examples to illustrate the power and flexibility of VISU: a biomedical imaging application and an industrial imaging application. In the first example, we used VISU to interactively segment CT image data of a human hand, which was later used to generate a 3-D mesh for use in FEM modeling of the bone structure[8]. **Figures 1 (a)-(d)** show two cross-sections of a CT volume displayed with different look-up tables, demonstrating the capability of both private and shared color maps. The concept of transparent overlays is described in **Fig. 2 (a)-(c)**. Finally, the use of overlays in semi-automatic segmentation is presented in **Fig. 2 (d)-(f)**.

Our industrial imaging application was to use VISU to analyze CT image reconstructions of vesicular basalt. This was a proof-of-principle experiment conducted in cooperation with the Complex Systems Research Center, Institute for the Study of Earth, Oceans, and Space at the University of New Hampshire (UNH). The goal of the analysis was to determine the size distribution of vesicles (voids) in a basalt sample supplied by UNH. Because of the non-spherical nature of the vesicles, the problem was to determine the volume distribution of the vesicles and their volume fraction. Currently, graduate
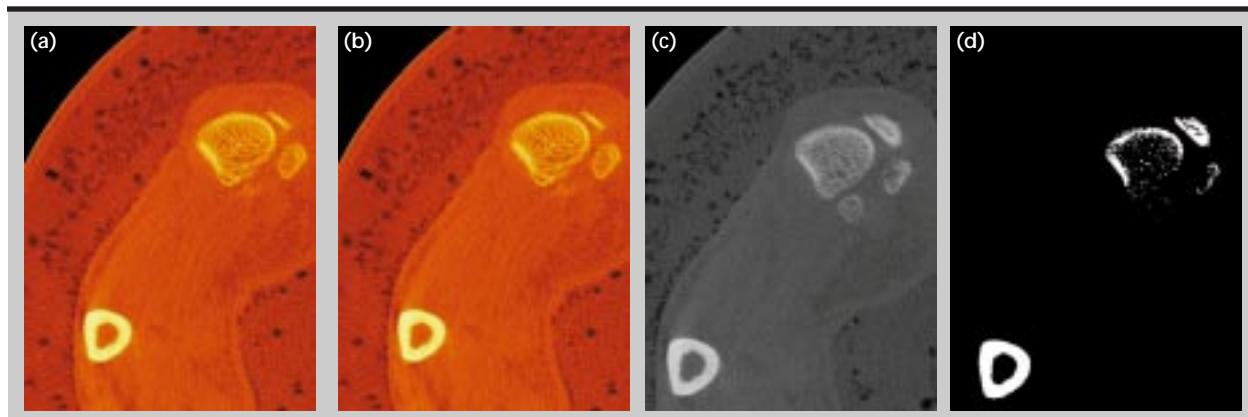
Figure 1: *Cross-sections of a thumb and index near joint: (a) corresponds to slice 308 displayed with a 'ct' private color map; (b) is also slice 308 with a shared color map; (c) corresponds to slice 314 with a 'gray' private color map; (d) is slice 314 with a shared color map. In this example, the private color maps are used as a reference while the colors are changed simultaneously in the two shared color maps.*

Figure 2: *Cross-section of a hand: (a) contours from an off-line segmentation; (b) contours overlaid on the image; (c) regions created from these contours; (d) coarse manual initialization of overlays; (e) interactive segmentation result after region-growing; (f) resulting segmented mask.*

students at UNH compute the volume fraction by (1) destructively slicing through basalt samples, (2) estimating the representative sizes of the voids in the slice using manual measurements, and (3) using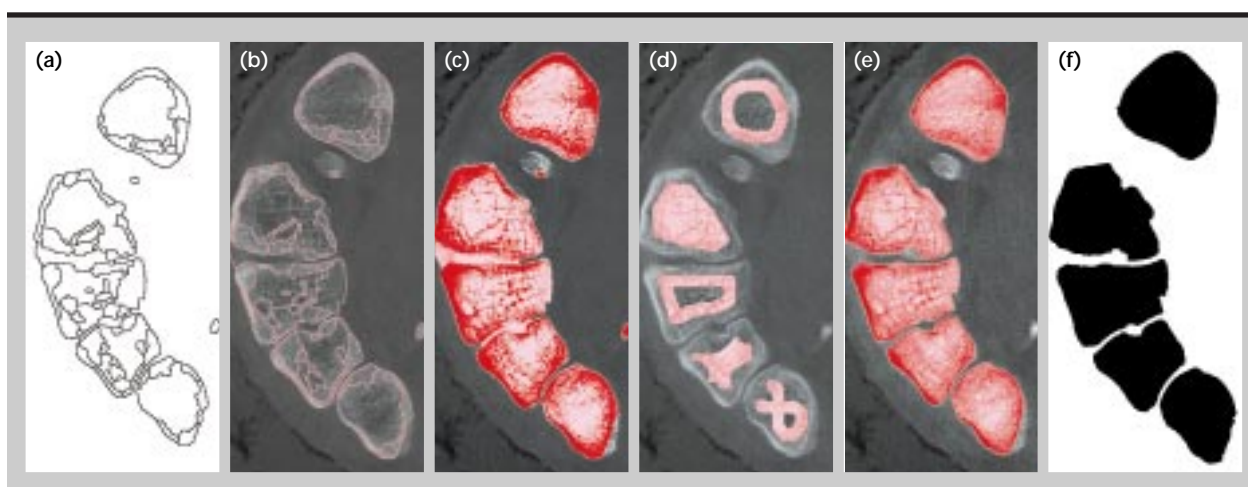 the volume of the original sample to determine the volume fraction of the voids. It is a time-consuming and error-prone process. We wanted to make the process nondestructive using CT, and automate it using the capabilities of VISU.

A photo of the basalt sample is shown in **Fig. 3**. It was taken from the Drakensburg Basalts of southern Africa. A 0.76-mm Al wire was taped to the sample for use as a fiducial in the image analysis. CT projection data were recorded using the LLNL PCAT CT scanner. 101 slices were reconstructed at a resolution of 60 μm. The analysis was performed in VISU by segmenting the image volume into two binary images, one representing voids and one representing basalt. Sample segmented image slices are shown in **Fig. 4** (voids) and **Fig. 5** (basalt). A connected component analysis was performed on the two volumes, showing that the voids are indeed complex structures connected in three dimensions. The volume of the voids was easily computed by summing over the void volume image and scaling by the pixel volume. The process was verified by computing the volume of the Al wire, which was known *a priori*. The volume fraction of the voids was shown to be 24%. The experiment was successful, showing that VISU can easily be used to accurately estimate the porosity of a volume image.

Future Work

Although Tcl/Tk is widely distributed in engineering and graphical applications, its use in imaging and visualization is recent. In this report, we described a new gray-scale widget and its use in visualization of large data sets. Although it relies on an interpreter, its speed compares favorably with typical X-motif self-contained applications, since we implemented an efficient management of data and colors. Designing high-level commands and combining hypertools helps users perform tedious tasks automatically, thus increasing the interaction efficiency. We plan to enhance our image processing library and to port the
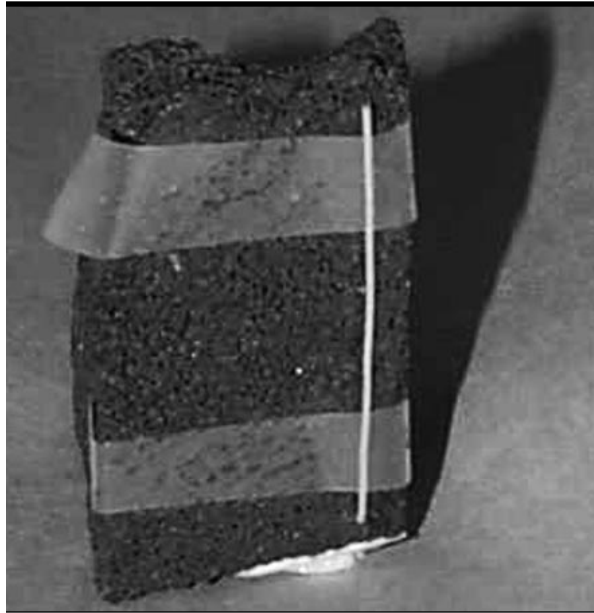

Figure 3: Photograph of the basalt sample used in the industrial imaging application. Note the very porous nature of the sample and the 0.76-mm-diameter aluminum fiducial wire taped to it.

Figure 4: A segmented image slice of the sample representing voids.
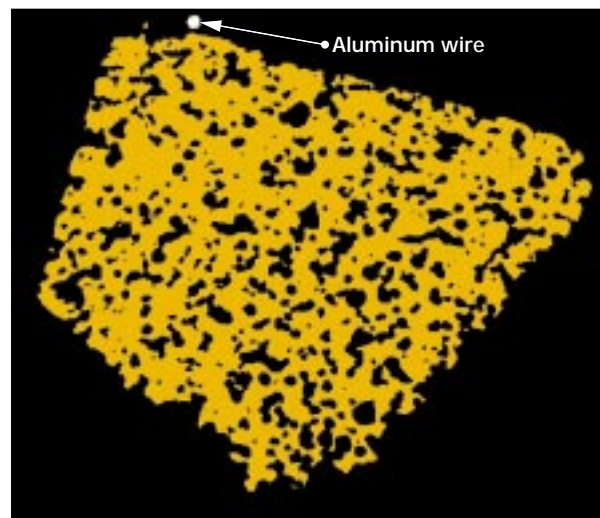

Aluminum wire

Figure 5: A segmented image slice of the sample representing the basalt. Note the cross-section of the aluminum wire at the top of the image.

code to Windows and Mac. The Tcl7.5 sockets will also be used to develop distributed visualization applications for the Internet.

Acknowledgments

References

1. Vincent, L., and P. Soille (1991), "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *I.E.E.E. Transactions on Pattern Analysis and Machine Intelligence,* Vol. **13**, (6), June, pp. 583–598.

2. Vincent, L. (1993), "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms," *I.E.E.E. Transactions on Image Processing,* Vol. **2**, (2), April, pp. 176–201.

3. Bossart, P-L. (1994), *Détection de contours réguliers dans des images bruitées et texturées: association des contours actifs et d'une approche multiéchelle,* Thèse, Institut National Polytechnique de Grenoble, France, Octobre.

4. Ousterhout, J. K. (1994), *Tcl and the Tk toolkit,* Addison-Wesley, Reading, Mass., (ISBN 0-201-63337-X).

5. Lorensen, W. E., and H. E. Cline (1987), "Marching Cubes: A High Resolution Surface Extraction Algorithm," *Computer Graphics,* Vol. **21**, (3), pp. 163–169.

6. Geiger, B. (1993), *Three-Dimensional Modeling of Human Organs and its Application to Diagnosis and Surgical Planning,* Ph.D. thesis, Ecole des Mines de Paris, France.

7. Lacroute, P. G. (1995), *Fast Volume Rendering Using a Shear-Warp Factorization of the Viewing Transformation,* Ph.D. thesis, Stanford University, Stanford, Calif. (CSL-TR-95-678).

8. Hollerbach, K. (1996), "Modeling Human Joints and Prosthetic Implants," *LLNL Science and Technology Review,* September, (UCRL-52000-96-9), pp. 19–21.

New Techniques in Laser Ultrasonic Testing

Diane J. Chinn, Robert D. Huber, David D. Scott, and Graham H. Thomas
Manufacturing and Materials Engineering Division
Mechanical Engineering

James V. Candy
Imaging and Detection Program
Laser Programs

James Spicer
Johns Hopkins University
Baltimore, Md.

Laser ultrasonic testing is an exciting new area of nondestructive evaluation (NDE). Using modern, up-to-date optical methods for generating and detecting ultrasonic waves in materials, laser-based ultrasonic testing is beginning to find a wide-range of novel applications. Several key disadvantages of conventional ultrasonic testing, a technique that has been in use for the last thirty years, are overcome with laser ultrasonics. Lawrence Livermore National Laboratory (LLNL) brings unique optics and signal processing capabilities to the development of laser ultrasonics and is poised to make major contributions to the field. As applications of NDE to LLNL programs become more demanding, with many different constraints, development of new NDE techniques such as laser ultrasonic testing is required. We are using LLNL technical expertise in collaboration with Johns Hopkins University to develop several areas of laser ultrasonic testing. In addition, we continue to support LLNL programs by developing laser ultrasonic testing solutions for a diverse set of applications.

Introduction

Technological advances in materials development, tool design and production methods present new challenges for NDE techniques. Many of these challenges are addressed by developments in laser ultrasonic testing. Conventional ultrasonic testing is commonly used in many applications to characterize material properties, find defects, measure thickness of components, study crack growth and monitor other quality parameters. However, conventional ultrasonic testing is limited to applications where high-frequency acoustic energy can be mechanically transmitted into a test specimen through a couplant. Laser-based ultrasound offers a method of coupling sound into and out of materials through optical generation and detection. Non-contacting optical testing of parts increases the flexibility of ultrasonic testing. Parts in hostile environments, in inaccessible areas, with complex geometries and in rapid production environments can all be evaluated using the non-contacting laser ultrasonic method. The major disadvantage of laser ultrasonic testing is its inherently poor sensitivity when compared to conventional ultrasonic testing.

Progress

This year we implemented both hardware modifications to our existing system and software signal enhancements to our processing algorithms to increase sensitivity and flexibility. Improvements to our path-stabilization circuitry increase the detection sensitivity level. Development of fiber-optic detection hardware permits remote sensing of ultrasonic wave fields in parts. Assembly of a data acquisition and scanning system allows the rapid collection of large amounts of data from which images can be generated.

Enhanced imaging and flaw localization algorithms increase our sensitivity by extracting more useful information from our collected data. Finally, our main effort this year focused on developing prototypes for several LLNL applications.

Hardware Modifications

Our present system uses a Nd:YAG pulsed laser to generate an ultrasonic wave and a He:Ne Michelson laser interferometer to detect the propagated wave. The generation laser delivers a 17 mJ pulse to heat a 3-mm spot on the surface of a part for 4 ns. **Figure 1a** shows the present configuration of our Michelson interferometric detection system.

Improved Stabilization Circuits. The Michelson interferometer, used to detect surface displacements resulting from the propagation of ultrasonic waves in materials, is a path-stabilized device. Ambient vibrations present in most measurements are considerably larger in amplitude than the bulk ultrasonic waves. Because Michelson interferometers are affected by ambient vibrations, a path-stabilization scheme is used to correct for these vibrations. Frequencies of ambient vibrations are generally much lower than ultrasonic wave frequencies. Path stabilization on the Michelson interferometer takes place through the reference leg and the piezoelectric tube shown in **Fig. 1a**. The path-stabilization

process involves the use of a feedback electronic circuit that drives a piezoelectric tube attached to the reference mirror of the interferometer. A correction circuit powers the piezoelectric tube, driving the reference mirror at frequencies that cancel out the effects of ambient vibrations.

Improvements in electronic devices have permitted us to make modifications to the correction circuit. We designed a new correction circuit that allows more efficient cancellation of the ambient vibrations, thereby improving the operation of the Michelson interferometer. The correction circuit uses state-of-the-art electronic components. The primary new component is the PA88 High Power Operational Amplifier circuit chip which is designed specifically to drive piezoelectric crystals and replaces several transistors. In addition, the piezoelectric tube is replaced by a stack of piezoelectric wafers that is more efficient in correcting for vibrations.

Fiber-Optic Detection. Fiber optics increase the flexibility of optical detection of ultrasonic wave propagation. Fiber-based detection permits large distances between equipment and test specimen, allows accessibility to hostile or hard-to-reach environments and facilitates changes to the test configuration. By permitting large distances between instruments and the material being tested, exposure of sensitive equipment to hostile environments can

Figure 1. Laser ultrasonic detection interferometer, modified to include fiber-based detection: (a) shows the original configuration for the Michelson interferometer; (b) shows optical fiber used for the signal leg of the Michelson interferometer; (c) displays direct beam generation; (d) shows direct beam generation with fiber-optic detection, reducing sensitivity by a factor of 4, but leaving comparable SNR.
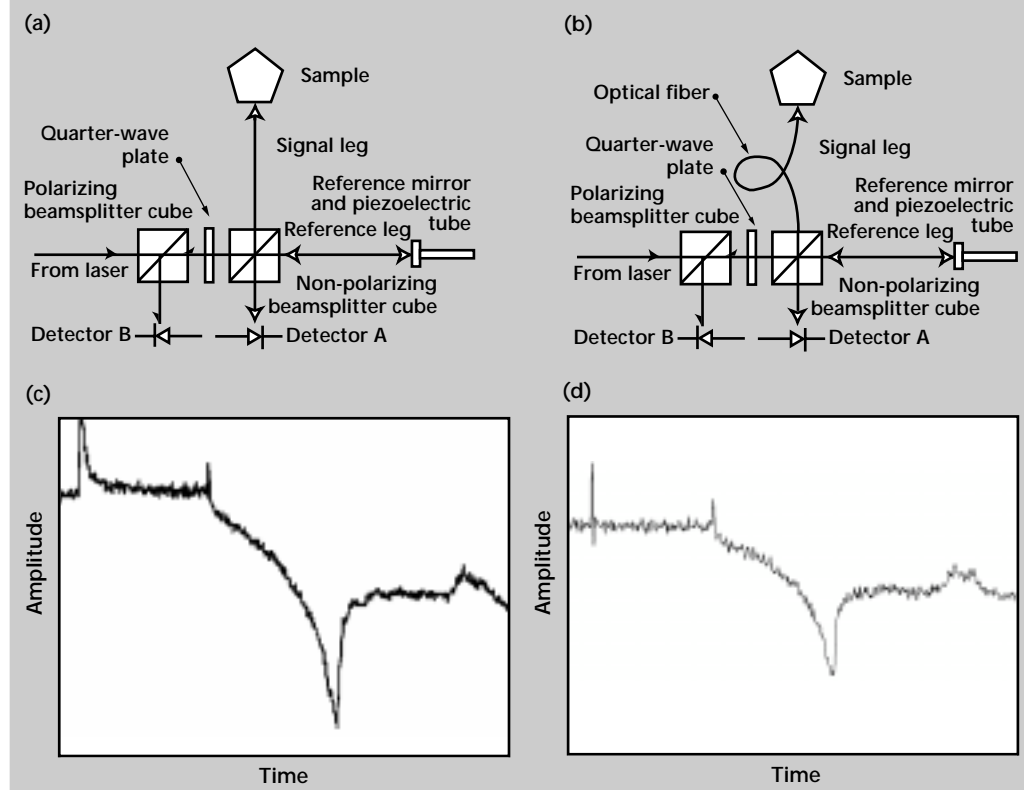
be eliminated. When scanning a part, only the fiber needs to be moved, decreasing the bulk and expense associated with staging equipment. Fibers are especially useful on applications that require testing in a glove box where large, heavy instrumentation may render testing impossible. When using light-weight fiber-optic cables, only the fibers need to be brought to the actual test location, and the electronics can remain in a remote vehicle or laboratory, or at a safe distance from the test surface.

The signal leg of the Michelson interferometer, shown in **Fig. 1b**, is placed in optical fiber. **Figures 1c** and **1d** show a comparison of direct beam (non-fiber) laser-generated ultrasonic signals in a polished, 12-mm-thick aluminum plate detected using a direct beam and a fiber-optic in the signal leg, respectively. While the sensitivity of the received signal is reduced by approximately four, the signal-to-noise ratio (SNR) remains about the same because of the smaller surface spot size. These preliminary results appear very promising for the use of fibers in LLNL applications.

Data Acquisition and Scanning System

High-resolution imaging is important to many applications where defect-sizing and mapping inhomogeneities are required. We have assembled a LabView-based data acquisition and scanning system featuring high-precision gigahertz sampling rates and scanning with micron resolution.

Signal Enhancing Software. We applied the model-based reference processor[1,2] developed last year to various experimental data sets. The processor enhances noisy measurements and provides SNR improvements. This year we investigated techniques to extend the data from our measurement system by enhancing the noisy data through the improved signal processing capability, creating a "synthetic aperture" to gather both spatial and temporal data, and developing imaging techniques that enable us to analyze the spatio-temporal displacement field in a test specimen. The goal is to use the spatio-temporal characteristics of the specimen displacement wave-field to isolate flaws or scatterers.

Imaging

To date most laser ultrasonic imaging techniques use a monostatic configuration for laser generation and detection. Both source and detector move together in tandem while the surface displacement signature is recorded. The spatio-temporal approach uses a bistatic configuration to create a synthetic receiving array. In the bistatic configuration, the

source remains fixed while displacements at the detector are recorded at multiple locations. This configuration is depicted in **Fig. 2a**, along with the expected response predicted by our thermoacoustic modeling program WAVER[3] (**Fig. 2b**). In this manner a complete spatio-temporal signal can be used to reconstruct the displacement wavefield throughout the specimen.

To demonstrate the viability of this approach, we generated a set of modeled data for a 12-mm-thick aluminum plate with a compressional wave speed of 6.32 mm/μs and a shear speed of 3.16 mm/μs. The modeled data predicts the surface displacement response from a 16-element linear array of interferometric detectors, spaced 1 mm apart, resulting from a single source sending an ultrasonic wave propagating throughout the medium.

We used this data to image the displacement field by developing a spatio-temporal filter, or beamformer.[4] A beamformer takes an array of spatially discrete responses and temporally filters each element in the array. With this filter, the
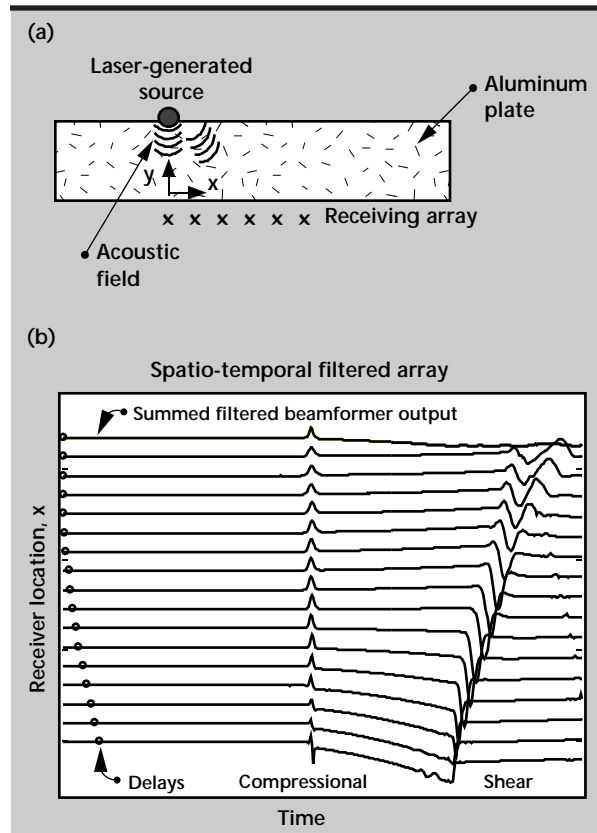

Figure 2. (a) An acoustic field in an aluminum plate, generated from a pulsed laser source. An array of surface displacements on the opposite side of the plate resulting from the thermoelastic source is modeled with the WAVER program. (b) The array of modeled responses, with both compressional and shear wave arrivals. Beamformer output is shown at the top.

displacement field is assumed to originate from a source at a particular spatial position in the test specimen (aluminum slab) and the appropriate time delays corresponding to this position are calculated. At each individual detector position, or channel, broadband signals are delayed and scaled appropriately. Finally, all of the signals are summed to calculate the location of the reconstructed source.

A typical beamformer output is shown at the top of **Fig. 2b**. The 16 delayed channel signals, summed to obtain the beamformer signal at the top, show the

compressional wave arrival. In this manner compressional displacement fields created by the laser ultrasonic source can be used to generate an image of the specimen. Each pixel in the image has the value of the maximum positive displacement from the beamformer output, derived from the spatial coordinates, channel delays and compressional wave speed, to the pixel position. The results of the modeled displacement field in **Fig. 2a** are shown in **Figs. 3a** and **3b** for both compressional and shear fields corresponding to peak positive and peak negative values. As expected,

Figure 3. Peak surface displacement images estimated from modeled data, showing the source location within the displacement field resulting from (a) the compressional wavefield (positive peak displacement) and (b) shear wavefield (negative peak displacement). Images generated from measured data show the displacement fields corresponding to (c) the compressional wavefield and (d) the shear wavefield.
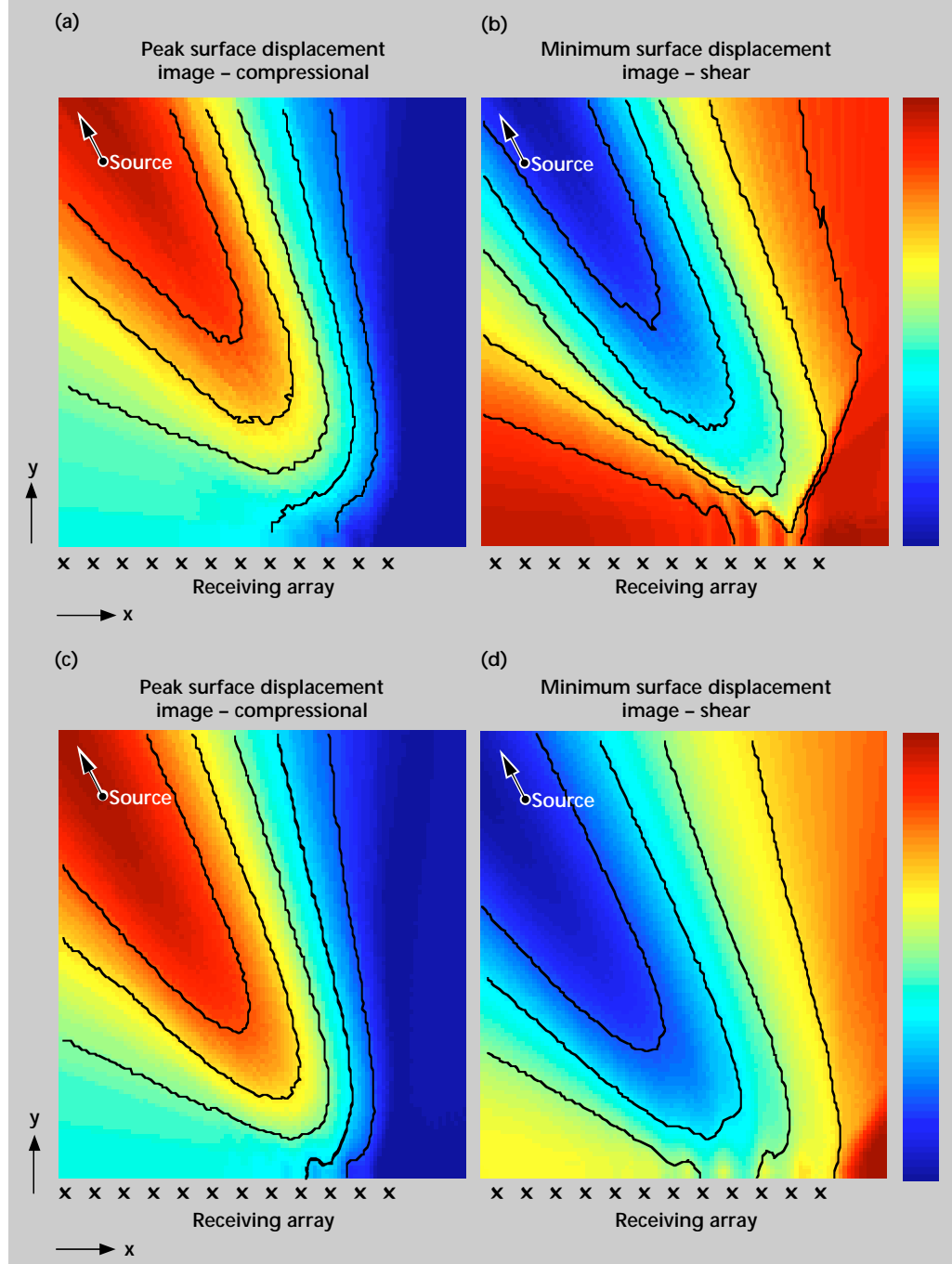
the maximum and minimum displacement values, corresponding to the source of compressional and shear waves, respectively, both occur at the origin (x = 0 mm, y = 12 mm).

Next, we validate the beamformer performance experimentally for the same material properties and dimensions under the same conditions as those in the simulation. The results are shown in **Figs. 3c** and **3d**. Here we see that the images correspond closely to images in **Figs. 3a** and **3b** that were generated from modeled data. The images for simulation and experiment are similar and demonstrate the viability of the beamformer processor.

Flaw Localization with Beamforming. We have also begun development of a flaw detector and localizer based on the beamformer imaging approach. We are extending our numerical simulator to predict scattered responses and developing validation experiments. However, to test the flaw detector we modeled the response from two ideal point scatterers located in a 12-mm aluminum plate, as shown in **Fig. 4a**. The scatterers are located between the source and the 16-element receiving array used in the previous beamformer configuration. The modeled noise-free spatio-temporal response, assuming scatter by a compressional wave, is shown in **Fig. 4b**. **Figure 5a** shows the image derived using the beamformer imaging technique. The point scatterer closest to the detectors (scatterer #1 in **Fig. 4a**) is better defined by the beamformer imaging than the point scatterer farther from the detector array (scatterer #2 in **Fig. 4a**).

Matched-filter Processing. To improve resolution as well as detection capability, we use information about the source of ultrasonic wave generation. For this modeled case the source function is used to develop an optimal spatio-temporal "matched-filter" processor. By using a multichannel time-reversal (TR) processor,[5] we filter the data with the time-reversed replicant of the scatterer propagated to the array from each assumed scatterer position. In this manner, an image proportional to the square of the surface displacement power is created, shown in **Fig. 5b**. The matched-filter image in **Fig. 5b** resolves scatterer #2 better than the beamforming image in **Fig. 5a**. We hope to continue exploring this processor next year to produce specimen images which enable reliable detection and localization.

Applications—Laser Cut Monitoring

High power, femtosecond lasers can be used to cut materials with minimal kerf and insignificant collateral damage. This type of femtosecond laser offers cutting of materials by ablation without heating the surrounding material. LLNL's Weapons Program presently has a project to study laser cutting techniques. We are applying laser interferometric detection to monitor ultrasonic waves resulting from laser cutting. Information about the ultrasonic wave propagating from the cut can be used as an important diagnostic tool for monitoring cut depth and tracking laser power.

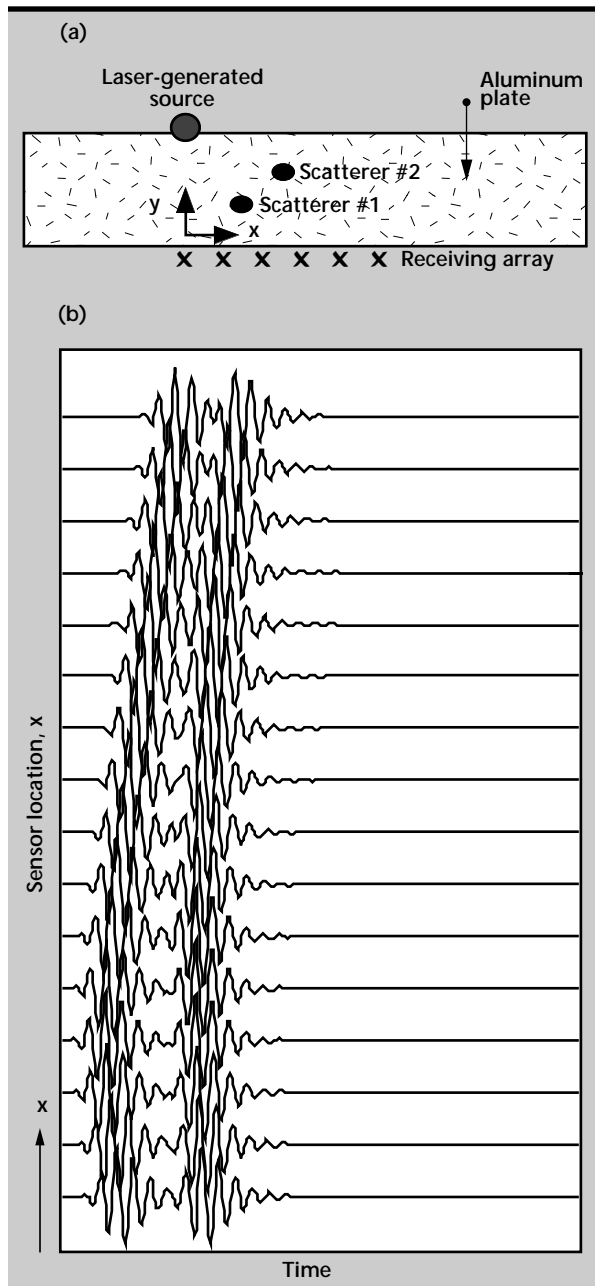*Figure 4. The response due to ideal scatterers in a plate, modeled with the WAVER program. (a) Scatterer #1 at (x, y) = (3 mm, 2 mm), and scatterer #2 at (x, y) = (4 mm, 8 mm) in an aluminum slab produce an array of sensor responses shown in (b).*

The Michelson interferometer successfully detected ultrasonic waves resulting from the cutting laser. Preliminary laser ultrasonic measurements taken during a cut of stainless steel plates by the femtosecond laser are shown in **Fig. 6**. Signal (a) is the epicentral response measured on the opposite side of the 8-mm-thick plate from the cutting laser. The compressional wave arrives as expected at 1.2 μs. Signal (b) is the epicentral response of the 4-mm-thick plate on the opposite side of the plate from the cutting laser.

The response and its multiple reflections from the plate surfaces are measured at 0.5 μs, 1.5 μs and 2.5 μs. Signal (c) is the measured response approximately 5 mm from the cutting laser on the same side of a 4-mm-thick plate. The low frequency response indicates that a plate wave has been generated by the cutting laser. The cutting laser's short duration pulse and high power generates a very broad band of frequencies with enough energy to excite several wave modes in thin plates. Generation of plate modes is advantageous because these types of waves can travel large distances through a part with minimal attenuation.

By analyzing the arrival time, amplitude and frequency of ultrasonic waves resulting from the cutting laser, cut depth information can be determined. Cut diagnosis can also be derived by the lack of ultrasonic waves propagating in a part, which would indicate cut-through. The amplitude of the ultrasonic signal depends on the amount of material ablated and thus the cutting efficiency. Monitoring the ultrasonic signal during a laser cut is a means of measuring the quality of the cutting laser beam. Both cut-through and laser quality can be used as feedback to the cutting laser. We are presently implementing a prototype system for the laser cutting project.

Summary

This year we made hardware and software improvements to increase the sensitivity and flexibility of our detection system. Fiber-optic sensing and an improved stabilization circuit increases the adaptability of our interferometer. New data acquisition and scanning capabilities, in addition to implementation of beamforming algorithms, have improved our imaging techniques. Application development for the laser cutting project has helped us become an integral part of LLNL programs.

Future Work

We will continue to develop laser ultrasonic hardware and software to enhance our imaging capabilities and increase adaptability to LLNL programs. Fiber-optic generation hardware will be designed to combine with our existing fiber-optic detection hardware, thereby completing a fully fiber-based laser ultrasonic system. With a fiber-based system, multiple fibers of different lengths will be used to generate a narrowband signal by sequential pulsing of the fibers. A narrowband signal is detected with better sensitivity than the wideband signals we presently use for generation.

Figure 5. Idealized scatterers from Fig. 4a are imaged using (a) the beam-former imaging technique and (b) the optimal time-reversal processor (matched filter).
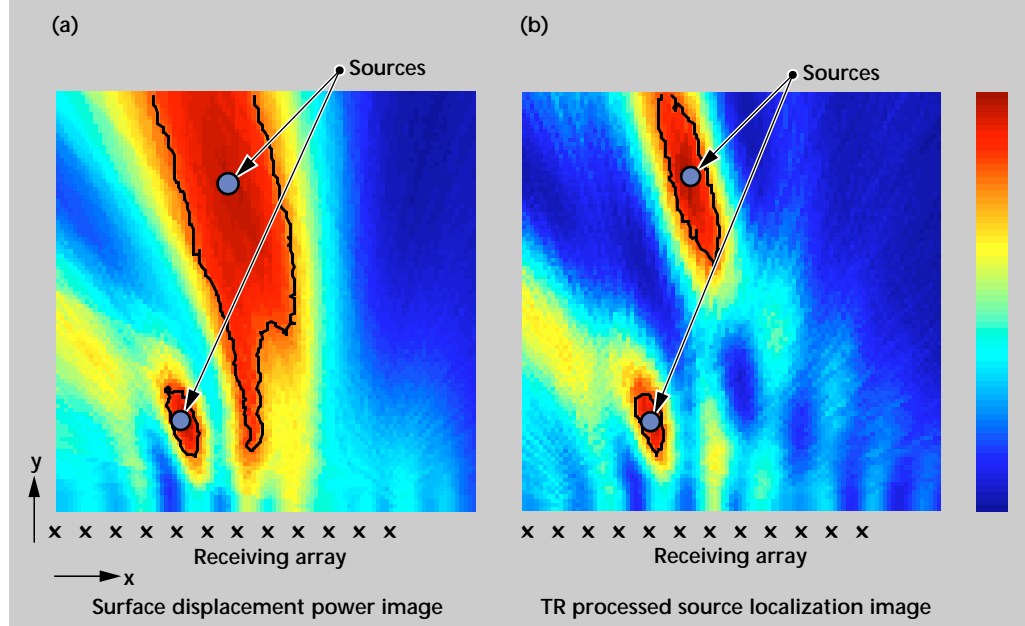
To develop our hardware capabilities further, we plan to procure a Fabry-Perot interferometer. Fabry-Perot interferometers capture diffuse reflections, yielding better sensitivity in the detection of ultrasonic waves on rough surfaces than the Michelson interferometer.

Software enhancements will expand our processing and imaging capabilities. Addition of a defect scatterer to the WAVER thermoacoustic wave propagation model will allow us to improve the model-based signal processor developed last year.

We will continue work on beamforming algorithms for enhanced imaging. Finally, we will use our expanded thermoacoustic model to determine the optimum collection configuration for sequential fiber-based pulsing.

Next year we will apply laser ultrasonic techniques to determine properties of aging nuclear materials. We will also develop rapid scanning capabilities to monitor tool wear and guide cutting tools in machining parts.

Figure 6. Ultrasonic signals generated by the femtosecond laser cutting a stainless steel plate, detected by the Michelson interferometer on (a) the opposite side of a 8-mm-thick plate; (b) the opposite side of a 4-mm-thick plate; and (c) at a surface distance of 5 mm from the cutting laser on a 4-mm-thick plate.

References

1. Candy, J. V., G. H. Thomas, D. J. Chinn, and J. B. Spicer (1996), "Laser Ultrasonic Signal Processing: A Model-Reference Approach," *J. Acoustic. Society Am.,* **100**, (1).

2. Huber, R. D., D. J. Chinn, G. H. Thomas, J. V. Candy, and J. Spicer (1996), "Model-Based Signal Processing for Laser Ultrasonic Signal Enhancement," Review of Progress in Quantitative Non-Destructive Evaluation, Plenum Press, N.Y.

3. Spicer, J. (1991), *Laser Ultrasonics in Finite Structures: Comprehensive Modeling with Supporting Experiment,* Ph.D. thesis, John Hopkins University, Baltimore, Md.

4. Johnson, D. and D. Dudgeon (1993), *Array Signal Processing: Concepts and Techniques,* Prentice-Hall, Princeton, New Jersey.

5. Candy, J. V. (1996), *Time-Reversal Processing: An Approach to the Scatterer Estimation Problem,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-JC-124942).
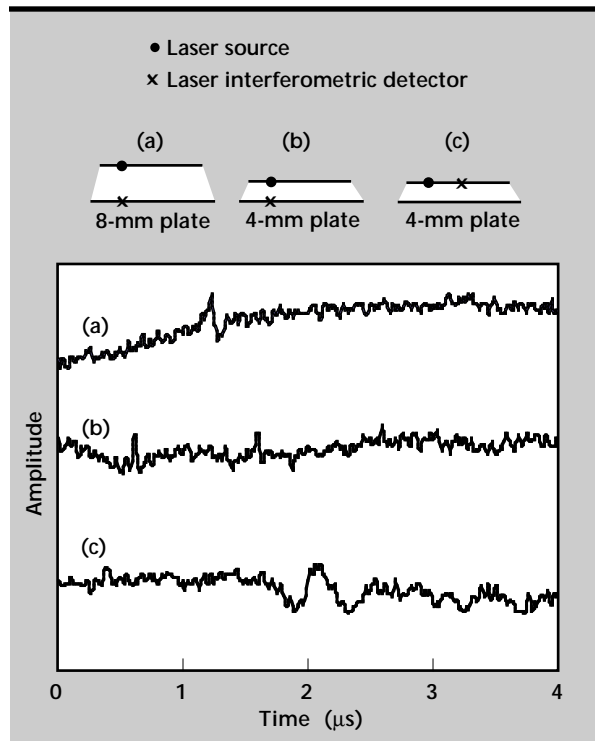
nfrared Computed Tomography for Thermal NDE of Materials, Structures, Sources, and Processes

Nancy K. Del Grande, Philip F. Durbin, and Dwight E. Perkins
Manufacturing and Materials Engineering Division
Mechanical Engineering

Paul C. Schaich
Defense Sciences Engineering Division
Electronics Engineering

Dennis M. Goodman and David H. Chambers
Lasers Engineering Division
Electronics Engineering

Thomas Milner
Beckman Laser Institute
University of California
Irvine, California

We conducted feasibility demonstrations that established infrared computed tomography (IRCT) as a unique tool for thermal nondestructive evaluation (NDE) of materials, structures, sources, and processes. The IRCT system produced precise, high-definition, time-resolved temperature and thermal inertia maps. These maps depicted thermal-absorption sites in KDP crystals; 3-D images of corrosion damage under aircraft wing fasteners; Zeolite heavy-ion source temperature gradients; U-6Nb temperatures during the forming process; and internal cracks within partially-formed Ti-6Al-4V parts. As a result of leverage from the Lawrence Livermore National Laboratory (LLNL)/Bales small-business CRADA, the IRCT system now provides a high-definition, clutter-free, evaluated, corrosion-defect map in minutes, compared to hours for previous methods. IRCT technology has supported several programs at LLNL and the Idaho National Engineering Laboratory (INEL).

Introduction

Limitations Of Thermal NDE

Interpretation of thermal NDE methods is often complicated by clutter, such as unwanted detail. For flash-heated aircraft skins, clutter (for example, from hot spots, sealants, and ripples) is difficult to distinguish from corrosion-thinned aircraft skins. Using a single-band infrared (SBIR) system, we could not distinguish structural defects from environmental clutter such as reflected IR backgrounds from light fixtures, windows, or sky radiation. Using a dual-band infrared (DBIR) system, we removed clutter associated with reflected IR backgrounds to clarify interpretation of structural defects.

Thermal NDE currently specifies the location of structural defects in two spatial dimensions. Three-dimensional spatial characterizations of defect types, volumes and depths require a unique approach to the detection, analysis and interpretation of thermal NDE. We call this approach IRCT.

IRCT Concept And Physical Rationale

IRCT maps precise surface temperatures and bulk-material thermal inertias.[1-4] Surface-temperature maps specify defect size and location. Thermal inertia maps specify defect thickness and depth.

The solution to the heat transfer equation for a thick target with an instantaneous surface heat flux is:[5]

$$T(x,t) = \frac{q}{\sqrt{4\pi k\rho ct}} \exp\left(-\frac{x^2}{4\alpha t}\right)$$

(1)

where T is temperature, x is depth beneath the surface, k is thermal conductivity, ρ is density, c is heat capacity, α is thermal diffusivity, t is time and q is surface heat flux. For a semi-infinite solid approximation, the surface temperature is proportional to the inverse square root of time and inversely proportional to the thermal inertia, P, which varies as $(k\rho c)^{1/2}$.

The IRCT System

The IRCT system uses natural (passively-heated) or pulsed (actively-heated) sources to heat-stimulate structural defects; a dual-band infrared camera to detect them; and rules-based defect-recognition algorithms to evaluate them. Our patented, dual-band IRCT temperature-mapping method,[6] has ten times better signal-to-noise than conventional thermography, which is a single-band IR temperature-mapping method. We use two IR bands to enhance surface-temperature contrast, and remove spatially-varying, surface-emissivity noise.

The IRCT system uses VIEW macros on SGI workstations to create masks for automatic defect recognition and clutter rejection. These masks exploit defect depth and volume information specified by time-sequences of thermal inertia maps. Shallow buried objects, defects, and clutter are seen on early-time thermal inertia maps. Deep buried objects, defects and clutter are seen on late-time thermal inertia maps.

Technology Transfer
for Aircraft Inspections

LLNL entered into a Cooperative Research and Development Agreement (CRADA) with Bales Scientific Incorporated (BSI) to commercialize DBIR corrosion-inspection technology.[1] As part of the CRADA, the LLNL/BSI scientific team produced a single, high-definition, corrosion-defect map for aircraft inspections. The defect map was free of clutter from excess sealants, ripples, and surface features.

The corrosion-defect mapping technique was incorporated into the BSI commercial DBIR scanner, making it easier to access and quantify corrosion damage. Accomplishing the objectives of the CRADA, we upgraded the prototype IRCT system for aircraft and other inspections.

Progress

Hardware Improvements
for Image Quality

Using a uniform flashlamp source and spectral hood, we improved the IRCT image quality for many of the FY-96 feasibility demonstrations. By removing hot spots equivalent to 1 to 3 °C temperature differences, we quantified 5% metal-thickness reductions from corrosion. Because the interior of the spectral hood is highly reflective, more heat reaches the sample from the flashlamps, which improves the signal-to-noise ratio. The hood acts as a blinder for the DBIR camera, by isolating the sample and IR detectors from the environment which surrounded them. This decreases reflected DBIR backgrounds, which reduces the clutter, and clarifies interpretation of the images.

The additional purchase of a 7X close-up lens, with a focal length of 5 cm, will enable us to resolve 25- to 50-μm (1- to 2-mil) defects. This will provide the IRCT system with the potential for high-definition defect maps, derived from thermal and thermal inertia images, which are unsurpassed by any other commercial DBIR system. This has significant implications for characterizing smaller defects in laser-optics components, weapons parts, and canned secondary assemblies. It enhances our capability to support LLNL Programs.

Enhanced Software
for Evaluated Defect Maps

We created an evaluated corrosion-defect map and incorporated the procedure for producing this map into the BSI commercial DBIR system which is used for the IRCT feasibility demonstrations. Also, we developed a unique, rules-based procedure to operate on DBIR temperature, emissivity, and thermal inertia maps. The procedure enables automatic defect recognition and clutter rejection. It tags corrosion defects which affect the structural integrity of the aircraft fuselage. It rejects excess sealants and ripples which have thermal signatures similar to those of corrosion defects.

INVERSE with IR Image Simulator

We studied a technique to localize absorbers in a medium, using images from an IR camera. This technique[7,8] has been used successfully to localize capillaries in the skin responsible for port-wine stains.[9] From a time sequence of IR images, it inverts the heat equation to obtain an estimate of

the temperature distribution in the material at the time of the first image. This ill-posed problem is solved using a regularization method.[10] The goals of this study were to adapt the INVERSE code for KDP, develop an IR image simulator, and make a preliminary evaluation of its effectiveness using simulated IR images. All of these goals were accomplished.

The IR image simulator was constructed in MATLAB, a matrix algebra and signal processing environment, which calculated the IR irradiance image as a function of time. It was built on the exact expressions developed by Milner, *et al.*[9] Simulations of images of point absorbers at various depths and distributions were performed. Typical simulated 3-µm IR images, using two point absorbers, separated by 6 mm, at a depth of 3 mm,

are shown in **Fig. 1**. Statistically white noise is added to the images at an rms level of 0.001 of the peak irradiance.

Note the blending of the separate patterns as time progresses. Applying the INVERSE code to the images gives a good estimate of the temperature distribution in the material at the time of the first image (0.2 s), for depths of 0 mm, 1.5 mm, and 3 mm (**Fig. 1b**).

Absorption Sites in KDP Crystals

In **Fig. 1**, we compare pulsed-laser stimulated (in (**c**)) and flashlamp-stimulated (in (**d**)), thermal images of absorption sites in KDP crystals. The laser-stimulated absorption sites were detected in
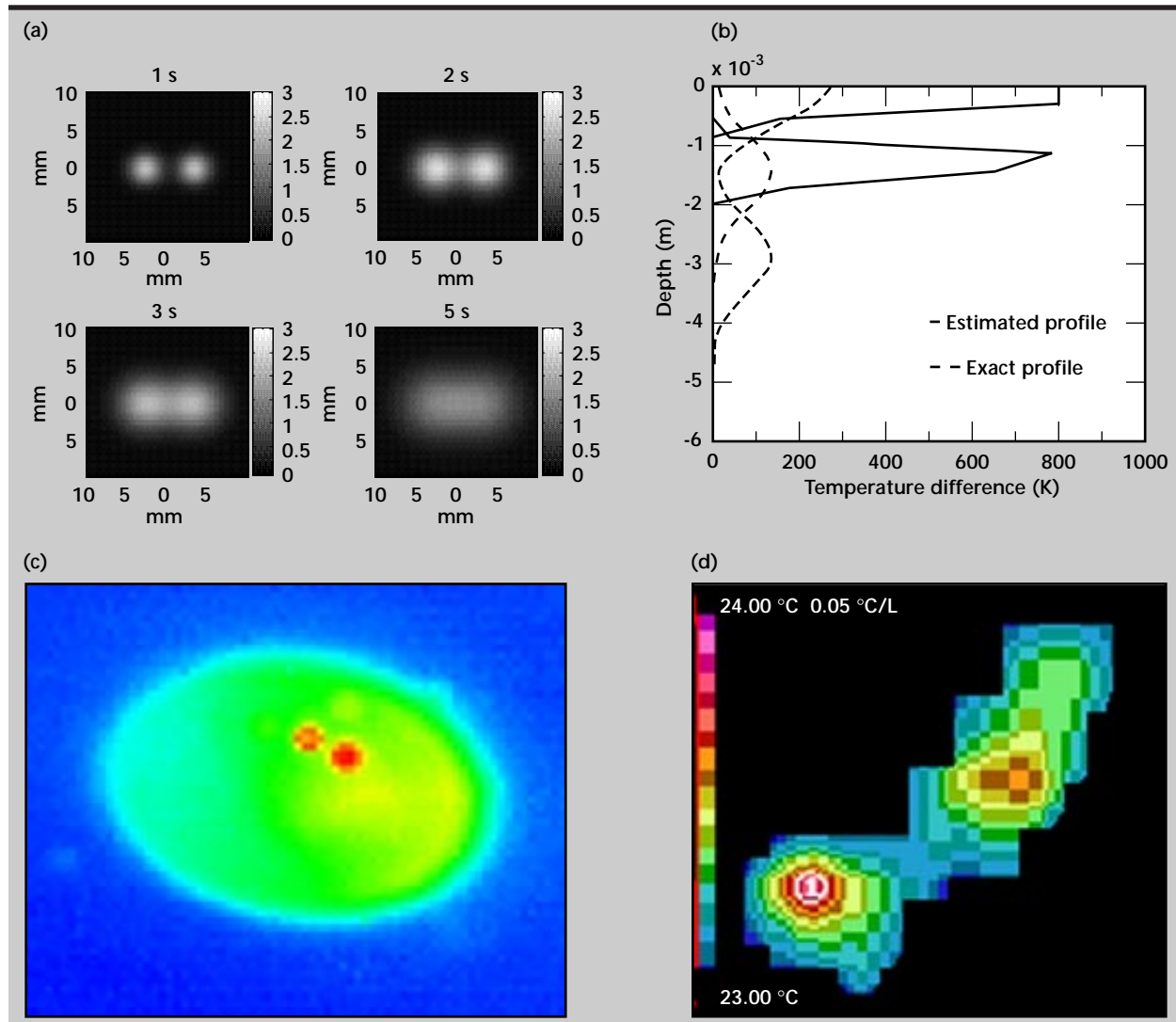

Figure 1. *(a) Simulated 3-mm deep point absorbers, separated by 6 mm, after 1, 2, 3, and 5 s; (b) estimated (solid) and exact (dashed) vertical profiles of the temperature distribution at t = 0.2 s, for initial absorber depths of 0 mm, 1.5 mm, and 3 mm, using the INVERSE code; (c) measured thermal images of laser-heated KDP absorption-sites, separated by 0.5 mm, at .002 s after the 150-µs heat flash; (d) flash-heated KDP absorption sites, separated by 4 mm, at .040 s after the 4-ms heat flash.*

collaboration with T. Milner at the Beckman Laser Institute, University of California at Irvine. Dr. Milner used an InSb 286 Hz focal-plane array (FPA) to detect 3- to 5-µm IR absorption-site signals. The thermal image was detected at 2 ms after the 150-µs, 1.5-mm diameter laser pulse heated the

KDP absorption site differently than the host KDP material. We measured two absorption sites, separated by 0.5 mm, as shown in **Fig. 1c**.

The IRCT system used a Hg-Cd-Te detector-array to detect the 3- to 5-µm and 8- to 12-µm DBIR absorption site signals at 40 ms after the 4-ms,

Figure 2.
(a) Photograph of spin-forming operating system. During the forming operation, U-6Nb part temperature gradients were measured in real time to benchmark code simulations. (b) After the partial-forming operation for Ti-6Al-4V, one flash-heated part had incipient internal cracks, depicted at spots 1, 3, and 4, left-to-right, within circled areas of the temperature map (c) and relative thermal inertia map (d).
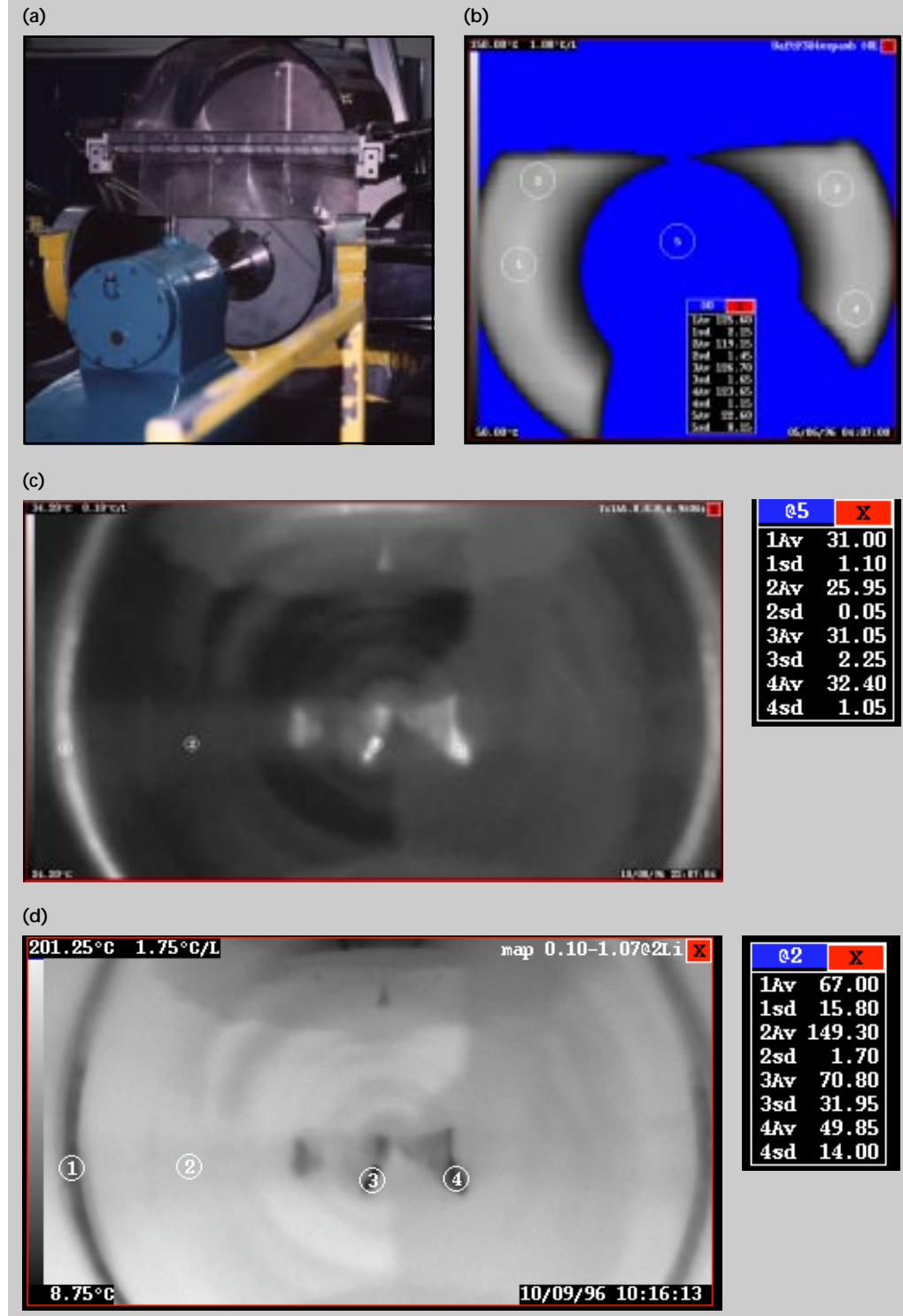
0.4-m wide flashlamp pulse which heated the KDP crystal absorption site differently than the host KDP material. The DBIR thermal image resolved two absorption sites, separated by 4.0 mm, as shown in **Fig. 1d**.

Resolution of 0.04-mm absorption sites should be possible using the new 7X close-up lens attachment for the IRCT camera system. We expect to explain vertical depth profiles for absorption-site images in KDP by adapting the INVERSE code for interpreting the experimental data.

Uranium Alloy-Forming Operations Benchmarked

We successfully monitored a U-6Nb uranium alloy plate while it was being shaped into a device part. **Figure 2a** shows the spin-forming apparatus. **Figure 2b** shows the real-time temperature data which helps scientists better understand changes which occur in the uranium alloy as it undergoes the forming operation. The precise IRCT temperature measurements are used to benchmark numerical simulations.

Internal Cracks Detected in Partially-Formed Titanium Alloy

We flash-heated a partially-formed Ti-6Al-4V part which had a smooth surface with no visible cracks. The thermal map in **Fig. 2c**, and thermal inertia map in **Fig. 2d**, show incipient internal cracks at the 14-in. diameter location where large visible cracks occurred in a similar, Ti-6Al-4V alloy sample. The thermal map depicted anomalous cooling rates at sites where the thermal inertia map had very little resistance to temperature changes. Incipient internal cracks may have developed during the forming operation that was conducted with the equipment shown in **Fig. 2a**. We also detected unexplained cooling-rate anomalies near the center of the uncracked sample.

Corrosion Damage Found under Aircraft Wing Fasteners

We measured corrosion damage under fasteners in a KC-135 aircraft wing panel from Tinker Air Force Base. Early-time (0.5 ± 0.4 s) minus late-time (1.7 ± 0.6 s) temperature-difference maps quantify the relative volume of corrosion damage under two aircraft wing fasteners, shown in **Fig. 3a**, and 13 aircraft wing fasteners, shown in **Fig. 3b**. Sites which have the most white areas have the highest peaks and the largest volumes of metal-thickness reduction from the corrosion damage.

Zeolite Source Temperature Gradients Measured

We measured the temperature of the recirculator heavy ion source at 500 to 1200 K (230 to 930 °C). The Zeolite target is the source of potassium ions produced in the small LLNL recirculator. The recirculator is a prototype recirculating induction accelerator under development by LLNL's Heavy Ion Fusion (HIF) Group. By imaging the source operating temperatures, and temperature gradients, information may be gleaned about the quality of the heavy ion beam. Our temperature measurements provided assurance that the recirculator heavy ion source was performing as expected.

We used a Ag mirror, and a ZnSe (vacuum barrier) window, which had a combined reflectance and transmittance of 75% of the signals in the shortwave (SW, 5 ± 0.5 μm), and longwave (LW, 10 ± 1.5 μm) IR bands. The emitting surface behaved like a selective radiator. It had emissivity differences in the LW and SW bands which may be related to silicate build-up from oxidation. Silicates have a broad molecular resonance at
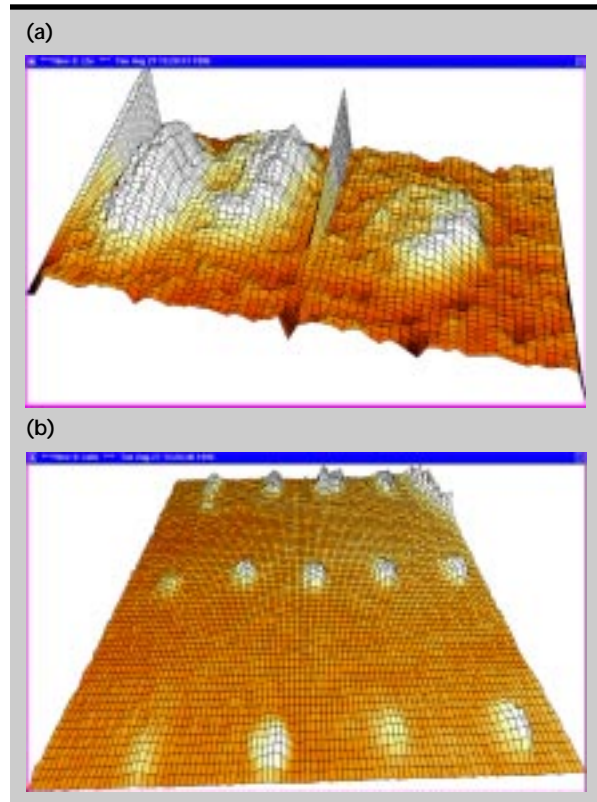

Figure 3. Three-dimensional displays of early-time minus late-time temperature-difference maps, which quantify the metal-loss volume from intergranular corrosion, where the damage is (a) 4× larger under the left, compared to the right aircraft wing fastener, and (b) varies under thirteen wing panel fasteners.

10 μm, which does not affect the SWIR temperature measurement. The LWIR temperature non-uniformities for the Zeolite emitting source are shown in **Fig. 4**.

Summary

IRCT applications conducted during FY-96 generated significant interest in the potential of IRCT to address key issues for major LLNL Programs.

The National Ignition Facility (NIF) Program is interested in the potential of high-resolution IRCT measurements, in conjunction with INVERSE simulations, to correlate KDP thermal absorption-site locations, sizes, and depths of origin, with known KDP defects, using the NIF laser-diagnostics table and damage testing system.

The Advanced Development and Production Technologies (ADaPT) Program is interested in IRCT technology to benchmark the forming operations for weapons parts. The initial tests succeeded in monitoring the spin-forming process temperatures in real time for U-6Nb parts.

The Enhanced Surveillance Program (ESP) is interested in the capability of thermal and thermal inertia maps to determine the type, location, size, and depth of dimensional changes in plutonium weapons components, and in corrosion-related dislocations, or migrations, in canned secondary assemblies.

Figure 4. Thermal images of the 1-in. diameter Zeolite heavy ion source SWIR images at 4.5 to 5.5 μm (left column), and LWIR images at 8.5 to 11.5 μm (right column). The measured average temperatures assume an ideal black body source with emissivity $\varepsilon = 1.00$ in (a), (b), (c), and (d); a window transmission $\tau = 1.00$ in (a) and (b); $\tau = 0.75$ in (c), (d), (e), and (f); a selective-radiator source with a SWIR emissivity, $\varepsilon = 0.99$ in (e); and LWIR emissivity $\varepsilon = 0.89$ in (f). The measured average temperature values and standard deviations apply to source sites depicted by the line profiles.
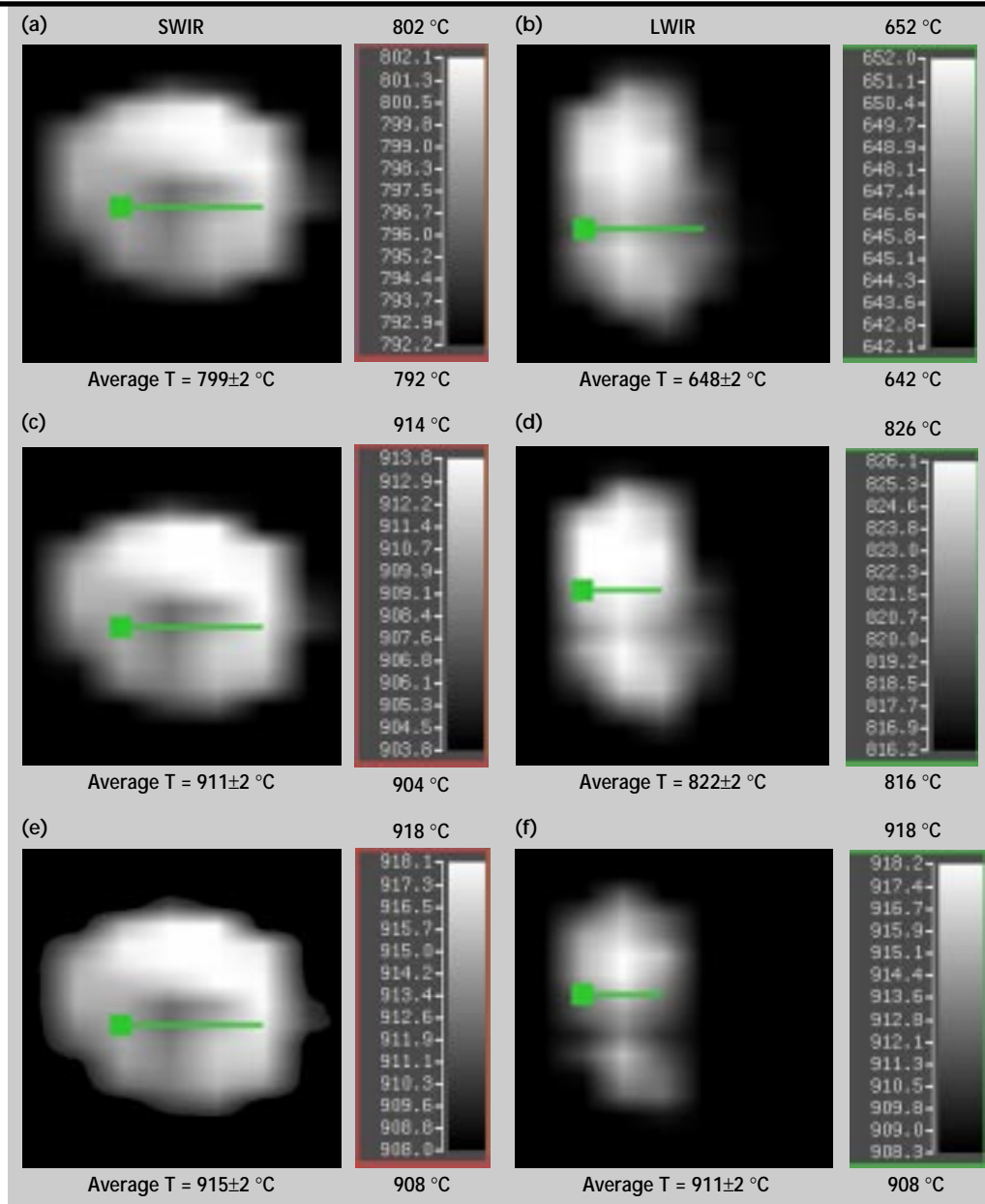
LLNL's and LBL's HIF teams need precise, high-temperature and temperature-gradient measurements to characterize the Zeolite heavy ion source. By monitoring aging and build-up of source oxides, the source can be replaced, as needed, to provide assurance of heavy-ion beam quality.

INEL is currently sponsoring a demonstration of the IRCT technology to measure wall-thickness reductions for corroded waste containers. Successful results would produce spin-off applications for a variety of environmental and hazardous-waste projects inside and outside INEL.

Future Work

We are planning additional proof-of-principle experiments to test the potential of new IRCT applications in support of LLNL programs. Our goals to address key program issues are discussed below.

For the NIF Program. We would test the capability of the high-resolution IRCT system to image flash-heated KDP absorption sites as small as 40 μm; adapt the IRCT system for imaging laser-heated KDP absorption sites in NIF environments; apply the INVERSE code to interpret their origin, develop a state-of-the art thermal diagnostic for the NIF damage-testing system.

For the ADaPT Program. We would test the predictability of design calculations for weapons-component forming operations; monitor forming operations in real time; develop image-processing procedures to improve the image display; and quantify the results.

For the ESP Program. We would make the existing IRCT system more rugged for operating in weapons system environments; test the capability of IRCT thermal and thermal inertia images to characterize internal gaps in canned secondary assemblies; and investigate dimensional changes in mock and actual plutonium weapons parts.

For the HIF Program. We would adapt the high-resolution IRCT system to measure precise heavy ion source temperatures and temperature gradients at 200 to 1000 °C; and provide co-registered, DBIR digital data for image analysis, display, and processing on suitable work stations.

Acknowledgments

References

1. Del Grande, N. (1996), "Dual-Band Infrared Computed Tomography: Searching for Hidden Defects," *Science and Technology Review,* University of California, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-52000-96-5), p. 23.

2. Del Grande, N. K., P. F. Durbin, and D. E. Perkins (1996), "Infrared Computed Tomography For Characterizing Structural Defects," *Nondestructive Evaluation Thrust Area Report, Engineering Research, Development and Technology FY95,* H. E. Martz, Ed., Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-95), February.

3. Del Grande, N. K., and P. F. Durbin (1995), "Mapping hidden aircraft defects with dual-band infrared computed tomography," Tobey Cordell, Ed., *Proceedings of SPIE Conference 2455: Nondestructive Evaluation of Aging Aircraft, Airports, Aerospace Hardware, and Materials,* Oakland, Calif., June, pp. 82–93.

4. Del Grande, N. K., and P. F. Durbin(1995), "Stimulated dual-band infrared computed tomography: a tool to inspect the aging infrastructure," B. F. Andreson and M. S. Scholl, Eds. *Proceedings of SPIE Conference 2552: Infrared Technology XXI, San Diego, Calif.,* July, pp. 292–301.

5. Carslaw, H. S., and J. C. Jaeger (1980), *Conduction Of Heat In Solids, Second Edition,* Oxford Univ. Pr., London, England, pp. 101, 112, and 259.

6. Del Grande, N. K., P. F. Durbin, K. W. Dolan, and D. E. Perkins (1995), *Emissivity Corrected Infrared Method For Imaging Anomalous Structural Heat Flows,* (Patent No. 5,444,241).

7. Milner, T. E., D. M. Goodman, B. S. Tannenbaum, B. Anvari, L. O. Svaasand, and J. S. Nelson (1996), "Imaging Laser-Heated Subsurface Chromophores in Biological Material: Determination of Lateral Physical Dimensions," *Phys. Med. Biol.*, **41**, pp. 31–44.

8. Milner, T. E., D. M. Goodman, B. S. Tannenbaum, and J. S. Nelson (1995), "Depth Profiling of Laser-Heated Chromophores in Biological Tissues by Pulsed Photothermal Radiometry," *J. Opt. Soc. Am. A*, **12** (7), pp. 1479–1488.

9. Milner, T. E., D. J. Smithies, D. M. Goodman, A. Lau, and J. S. Nelson (1996), "Depth Determination of Chromophores in Human Skin by Pulsed Photothermal Radiometry," *Appl. Opt.*, **35**, (19), pp. 3379–3385.

10. Goodman, D. M., E. M. Johansson, and T. W. Lawrence (1993), "On Applying the Conjugate-Gradient Algorithm to Image Processing Problems," *Multivariate Analysis: Future Directions*, C. R. Rao, Ed., pp. 209–232.

Automated Defect Detection for Large Laser Optics

Laura N. Mascio
Engineering Research Division
Electronics Engineering

Clint M. Logan
Defense Technologies Engineering Division
Mechanical Engineering

Harry E. Martz
Manufacturing and Materials Engineering Division
Mechanical Engineering

We have developed and implemented automated defect detection software for round large laser optics imaged by the Beamlet Lens Inspection System. In pursuit of this goal, we encountered several extensions to the original problem. As a result, progress has been made towards an automated inspection of more complicated square lens images on Beamlet, and towards estimating the image processing needs for the Lasers Program and the National Ignition Facility project at Lawrence Livermore National Laboratory.

Introduction

The National Ignition Facility (NIF) at Lawrence Livermore National Laboratory (LLNL) faces enormous challenges in the inspection and maintenance of thousands of large laser optics. Optics inspection must occur at several stages: 1) off-line contamination detection to meet cleanliness standards before assembly; 2) on-line (*in situ*) damage detection, especially at vacuum barriers, to ensure safety during operation; and 3) surface quality inspection and defect characterization for quality control.

This FY-96 project was originally proposed to address a more modest and specific problem of automating the tedious manual inspection process used to document the laser damage and other defects on a round Beamlet lens between operations (off-line). That goal was quickly addressed by developing and implementing automated inspection software to analyze video camera images of the lens. Because a digital acquisition system had been installed on-line, these images could be acquired and analyzed without removing the lens from the beamline. The defect detection software is currently operational and can be executed from the Beamlet control room.

During development of that software, many extensions of the original problem became apparent, and this project began to address them. The first extension involved analyzing newly available on-line image data acquired from a square lens on Beamlet. These images are more complicated in that reflections and peripheral hardware are visible within the area of the lens. Algorithms for background suppression and defect enhancement have been developed to assist in analyzing these images.

Further extensions of the original goal included investigating various future requirements of LLNL's Lasers program and the NIF project in the area of optics inspection. This investigation was done in conjunction with personnel who are designing the contamination control, optics damage inspection and quality control in support of NIF.

Progress

Automating the Manual Inspection Process

The original goal of the project was to automate the tedious manual inspection of large laser optics,

which was done by a technician in a dark room using a flashlight and clipboard. Defects were laboriously measured and recorded as shown in **Fig. 1**.

When this project got under way, an on-line image acquisition system (called the "Beamlet Lens Inspection System") had already been installed such that a round (58 cm diameter) sol-gel AR coated fused silica lens in the L3 position on Beamlet could be imaged in place between laser shots. A video CCD camera (640×480 pixels with 8 bits of grayscale per pixel) was fixed in place below the optic and out of the beam path (**Fig. 2**). Between 3 and 8 quarter-inch fiber optic bundles were used to inject light into the edge of the optic. Images of this optic benefit from the dedicated acquisition conditions and from minimal clutter in the field of view.

Algorithm for Detecting the Optic

Any defects in the optic manifest as irregularities in the acquired image, due to the light scattered at the defect site. **Figure 3** shows defects that have been computationally detected by the method to be described. The first step towards defect detection was to write algorithms to isolate the optic from the rest of the image. In the case of the round lens, the image shows an elongated lens because the camera position is below the optic and the line of view is not at normal incidence to the optic. The Circle Hough Transform (CHT)[1] was applied because it is robust in the face of noisy, partially obscured or partially missing data, and is significantly faster than the equivalent transform for ellipses.
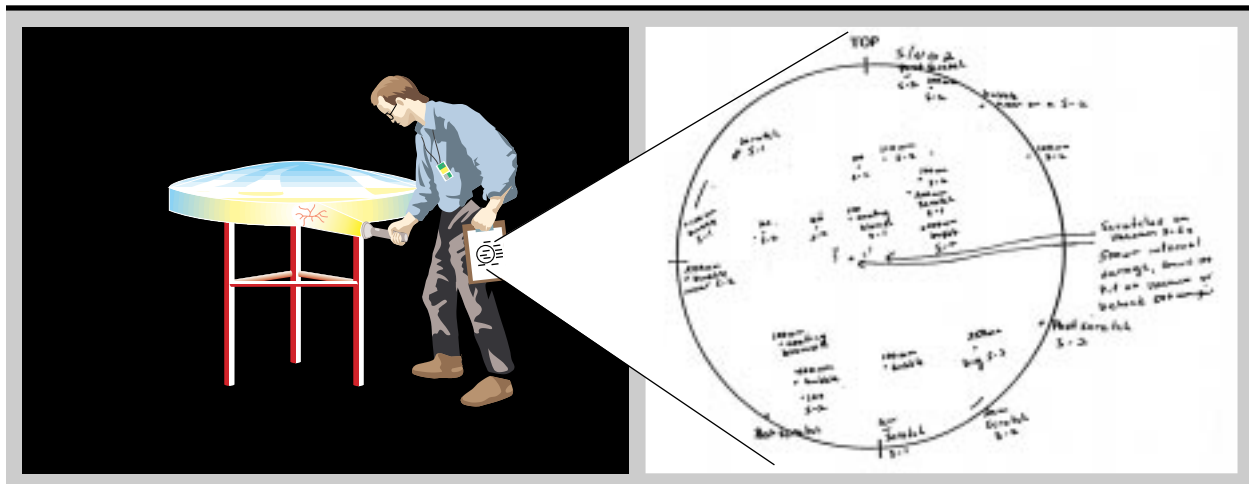

Figure 1. *Illustration of the tedious manual inspection process, where defects were laboriously measured and recorded over time.*
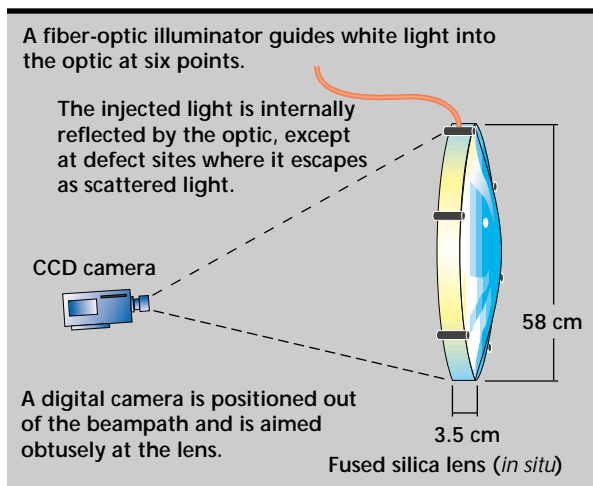

Figure 2. *Schematic of CCD video camera providing dedicated acquisition of the entire L3 round lens on Beamlet. The camera is positioned below the lens and at an angle. Fiber optic guides feed white light into the optic for illumination.*
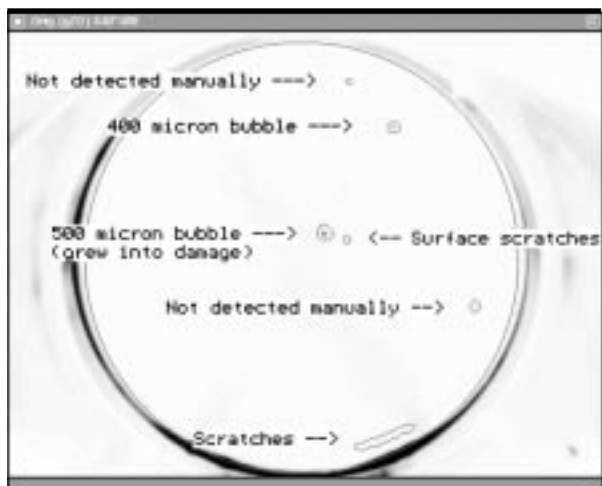

Figure 3. *Automatic detection of the boundary of the optic (large circle) and defects within the optic (small, irregular contours). Manual inspection indicated the size and type of each detected defect as labelled.*

The basic premise of the CHT is that each point (from an edge magnitude image) along a circle can be transformed into a line in Hough parameter space (3-D: x coordinate, y coordinate and radius) such that each point of the line "votes" for the location (center x and y coordinates and the radius) of the circle. When the transformation is complete, the maximum number of votes in Hough space reveals the center and radius of the circle. The CHT is summarized in **Fig. 4**.

The CHT consistently and reliably found the best circle that fit the elongated lens data, and effectively isolated the area of the optic from the surrounding image data for all 83 available test images, from December, 1995 to September, 1996. An example of the CHT result is illustrated as the large complete circle in **Fig. 3**.

Algorithms for Detecting Defects

Once the region of the lens was isolated, algorithms were applied within that region for defect detection. The developed method is based on edge magnitude, since gradients are detectable independent of the choice of illumination (dark field or bright field). First, a Sobel differential filter is applied to the image by computing the square root of the sum of the squares of the gradient in the horizontal and vertical directions. Then, a histogram of the Sobel result is computed, which charts the function of intensities (edge magnitude) versus the rate

of occurrence for each intensity in the image. This histogram is computationally analyzed using a triangular construction method to determine the threshold level, a single intensity value below which there is no useful defect information. The triangular construction algorithm computes the clipping level to be a constant offset below the knee of the "triangular" peak formed by the background at the low intensity end of the histogram.

Applying the computed threshold level to the image yielded a binary mask that has "on" pixels where the defects lie and "off" pixels elsewhere. This binary image is then manipulated using binary morphology functions to fill in any holes within the area of defects. These areas can be measured for various features, which will be addressed later. Finally, the defect areas are dilated and contoured so the contour can be placed over the original image for displaying the results that were shown in **Fig. 3**.

Algorithms for Visual Enhancement of the Area

Figure 5 shows the same optic using a severe local contrast enhancement. This is useful in that it is otherwise extremely difficult to discern defects visually in the original image. It provides a visual aid for the operator, and acts as a quality check when the optic cannot be inspected manually. The defects that were computationally detected and marked in **Fig. 3** can be visualized as abnormalities
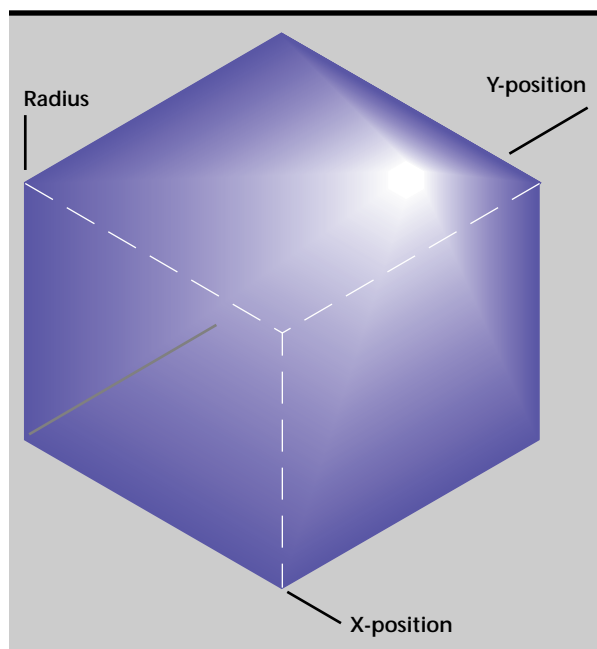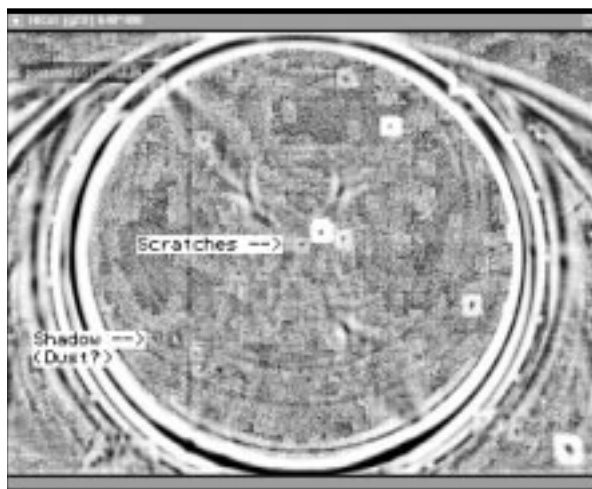
Figure 4. Summary of the Circle Hough Transform.

Figure 5. Visual quality check for fine detail that is not visible in the original image. The two subtle abnormalities indicated are visible here but were not detected by the automated detection scheme demonstrated in Fig. 3.

here, where they were not necessarily visible in **Fig. 3**. The additional defects visible in **Fig. 5** (a shadow and some scratches) are small subtleties which had been detected by manual inspection but were not considered alarming. Other smaller blemishes that had been detected by manual inspection were not resolvable with any of the developed enhancement of detection schemes.

The method developed to create this high contrast image consists of a 13×13-pixel moving kernel that travels across the image in a raster fashion. At each placement of the window, the 169 corresponding intensities in the image are ranked in order from darkest to brightest. These are then mapped so that the darkest corresponds to 0, the brightest corresponds to 255 (to cover the 8 bit dynamic range of the display) and the rest are linearly mapped in between. The result from each position is that the intensity of (only) the center pixel in the result image is reassigned to its corresponding "stretched" value from the linearly mapped values. The other pixel values are unaltered until they are at the center of the moving kernel.

Defect Detection Results

For these images, manual inspection results were available to compare with the computer detection results. The manual inspection results recorded all minute blemishes on the lens and revealed that the defects on this lens ranged in size up to 500 μm.
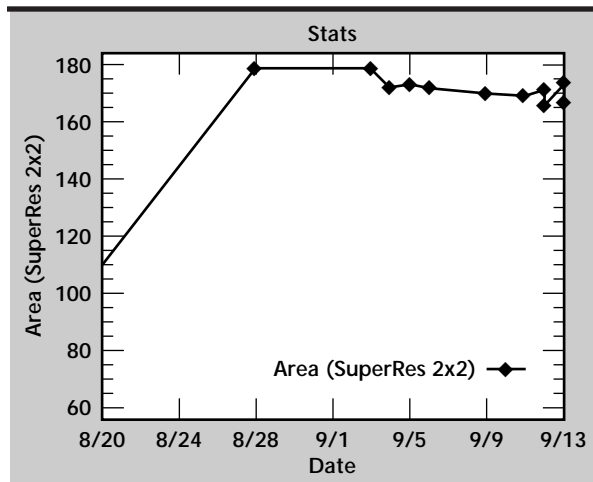

Figure 6. *Plot of the area of the detected signal for the largest defect on the round lens vs time. The initial jump in the signal area was due to a scheduled camera focus adjustment. The relatively stable area measurement on subsequent days indicates little change in area of the physical defect and concurs with control room records that acquisition conditions were kept stable.*

The required critical detection size for Beamlet is on the order of 1 to 2 mm. The critical flaw size for NIF will likely be even larger.

The computer detection scheme consistently detected the larger of these objects, but inconsistently detected the smaller of them. Some of the smaller defects would be detected, then "disappear" in a subsequent image and then later "reappear". The reason for this is that the quality of the images varied due to different acquisition parameters, especially camera focus, and so their amenability to analysis varied as well. The fine detail information content among images was inconsistent. Since this project only worked retrospectively on archived images, it was not possible to control the image acquisition parameters. Nonetheless, the developed algorithms were able to consistently detect the 400 to 500 μm bubbles and some scratches in all available test images.

It should be noted that computer detection of defects is dedicated to locating defects. It does not necessarily preserve their shape or size. To measure the defects, they must be analyzed quantitatively after detection, as described below.

Defect Measurements

The camera resolution for these images is about 1.3 mm per pixel, varying across the image due to various optical distortions. But due to defocus and saturation during acquisition, and due to diffraction at the defect site, the visible signal (spot) on the image can be much larger than the physical defect, which itself may not be resolvable. For example, a defect on the order of 500 μm in diameter scattered light to form a spot that was 7 pixels (approximately 9 mm) in diameter in the image. Because of this effect, measurements made on the resulting spot in the image only serve to characterize the distorted/diffracted light signal and not necessarily the absolute physical properties of the defect. The relationship between the scattered light signal and the absolute physical properties of the defect were not studied here. Relative measurements, however, on the defect signal over time reflected relative physical changes in the environment.

Figure 6 shows how a change in camera focus drastically altered the area of the defect signal for the 500 μm bubble near the center of the round lens around August 23. It also shows how the visible signal did not change size over the next subsequent shots, assuming the acquisition conditions stayed constant over that time, which is in concurrence with acquisition records from the Beamlet control room.

The plotted measurement values were attained by first "reconstructing" the original analog signal by multiplying each pixel of the sampled data by a sinc function.[2] This increases the resolution of the data so that each pixel becomes 2×2 pixels and the number of pixels across the diameter of the object is approximately doubled. For small defects, this has the effect of decreasing the error of the area measurement from 10% to 1%.[2] The next step is to use grayscale morphology to subtract the background from the image so variations in illumination across the image will be leveled. Then, a triangular threshold algorithm, described above in the section on defect detection, was applied to determined the cutoff intensity that separates the defect from its surroundings. The area is then determined by counting the number of pixels resulting from the threshold. This area can be converted to physical units if the resolution at the site of the defect is known.

Extension to a Final Focus Square Lens

Later in the year, images from a square (40 cm $\times$ 40 cm) sol-gel AR coated wedged fused silica lens became available from the final focus lens of the Beamlet Lens Inspection System. These images also benefited from having dedicated illumination and camera, but the image is acquired through a mirror which is moved in and out of position between shots, which may cause sub-pixel registration offsets between images. More light is pumped into this square lens, and thus more light spills out of it, illuminating the surroundings. Bevels, mirrors, debris shields, hardware, and the beamline are all visible in these images and defect detection becomes more complicated. In addition, the dynamic range of these images is particularly limited. Much of the data is either near saturation, or concentrated near zero intensity, and there seems to be room for improvement in the current image acquisition techniques.

Algorithm for Detecting the Optic

Again, as with the round lens, the first step is to develop an algorithm that will isolate the optic from the rest of the image, so the defect detection can be concentrated in the area of the optic. The General Hough Transform (GHT)[3] was applied for this task, since the square optic appears as a trapezoid and cannot be represented parametrically. The GHT permits the detection of any previously specified arbitrary shape, so that it can be detected in images where parts of the object are obscured or missing. Again, the GHT was implemented for the skewed square lens shape and it consistently detected the optic.

Algorithm for Detecting Defects

The more complicated background in the square lens images required a background suppression step before attempting to detect the defects. This was achieved by applying an "unsharp masking" operation. Effectively, the lower frequencies in the image are subtracted from the original image, leaving only the high frequencies. While this helps to suppress the slow-changing background, it amplifies many of the edges in the image, as well as the desired defects. Therefore, the next step is to apply algorithms that make use of morphology and texture to single out the defects. **Figure 7** shows the effect of
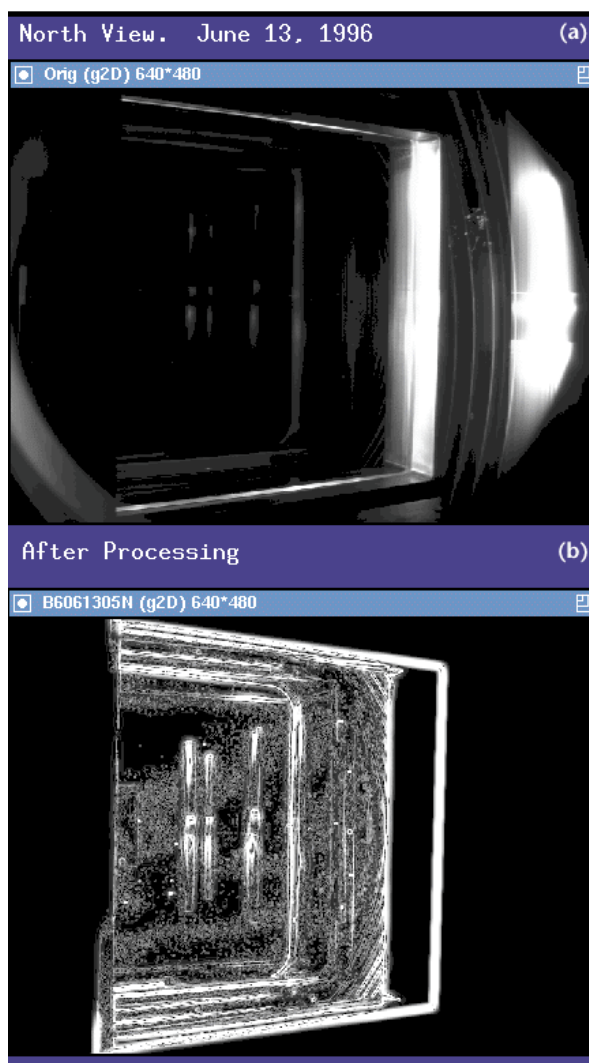

Figure 7. (a) Raw image as acquired by the Beamlet Lens Inspection System; (b) the same image data after processing. A General Hough Transform was applied to zero the surroundings of the lens. Morphological and other operators were then used to suppress the complicated background and enhance small defects and edges. The resulting enhanced image improves visualization for an observer.

isolating the optic after applying the General Hough Transform and applying the defect enhancement algorithm. Sobel operators, top hat filters and other convolution kernels have been applied to this end.[4,5,6] While each has its merits, none is capable of extracting just the defects. Therefore, the results of these operators should be input to a classification scheme whereby a binary decision tree could combine information from these operators to decide which pixels belong to defects and which do not. Operators that contribute information about the surroundings are just as valuable as those that contribute data about the defects, because rejecting an area is as important as classifying it as a defect. This classification step has not been implemented by this project.

As a result, automated defect detection has not been achieved for the more complicated square lens, but algorithms have been developed to isolate the optic and selectively enhance the defects so they are more readily visible to an observer as shown in **Fig. 7**. The defects on this lens range in size up to a maximum of 150 μm according to limited manual inspection.

Defect Tracking

Once a defect is detected in any kind of lens, the defect location, area and intensity can be computed. These values are to be recorded in a database, so that the detected defects can be studied in context with prior knowledge. The database should provide the platform for documenting, addressing, and tracking defects over time.

Future Work

The NIF Project will require automated inspection of thousands of large laser optics at various stages:

1) Currently, efforts are under way to devise appropriate cleaning methods for the optics to meet contamination requirements. Cleaning methods are being applied to small pieces of optics and examined microscopically. Image analysis algorithms are needed to detect and measure remaining contamination so as to objectively evaluate the success of the various cleaning methods.

2) During assembly, transport and installation, cleaned lenses must be inspected off-line for contamination, detecting particles as small as 5 μm. While the off-line conditions allow dedicated image acquisition, there is a need to minimize the time for acquisition. Efforts are under way to develop an acquisition system with sufficient resolution and speed to provide adequate data for image analysis.

3) Once on-line, optics will need to be monitored to detect defects approaching the critical flaw size which may be as large as 3 mm on NIF. But the large numbers of optics on NIF will preclude the use of dedicated internal illumination and camera. Dark field Schlieren imaging is proposed for acquiring data from multiple optics along the beam path with one camera. Hardware and software must be developed jointly to detect defects in specific lenses by manipulating the camera's plane of focus.

4) In the meantime, studies of laser damage physics on optical materials are being conducted on a micro (nm to 100 μm), macro (mm to cm) and full-sized (m-sized) scale. Visual inspection and evaluation of these data is limited. Consistent and repeatable image processing and analysis are needed to objectively evaluate defects at these scales with high sensitivity. The results of these studies will guide future research and development leading to the actual implementation of NIF.

5) Lastly, tracking defects will require sophisticated acquisition, calibration and image analysis methods which can align and reliably relocate objects in subsequent images. The location, area and intensity information resulting from image analysis should be recorded into a dedicated database which allows the documenting, addressing, and tracking of defects over time. This database should signal an alarm when defects approach critical flaw size.

References

1. Illingworth, J., and J. Kittler (1988), "A survey of the Hough transform," *Computer Vision, Graphics and Image Processing*, **44**, pp. 87–116.

2. Young, I. T. (1988), "Sampling Density and Quantitative Microscopy," *Analytical and Quantitative Cytology and Histology*, pp. 269–275.

3. Ballard, D. H. (1981), "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, **13**(2), pp. 111–122.

4. Laws, K. I. (1979), "Texture energy measures," *Proc. Image Understanding Workshop*, (Nov.), pp. 45–51.

5. Laws, K. I. (1980), "Textured image segmentation," *Image Processing Institute*, University of Southern California, Rept. **940**, (Jan.).

6. Pietikainen, M., A. Rosenfeld, and L. S. Davis (1982), "Texture classification using averages of local pattern matches," *IEEE Sixth International Conference on Pattern Recognition*, pp. 301–303.

Multistatic Micropower Impulse Radar Imaging for Nondestructive Evaluation

Jeffrey E. Mast
Defense Sciences Engineering Division
Electronics Engineering

Brooks Johnston and Stephen G. Azevedo
Laser Engineering Division
Electronics Engineering

We have developed multistatic microwave imaging tools that provide an integrated platform for performing nondestructive evaluation (NDE) in a wide variety of applications. Micropower Impulse Radar (MIR) technology provides a low-cost, flexible radar that is easily adapted to multistatic modes of operation. We have expanded our MIR techniques to include bistatic and multi-bistatic imaging in both planar and circular geometries. In this report we discuss the development of these new multistatic modalities, describe the corresponding imaging systems, and show results from experiments demonstrating their performance and potential usefulness.

Introduction

Micropower Impulse Radar (MIR) technology has demonstrated its usefulness as a powerful NDE tool that is low-cost and highly portable. Successful NDEs have been performed using MIR, including the imaging of subsurface objects and the interior of civil structures.[1–3] We have developed a Calibrated Autonomous Radar Testbed (CART) shown in **Fig. 1**, a mobile platform consisting of an MIR sensor, a linear translational stage, and a portable computer for control, acquisition and image processing. The translational stage provides antenna motion in one dimension while a computer-controlled motor drives the CART forward, providing a platform for accurate 2-D planar aperture synthesis with data suitable for 3-D imaging.

Limitations of our FY-95 MIR CART system prevented it from being used in many NDE scenarios, mainly due to geometry and apparatus constraints. The system could only handle planar geometries imaged in monostatic reflection mode, and could only use a single view of the object under inspection. Acquisition was limited to downward-looking (subsurface) acquisition, preventing other operating modes such as forward-looking or side-scanning techniques. These constraints made it very difficult to image the interior of objects or take advantage of situations in which access to more than one surface or side of objects is available. Additionally, a multistatic system capable of independent motion of the transmitter and receiver about the object under inspection provides better spatial diversity and improved visualization of interior features.

Multistatic systems consist of combinations of any of the following transmitter and receiver configurations in planar, circular, or arbitrary aperture geometries:

1) monostatic: transmitter and receiver collocated during formation of synthetic aperture;
2) bistatic: single transmitter separated from synthetically formed receiver aperture;

Figure 1. Calibrated Autonomous Radar Testbed (CART) for data acquisition using the MIR.

3) multi-bistatic: multiple separated transmitters and receivers;

4) reflection mode: transmitter and receiver on same side of object; and

5) transmission mode: transmitter and receiver on different sides of the object.

There are several trade-offs in system complexity and performance for each of these implementations. Multi-bistatic systems offer a potential for improved performance over strictly monostatic or bistatic systems, but typically require longer, more complex data acquisition and more computationally intensive imaging algorithms, especially for 3-D implementations. The data acquisition rate can be improved by using arrays of elements to reduce or remove the amount of mechanical scanning required. Expansion of the MIR CART imaging system to include the hardware and apparatus capabilities to do multistatic acquisition and reconstructions greatly increases the scope and type of objects that we can nondestructively evaluate. Furthermore, a system capable of multistatic configurations is important for optimizing geometry and antenna array configurations, and determining overall imaging system parameters and performance.

Addition of these multistatic imaging capabilities not only requires development of data acquisition apparatus, but each mode also requires modified imaging techniques to accommodate the new geometrical constraints. Therefore, development and implementation of signal and image processing algorithms are required to process the data from each expanded mode of operation. These algorithms are based upon generalized methods for holographic and diffraction tomographic imaging using wideband sources.[4–7] In this report we discuss our implementation of these multistatic imaging systems, as well as experiments demonstrating the performance and usefulness of these systems for microwave NDE.

Progress

In FY-96 we developed and integrated several new imaging modes described above, including multi-bistatic planar reflection mode and multi-bistatic circular transmission and reflection mode. This development required modifications to the FY-95 CART system to facilitate 2-D multi-bistatic linear aperture data acquisition. Other modes of operation were incorporated into the CART in the course of
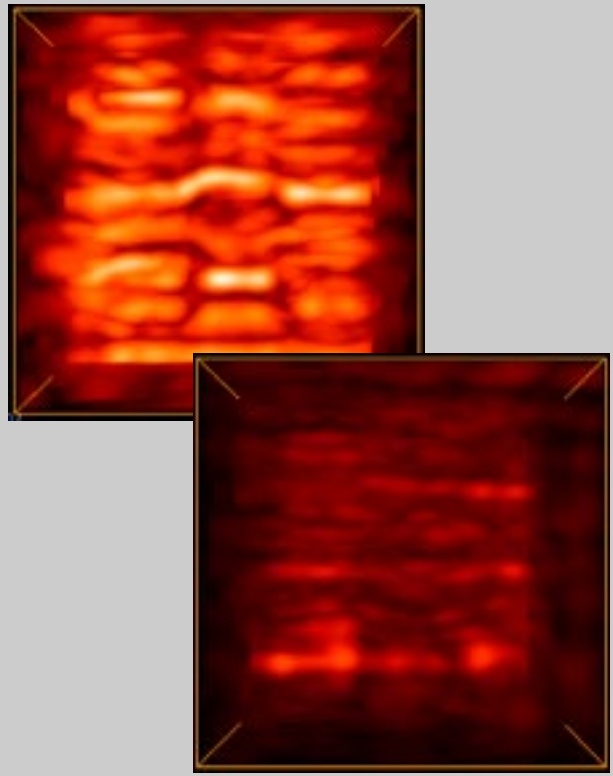

Figure 2. The CART in vertical mode, used for scanning large objects such as crates or containers, walls and other vertical objects. This mode also provides side-scanning SAR capability. The 3-D rendered images on the right are from cinder-block walls backfilled with pebbles (top) and reinforced concrete (bottom).

making these modifications. For circular mode imaging, development of a new testbed was required. A small rotational stage with appropriate antenna mountings was developed to acquire data from all angles around an object. For each of the new modes of acquiring data, the corresponding 2-D imaging algorithms were developed, integrated and tested using the new apparatus.

2-D Multi-bistatic Reflection Imaging

The antenna mounting on the CART was modified to a boom mount, extending it in front of the chassis, giving added flexibility and reducing unwanted reflections and interference from the frame. This extension also enabled forward-looking and limited side-scanning capabilities which are useful modes for subsurface evaluations. The translational stage was hinged, allowing mobile vertical 2-D aperture formation (**Fig. 2**) and "unlimited" side-scanning distances. The vertical mode is useful for scanning large objects such as crates, containers, or walls. For example, the 3-D rendered images in **Fig. 2** are from cinder-block walls backfilled with pebbles and reinforced concrete.

The stage was also upgraded for improved speed and positional accuracy. A second translational stage was added to facilitate independent linear motion of two antennas. Finally, the data acquisition and control software of the CART was upgraded to provide linear multi-bistatic aperture formation along the direction of the two stages.

Modular MIR[8,9] was integrated into the system and provided the flexibility required for multistatic implementations (**Fig. 3**). Separate transmitter and receiver modules provide bistatic capability, allowing the transmitter and receiver to be positioned independently. The operational bandwidth of this radar is approximately 1 to 6 GHz with 10 MHz pulse repetition frequency (PRF), a peak power of 100 mW and a scan rate of 40 Hz.

Figure 3. The modular MIR system: transmitter module, receiver module (front), two antennas (sides), and the timing and control unit (back).

We set up a target scene from which to acquire linear 2-D multistatic data consisting of a collection of metallic objects and spanning approximately 1 m (**Figs. 4 and 5**). The scene was scanned using the CART and a 1.5-m multi-bistatic synthetic aperture with 1 cm spacing for both transmitter and receiver. The set of monostatic equivalent data was extracted from the data set and is shown reconstructed in **Fig. 5c**. Also shown is a bistatic image with a single transmitter located in the center of the scene (**d**) along with two multi-bistatic images with combined information from multiple transmitters ((**e**) and (**f**)). The bistatic image was limited, due to the illuminating beamwidth of the radar, and therefore the entire scene is not visible. However, with as few as ten equally-spaced transmitters across the scene, the entire scene is adequately illuminated and is comparable to, if not better than the monostatic image.

To make a quantitative comparison of the different imaging modes, we must compare the clutter levels of the images. Generally, clutter is defined as anything in the image that is not a target. Two types of clutter are visible in the images of **Fig. 5**. The first type is side-lobe clutter resulting from the radar signal not focussed directly onto a target. The second type is multipath clutter resulting from non-linear multiple target interactions with the radar signal not accounted for in the imaging algorithm. From the images it is clear that the multi-bistatic data has less clutter as well as better image quality for shadowed objects, such as the cylinder located beneath the four thin rods.
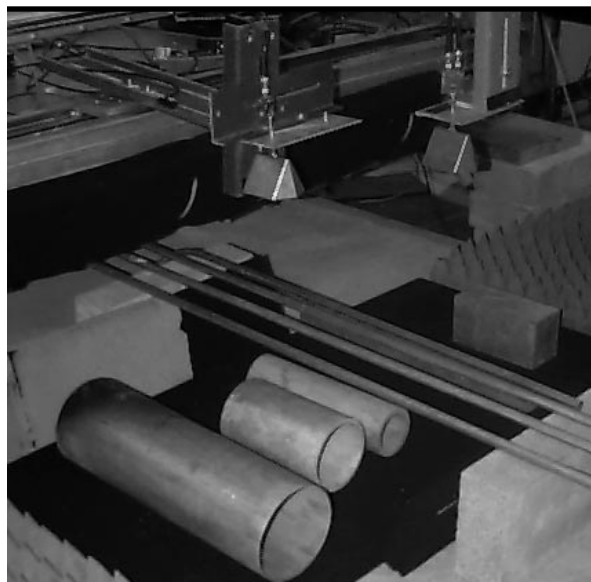

Figure 4. Target scene consisting of various metallic objects for multistatic data acquisition using modular MIR. Scanning was done horizontally across the center of the scene. Multi-bistatic data was acquired, consisting of 148 equally spaced transmitter locations.

Also note that the multi-bistatic image is less susceptible to multipath clutter. A multipath artifact is visible between the right two cylinders in the monostatic image; however, in the multi-bistatic images the multipath artifacts are diminished. A quantitative measure of peak signal-to-clutter ratios (SCR) was computed for the images in **Fig. 5** and is plotted in **Fig. 6**. Clearly only a few transmitter positions are required to achieve significant SCR improvement over the monostatic case.

The multi-bistatic images also offer improved signal-to-noise ratio (SNR) over strictly monostatic systems. This improvement results from integration of the scattered signal over all receiver positions for each transmitter used. The exact amount of improvement depends on the antenna beamwidth, because not all receivers collect energy from the scattering induced by each transmitter. This is highly dependent upon the relative position of the transmitters and receivers.

2-D Circular-Mode Imaging

A rotational stage testbed (RST) was developed to image small objects in multistatic modes (**Fig. 7**). Data acquisition and control software from the CART were used to acquire multistatic data from the RST. However, new imaging algorithms were developed that accounted for the circular geometry of the resulting data. Applications include small objects that can fit within the 40-cm diameter of the stage, or more specifically, cylindrical objects such as utility poles and small concrete columns. We chose a portion of a utility pole for initial testing and evaluation. In **Fig. 7**, a section of a utility pole is imaged on the rotational stage using a monostatic configuration. Data was acquired from a single rotation of the utility pole while illuminating it with a stationary monostatic MIR configuration, obtaining 437
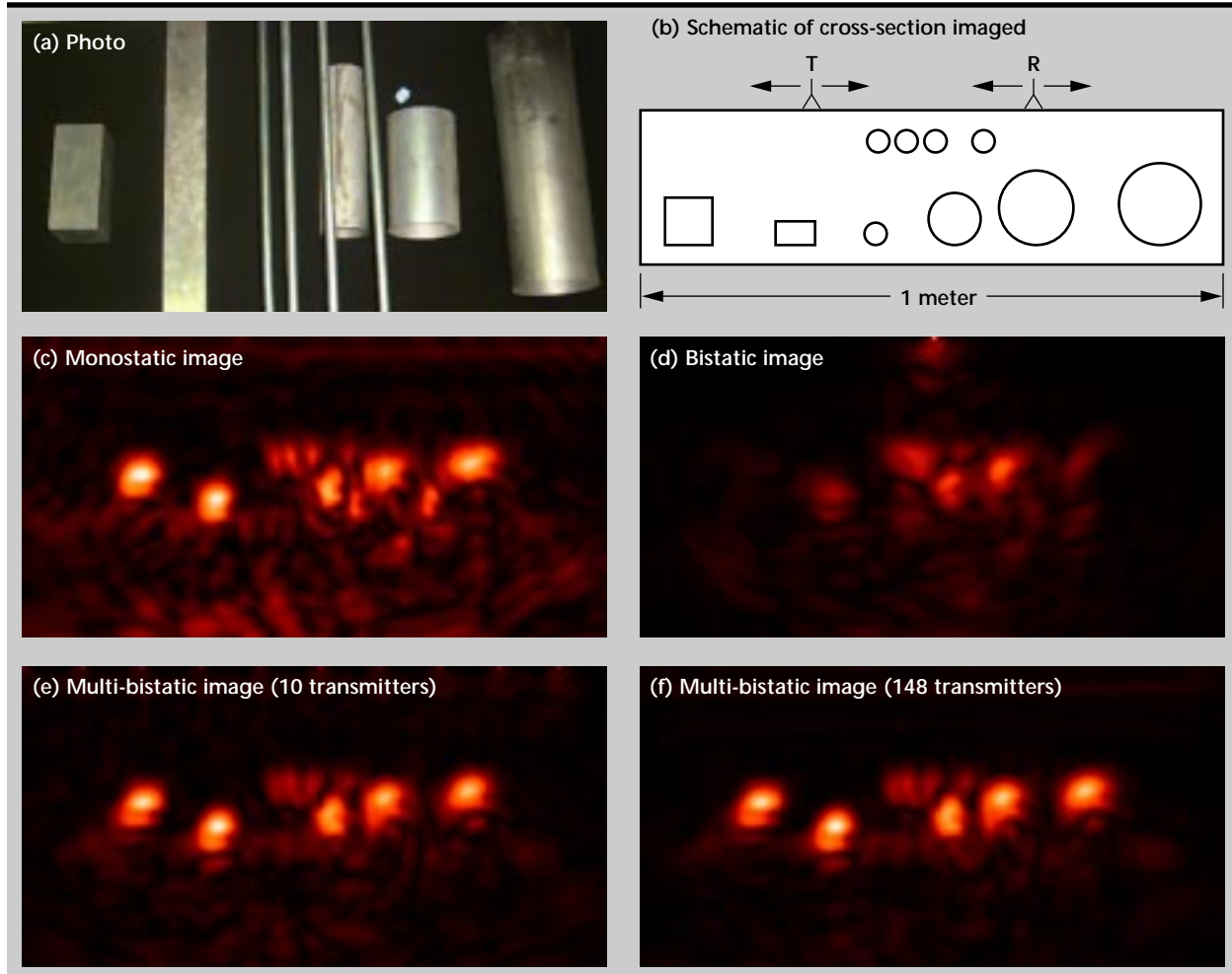

Figure 5. Images of the target scene using different imaging modes. Each image was obtained using 148 receiver positions separated by 1 cm. The monostatic data was acquired from a single T/R forming a synthetic aperture. For the bistatic image, the transmitter was positioned directly over the center of the scene with the receivers spanning the entire aperture. For the multi-bistatic cases, the transmitters were equally spaced across the scene.

equally spaced samples around the circular aperture. This technique combined with subsequent image processing easily images the 2-cm-wide fissure within the utility pole.

We also performed multi-bistatic data acquisition from the utility pole. A transmitter was attached to the stage such that it rotated with the object while an MIR receiver was held stationary. Data was then acquired from a single rotation of the pole and the process was repeated for each of 36 equally spaced transmitter locations. The multi-bistatic technique, with only a single T/R pair, is significantly more time-consuming than the monostatic data acquisition, because a full rotation of the stage must be done for each transmitter location. However, similar to the linear multistatic method, it provides better SCR than the monostatic scenario, as well as
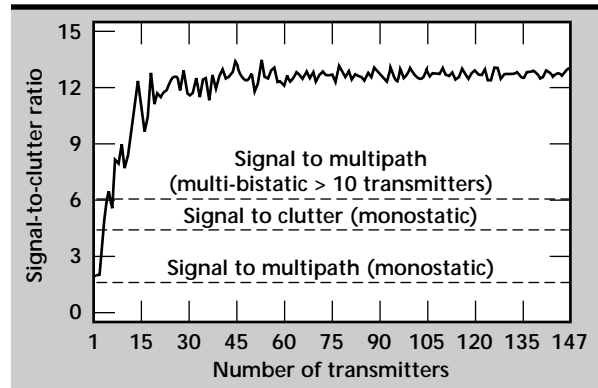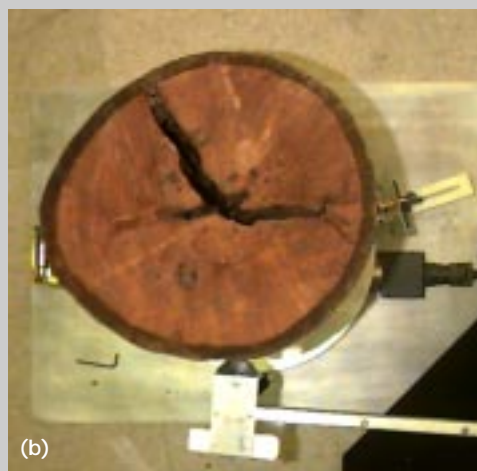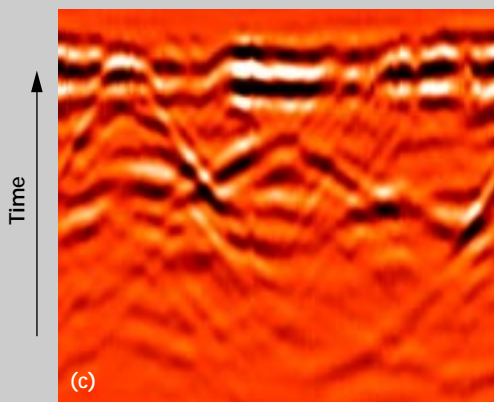

Figure 6. Signal-to-clutter ratio for multi-bistatic image. Dashed lines represent signal-to-clutter and signal-to-multipath levels for the multi-bistatic and monostatic images. Multipath artifacts represent multiple bounces of the radar signal which are nonlinearities unaccounted for by the imaging algorithm.
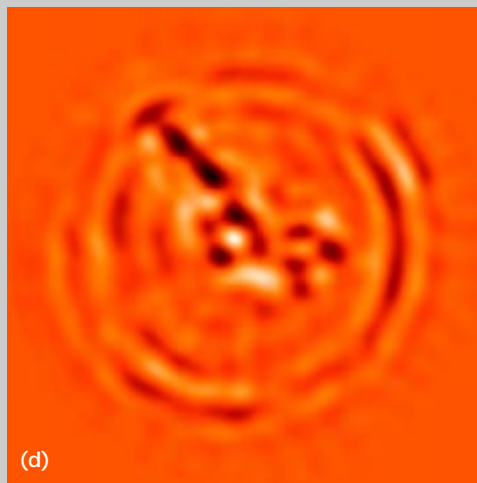
Figure 7. Rotational stage used to acquire monostatic circular reflection profile data from a utility pole (a). Impulse radar data was acquired from 437 equally-spaced sample points around the circumference of the pole (b). The data (c) was reconstructed to produce an image of the 2-cm fissure in the pole (d).

(a) Rotational stage testbed

(b) Utility pole containing a fissure

(c) Radar data from a 360 degree circular monostatic scan of the utility pole

(d) Reconstructed slice of utility pole clearly shows the fissure.

better visualization of some of the salient features of the fissure. This improvement is clear in the multi-bistatic reconstructions of the fissure shown in **Fig. 8**.

In the full reconstruction, 36 bistatic views are combined to image the features within the utility pole, but 9 views are sufficient to reconstruct the image without significant degradation. In general, the required number of transmitters will vary, depending on the size of the object and the beamwidth of the antenna.

Future Work

Our future work involves concentration in three areas:
1) improving radar characteristics, including improved antennas, higher frequencies, and higher PRF;
2) modifying data acquisition to involve antenna arrays and custom apparatus for hand-held and other specialized applications; and
3) expanding imaging capabilities to include 3-D multistatic and circular geometries as well as accounting for object shape and other application-specific signal processing considerations.

In this report we discussed non-coupled data acquisition in which there was an air gap between the antenna and the object. Improved SNR and resolution can result from matching the antenna characteristics more closely to the dielectric properties of the material under inspection. This matching requires dielectrically-loaded antennas that couple the radar directly to the material. The resulting antenna would be more compact and would induce less ringing between the surface of the object and the antenna than the air-launching antennas. Furthermore, optimal frequency selection for the radar could result in improved performance.

The exact frequency band is highly application-dependent, due to the wide variety of attenuation characteristics of different media types. For the utility pole application, increasing the frequency to 10 GHz could improve resolution, detection, and visualization of interior features such as cracks, rot, and water damage. Further testing is required to determine the optimal radar frequency band for inspection of utility poles.

Data acquisition times for multistatic scenarios can be quite large, especially when taking data using only a single T/R pair. Future work would incorporate antenna arrays or other techniques which could reduce acquisition times significantly. With multiple element systems, mutual coupling between antenna elements and optimization of array performance will have to be considered. The specific implementation details will be dependent on the given application. However, in general, increasing the radar effective PRF and using multiple T/R pairs will provide increased acquisition rates.

The technique we used for imaging the utility pole did not account for the non-circular shape of the pole. Similarly, the planar imaging method did not compensate for non-planar surfaces. If not compensated for, these nonlinearities cause phase distortion, resulting in image degradation. Future work involves accounting for these nonlinearities, at least to first order, to help improve resolution and visualization of interior object features.

Acknowledgments

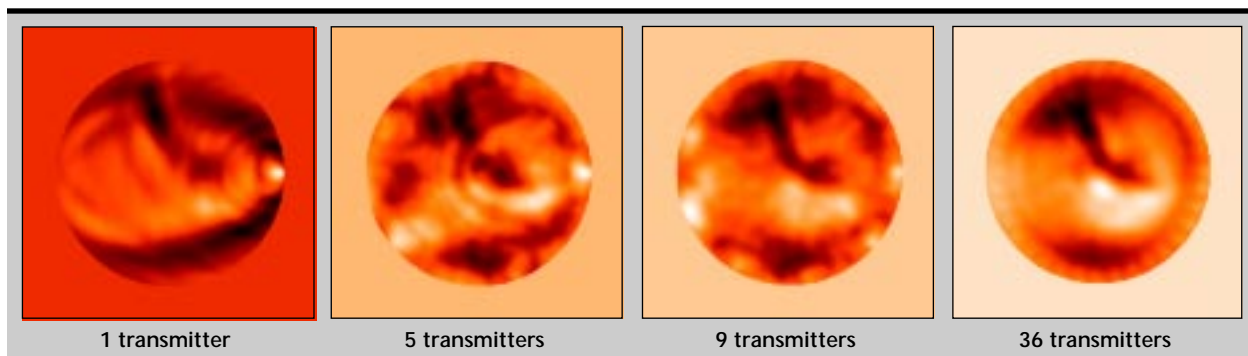| 1 transmitter | 5 transmitters | 9 transmitters | 36 transmitters |

Figure 8. Multi-bistatic reconstructions using indicated number of equally-positioned transmitters around the utility pole. In all cases, 437 receiver positions were used. Note that the multi-bistatic image has better signal-to-clutter than the monostatic image of Fig. 7.

References

1. Mast, J. E., and S. G. Azevedo (1996), *Applications of Micropower Impulse Radar to Nondestructive Evaluation*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-95).

2. Johansson, E. M., and J. E. Mast (1995), *Ultra-wideband radar imaging for the nondestructive evaluation of bridges*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-94).

3. Johansson, E. M., and J. E. Mast (1994), *Imaging algorithms for synthetic aperture ultra-wideband radar*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-93).

4. Devaney, A. J. (1982), "A filtered backpropagation algorithm for diffraction tomography," *Ultrasonic Imaging*, Vol. **4**, pp. 336–350.

5. Porter, R. P., and A. J. Devaney (1982), "Generalized holography and computational solutions to inverse source problems," *Journal of the Optical Society of America*, Vol. **72**, pp. 1707–1713.

6. Porter, R. P. (1989), "Generalized holography with application to inverse scattering and inverse source problems," *Progress in Optics XXVII*, E. Wolf, Ed., Elsevier Science Publishers B.V., pp. 317–397.

7. Chew, W. C. (1990), *Waves and Fields in Inhomogeneous Media*, Van Nostrand Reinhold, New York, N.Y.

8. Azevedo, S. G., and T. E. McEwan (1997), *Millimeter-wave MIR*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-96).

9. Azevedo, S. G., T. E. McEwan, and J. P. Warhus (1996), *Microradar Development*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-95).

Multi-modal NDE for AVLIS Pod Shielding Components

Diane J. Chinn, Nancy K. Del Grande, and Dwight E. Perkins
Manufacturing and Materials Engineering Division
Mechanical Engineering

We are assessing several non-destructive evaluation (NDE) techniques as possible replacements for existing pod component inspection methods for the Advanced Vapor Laser Isotope Separation (AVLIS) program at Lawrence Livermore National Laboratory (LLNL). The present inspection method uses dye-penetrant testing to inspect surfaces of components for cracks. Because this type of testing is labor-intensive, subjective, and evaluates only the surface of components, five other methods are under consideration. Each technique in our assessment, digital radiography, ultrasonic testing, optical imaging, dual-band infrared imaging and dye-penetrant testing, is evaluated for performance criteria that will be important to production runs.

Introduction

The AVLIS program at LLNL is presently designing a pilot demonstration facility. When demonstration runs begin, high through-put will require all aspects of the production process to be streamlined and robust. Shielding components in the AVLIS pod undergo mechanical and thermal loading during use. Components with stress-induced cracking caused by these loads must be replaced upon detection of cracks.

During the conceptual design phase, visual dye-penetrant testing was used to inspect the shielding components for cracks. Although dye-penetrant testing is especially labor-intensive, the low through-put in the conceptual design phase allowed designers to use the technique because it minimized capital costs. Other NDE techniques may be more suitable for large volume inspection.[1]

We are currently assessing several techniques for NDE of pod components for the AVLIS pilot facility design as possible replacements for existing pod component inspection methods.

AVLIS designers are seeking faster inspection methods that are reliable, recordable, capable of finding internal cracks and detecting cracks occurring on inaccessible surfaces. The present dye-penetrant inspection method takes as long as 5 to 6 days for a complete evaluation of all the pod components. Because it is a surface inspection method, dye-penetrant testing cannot find internal cracks, nor can it find cracks on the inner diameter of tubular sections.

The present dye-penetrant technique does not include digital archiving of the results. Digital archiving is important for production runs because components with cracks occurring at non-critical locations are often not replaced. Archiving of the inspection results allows inspectors to track any crack growth which may occur in non-critical locations.

We are evaluating five different NDE techniques for detecting cracks in shielding components of the AVLIS pod. Each technique, digital radiography, ultrasonic testing, optical imaging, infrared imaging and dye-penetrant testing, is evaluated for (1) ability to find surface-breaking and internal cracks as small as 1.6 mm in length; (2) ability to inspect the inner diameter of tubular components; (3) potential for automation; (4) digitization of inspection results; (5) reliability and accuracy in correctly distinguishing cracks from other defects; (6) ability to be handled in a glove box; and (7) rapid evaluation of complex shapes.

Progress

The pilot facility will contain approximately 120 pod shielding components. The three major component shapes are plates, tubes and troughs. A few odd-shaped components will also need to be

inspected. We tested several components, of which one plate contained a crack. Because of the stress profile on the components, cracks almost always propagate through the thickness, rather than circumferentially or in a plane.

Shielding components are contaminated during normal AVLIS operations and must be inspected in a glove box with appropriate health and safety precautions. Some of the inspection methods are appropriate only for certain shapes, while other methods would be difficult to use in a glove box.

So far, we have had success detecting cracks and material variations with ultrasonic testing, infrared imaging and dye-penetrant testing of several components. Digital radiography and optical imaging do not appear suitable for pod components.
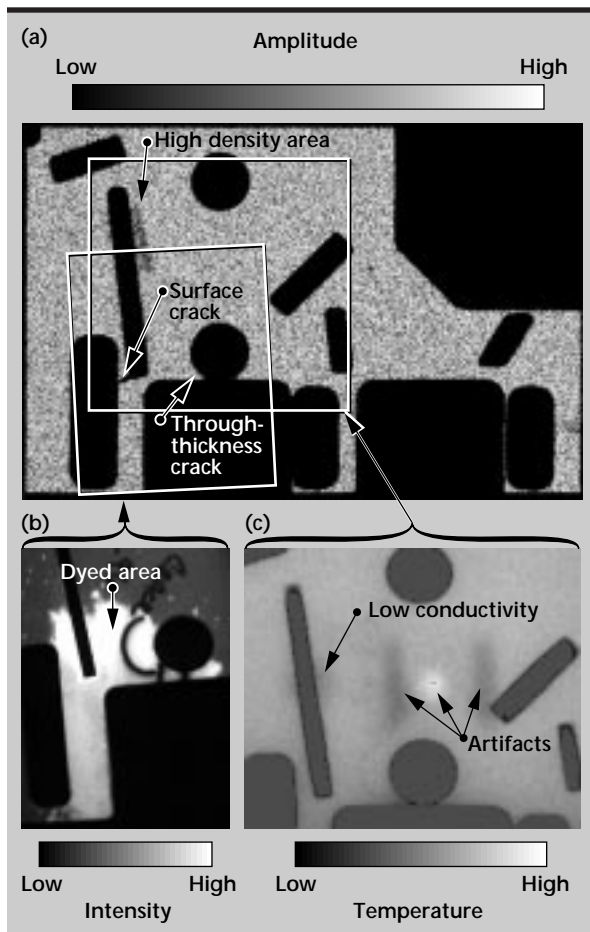

Figure 1. Imaging of AVLIS pod components using three techniques. (a) Ultrasonic imaging of the plate component shows cracks and density changes. Areas included in (b) and (c) are shown on the ultrasonic image. (b) Optical imaging of dye-penetrant testing can be used to archive results of the inspection. The arrow shows the area where dye penetrant has been applied. (c) Infrared imaging shows the progression of heat conduction through the part. Artifacts in the image result from the collection system.

Ultrasonic Testing

Ultrasonic inspection uses acoustic waves traveling through the specimen to interrogate the surface and interior of the specimen for defects. Changes in homogeneity of the material alter the characteristics of the ultrasonic wave. By scanning the specimen, we obtain an ultrasonic image of the specimens showing variations in homogeneity.
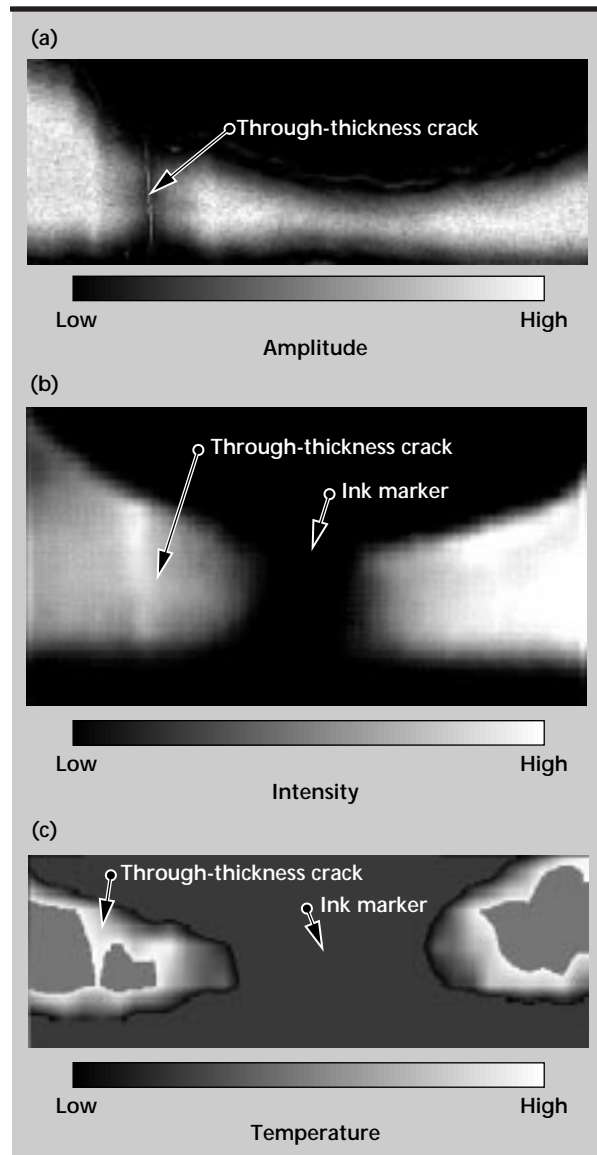

Figure 2. A through-thickness crack imaged using the three inspection techniques. (a) Ultrasonic imaging of surface wave propagation shows that very little energy propagates across a through-thickness crack. (b) Optical imaging of dye-penetrant testing records the location of the crack. (c) Infrared imaging shows discontinuity in thermal transfer across the through-thickness crack. The dark line to the right of the crack is from pen markings.

Figure 1a shows an ultrasonic image of a plate component. Areas with density variation, a through-thickness crack and a surface crack are indicated in the image. **Figures 2a** and **3a** show high-resolution ultrasonic images of the through-thickness crack and the surface crack, respectively. The images in **Fig. 2a** and **3a** plot the amplitude of a surface wave traveling along the surface of the plate. The surface wave has very low amplitude traveling across a through-thickness crack (**Fig. 2a**) and slightly decreased amplitude traveling across a surface crack (**Fig. 3a**) depending upon the depth of the crack. These images demonstrate the capabilities of ultrasonic testing in finding and characterizing defects.

The disadvantages of using ultrasonic testing include the need for a couplant such as water to introduce acoustic waves into the material. Different scanning configurations are required for each shape making the ultrasonic technique expensive if multiple scan systems are required for each shape. In addition, high-resolution scanning generally takes 15 to 30 min to scan specimens of the size of pod components.

Dye-penetrant Testing

This method is presently used to manually inspect the pod components. A fluorescent dye is painted onto the surface of the component and is absorbed by surface-breaking cracks. Filled with dye, these cracks are easily detected by visual inspection of the part under ultraviolet light. Many different shapes are easily inspected with this method.

Internal defects can not be detected with dye-penetrant testing. The disadvantage of this technique is that it is time-consuming and difficult to implement when inspecting the inner diameter of tubular sections. This technique can be archived easily using an optical recording system. An optical image of the pod component, through-thickness crack and surface crack filled with dye are shown in **Figs. 1b**, **2b**, and **3b**, respectively.

Infrared Imaging

Infrared imaging uses the conductivity of heat through the specimen to evaluate the surface and interior of the specimen. After flashing the material with a few milliseconds of intense uniform light, variations in thermal conductivity of the part are revealed by imaging the surface with an infrared camera.[2] One feature of infrared imaging is that it can inspect large areas of specimens with complex shapes very quickly. **Figures 1c**, **2c**, and **3c** show infrared images of the pod component.

Radiographic Testing

Radiography uses the transmission of x rays through a material to image differences in density. Cracks are generally imaged best with the source transmitting parallel to the plane of the crack. Even under ideal source-crack configurations, cracks must still have sufficient size in order to be resolved by film or digital radiography. For this reason, we found that radiographic
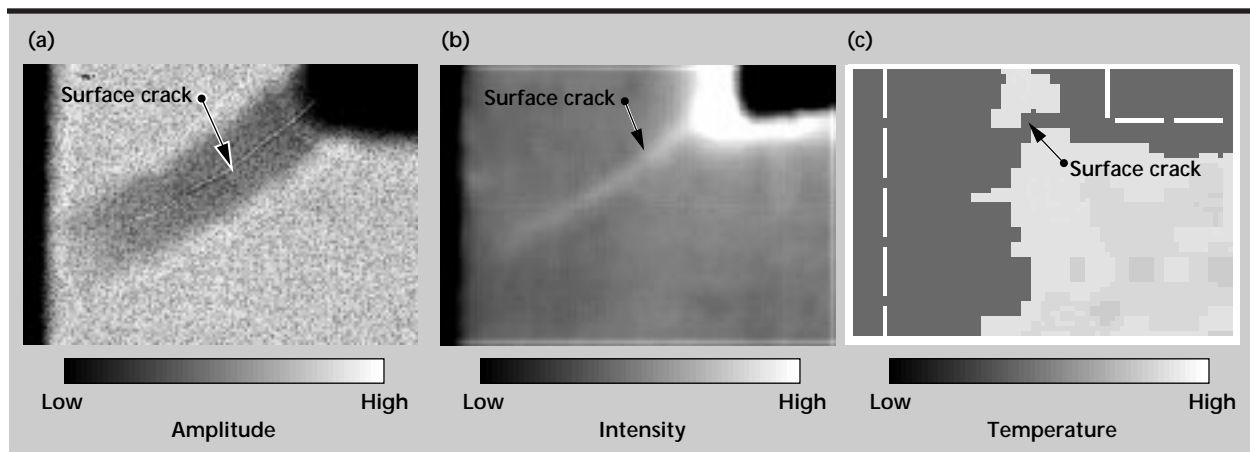

Figure 3. Imaging of a surface crack by the three inspection techniques under consideration. The arrow shows the location of the surface crack indication in each image. (a) Ultrasonic image of surface wave propagation shows partial attenuation of the wave as it passes over the surface crack, transmitting some of its energy. (b) Dye-penetrant testing easily detects surface cracks. (c) Infrared imaging indicates a decrease in the thermal conductivity at the location of the surface crack. Dotted lines show the edges of the test specimen near the crack.

Table 1. Evaluation of inspection techniques. Each technique is evaluated according to its ability to satisfy the performance criteria.

Performance criteria	Ultrasonic testing	Infrared imaging	Dye-penetrant imaging	Radiographic testing	Optical imaging
Surface cracks	√	√	√	×	×
Internal cracks	√	√	×	×	×
Inside tubes	√	√	√	×	√
Automation	√	√	√	√	√
Digital archiving	√	√	√	√	√
Defect classification	√	?	√	×	×
Complex shapes	×	√	√	√	√
Glove box inspection	√	√	√	×	√
Inspection time (min), 1 plate	15	2	5	3	3

√ - satisfies criteria   × - unacceptable or not possible   ? - to be determined

testing is not suited to this application because cracks in the pod components are not large enough to be detected.

Optical Imaging

Optical imaging records the transmission of light through the part to detect inhomogeneities. As with radiographic testing, optical transmission requires that a crack be large enough relative to the detector sensitivity in order to be detected. We found that optical imaging is not suitable for pod components because of the small crack sizes.

Preliminary Assessment

Ultrasonic testing, dye-penetrant testing and infrared imaging show promise as pod component inspection methods. We have begun to assess these three techniques for different component shapes based on AVLIS inspection needs. So far, we have performed ultrasonic testing on plate and tubular members, infrared imaging on plate and trough components, and optical imaging of dye-penetrant testing on plates. The matrix in **Table 1** summarizes our preliminary findings on each technique for AVLIS inspection performance criteria.

Radiography and optical imaging have been eliminated from consideration because they cannot find surface or internal cracks easily. Of the remaining

techniques, ultrasonic testing takes the longest time to inspect a plate component. Infrared imaging has the fastest inspection time, however, we are still working on interpretation of the images for defect classification.

Future Work

Comprehensive NDE of AVLIS pod components will most likely involve several modes of inspection. We have begun to collect images, perform image analysis and processing on each of the remaining three NDE methods. We will continue our work on other shapes and other types of defects. This work provides guidance for the selection of improved inspection methods that will significantly impact the operations of the AVLIS system.

References

1. Chinn, D. J., D. J. Schneberk, N. K. Del Grande, and G. H. Thomas (1996), "A comparison of quantitative imaging techniques for lightweight materials," *J. Acoust. Soc. Am.,* **100** (4), October.

2. Del Grande, N. K., P. F. Durbin, and D. E. Perkins (1996), "Infrared computed tomography for characterizing structural defects," *Engineering Research Development and Technology Thrust Area Report*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL 53868-95).

David N. Hunt, Thrust Area Leader

The Information Engineering thrust area develops information technology to support the programmatic needs of Lawrence Livermore National Laboratory's (LLNL) Engineering Directorate. The mission of the thrust area is to maximize the information bandwidth between the system and its users.

A major requirement to fulfill this mission is "intelligent" access to the information served by the system, calling for advances in compute platforms, networks, user interfaces and applications.

In particular, the development of systems that can act as intelligent agents for their users is critical to maximizing the benefits of this technology.

Our approach has two major components. First, we are matching the needs of the programs at LLNL with developments in leading edge information technology. Our efforts include work in high-speed computing and networking, data acquisition and control, and the collection, storage, collation, and analysis of massive amounts of scientific data. Progress in this area requires collaboration with programs and other disciplines at LLNL, as well as with the industrial and academic sectors.
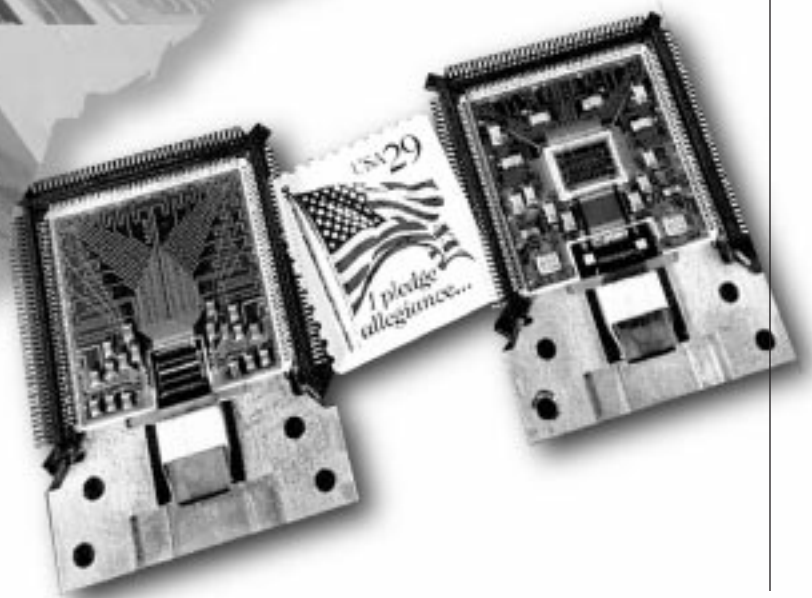
The second major component of the thrust area effort is focused on leveraging information technology by applying it to the LLNL infrastructure. Much of the technology used to meet programmatic needs can be used to increase the efficiency of day-to-day operations. In these days of fiscal consciousness, proper application of information technology can conserve both time and money.

As we close FY-96, we see benefits of the work we have pursued over the past two years: we have produced patentable work in subcarrier multiplexed networks, and LLNL's Procurement Department is embracing the Zephyr concept in rapid prototype procurement and is investigating its broader applications.

The thrust area activities for FY-96 are described in the following five papers: (1) Three-Dimensional Object Creation, Manipulation, and Transport; (2) Zephyr: A Secure Internet-Based Process to Streamline Engineering Procurements; (3) Subcarrier Multiplexing: Optical Network Demonstrations; (4) Parallel Optical Interconnect Technology Demonstrations; and (5) Intelligent Automation Architecture.

Information
Engineering

8

8. Information Engineering

Three-Dimensional Object Creation, Manipulation, and Transport

Scott D. Nelson
Defense Sciences Engineering Division
Electronics Engineering

The effective creation, manipulation, communication, and transport of 3-D objects has historically been a problem because of competing or specialized data formats, the requirements for sophisticated tools on the end user's part, or a steep hardware requirement for the display and transport of objects. This effort combines several parallel activities into a unified medium, taking advantage of recent economies of scale in the commercial sector that allow for the efficient distribution of tools to a broad class of users.

Introduction

There are several kinds of 3-D data types used to convey information about solids. The newer versions of these are specific to various programs, and at Lawrence Livermore National Laboratory (LLNL) they are used in a series of computational modeling codes. The key to the approach presented here is the creation of a standard 3-D infrastructure to aid in the creation, manipulation, and transport of 3-D data objects.

The exchange and definition mechanism was dealt with by creating the new high-level MIME format for 3-D data types. The software portion of the transport issue is handled as part of this new primary MIME type. The methodology used for the creation aspect is different from what others are working on.[1] Our effort has the advantage of using general images taken from positions surrounding the body of interest, with little information or control used during the acquisition phase. The difference is that more information is known about the object and less about the camera positions.

By contrast, this approach is different from simplistic commercial systems[2] that require the user to specify object geometries in coordinate planes. For complex objects such as missiles and helicopters, this is beyond the scope of such software.

Progress

New Primary MIME Type Registered

The Glossary in Reference 3 contains definitions and abbreviations.

As part of this activity, the MIME type for the general class of problems was created and is an Internet standard as specified in Reference 4. This allows the easy conversion of various 3-D types by grouping the data types in a common framework of the model/*MIME type. This type presently includes the VRML 3-D data type, the IGES data type, and a generic mesh data type.

The principal focus of this effort was to create a standard "umbrella" mechanism that would allow other 3-D/4-D data types to be encapsulated underneath. As a consequence, the base set of parameters for these data types was also created and includes requirements on scale, interaction, abstraction, and the flow of time.

A tree diagram and outline for the model type is as follows:

```
model/
    vrml
    iges
    mesh
```

A model primary MIME type is an electronically exchangeable behavioral or physical representation within a given domain. Each sub-type in the model structure has unique features, just as does each sub-type in the other primary types. The important fact is that these various sub-types can be converted between each other with less loss of information than by converting to other primary types. This fact groups these sub-types together into the model primary type.

All of the expected sub-types have several features in common that are collectively unique to this primary type. To loosely summarize: models are multi-dimensional structures composed of one or more objects. If there are multiple objects, then one object defines the arrangement, setting, and relationship of the others. These objects all have calibrated coordinate systems, but these systems need not be in the same units, nor need they have the same dimensionality. In particular:

1. They have three or more dimensions that are bases of the system and form an orthogonal system (any orthogonal system is sufficient). This system is specifically defined in terms of an orthogonal set of basis functions (for a sub-space of the L2 function space) over a coordinate system of dimension three or more. Note that this does not preclude regular skewed systems, elliptical coordinates, or different vector spaces.
2. They contain a structural relationship between model elements.
3. They have scaling or calibration factors that are related to physical units such as force, momentum, time, velocity, acceleration, and size. Thus, an IGES file will specify a building of non-arbitrary size; computational meshes and VRML models will have real spatial/temporal units. This allows for differing elements to be combined non-arbitrarily.
4. The models can be single objects or composed of a collection of objects. These normally independent objects are arranged in a master/slave scenario so that one object acts as the reference, or primary object, and defines how the other objects interrelate and behave. This allows for the creation of mathematical, physical, economic, or behavioral models that typically are composed of different elements.

The key is in the description: these types describe how something "behaves," in contrast to typical data types which describe how something "is." The inclusion of this "collective" system works in a way similar to the e-mail system's multi-part/related type which defines the actions of the individual parts. Further specification of the model/*subtypes using these properties is left to the sub-type authors.

With these assumptions:

1. The default dimensionality will be spatial and temporal (but any are allowed).
2. Models will contain underlying structure which may or may not be immediately available to the user (for example, fluid dynamics vector fields, electromagnetic propagation, interrelated IGES dimensional specifiers, VRML materials and operators).
3. Basis set conversion between model domains is lossless. The interpretation of the data may change but the specification will not, that is, convert the model of the U.S. Gross Domestic Product into a VRML model and navigate it to explore the variances and interrelationships. The model has many dimensions but also "passages" and "corridors" linking different parts of it. A similar situation is true for meshes and CAD files. The key is identifying the basis set conversion that makes sense.
4. Models are grouped to assure less loss of information between the model sub-types than to sub-types of other primary types.

Data Conversion

The conversion part of this activity grew out of standard activities and included the SBIR[5] made by NIST on display and conversion programs. Additional efforts were made by the community at-large to develop the appropriate suite of conversion tools for VRML, and these are mostly in place and operational. At the beginning of the project, quite a bit of effort was put into a set of functions that operated with the LLNL mesh format, but since then the broader class of STEP[6] mesh descriptions is going through the International Standards Organization (ISO) process and is expected to be the more generic alternative. Conversion functions for this format will be created as part of future program-specific activities as needed.

Examples of the existing formats include a mesh model of a shipping container,[7] a VRML model of an aircraft carrier, and an IGES model of a bone (**Fig.1**).

Object Creation Definition and Approach

This problem involves locating the coordinates of N points in 3-space contained in a collection of M 2-D images taken from different perspectives. Not

all points must be visible in each of the images; however, each point must be visible in at least two images. It is also assumed that the field of view (FOV) of each image and the point correspondence between images are known.

To make acquiring the images in the field as simple as possible and to reduce the effects of measurement error, it is desirable to solve for the N point locations with the minimum required information about the camera positions and orientations, while using the measured data to gain as much information as possible. The orientations and locations of all cameras are solved for whose images contain these points.

Next, all uncomputed points in each of the cameras' images are computed. This process is then repeated until all camera information is known and the locations of all points are known. Note that

depending on how many images a point appears in, the location of a point may be computed many times. These locations are averaged to provide the ultimate estimate of the point location.

Problem Solution

We define the geometry for an image as shown in **Fig. 2**. In local coordinates, x', y', z', the normal vector of the image plane $\hat{n}' = \hat{z}'$ originates at the center of the image plane. Each image is $w \times h$ pixels in size, with known angular field of view in both the x and y directions. The coordinate location (x_{ij}, y_{ij}), of the jth point in the ith image, along with point correspondence between points in the images, are given.

From these definitions, the direction vector from the ith camera location to the jth point, $\hat{u}'_{ij}$, follows from
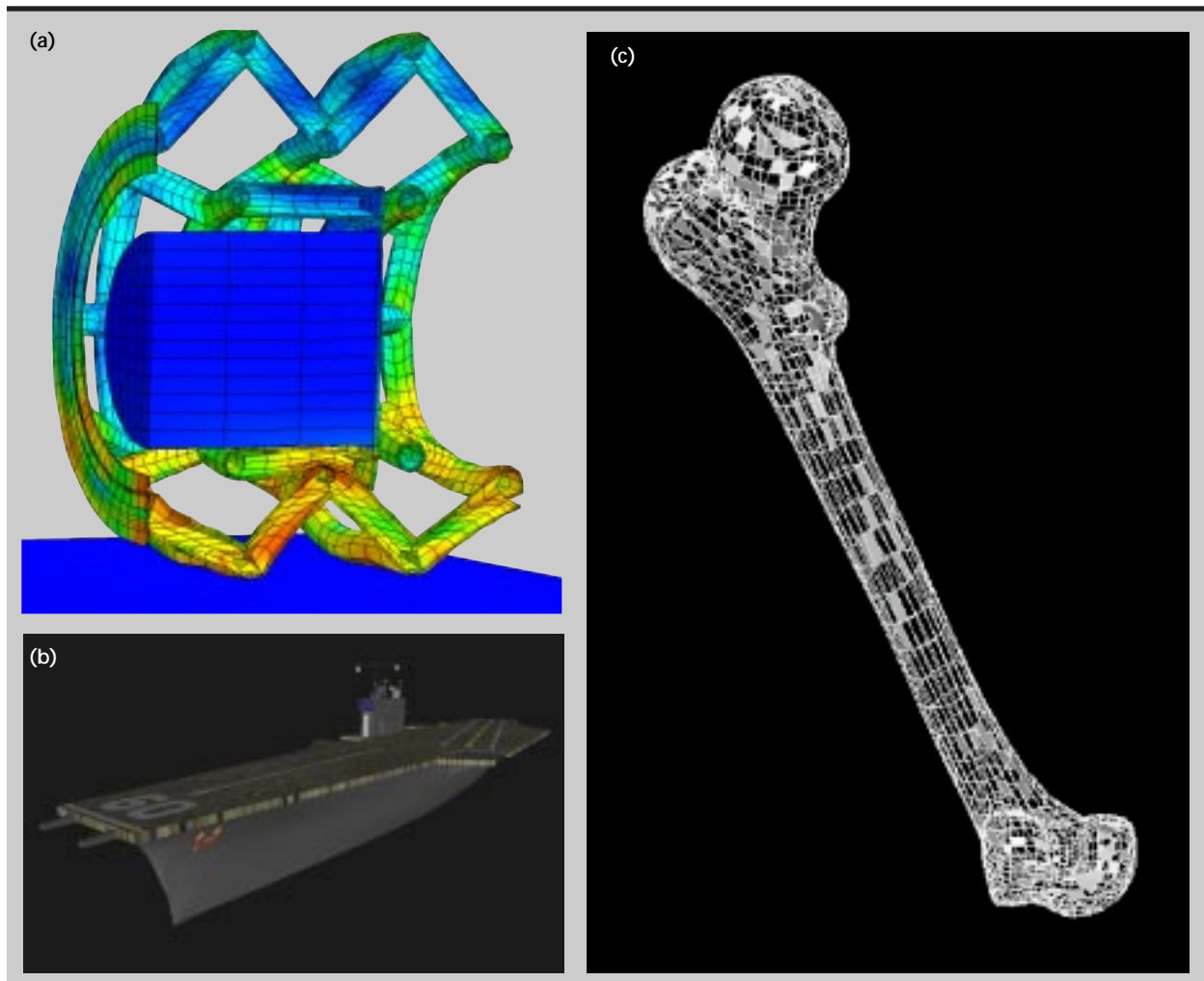

Figure 1. Mechanical model of (a) a shipping container; (b) a VRML model of an aircraft carrier; and (c) an IGES model of a bone.

$$\theta_{xz} = sign(x_{ij}) \sqrt{\left(\frac{FOV_x}{\omega} x_{ij}\right)^2 + \left(\frac{FOV_y}{h} y_{ij}\right)^2} \quad,$$

$$\theta_{xy} = \begin{cases} \tan^{-1}\left(\dfrac{y_{ij}}{x_{ij}}\right) & x_{ij} \neq 0 \\ 90° & x_{ij} = 0 \end{cases}$$

giving

$$\hat{u}'_{ij} = \begin{bmatrix} \cos(\theta_{xy}) & \sin(\theta_{xz}) \\ \sin(\theta_{xy}) & \sin(\theta_{xz}) \\ & \cos(\theta_{xz}) \end{bmatrix}.$$
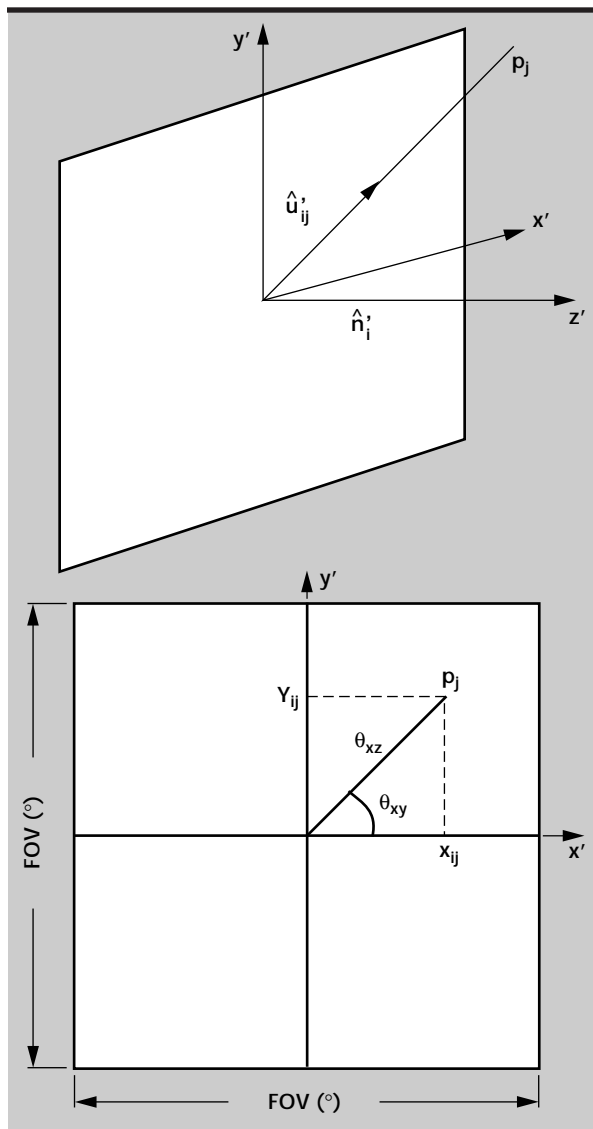

Figure 2. Point location defined in image geometry and in 3-D space relative to camera location and direction. The direction vector of the camera is normal to the image plane and originates at the center of the image. The image provides angular information to the point relative to the direction vector of the camera.

In global coordinates, the location of the direction vector to the jth point from the ith camera is

$$\hat{u}_{ij} =$$

$$\begin{bmatrix} \cos(\theta_{xy})\left[n_{xi}\cos(\theta_{xz}) + n_{zi}\sin(\theta_{xz})\right] - n_{yi}\sin(\theta_{xy}) \\ \sin(\theta_{xy})\left[n_{xi}\cos(\theta_{xz}) + n_{zi}\sin(\theta_{xz})\right] + n_{yi}\cos(\theta_{xy}) \\ -n_{xi}\sin(\theta_{xz}) + n_{zi}\cos(\theta_{xz}) \end{bmatrix}$$

where

$$n_i = \begin{bmatrix} n_{xi}, n_{yi}, n_{zi} \end{bmatrix}^T.$$

Defining the translation vector, $\bar{t}_i$, denoting the position of the ith camera we have the relationship between corresponding points contained in two images as depicted in **Fig. 3**. The two angles θ_1 and θ_2 follow straightforwardly from the fact that the direction vectors from the cameras to the point and the location vectors are known. The distances to the point from the respective cameras then follow from the Law of Sines. Each point location is computed in this manner for each corresponding stereo image pair.

We now use the information about the location of the points to compute the unknown position and direction information for all images which contain these points. For a given point in an image, the relationship between the point location and the camera position and direction is

$$\bar{p} = \bar{t} + R\alpha\hat{u}',$$

where R is defined in terms of the three Euler rotation matrices

$$R = R_{xy}R_{yz}R_{zx}$$

$$R_{xy} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R_{yz} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\rho) & -\sin(\rho) \\ 0 & \sin(\rho) & \cos(\rho) \end{bmatrix},$$

$$R_{zx} = \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix}.$$

With R having three degrees of freedom, the translation vectors being unknown, and the distance from the image plane to the points being unknown, this vector equation contains seven unknowns. If we take two additional points within the image such

that the three points are not collinear, then we obtain three linearly independent vector equations in 3-space with nine unknowns

$$\overline{p_1} = \bar{t} + R\alpha_1 \hat{u}_1',$$

$$\overline{p_2} = \bar{t} + R\alpha_2 \hat{u}_2',$$

$$\overline{p_3} = \bar{t} + R\alpha_3 \hat{u}_3' .$$

To solve the above equations, we first subtract pairs of equations to remove dependency upon the translation vector $\bar{t}$. Using the fact that magnitude is invariant to rotation, we can remove the dependency of R, giving:

$$\left| \overline{p_2} - \overline{p_1} \right|^2 = \left(\alpha_2 \hat{u}_2' - \alpha_1 \hat{u}_1' \right)^T \left(\alpha_2 \hat{u}_2' - \alpha_1 \hat{u}_1' \right),$$

$$\left| \overline{p_2} - \overline{p_3} \right|^2 = \left(\alpha_2 \hat{u}_2' - \alpha_3 \hat{u}_3' \right)^T \left(\alpha_2 \hat{u}_2' - \alpha_3 \hat{u}_3' \right),$$

$$\left| \overline{p_3} - \overline{p_1} \right|^2 = \left(\alpha_3 \hat{u}_3' - \alpha_1 \hat{u}_1' \right)^T \left(\alpha_3 \hat{u}_3' - \alpha_1 \hat{u}_1' \right).$$

These three equations are then solved numerically for α_1, α_2, and α_3.

At this point, we need to solve for the rotation and translation unknowns. We assume the two vector sets $\{ \overline{p_1}, \overline{p_2}, \overline{p_3} \}$ and $\{ \overline{\alpha_1 u_1}, \overline{\alpha_2 u_2}, \overline{\alpha_3 u_3} \}$ are linearly independent and not orthogonal. This assumption is reasonable, given the geometry and the way the points were selected. Noting that both multiplication by an orthogonal matrix and translation are angle-preserving transformations and that the two sets of vectors form congruent triangles, we proceed, letting $\overline{q_1} = \overline{p_2} - \overline{p_1}, \overline{q_2} = \overline{p_3} - \overline{p_1}, \overline{q_3} = \overline{q_1} \times \overline{q_2}$, and $\overline{q_4} = \overline{q_3} \times \overline{q_1}$. Similarly, we create $\overline{r_1}, \overline{r_2}, \overline{r_3}, \overline{r_4}$, from the vectors $\alpha_1 \hat{u}_1$, $\alpha_2 \hat{u}_2$, and $\alpha_3 \hat{u}_3$ respectively. Normalizing, we now complete the solution for the rotation and translation of the camera by solving

$$R = \begin{bmatrix} \hat{q}_1 & \hat{q}_4 & \hat{q}_3 \end{bmatrix} \begin{bmatrix} \hat{r}_1^T \\ \hat{r}_4^T \\ \hat{r}_3^T \end{bmatrix},$$

and

$$\bar{t} =$$
$$\frac{1}{3} \left(\overline{p_1} + \overline{p_2} + \overline{p_3} - R\left(\alpha_1 \hat{u}_1 + \alpha_2 \hat{u}_2 + \alpha_3 \hat{u}_3 \right) \right).$$

The algorithm for determining the collection of M points is as follows:

Step 1. For each image with known position and direction taken two at a time, compute the location of each point in the images. Iterate through all possible combinations of pairs of images.

Step 2. Compute the average value for each computed point.

Step 3. For each image without known location or rotation and at least three common known points, compute R and t for these images. Repeat for each group of three points in each image.

Step 4. Compute average value for each image of the direction vector and translation vectors computed in Step 3.

Step 5. Repeat Steps 1 through 4 until all points in all images are computed.

This solution is being applied to ideal and realistic 3-D objects with compound curved surfaces. The key is that information about the camera locations is limited or unknown thus allowing the surveyor more latitude in acquiring the data. Also note that this is applicable post-fact from a series of images, assuming that they meet certain criteria.
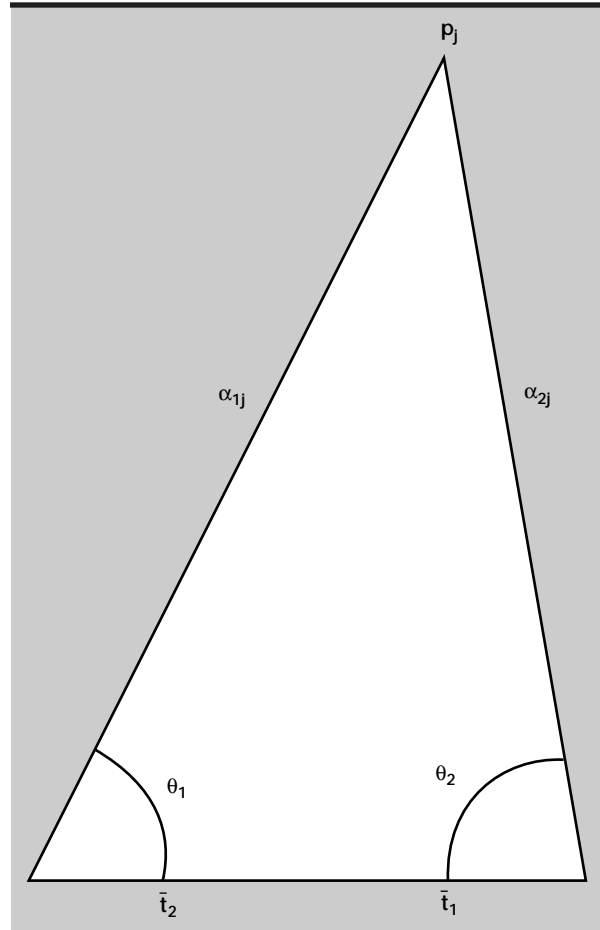

Figure 3. Stereo image geometry for solving for distance to point camera locations.

Acknowledgments

References

1. Lu, S-Y., J. E. Hernandez, and R. K. Johnson (1994), *Real-time Computer Vision,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-53868-94).

2. 3D Construction Company, (http://www.3dconstruction.com/).

3. Nelson, S. D., J. Kallman, and E. Brugger (1995), *Design and Implementation of 3D MIME Data Access and Exchange,* Lawrence Livermore National Laboratory, Livermore, Calif., (http://www-dsed.llnl.gov/documents/vis/sdnieta95.html), (UCRL-53868-95).

4. Nelson, S. D., M. Mitra, and C. Parks (1996), *The Model Primary Content Type for Multipurpose Internet Mail Extensions,* Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-ID-123343).

5. National Institute of Standards and Technology (1995, 1996), *Small Business Innovation Research (SBIR) for IGES Figure Viewer,* (http://www.eeel.nist.gov/iges/viewerSpec.html).

6. International Standards Organization (1996), *Integrated Application Resources: Finite Element Analysis,* (ISO 10303-104).

7. Hoover, C. (1994), *Shipping Containers for Hazardous Materials—DYNA3D Crash Simulation,* Lawrence Livermore National Laboratory, Livermore, Calif., (http://www.llnl.gov/liv_comp/meiko/apps/-dyna3d/dyna3d.html), (UCRL-MI 118882).

Zephyr: A Secure Internet-Based Process to Streamline Engineering Procurements

Cecil W. Jordan, Rudy E. Cavitt, and William A. Niven
Laser Engineering Division
Electronics Engineering

Nikola Mitschkowetz
New Technologies Engineering Division
Mechanical Engineering

Timothy M. Sharick and Sandra S. Taylor
Administrative Information Systems

Donald L. Vickers and Fred E. Warren
Computations Organization

Richard L. Weaver
Distributed Procurement Program
Procurement and Materiel Department

Lawrence Livermore National Laboratory (LLNL) is piloting an Internet-based paperless process called "Zephyr" to streamline engineering procurements. Programs at LLNL are benefiting from the efficiencies introduced since implementing Zephyr's engineering and commerce on the Internet. Major benefits include reducing procurement time, speeding the engineering development cycle, facilitating industrial collaboration, and reducing overall costs.

Introduction

Zephyr is being developed by LLNL's Engineering Directorate in partnership with its departments of Procurement, Administrative Information Systems, and Computations. Procurements that use Zephyr can move 60 to 90% faster from engineering design release, through contract award, to product delivery. Zephyr links people in LLNL's Engineering, Procurement, and Finance Organizations to pre-qualified small and medium enterprises (SME's) throughout the nation in a practical, simple, secure way. Procurements that used to take 30 to 60 days are now being done in 2 to 3 days, from initial request to delivery.

All the process steps are conducted by a World-Wide Web browser (Netscape or Internet Explorer) and e-mail. These steps are secure, that is, encrypted and password-protected. Zephyr's Web-based workflow process moves projects away from the sluggish paperwork path. LLNL's programs benefit from a compressed time-frame for engineering deliverables by ensuring valid designs, minimizing schedule impact, and reaching project goals sooner.

Concept

Zephyr is part of a system concept called CERPS, for Concurrent Engineering Rapid Prototyping System[1,2] (**Fig. 1**). CERPS integrates LLNL's engineering and business systems and links them to multiple development partners to speed all aspects of the engineering development cycle (design, procurement, fabrication, testing, and evaluation) independent of geographic location. Zephyr specifically focuses on speeding engineering procurements.

The Zephyr concept builds upon the ready availability of computer networks and new browser software technology, the Internet and the World-Wide Web. Engineering's computer-aided design systems are integrated with institutional business systems that support electronic commerce, linking procurement and

finance, commercial vendors, and end-users. Zephyr's systems integration and workflow facilitate a very rapid procurement and delivery cycle that meets all DOE purchasing requirements. With centralized, sharable data sources, and Zephyr's e-mail workflow process, people in LLNL's programmatic, engineering, procurement and finance organizations and our development partners are assured of accurate and up-to-date drawings, manufacturing models, and financial information.

Zephyr's workflow process (**Fig. 2**) is initiated upon a user request to procure engineering goods or services once a design package is approved for engineering release. Soon after engineering design release, a buyer prepares the electronic request for quotation (RFQ) package. Then Zephyr's workflow process e-mails a vendor announcement indicating the presence of the RFQ in a secure section of the Zephyr World-Wide Web "home page." After the announcement, solicitations for bid-quote, award,

technical data exchange, payment, delivery tracking, and record-keeping are provided by Zephyr. At all transactional levels, the benefits of compressed cycle time, ease of use, and secure business information are realized.

RSA encryption through the use of Netscape's Commerce Server allows secure exchange of engineering drawings, models, specifications, project planning information, bid-quote procurement, and corporate credit card information. Computer databases provide the record-keeping and audit trails required by LLNL programs, industrial partners, and DOE.

The Zephyr public home page is at http://zephyr.llnl.gov/.

Progress

One of Zephyr's early project milestones was the first meeting of LLNL's Engineering and Commerce on the Internet (ECI) Working Group. This working
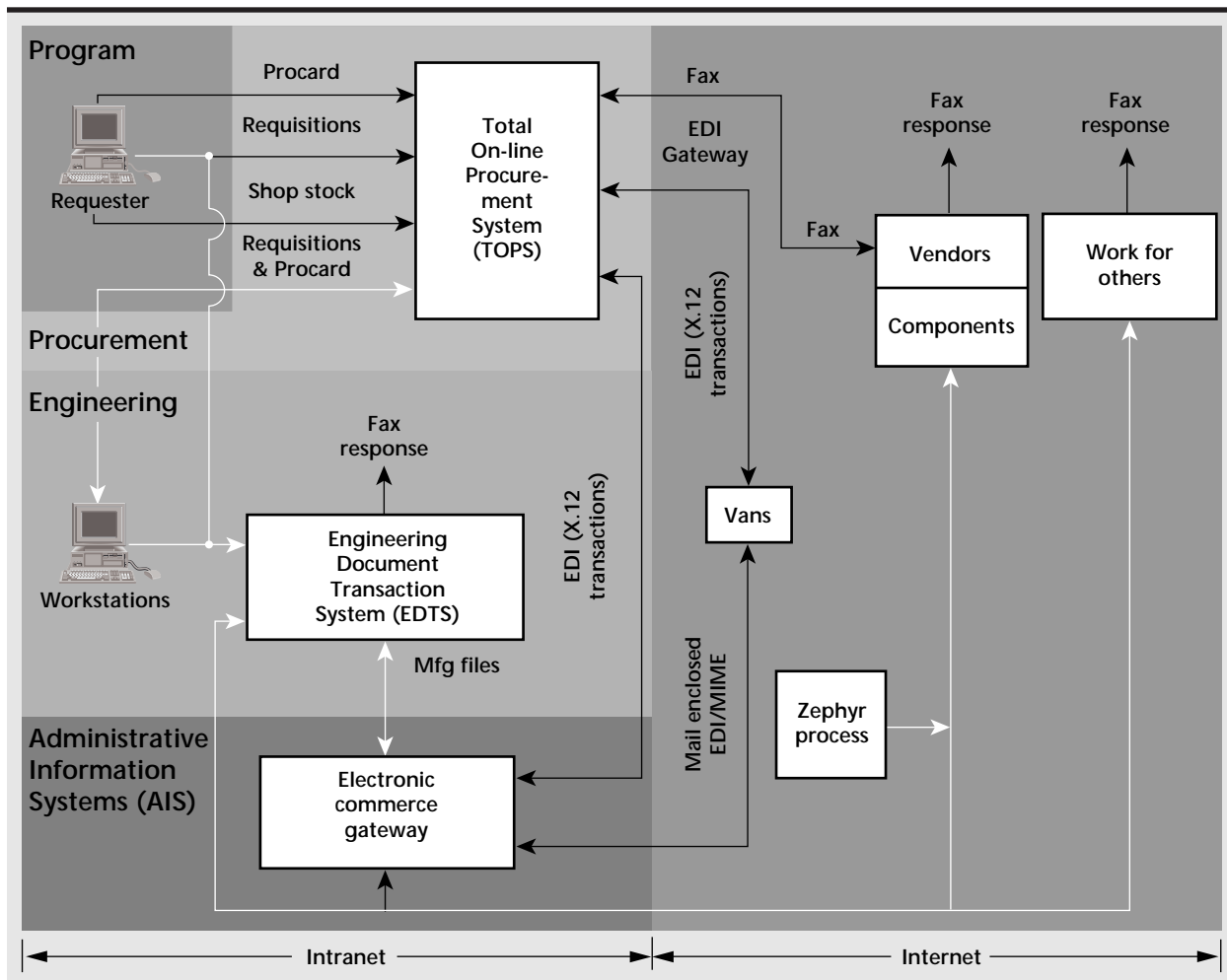

Figure 1. Concurrent Engineering Rapid Prototyping System (CERPS). CERPS integrates LLNL's Engineering Document Transaction System (EDTS) and business systems Total On-line Procurement Systems (TOPS) and Electronic commerce gateway, linking them to multiple development partners to speed all aspects of the engineering development cycle.

group consisted of representatives from Procurement, Engineering, Computations, Administrative Information Systems, and the Technical Information Department.[3] We reached an agreement to work together to promote Internet-based approaches to lower LLNL's cost of doing business. Corporate-wide "working group" agreements are critical to implementing institutional cultural change. Without such agreements, many excellent technical solutions lose out to "time honored" techniques.[4]

The Zephyr project has received three years of funding through DOE. It is built on a decade of work funded by the Department of Defense (DoD). Projects that have contributed to its success include Lawrence Berkeley National Laboratory's Web-based computer-aided design system, the U.S. Air Force's Computer-Aided Acquisition and Logistics Support (CALS) Program, the Air Force's Government Acquisition Through Electronic Commerce (GATEC) System, and the Department of Commerce (DOC) Technology Reinvestment Program's CommerceNet.

Zephyr has been in constant operation since its inception as a prototype in March, 1995, providing immediate access to a wide variety of vendors in many locations across the country.

Metrics

In fiscal years 1995 and 1996, LLNL's engineering programs saved time dramatically in a series of prototype fabrication tests. For example, Internet-based procurement of an assembly fixture dropped to a remarkably low 5.5 days, from a traditional (paper-based) cycle of 56 days, representing a 90% reduction. In other tests, we purchased printed circuit boards in less than two days as opposed to our customary 20 to 25 days. We conducted these tests for procurements of less than $5,000. LLNL's corporate purchasing card, ProCard, can complete the transaction with electronic payment to the vendor in only one or two days, compared with the traditional one or two months.
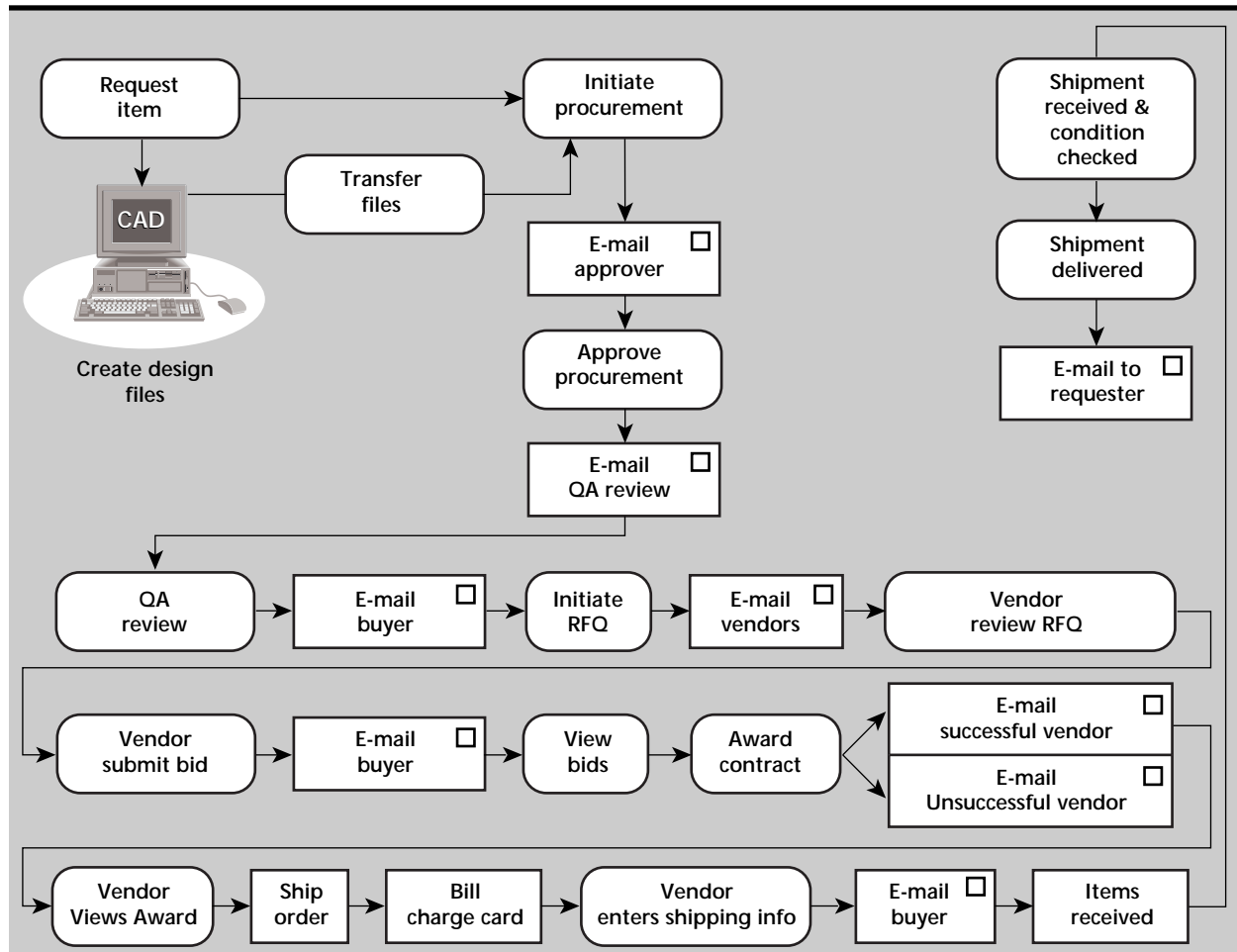

Figure 2. Zephyr's e-mail driven Web-workflow process to speed engineering procurements by communicating with users.

In a recent series of pilot procurements, Zephyr's procurement officer purchased non-standard optical components using Zephyr's paperless bid-quote process and ProCard. The bid package, RFQ, and award-process cycle took less than five days, as opposed to the traditional 30 to 60 days in a paper-based process. Actual procurement processing time was less than 15 minutes, reduced from the typical five to seven days.

Because of ProCard's prompt payment, our procurement officer negotiated "same as cash" discounts amounting to $120,000 in just two months. Zephyr's workflow process allows rapid accrual of such savings.

Zephyr's workflow process has been successfully used to procure engineering goods and services for several projects and programs at LLNL. Our early testing involved the production of a part designed by EG&G in Nevada, giving us a chance to verify that our rapid prototyping process works in collaboration with another DOE site. Zephyr eliminated paper from every step of the design and procurement process while using engineering models to directly manufacture parts.

Future Work

LLNL will continue to improve its internal business and engineering practices. Procurement and Administrative Information Systems will simplify both administrative and technical support systems to gain increased productivity while reducing overall costs. Zephyr's pilot engineering procurement phase will be expanded to explore synergistic ways to complement ongoing programs like the National Ignition Facility (NIF) and Nuclear Weapons Stockpile Stewardship.

LLNL's Zephyr experience is being applied to two far-reaching external initiatives.

In the first initiative, the DOE Nuclear Weapons Complex (NWC), with LLNL participation, is developing new strategies for operations that automate traditional engineering and business processes, while significantly reducing cycle time and costs within. Several DOE projects are tasked with implementing these strategies, including Advanced Manufacturing using National Information Infrastructure (AM-NII).

The NWC's electronic concurrent engineering and business processes will use the Internet to connect the AM-NII national laboratories (LLNL, Sandia National Laboratory, Oak Ridge National Laboratory, Los Alamos National Laboratory, and Allied Signal, Kansas City, Mo.) and their vendors. Given the sensitive nature of the work, the infrastructure must provide strong authentication, non-repudiation, message integrity, and privacy for information being exchanged over the Internet. The technology and strategies for implementing these capabilities over a large community of users has only recently become available.

AM-NII's Public Key Infrastructure (PKI) project will deploy a solution that spans five DOE sites. This solution will enable secure, authenticated exchange
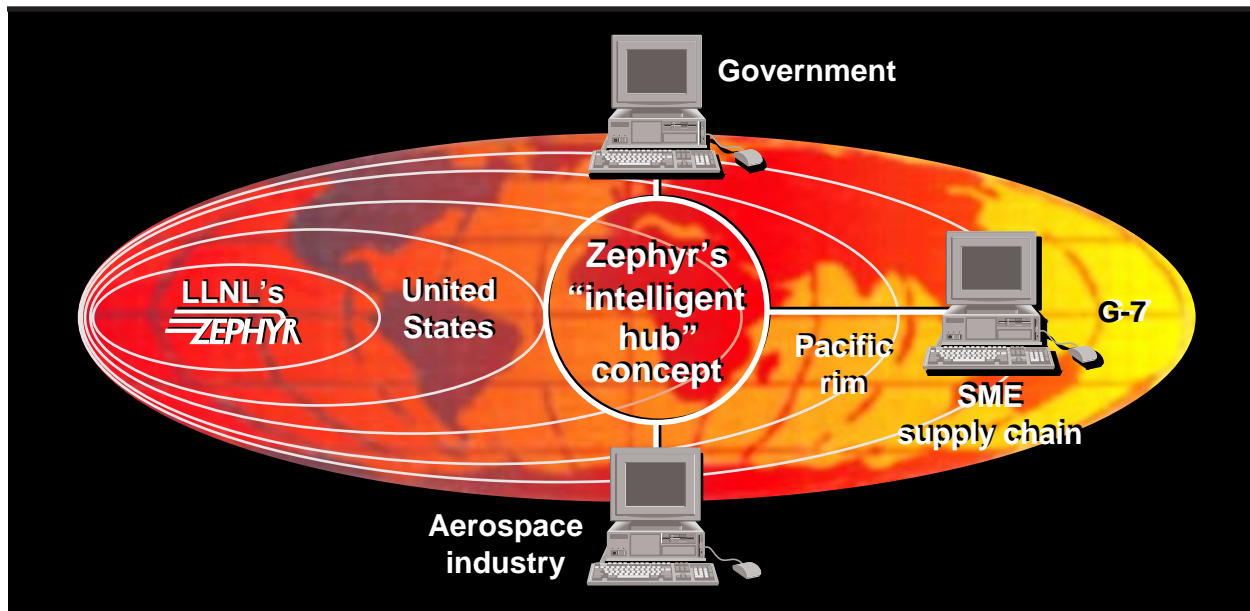

Figure 3. Zephyr's "intelligent hub" concept, linking the government, the aerospace industry, and the aerospace small and medium-sized enterprise supply chain. The goal is the support of ad hoc trading partnerships for an aerospace industry virtual supply network using collaborative engineering practices via the Web.

of e-mail, technical data, electronic forms, and other information required to conduct complex engineering and business processes over the Internet.

In the second initiative, LLNL has joined the CommerceNet-CALS Working Group pilot to develop and test the architecture for an intelligent hub. This server solution will eventually support information and operation of *ad hoc* trading partnerships for an aerospace industry virtual enterprise supply chain (with small and medium-sized enterprises) using collaborative engineering practices via the Web.

The CommerceNet-CALS pilot has received "global economic powers" (G7) endorsement as a potential candidate for world-wide trade and commerce. Zephyr's participation will provide LLNL with experience in collaborative commerce relationships with other members of the CommerceNet-CALS Working Group. It will also provide a way to access an additional base of SME vendors that will be important to growing programs such as NIF and NWC.

Acknowledgments

The authors wish to thank L. Gottlieb, L. Reed, and LLNL's Technical Information Department for communications support (for example, home page development and graphic design); S. Hawkins for demonstrating solid modeling collaborative exchanges between LLNL, EG&G/North Las Vegas, Allied Signal/Kansas City, and commercial vendors; W. Edmonson (recently deceased) for beta-testing ProCard and Unix Systems Management; B. Weber for code supporting file transfers to LLNL's Engineering Records Center; J. Miller for our printer/plotter capabilities; and R. Moore for benchmarking CERPS network performance with ISDN using IBM/PC clones from a remote location.

References

1. Jordan, C. W., W. A. Niven, R. E. Cavitt, G. Pavel, G. Obegi, F. E. Warren, C. J. Buchanan, and J. M. Taylor (1996), *Using the World-Wide Web as a Medium for Concurrent Engineering and Rapid Prototyping*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-ID-122242).

2. Carter, D. E., and B. S. Baker (1992), *CE Concurrent Engineering, The Product Development Environment for the 1990's*, Addison-Wesley Publishing Company, Reading, Mass.

3. Palmer, C. (1996), *Bringing Design and Procurement Together*, Mechanical Engineering, Lawrence Livermore National Laboratory, Livermore, Calif, April.

4. Walton, R. E. (1989), *Up and Running: Integrating Information Technology and the Organization*, Harvard Business School Press, Boston, Mass.

Subcarrier Multiplexing: Optical Network Demonstrations

Paul D. Sargis and Bruce D. Henderer
Defense Sciences Engineering Division
Electronics Engineering

We have increased the channel capacity of our fiber-optic transport technology. We have also conducted field demonstrations on commercially-installed fiber to prove its viability and to discover its limitations. Our computer model now includes the effects of noise and cross-talk to better simulate network performance.

Introduction

Subcarrier multiplexing (SCM) is a multi-channel, multi-gigabit per channel transport technology that breaks the traditional bit rate $\times$ distance barrier of conventional multi-gigabit, long-haul fiber-optic transmission systems. Hybrid transmission systems using SCM and wavelength-division multiplexing (WDM) technology may allow for 500 Gbit/s-plus data transmission over long distances (1000 km).

In FY-95, we introduced our phase one prototype of SCM, which was capable of carrying two 2.5 Gbit/s data streams. Most of our efforts were then focused on studying SCM's inherent dispersion-reduction feature using laboratory demonstrations and computer models. We showed that our dispersion model was in close agreement with laboratory measurements.[1]

When we realized that SCM could have a revolutionary impact on long-distance telecommunications, we considered the issue of technology transfer. We quickly discovered that the telecommunications industry is not impressed with demonstrations which are confined to a sanitized laboratory environment. SCM would have to prove its viability on existing telephone company fiber routes, complete with lossy fiber and poor connections.

Progress

In the past year, our efforts have focused on hardware and software improvements and on network demonstrations.

In our second hardware phase, we designed a four-channel SCM prototype, but were limited to a three-channel implementation because neither of
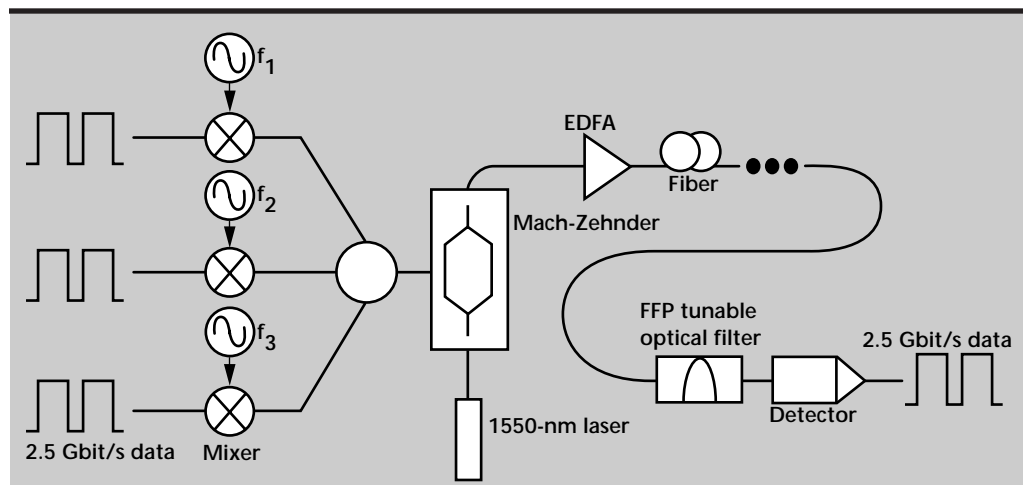

Figure 1. Block diagram of SCM system.

our suppliers of high-bandwidth Mach-Zehnder external modulators was able to deliver a device. Our 7.5 Gbit/s SCM system is diagrammed in **Fig. 1**.

In our second modeling phase, we added the capability to simulate transmission through multiple spans of fiber and multiple erbium-doped fiber amplifiers (EDFAs). Also, we have expanded the model to include the effects of adjacent-channel cross-talk, four-wave mixing, and stimulated Raman scattering. Inclusion of these noise sources and non-linear effects allows us to more accurately model the performance limitations of SCM under a variety of network conditions. The SCM model was created using LabVIEW.

We conducted demonstrations over the National Transparent Optical Network (NTON) around the San Francisco Bay. A map of the network is shown in **Fig. 2**. SCM occupied one of the four wavelengths present on the NTON. The network was configured so that we could loop back the SCM wavelength at San Ramon (100 km round trip) or at Burlingame (500 km round trip).

For our first network demonstration, we wanted to challenge the distance capability of SCM by adding 200 km of spooled fiber to the 500 km round trip over the network. Theoretically, fiber dispersion should begin to cause signal degradation at a fiber length of 640 km.[2] The results of this test are shown in **Fig. 3**. **Figure 3(a)** indicates slight degradation of the waveform due to dispersion. In

contrast, **Fig. 3(b)** shows what happens to the data when the dispersion-reduction enabling component is left out of the SCM receiver.

Next, we used the network to validate our enhanced computer model. The network was configured to loop back the SCM wavelength at San Ramon. We varied the optical attenuation at the SCM receiver to generate bit-error rate (BER) curves. Data was collected on a 2.5 Gbit/s subcarrier by itself, with cross-talk from one other subcarrier, and with cross-talk from two other
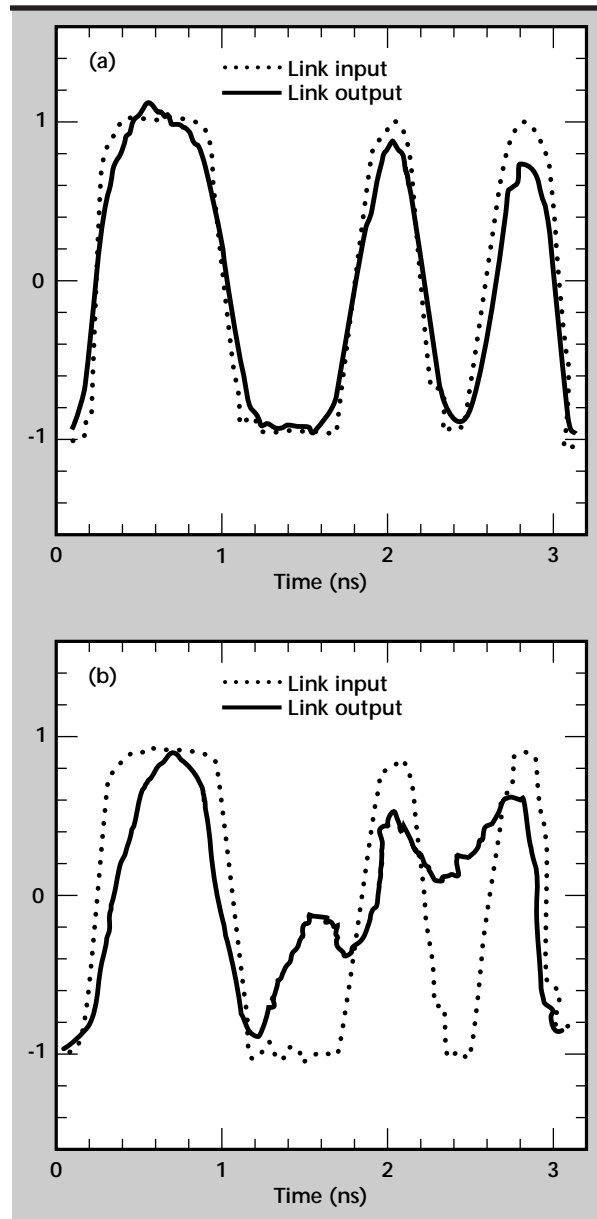

Figure 3. 700-km-SCM link results with (a) dispersion reduction and (b) no dispersion reduction.

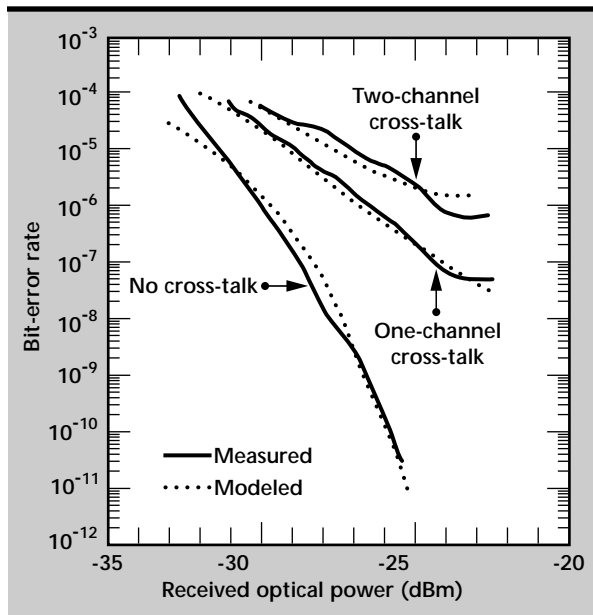*Figure 2. NTON fiber route for SCM demonstrations.*

Figure 4. Results from SCM field trial.

subcarriers. Results from the computer model are compared with measured data in **Fig. 4**. High fiber loss between Livermore and San Ramon caused a low signal-to-noise ratio on the link, resulting in BER noise floors in the data. Minor differences between modeled and measured data can be attributed to measurement error and repeatability of fiber connections.

Representatives from Northern Telecom, GTE Sprint, Pacific Bell, and Uniphase Telecommunications Products witnessed a demonstration of SCM on the NTON. We are in the process of following up on those who expressed an interest in the technology.

Future Work

When we are able to obtain a higher-bandwidth Mach-Zehnder external modulator, we will expand SCM to four 2.5 Gbit/s channels. In addition, we plan to move our SCM receiver to a remote location on the optical network and demonstrate remote channel selection. Also, we will continue to focus on establishing industrial partnerships with telecommunications companies and soliciting interest from government agencies.

References

1. Sargis, P. D., R. E. Haigh, and K. G. McCammon (1995), *Dispersion-reduction technique using subcarrier multiplexing*, SPIE Vol. **2614**, pp. 244–251.

2. Sargis, P. D., R. E. Haigh, and K. G. McCammon (1995), *Subcarrier Multiplexing for Fiber-Optic Networks*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-95).

Parallel Optical Interconnect Technology Demonstrations

Robert F. Hills, Ronald E. Haigh, and James M. Zumstein
Defense Sciences Engineering Division
Electronics Engineering

We report on the development of functional prototype communication and computing systems that use parallel optical interconnect (POI) technology developed under DARPA sponsorship. These prototype systems were developed and tested in field demonstrations with industrial partners to accelerate the adoption of POI technology.

Introduction

Parallel optical interconnect technology uses an array of multimode fiber optic waveguides to transmit light from semiconductor laser diodes, usually vertical surface emitting lasers (VSCELs), to an array of photodiodes that can be monolithically integrated with pre-amplifiers and decision circuitry.[1] The multimode ribbon fiber and an array connector simultaneously interconnect all of the fiber waveguides in a single ribbon connector.

The size of the multimode fiber, usually 62.5 μm, allows for low cost, passive alignment packaging techniques. High levels of device integration allow compact, efficient optoelectronic transceiver modules to be realized.[2] Typical channel bandwidths in a POI module are 1 to 2 Gbits/s. The information capacity of a single POI module can approach several Gbytes/s.

The adoption of high-performance POI technology is essential to enabling high-performance communication and computing systems envisioned by several programs in the national interest; for example, Joint Strike Fighter (JSF) and the Accelerated Strategic Computing Initiative (ASCI).

For ASCI, microprocessors are seen as the building blocks of high-performance, parallel processor computer systems. As the number of processors on a system bus, and the instruction rate of each processor increases, electronic buses interconnecting these systems become a transaction-rate-limiting bottleneck for increasing computational performance. Optical interconnects will allow new bus interconnect fabrics to eliminate this bottleneck and will supersede the traditional electronic buses found in today's computers. An optical interconnect eases the distance constraints between parallel processing nodes, and removes many of the limitations of copper interconnects at high data rates.

Functional prototypes of optically interconnected parallel processing and communication systems were co-developed with industry during this project. The POI technology was demonstrated with Lockheed-Martin under the DoD JSF technology development and maturation program, FORE Systems, with a prototype medical communications network at the Mayo Clinic, and a workstation cluster with Dolphin Interconnect Solutions. We have worked jointly with the DARPA-sponsored Optoelectronic Technology Consortium (OETC), as a neutral partner, to demonstrate optical interconnects in industry, provide testing and evaluation, and enable industry device developments.

Progress

The POIs used in the technology demonstrations with industry (**Fig. 1**) are 32 channels wide with a total aggregate communication bandwidth of

Figure 1. 32-channel OETC Parallel Optical Interconnect transceiver modules.

2 Gbytes/s. The devices were developed by the OETC, a DARPA-sponsored consortium of IBM, Honeywell, Lockheed-Martin, and AT&T. The links were produced with two electrical interfaces, ECL and low-voltage differential signals (LVDS). LVDS is an IEEE standard (1596.3-1994), and specifies a low-voltage differential signal (250 mV swing) that is compatible with CMOS, BiCMOS, and GaAs circuitry. Bit-error tests on the LVDS POI's yielded a bit error rate of better than 10^{-15}, which is acceptable for most applications. Since no errors were detected, the error rate may be better than this.

Three cooperative demonstrations with industry were the primary deliverables of this project. The necessary hardware to interface the POIs to the various applications was developed at LLNL. Each demonstration had unique system requirements, electrical interfaces, and transmission protocol. A common interface to POI was developed to allow the interface to be re-used in multiple demonstrations.

Figure 2. Parallel Optical Interconnect with 50 m of fiber optic ribbon cable configured for insertion into Joint Strike Fighter demonstration.

Application-specific interfaces were developed and served as the glue logic between the interconnect and the custom applications environments.

DoD JSF Technology Insertion

The POIs were integrated into an SCI-based parallel processing system developed by Lockheed-Martin for JSF avionic applications.

In this demonstration the POIs were integrated into the SCI backplane to allow parallel optical data transmission of 18 data and control bits at data rates of 250 Mbits/s per channel. POIs have application in JSF avionics by allowing processing and sensor resources to be distributed strategically across the air frame. During this demonstration, critical processing resources were physically isolated by 50 m of fiber optic ribbon cable with no degradation in computing performance (see **Fig. 2**).

ATM Switch Transceiver Integration

The POIs were integrated into a commercial ATM communications switch developed by FORE Systems. The ATM backbone at the Mayo Clinic served as the site for the field trial. Essential telemedicine services are being developed by researchers at the Mayo clinic to allow physicians real-time access to digitized medical records, for image enhancements, and to enable distributed collaborations and consultation between physicians.

Due to the nature of the ATM protocol and the transmission characteristics of the POIs, an encoder/decoder circuit was developed to allow multiple ATM streams from asynchronous systems, to be transmitted across the interconnects. The

Figure 3. ATM Manchester encode/decode circuitry.
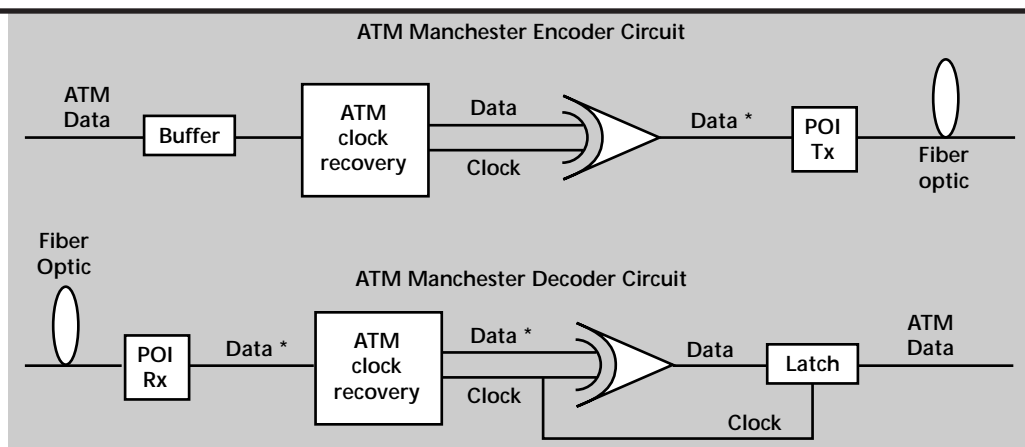
circuit Manchester encodes the data so that a data transition occurs at least once every bit period, and a complimentary decode circuit recovers the original data (**Fig. 3**).

Optically Interconnected Computing Cluster

An optically interconnected computing cluster has been assembled at LLNL. The cluster consists of two SPARC workstations with SBus to SCI (Scalable Coherent Interface) adapter cards developed by Dolphin Interconnect Solutions. The SCI protocol allows applications to allocate and share memory across workstations, thus creating a powerful multiprocessing environment with commercially available workstations, a computing paradigm commonly referred to as cluster computing. The integration of POI technology into the physical layer allows these workstations to be distributed among offices and buildings, and across small business complexes. Logic translation from ECL to PECL was successfully implemented to interface the POIs to the SPARC adapter cards.

Future software engineering will reduce the communications overhead in the operating system, device drivers, and protocol stacks. Next generation silicon will push the interconnect bandwidth to beyond 500 Mbytes/s. Applications are being ported to this environment and we expect the optically interconnected computing cluster to be showcased at Supercomputing '96.

Future Work

We are continuing this work by developing first generation wave division multiplexing (WDM) POI technology. WDM transceivers enable many channels of data to be transmitted on a single optical fiber. Unique WDM technologies developed at LLNL will enable many interconnect fabrics that are not possible with copper interconnect. It is expected that these novel fabrics will be useful in advanced communication and computing systems.[3]

The objectives of these technology demonstrations were to accelerate the adoption and insertion of optical interconnect technology into the future communication and computing systems. Market acceptance and development, and insertion of this technology into COTS components are critical to the missions of DOE ASCI and DOD JSF programs and others.

References

1. Choi, J., D. L. Rogers, D. M. Kuchta, F. H. Kwark, H. Ainspan, K. G. Stawiasz, and J. D. Crow (1996), "High-performance high-yield, uniform 32 channel optical receiver array," *Optical Fiber Conference (OFC) '96, Maui, Hi.,* **FC4**, October.

2. Wong, Y.-M., D. J. Muehlner, C. C. Faudskar, D. B. Buchholz, M. Fishteyn, J. L. Brandner, W. J. Parzygnat, *et al.* (1995), "Technology development of a high-density 32-channel 16-Gb/s optical data link for optical interconnect applications for the optoelectronic technology consortium (OETC)," *IEEE Journal of Lightwave Technology,* **13**, pp. 995–1016.

3. DeGroot, A. J., R. J. Deri, R. E. Haigh, F. G. Patterson, and S. P. DiJaili (1996), "High-Performance Parallel Processors Based on Star-Coupled WDM Optical Interconnects," *1996 Conference on Massively Parallel Processing with Optical Interconnects, Maui, Hi.*

Intelligent Automation Architecture

Scott A. Couture
Applied Research Engineering Division
Mechanical Engineering

This short-term exploratory project focused on identifying an expedient and efficient mechanism to encourage collaboration among a wide variety of Lawrence Livermore National Laboratory (LLNL) researchers working on various aspects of intelligent computing machines. The primary project goals were to identify the set of applicable competencies at LLNL, review the literature in the field, and determine the most practical way to merge the results to encourage collaboration. A careful review of the subsumption architecture, coupled with a review of critical theoretical work in the fundamentals of system complexity lead to the conclusion that the subsumption architecture differs in focus and distribution of system complexity, but not in the overall level of complexity required to achieve intelligent behavior. This facilitates less rigorous approaches to collaboration and will allow the integration of a wider set of competencies than would be practical otherwise.

Introduction

Constructing machines that act intelligently has been a pursuit of researchers around the world for many years. Expert systems have been used to capture rules and knowledge that human experts use when solving problems in specific areas. Research in artificial intelligence and reasoning has been used to develop "intelligent" planning systems. However, little of this work has made it into practical use for real-time control of autonomous systems. Currently available controller technology is insufficient to build "intelligent agents" for even simple tasks in the real, unstructured world.

A number of researchers at LLNL have been advancing the state of the art in many key technologies needed to develop more intelligent machines. These include machine controls; high-speed collision avoidance and real-time path planning; high-speed 3-D image processing; neural nets; fuzzy logic; task planning; parallel processing and supercomputing; intelligent sensor development; data representation; and data mining. This research has been supported by many different sponsors and has been applied to many different programs. One of the key missing pieces has been a unifying architecture or "backbone" that facilitates integration of this wide variety of competencies. Determining an expedient and effective method to develop that backbone is the objective of this research.

In the short term, we plan to identify LLNL researchers in all the key areas needed to pursue development of an interactive intelligent computer system, to identify the state of the art in intelligent machine architectures, and to select an architecture that will facilitate the researchers' collaboration toward implementing intelligent behavior in a particular application. In the longer term, our goal is to develop a uniform computing paradigm for autonomous systems that allows tasks to be described in a high-level goal definition language suitable for use in an industrial environment, while incorporating advanced control features being developed at LLNL and other research institutes.

The definition of the architecture and interfaces will help foster collaboration across diverse research groups and facilitate integration of different research results. Hopefully, the process will also serve to kindle increased interest in intelligent systems and spawn follow-on work to implement a core intelligent "backbone."

Progress

Problem Definition

The lack of widespread use of traditional artificial intelligence approaches in real-time control is primarily due to their inefficiency and their failure to meet the fast reaction times needed for real-time

systems. Reaction times suffer from the vertical (or hierarchical) control approach that is the basis for current machine controllers. Passing data up and down through control layers requires an incredible amount of vertical bandwidth to communicate input data to the high-level model and output data back to the actuator. Similarly, nearly all of the data interpretation and behavior rules reside at the highest level of the control, leading to very complex context-dependent knowledge representation schemes and rules.

A symptom of this focused complexity is difficulty maintaining real-time performance when reacting to these types of situations. A quick comparison of our computer systems' sensor/processor/actuator bandwidth with some of nature's biological counterparts shows something is awry. The biological systems have much less vertical communication bandwidth and less processing, but perform much more intelligently. Hence, a possible change in approach seems warranted.

In fact, a number of different approaches have been proposed to address the deficiencies in current "intelligent" machines. These include the biologically-inspired subsumption architecture (SA) by Brooks,[1] the Perception-Action model by Kaelbling,[2] a number of other biologically inspired approaches[3,4,5] and a variety of hybrid architectures that blend goal- and behavior-oriented approaches.[6,7] Of these, the biggest stimulus to move away from goal-oriented, plan-based, "good-old-fashioned AI" systems towards behavior-oriented "new AI" was provided by Brook's work in SA at MIT beginning in 1986. Since that time, the SA approach has been widely applied to (primarily) low-level control of autonomous agents in real-world situations.

Behavior-based approaches still dominate the literature and are the focus of many related approaches. Therefore, our work began with developing an understanding of the SA approach, its strengths and weaknesses, and its applicability to LLNL's needs. As it turns out, careful examination of the SA approach as it has evolved over the past ten years reveals telling similarities to other biologically-inspired research. Further, a theoretical understanding of the conservation of complexity principle points to reasons why the architecture yields robust systems in rich environments and also bounds the possible complexity of the "emergent" intelligent behavior.

Subsumption Architecture

SA is a modular, behavior-based approach to building machine controllers to work in the real-world. The inspiration for SA comes, in part, from a study of insect behavior and a realization that even very neurologically simple insects behave "intelligently," or at least appropriately, in the real world. In general, each module in a behavior-based system exploits the physics of the environment as much as possible, describes an interaction between the agent and the environment, and attempts to use emergent behavior whenever possible.[8] As shown in **Fig. 1**, modules have very simple interfaces with other modules and the world. Modules are reactive—they read the state of the world as input, determine an appropriate behavior, and output that behavior. What is appropriate behavior is dependent upon the internal state and sensors measuring the state of the environment. There is no global symbolic representation of the world or plan.

Figure 1. Basic behavior module arrangement in Brook's subsumption model.
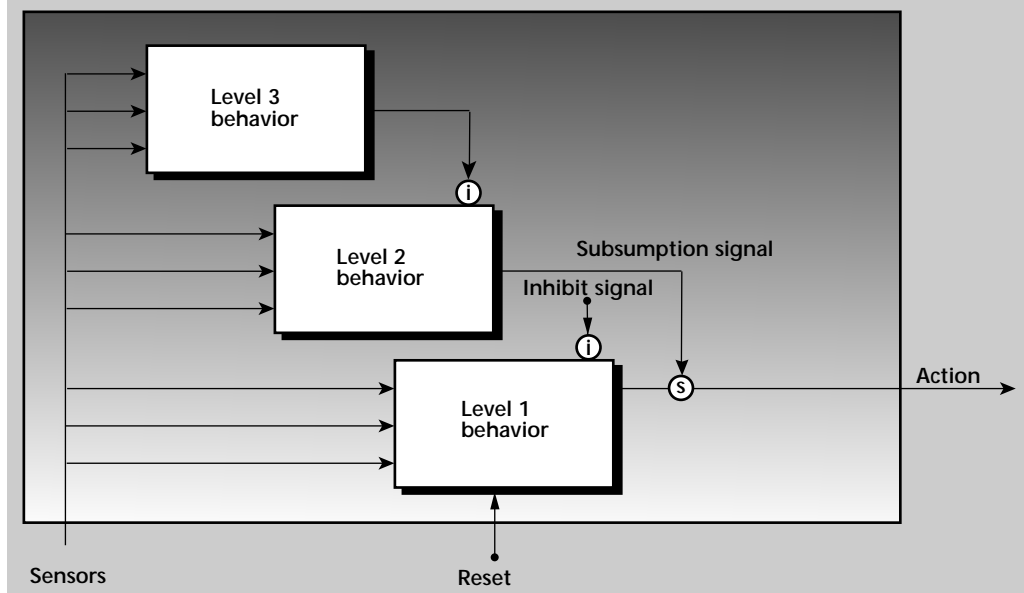
Behavior modules are implemented as augmented finite-state-machines (FSMs) with timers. Additional control signals allow the module's behavior to be integrated with other behaviors. The primary control signals are a suppression signal, which causes the output of the module to be suppressed while it is active, and a subsumption signal, which causes the output of the module to be replaced, or subsumed, by the behavior supplied by another module. More complex behaviors build upon simple behaviors. For example, a *walk* behavior may subsume the *stand* behavior's "leg down" action, to allow the insect to walk. Each module implements a specific behavior. Collections of behaviors interacting with the environment determine the agent's collective competence level.

Since Brooks introduced SA in 1986, numerous systems have been built using its basic principles, with applications varying from mobile robots to vision systems and theorem proving.[9–12] Results have shown that when SA is applied to low-level reactive control problems, like collision avoidance during mobile vehicle navigation or wall following (where the physics of the environment can be exploited and sensor information is rich), the approach is more compact (requires less computer code) and generally more robust than monolithic control approaches.

However, a number of shortcomings must be addressed before SA approaches are likely to address higher level, reasoning behaviors.[13] In particular, strict SA approaches have no representation of the world, which will make higher level reasoning based on symbolic knowledge representations difficult. In SA, each layer works independently without consideration of the total strategic plan, so systems are often deadlocked trying the same reaction many times.[14] A number of these drawbacks have been addressed through refinement and extension of the SA approach.

Subsumption Architecture Variants

Gottschalk addressed the deadlock problem that occurs when a reactive SA system is caught between two opposing stimuli using a hybrid architecture that resulted in a system which mimics the brain's instinctive (brain stem), behavioral (limbic system) and goal (cerebral cortex) levels.[3] The behavior level uses situational if/then rules to resolve the deadlock and the goal level to address longer term goals.

One of the factors that inhibits extending the SA approach to higher level reasoning is that the state information about the world is distributed throughout the behaviors and is not visible to other behaviors.

To communicate the information between modules, more elaborate mechanisms will need to be developed than the simple control signals now in use.

One of the current favorite mechanisms to communicate global information in SA systems is the use of hormones—another concept borrowed from the biological community.[15,16] Hormones are global state variables that are influenced by the environment and are responsible for establishing which behaviors are dominant in a particular situation. For example, when sensors determine that a walking surface is uneven, a hormone related to *fear* increases. In response, the *walk* behavior causes the walking rate to slow so the system is more "careful." Researchers have noted similarities between biological systems in panic and fear situations to their mechanical equivalents after hormones have been added.

Another severe criticism of SA is that there is no way to arrive at a compromise solution. When one module inhibits the behavior of another, the concerns of the inhibited module are ignored, so there is no way to arbitrate between them.[17] The inhibit function often undermines the architecture by shutting off low-level survival type behaviors. Instead of forcing one behavior to be dominant and subsume another behavior altogether, an alternative approach researchers have taken is to modify the SA architecture connectors to allow cooperation between behaviors.

The Action Selection Dynamics architecture allows all behaviors to be active at the same time, but applies variable priority (or weighting) to competencies depending on the situation.[18,19] The weightings can also be changed based on reinforcement to accomplish learning, similar to neural nets. With outputs weighted between –1 and 1, one behavior can override another completely, but the system responds in a more continuous way, and the approach allows a sharing of information between modules that is not possible with the SA inhibit signal.

Kaelbling's perception-action model, shown in **Fig. 2**, is also behavior-based inasmuch as it maps input information and agent state to an output action.[2] However, Kaelbling explicitly addresses information in the context of reaching a goal. As more state information becomes available, perception of the state of the environment improves and the action is more specific to the situation at hand. The perception/action model can be implemented in an SA approach, where each perception/action pair is a behavior module, or in a more traditional vertical approach where information acquisition (perception) is separate from action.

Kaelbling's theoretical treatment of information, goals and behaviors applies whether a vertical or horizontal implementation of the perception-action model is pursued. If that is the case, how does "emergent" intelligent behavior arise from a SA approach when it presumably does not in traditional systems?

Conservation of Complexity

Proponents of the SA approach claim that intelligent behavior "emerges" from collections of simple behaviors. How complex the resulting behavior can be is a key question that is rarely answered in application papers. This is an issue that goes to the core of behavior-based approaches—if there is no "emergent" intelligence in a behavior-based approach, then the difference between goal and behavioral approaches is one of structure and focus and not substance.

This issue is addressed in detail by Werner in his theoretical analysis of the complexity of systems.[20] As Kaelbling also points out, agents and their environments must be treated together as a system. The complexity of a task depends in large part on the complexity of the environment that the task is taking place in. For example, if all parts in an assembly are placed in the environment in their correct relationships, an agent does not have to have the internal complexity necessary to figure out their relationships prior to assembly.

The first SA design principle—take advantage of the physics of the environment—is a conscious effort to reduce agent complexity by transferring agent complexity to the environment. If the designer

knows there will always be a line on the floor to follow, the navigation agent can be very simple even if the pattern on the floor is very complex. Given that the assumptions about the physics of the environment always hold true, the agent is competent to behave intelligently whenever it is in such an environment.

The fundamental conservation of complexity principle states that the complexity of an event generated by a set of agents cannot be greater than the complexity of the agents plus the complexity of the influencing environment. Complexity is additive, so the greater the number of agents with distinct roles (competencies), the greater the possible complexity of the emergent behavior. Again, behavior-based approaches that use a broad collection of specific agents can be expected to have fairly complex "emergent" behavior. However, collections of identical agents, as some have proposed, cannot exceed the complexity level of the individual agent unless the complexity arises from information "encoded" in the environment or a plan (program). For some agents, the environment is the program.

Information reduces possibilities and increases abilities of agents. The representation the agent has of the world increases in complexity with increasing information, allowing the agent to generate more complex output. Reactive, behavior-based systems are sensor- and information-rich compared to many of their traditional AI counterparts. The information is distributed, perhaps appropriately, much closer to the behavior that is influenced by the information, leading to faster reaction times than traditional global model based systems.

Figure 2. Division between the perception and action components in Kaelbling's model.

State

Perception

Action
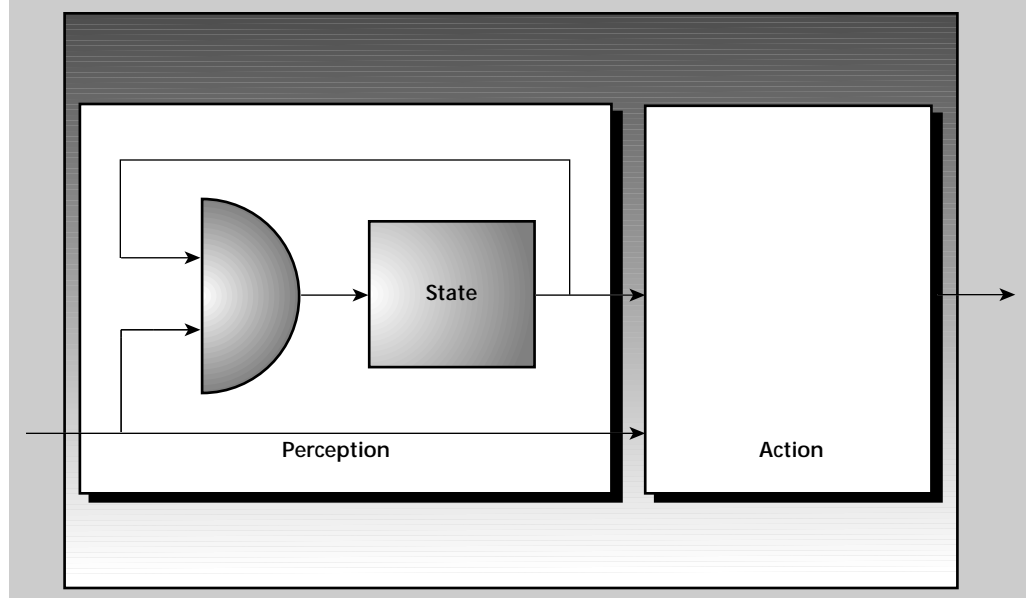
After careful consideration of the conservation of complexity principles, it appears that "emergent" behaviors from SA are intelligent only insofar as the behaviors are designed to properly account for the physics of the environment and the expected interaction between the collection of agents. It is very likely that behavior-based systems can, and often do, exhibit inappropriate, and not-so intelligent behavior, given either poorly defined behaviors, or improper assumptions regarding the environment.

One inherent benefit from the SA approach is the established design discipline that focuses the designer on the environment, the agent's behavior, and the agent's interaction with the uncertainties of the environment. Given the same analysis and the same set of information-gathering tools (sensors), alternate architectures will likely behave as intelligently as SA approaches. In addition, alternative approaches may benefit from a more extensive set of communication and symbolic data representation capabilities than the restrictive set provided by the classic SA approaches.

There is no fundamental reason why higher level symbolic reasoning systems and low-level reactive systems cannot be used together to achieve more complex goals. In fact, according to the conservation of complexity principle, improved performance is certain.

Conclusions

By understanding how evolution has successfully encoded environmental complexity and data interpretation into our biological systems, computer analogs can potentially use similar tricks to avoid a lot of agent complexity. The jury is still out regarding how far up the reasoning chain the SA approach can be taken without encountering serious stumbling blocks. Perhaps current work in biological equivalents will assist in the transition from reactive to symbolic reasoning paradigms when that bridge must be crossed.

Future Work

This exploratory effort was funded during the last part of FY-96 and is complete. A task analysis identified a core set of sensors and behaviors needed to implement an intelligent system in the selected task and environment scenario. Given available funding, there is sufficient interest and competency at LLNL to do further work this area.

Acknowledgments

References

1. Brooks, R. A. (1986), "A Robust Layered Control System for a Mobile Robot," *IEEE Journal of Robotics and Automation*, Vol. RA-**2**, (1).

2. Kaelbling, L. P. (1988), "Goals as Parallel Program Specifications," *Proceedings of the Seventh National Conference on Artificial Intelligence, St. Paul, Minn.*

3. Gottschalk, M. A. (1996), "Intelligent motion controller mimic's brain's structure," *Design News*, March, pp. 119–121.

4. Mura, F., and N. Franceschini (1994), "Visual control of altitude and speed in a flying agent," *Proceedings of the Third International Conference on Simulation of Adaptive Behavior, (From Animals to Animats), Cambridge, Mass.*, **3**, pp. 91–99.

5. Murray, D. (1995), "Developing reactive software agents," *AI Expert,* Vol. **10**, (3), March, pp. 26–29.

6. Innocenti, C., G. Mondino, P. Regis, and G. Sandini (1994), "Trajectory planning and real-time control of an autonomous mobile robot equipped with vision and ultrasonic sensors," *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems (IROS '94), New York, N.Y.*, Vol. **3**, pp. 1861–1866.

7. Lindley, C. A. (1993), "Autonomous satellite architecture integrating deliberative reasoning and behavioural intelligence," *Telematics and Informatics, United Kingdom*, Vol. **10**, (3), pp. 231–250.

8. Steels, L. (1994), "A case study in the behavior-oriented design of autonomous agents," *Proceedings of the Third International Conference on Simulation of Adaptive Behavior, (From Animals to Animats), Cambridge, Mass.*, **3**, pp. 445–452.

9. Stein, M. R., and R. P. Paul (1994), "Operator interaction, for time-delayed teleoperation, with a behavior-based controller," *Proceedings of the 1994 IEEE International Conference on Robotics and Automation, Los Alamitos, Calif.*

10. Gomi, T., and K. Ide (1994), "Vision based navigation for an office messenger robot," *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems, (Advanced Robotic Systems and the Real World), (IROS '94), New York, N.Y.*, Vol. **3**, pp. 2015–2022 .

11. Pinhanez, C. S. (1994), "Behavior-based active vision," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. **8**, (6), pp. 1493–1526.

12. Brady, A. (1994), "A subsumption architecture for theorem proving?" *Philosophical Transactions of the Royal Society, Series A (Physical Sciences and Engineering), United Kingdom,* Vol. **349**, (1689), pp. 71–84.

13. Arnold, J. E. (1989), "Experiences with the subsumption architecture," *Proceedings of The Fifth Conference on Artificial Intelligence Applications, Washington, D.C.*, pp. 93–100.

14. Yamamoto, M. (1994), "'SOZZY': a hormone-driven autonomous vacuum cleaner," *Proceedings of the SPIE—The International Society for Optical Engineering*, Vol. **2058**, pp. 211–223.

15. Brooks, R. A., and P. A. Viola, (1990), "Network based autonomous robot motor control: from hormones to learning," *Advanced Neural Computers, Amsterdam, Netherlands*, pp. 341–348.

16. Gomi, T., and J. Ulvr (1993), "Artificial emotions as emergent phenomena," *Proceedings of 1993 2nd IEEE International Workshop on Robot and Human Communication, New York, N.Y.*, pp. 420–425.

17. Rosenblatt, J. K., and D. W. Payton (1989), "A fine-grained alternative to the subsumption architecture for mobile robot control," *International Joint Conference on Neural Networks, New York, N.Y.*, Vol. **2**, pp. 317–323.

18. Webber, A. D., and D. L. Bisset (1994), "Competition and co-operation-a model for behaviour-based robot controllers," *Proceedings of PerAc '94, (From Perception to Action), Los Alamitos, Calif.*, pp. 384–387.

19. Gomi, T., and K. Ide (1994), "Emulation of emotion using vision with learning," *Proceedings of 1994 3rd IEEE International Workshop on Robot and Human Communication, New York, N.Y.*, pp. 210–215.

20. Werner, E. (1994), "What ants cannot do," *Proceedings of 6th European Workshop on Modelling Autonomous Agents in Multi-Agent World, Distributed Software Agents and Applications*, pp. 19–39.

Author Index