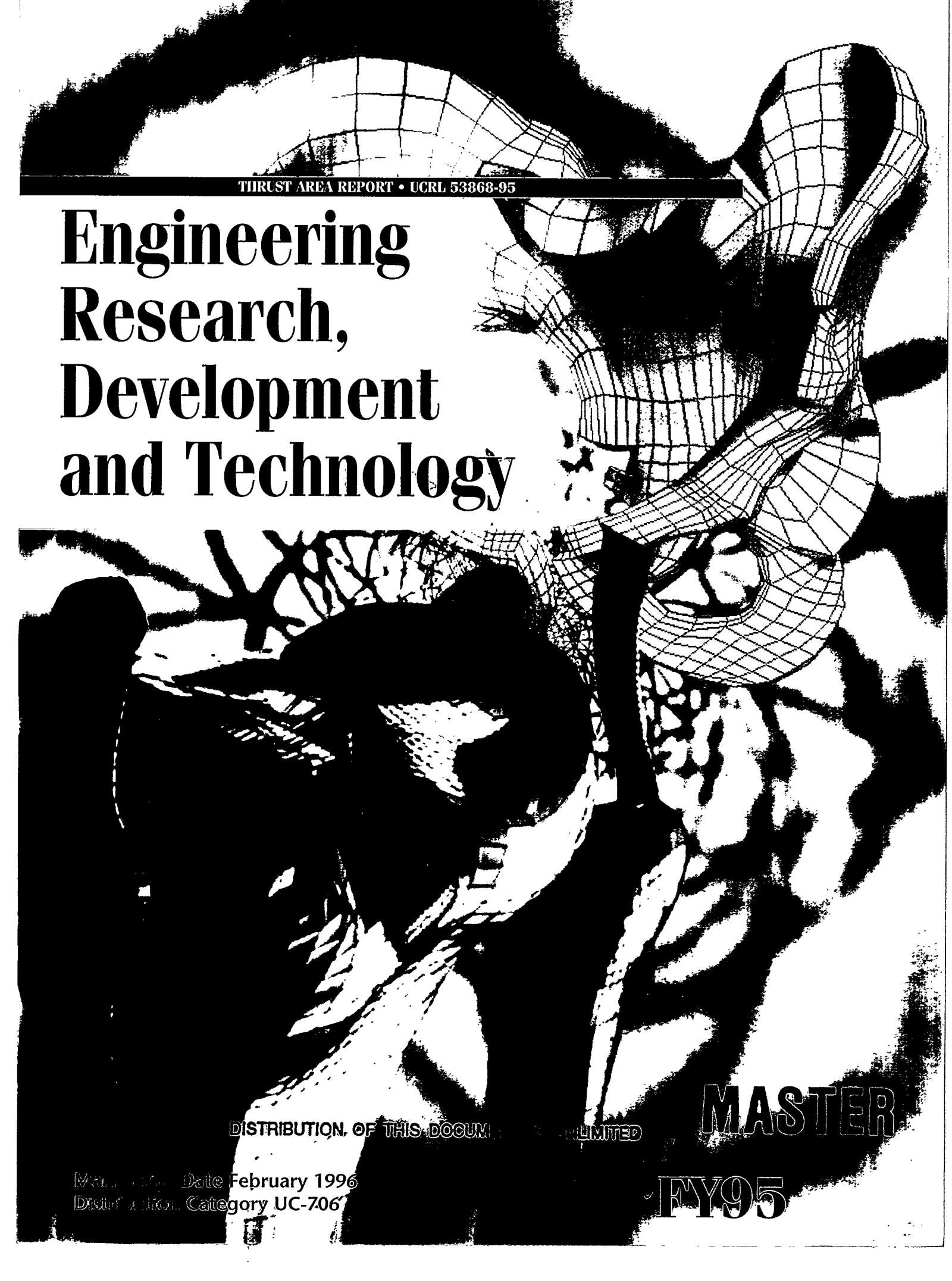


THRUST AREA REPORT • UCRL 53868-95

Engineering Research, Development and Technology

DISTRIBUTION OF THIS DOCUM LIMITED

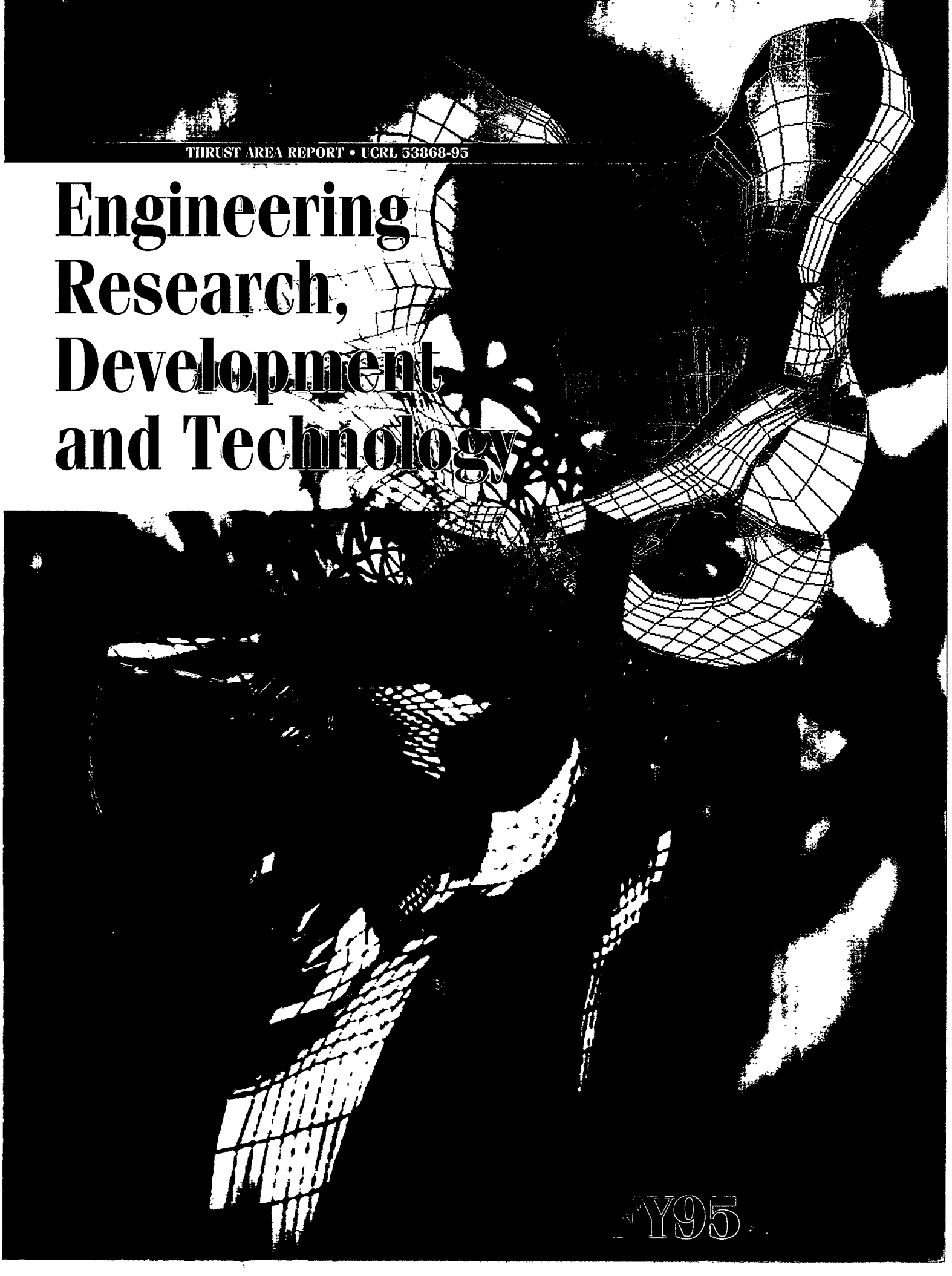
MASTER

Ma Date February 1996
Distri tion Category UC-706

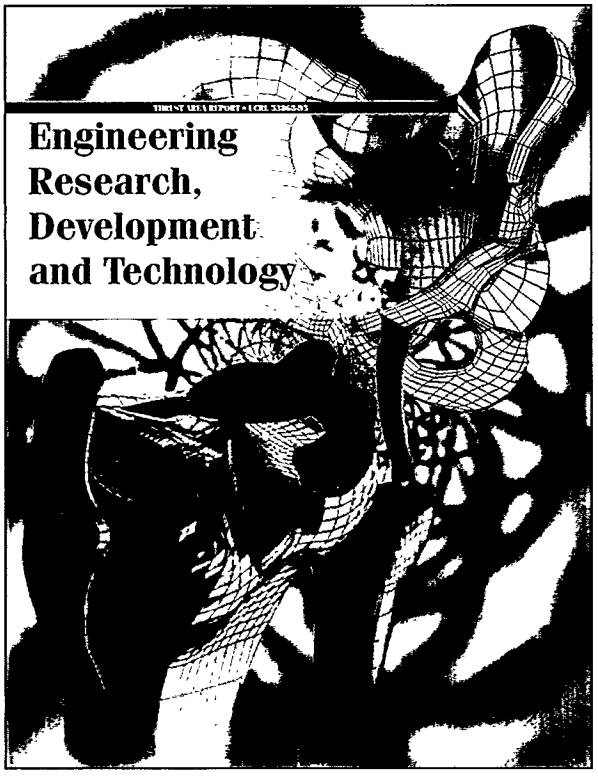
FY95

THRUST AREA REPORT • UCRL 53868-95

Engineering Research, Development and Technology



FY95



On the cover

A finite element mesh for two different human skeletal structures. The left foreground depicts the tibia and fibula, with soft tissue representations of the collateral ligaments on both sides, and the menisci, the anterior and posterior cruciate ligaments on the top. At the right background is a model of the pelvis. The crater-like cavity, located just above and to the right of the circular hole, is the hip socket. (See article, p. 2-1.)

Acknowledgments

Scientific Editing
Robert Langland
Camille Minichino

Composition
Sharlene Markow

Graphic Design
Pamela A. Allen

Art Production/Layout
Jeffrey B. Bonivert
Dennis Chan
Irene J. Chan
Danial A. Fletcher
Penny L. Keenan
Antonio C. Sanchez
Kelly A. Spruiell

Publication Services
Jody A. Johnson

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California and shall not be used for advertising or product endorsement purposes.

Work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Introduction

Spiros Dimolitsas, Associate Director for Engineering

1. Computational Electronics and Electromagnetics

Overview
Clifford C. Shang, Thrust Area Leader

Photonics Modeling and Design
Jeffrey S. Kallman and Richard P. Ratowsky *1-1*

Electro-Mechanical Modeling of Nodular Defects in High-Power Multilayer Optical Coatings
Nicole E. Molau, Clifford C. Shang, Francois Y. Genin, Mark R. Kozlowski, Hal R. Brand, and Richard H. Sawicki *1-7*

Development of an Electromagnetic Electron and Ion Fluid Computer Code
David J. Mayhall, Jick H. Yee, and Dale E. Miller *1-13*

Single-Pass Surface-to-Borehole Reconstructive Imaging
Scott D. Nelson and Erik Johansson *1-19*

Wakefield Modeling of Advanced Accelerator Components
Brian R. Poole, Wang C. Ng, Clifford C. Shang, and George J. Caporaso *1-25*

EIGER: A New Generation of Computational Electromagnetics Tools
Robert M. Sharpe, J. Brian Grant, William A. Johnson, Roy E. Jorgenson, Donald R. Wilton, David R. Jackson, and John W. Rockway *1-33*

2. Computational Mechanics

Overview
Gerald L. Goudreau, Thrust Area Leader

Biomechanical Modeling of the Human Knee
Dale A. Schauer, Jeffrey A. Weiss, Bradley N. Maker, and Michael A. Puso *2-1*

A Hyperelastic Shell Model for Elastomers and Biological Soft Tissues
Michael A. Puso *2-7*

Investigation of the Radiative Properties of Ceramic Casting Materials
James D. Maltby, Barbara T. Kornblum, Charles S. Landram, and Charles W. Dingus *2-13*

Performance of NIKE3D with PVSOLVE on Vector and Parallel Computers
Bradley N. Maker, Qin Jiangning, and Duc T. Nguyen 2-17

GEMINI Re-Certification and Development
Stan Bumpas, Michael A. Gerhard, and Stanley C. Sommer 2-25

Material Model Development for Thermal Forming of Anisotropic Materials
Roger W. Logan 2-31

Multilength-Scale Modeling: Crystal-Plasticity Models in Implicit Finite Element Codes
David D. Sam 2-39

3. Microtechnology

Overview
Raymond P. Mariella, Jr., Thrust Area Leader

Optical Coating, Testing, and Etching for Advanced Semiconductor Optical Amplifier Technology
Sol P. DiJaili, Frank G. Patterson, and Robert J. Deri 3-1

Micromachined Sample Injector for a Portable Hand-Held Gas Chromatograph
Conrad M. Yu 3-3

Fiber-Optic DNA Detector
Laurence R. Brewer and Donald A. Masquelier 3-5

Silicon Dioxide Film Deposition and Characterization
Norman F. Raley and Dino R. Ciarlo 3-7

Interconnections for Microflow Systems
James A. Folta and William J. Benett 3-11

Shape Memory Film Microactuators
Peter A. Krulevitch, Abraham P. Lee, Philip B. Ramsey, and M. Allen Northrup 3-15

Polyimide Electrostatic Microvalve Development
Abraham P. Lee, Julie Hamilton, and Jimmy Trevino 3-19

Microinstrumentation for Analysis of Biological Samples
Raymond P. Mariella Jr., Gerald G. Eveleth, Richard G. Langlois, Donald A. Masquelier, Dino R. Ciarlo, Mukund C. Venkatesh, and Shadi Shakeri 3-21

SMA-Actuated Silicon Microgripper
Abraham P. Lee, Peter A. Krulevitch, Jimmy Trevino, Stacy Lehew, Dino R. Ciarlo, and M. Allen Northrup 3-23

4. Manufacturing Technology

Overview
Kenneth L. Blaedel, Thrust Area Leader

Precision Saw for Slicing Wafers of Crystalline Materials
Bradley L. Fix *4-1*

Reactive Sputter-Deposition of Boron Nitride
Daniel M. Makowiecki, Alan F. Jankowski, Jeffrey P. Hayes, and Mark A. McKernan *4-5*

5. Materials Science and Engineering

Overview
Donald R. Lesuer, Thrust Area Leader

Strength and Fracture Toughness of Material Interfaces
Robert Riddle, David S. Hiromoto, Nhan Nguyen, Wayne King, Geoffrey Campbell, and Walter Wien *5-1*

Dynamic Mechanical Behavior of Fiber Composites
Steve J. DeTeresa and Scott E. Groves *5-7*

Participating Media Radiant Transport at High Temperature: NIF Flashlamps and Investment Casting Molds
Mark A. Havstad and Charles W. Dingus *5-15*

Synthesis and Processing of Nanocrystalline Aluminum and Hydroxyapatite
T. G. Nieh and Donald R. Lesuer *5-21*

Fatigue Behavior of Al Metal Matrix Composite at Elevated Temperatures
Chol K. Syn, Donald R. Lesuer, and David S. Hiromoto *5-27*

Pearlite in Ultrahigh Carbon Steels: Heat Treatments and Mechanical Properties
Eric M. Taleff, Chol K. Syn, Donald R. Lesuer, and Oleg D. Sherby *5-35*

6. Power Conversion Technologies

Overview
Mark A. Newton, Thrust Area Leader

Development of Advanced Modulator Technology for Heavy Ion Recirculators
Mark A. Newton, Hugh C. Kirbie, and Roy L. Hanks *6-1*

Modeling and Testing of Nanostructure Multilayer Capacitors
Gary W. Johnson and Wang C. Ng 6-7

Microradar Development
Stephen G. Azevedo, Thomas E. McEwan, and John P. Warhus 6-17

7. Nondestructive Evaluation

Overview
Harry E. Martz, Thrust Area Leader

Infrared Computed Tomography For Characterizing Structural Defects
Nancy K. Del Grande, Philip F. Durbin, and Dwight E. Perkins 7-1

Techniques to Improve the Sensitivity of Laser-based Ultrasonics
Diane Chinn, Graham Thomas, James Candy, and James Spicer 7-13

Progress in 3-D Quantitative DR/CT
Hal R. Brand, Daniel J. Schneberk, Harry E. Martz, Pierre-Louis Bossart, and Stephen G. Azevedo 7-19

Computed Tomography Software and Camera Upgrade
Hal R. Brand, Daniel J. Schneberk, and Michael R. Gorvad 7-31

Applications of Micropower Impulse Radar to Nondestructive Evaluation
Jeffrey E. Mast and Stephen G. Azevedo 7-39

8. Information Engineering

Overview
David N. Hunt, Thrust Area Leader

Using the World-Wide Web as a Medium for Concurrent Engineering and Rapid Prototyping
Cecil W. Jordan, William A. Niven, Rudy E. Cavitt, George Pavel, Goffril Obegi, Fred E. Warren, C. James Buchanan, and Janine M. Taylor 8-1

Subcarrier Multiplexing for Fiber-Optic Networks
Paul D. Sargis, Ronald E. Haigh, and Kent G. McCammon 8-7

The World-Wide Web as a Medium For Presenting LLNL Training Courses
Darrel Lager and Sue Koopman 8-11

Design and Implementation of 3-D MIME Data Access and Exchange
Scott D. Nelson, Jeffrey S. Kallman, and Eric Brugger 8-17

Author Index

Author Index.....*AI-1*

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.

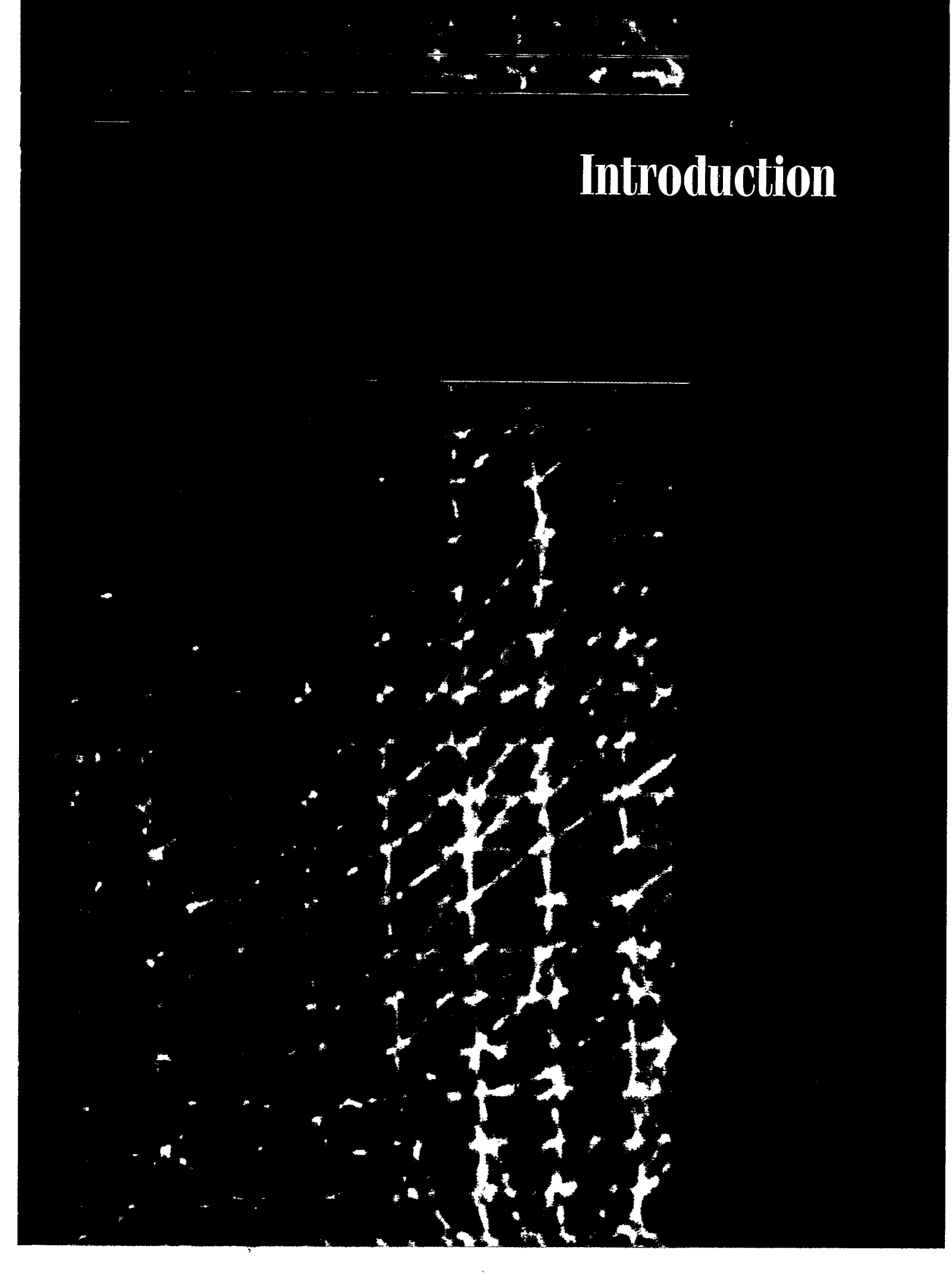
The thrust area leaders are accountable to me for the quality, programmatic relevance and progress of their activities, but have sufficient latitude to manage resources allocated to them. Thrust area leaders are also expected to establish strong links with LLNL programs and industry, be knowledgeable 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 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 1995. 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 is a focal point for computer modeling activities in electronics and electromagnetics for engineering at Lawrence Livermore National Laboratory (LLNL). Traditionally, we have focused our efforts in technical areas of importance to existing and developing LLNL programs, and this continues to form the basis for much of our research. Another important emphasis for the thrust area is the formation of partnerships and the application of our simulation technology and expertise to the solution of problems faced by industry.

Our FY-95 work is described in the six reports that follow: 1) Photonics Modeling and Design; 2) Electro-Mechanical Modeling of Nodular Defects in High-Power Multilayer Optical Coatings; 3) Development of an Electromagnetic Electron and Ion Fluid Computer Code; 4) Single-Pass Surface-to-Borehole Reconstructive Imaging; 5) Wakefield Modeling of Advanced Accelerator Components; and 6) EIGER: A New Generation of Computational Electromagnetics Tools.

Computational Electronics and Electromagnetics

1

1. Computational Electronics and Electromagnetics

Overview
Clifford C. Shang, Thrust Area Leader

Photonics Modeling and Design
Jeffrey S. Kallman and Richard P. Ratowsky 1-1

Electro-Mechanical Modeling of Nodular Defects in High-Power Multilayer Optical Coatings
Nicole E. Molau, Clifford C. Shang, Francois Y. Genin, Mark R. Kozlowski, Hal R. Brand, and Richard H. Sawicki 1-7

Development of an Electromagnetic Electron and Ion Fluid Computer Code
David J. Mayhall, Jick H. Yee, and Dale E. Miller 1-13

Single-Pass Surface-to-Borehole Reconstructive Imaging
Scott D. Nelson and Erik Johansson 1-19

Wakefield Modeling of Advanced Accelerator Components
Brian R. Poole, Wang C. Ng, Clifford C. Shang, and George J. Caporaso 1-25

EIGER: A New Generation of Computational Electromagnetics Tools
Robert M. Sharpe, J. Brian Grant, William A. Johnson, Roy E. Jorgenson, Donald R. Wilton, David R. Jackson, and John W. Rockway 1-33

Photonics Modeling and Design

Jeffrey S. Kallman and Richard P. Ratowsky
Engineering Research Division
Electronics Engineering

In FY-95 we developed design codes for ball lens mode converters, two-dimensional integrated optic waveguide structures, and anisotropic devices. We also continued our work on hybrid modeling codes.

Introduction

Photonic devices promise to be very important to the future of high speed communications and computation. Signals in photonic networks can travel long distances at the speed of light, with very little power loss. These networks are made up of fibers, waveguides, sources, receivers, converters, and a host of other devices. All of these devices must be carefully designed if they are to work well together. This is a problem, because these devices are difficult to design and expensive to fabricate. Rather than fabricate every design, it is better to simulate them first, to see if they meet the required specifications.

We are actively pursuing the generation and use of tools for the design and modeling of photonic devices. We have already developed a number of tools, such as BEEMER, TSARLITE, MPM, CYLINDER, and GREAT_CIRCLE, and we continued this process in FY-95. This year, our efforts were spent on implementing and refining a number of simulation techniques: quasi-analytic modeling of simplified geometries, finite difference beam propagation techniques, and hybrid modeling codes. We also added electrical anisotropy to our finite difference time domain codes. This report will discuss these advances.

Progress

Hybrid-Space Code

As the field of integrated optics progresses, the optical devices become larger and more sophisticated. In prior years our two-dimensional simulation tools were able to deal with two modeling extremes. Beam propagation method (BPM) simulators could handle large devices (thousands of wavelengths on a side), as long as the light was propagating in a single direction and there were no large changes in index of refraction; and finite difference time domain (FDTD) simulators could accurately simulate light in any linear device, as long as it was small (tens of wavelengths on a side). We were unable to simulate any device that was hundreds of wavelengths in length with sub wavelength-sized features, or involving reflections. Since we wished to simulate such devices, a means of hybridizing these methods was required.

We started development of a hybrid code to integrate these methods in FY-94, and completed the code this fiscal year. The code uses a BPM where it is applicable, and a finite difference technique where it is necessary. The code is graphical-user-interface (GUI)-driven and can simulate two-dimensional devices that can be modeled using a single FDTD section between two BPM sections, for instance, a stubbed taper (**Fig. 1**).

The most significant challenge in hybridizing these simulation methods was treating the interfaces between the methods. The interface encountered going from the BPM section to the FDTD section required decomposing the output field from the BPM section into modes of the FDTD section. The main reason for this is the requirement of a one-way launch of light into the FDTD section, in case we want to look at reflections. The launching of a unidirectional wave requires matching the propagation constant and grid dispersion of every mode of interest in the input waveform. The interface encountered going from the FDTD section to the BPM section was much simpler, and only required time integration of a column of FDTD grid points to obtain amplitude and phase values that could be fed to the BPM section.

Finite Difference Beam Propagation

As previously mentioned, one of the techniques we have been using for a number of years is the BPM. Since its origination, a wide variety of propagation techniques have been developed. This year we implemented a finite difference beam propagation method (FDBPM) and integrated it into the BEEMER design code.

The FDBPM has several advantages over the FFT-based BPM previously used. The primary advantages are speed and boundary conditions. The FFT-based BPM algorithms take an amount of time proportional to $n \ln(n)$ for each propagation step, where n is the number of waveform samples, while the FDBPM algorithm takes time proportional to n for each step. FFT BPM algorithms have a hard time dealing with radiation that leaves the problem boundaries, while FDBPM algorithms can be used with "transparent" boundary conditions[1] (**Fig. 2**).

Figure 2 depicts a Gaussian beam launched onto a waveguide which sheds radiation until only the waveguide mode remains. In the transparent boundary condition case, the radiation shed by the waveguide passes through the boundaries, while in the absorbing boundary conditions most of the radiation is absorbed, but some reflects back into the problem space. The radiation reflected from the boundaries is an inaccuracy in the simulation, and causes it to yield incorrect results.

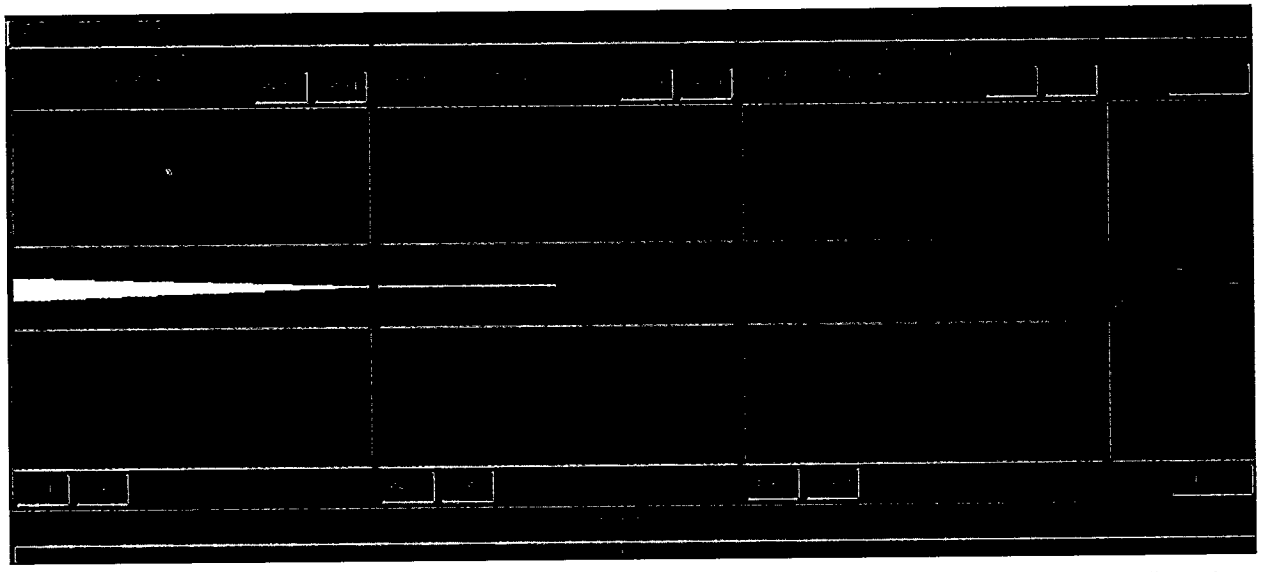


Figure 1. The GUI of our hybrid BPM-FDTD-BPM method, getting set to work on a taper problem. The BPM method handles regions where nearly all the light can be expected to propagate in the forward direction, while the FDTD method handles the region where we expect reflections or other deviations from forward propagation. Note that the aspect ratios of the sections of the display are all very different.

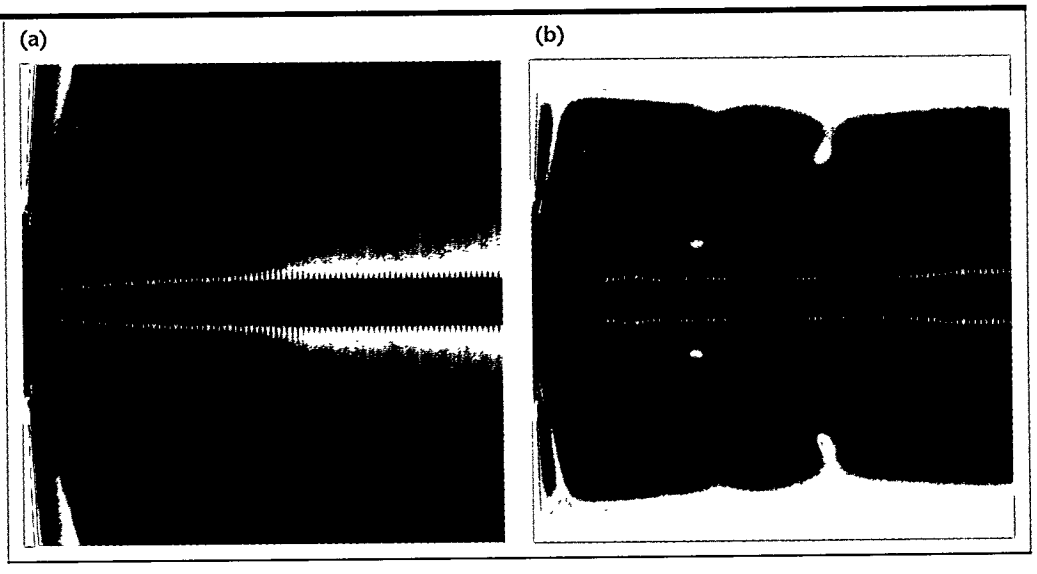


Figure 2. Comparison of (a) the FDTD BPM's transparent boundary conditions with (b) the absorbing boundary conditions of the FFT BPM.

In addition to correctly dealing with exiting radiation, the transparent boundary conditions allow the problem domain to be smaller, which reduces the run time for many problems.

SPHERE

The SPHERE[2] code can be used to calculate propagation of laser light through coated spherical lenses (**Fig. 3**). This code relies on the fact that Maxwell's field equations can be solved exactly for the case of scattering from objects with spherical symmetry. The solution for the scattering of a plane wave from a cylinder or a sphere is well-known, going back to work by Rayleigh and Mie[3]. Since we are interested in propagating localized laser light, we wish to propagate a beam rather than a plane wave. Because Maxwell's equations are linear, we can expand a Gaussian beam (or an arbitrary beam profile) in a basis of plane waves, and calculate the scattering from each plane wave separately. The code represents the solution inside the lens, inside the coating, and outside the lens in terms of vector spherical harmonics. Satisfying the boundary conditions leads to specification of the field everywhere. An averaging procedure is used to account for the imperfections in a real lens not present in our idealized model. This code is quasi-analytical because, although the solution can be expressed explicitly as a sum over special functions, the sum involves a delicate numerical evaluation of a potentially large number of terms.

SPHERE is a powerful tool because it can accurately handle lenses on the scale of hundreds of wavelengths, e.g., commercially available ball lenses, where purely numerical codes such as FDTD are useless. And, because no paraxial expansion is used, we can accurately represent a wavelength-scale Gaussian beam for modeling a laser diode (LD) spot. All of the scattering physics is included, i.e., polarization and coherent effects due to multiple reflections. The principal limitation of this code is its inability to treat departures from spherical symmetry of the lens, although this may be included perturbatively at a later point. The code can also be generalized to include other separable geometries to model a wider class of aspheric lenses.

SPHERE was benchmarked[4,5] and was used this past year for design work. We used SPHERE to model the back reflection into an LD from a ball lens, as well as the coupling efficiency for a ball lens between an astigmatic LD and a single mode optical fiber (SMF) **Fig. 4**. The latter included modeling the effect of offsets of the LD and SMF from the optical axis. We also used SPHERE to optimize LD to SMF coupling for a two-ball lens system. Finally, a GUI was built as a front end for the code (**Fig. 5**).

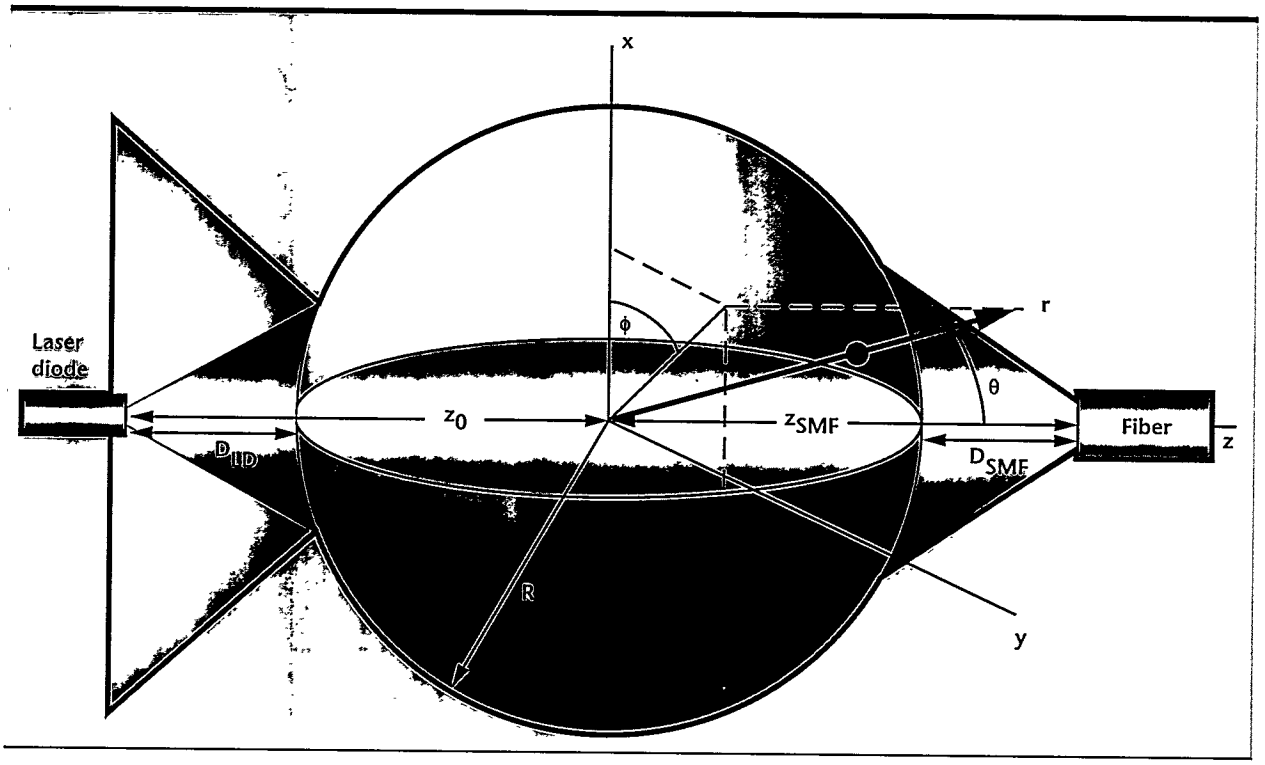
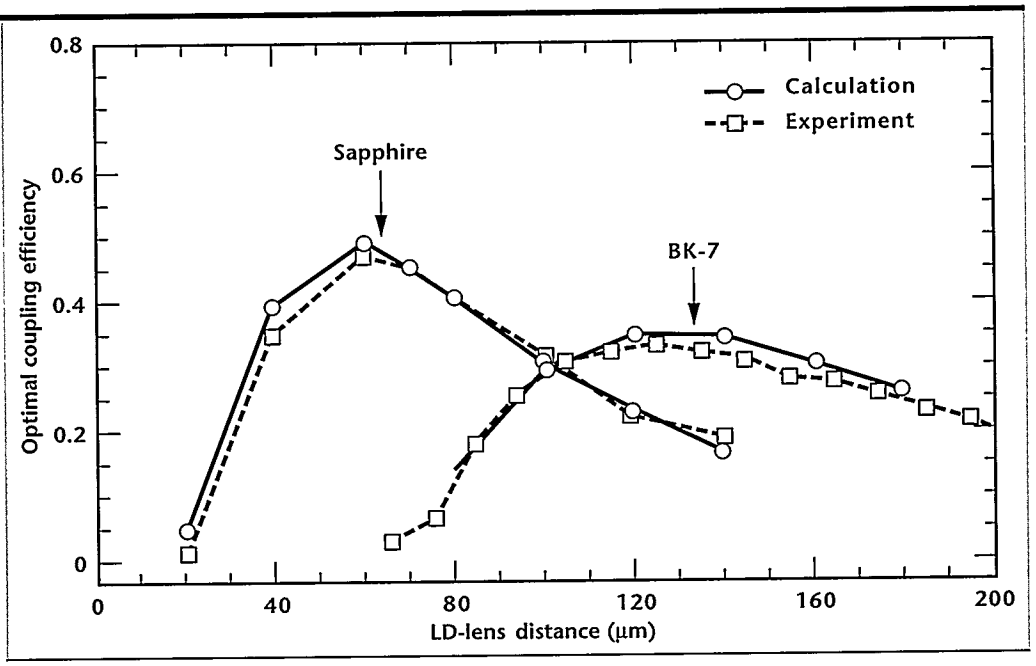


Figure 3. Problem geometry for the SPHERE code. A Gaussian beam from a laser diode is incident on an anti-reflection coated spherical lens, here shown in cross section. The vector electric field can be calculated at any point inside or outside the lens.

Figure 4. The GUI now available for use with the SPHERE code. This GUI makes every parameter in the code available on the front panel of the display.



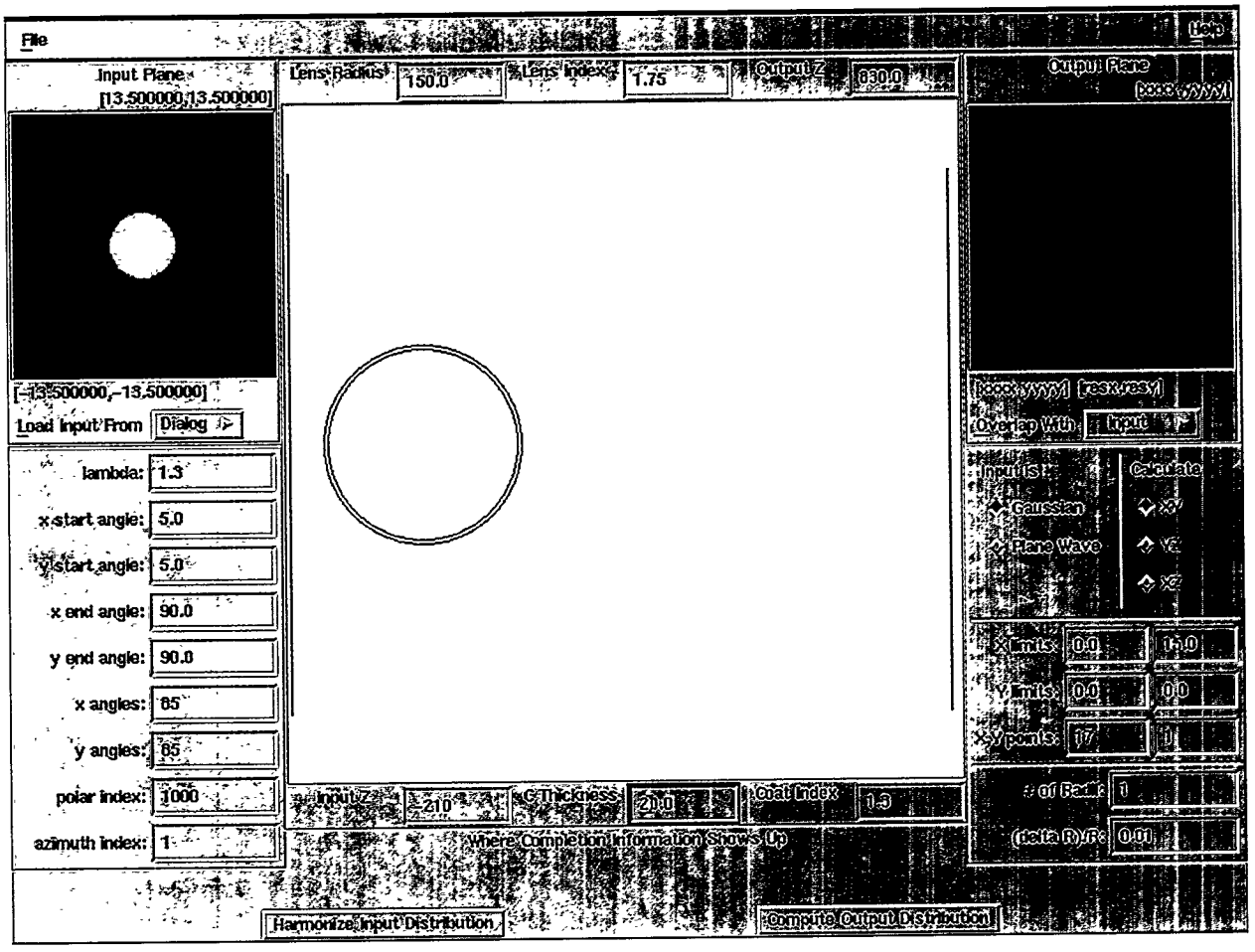


Figure 5. Illustration of the ability of the SPHERE code to predict measured laser diode to single-mode fiber coupling efficiencies. Peak coupling efficiency against laser diode-ball lens distance is shown for two different ball lenses.

Anisotropic Finite Difference Time Domain

One of the devices we were interested in simulating this year was a quantum well infrared photodetector (QWIP). This device works by absorbing long wave (9.5 μm) infrared radiation in an assembly of quantum wells. The quantum well absorber is anisotropic, so we had to build an anisotropic FDTD code to model this device. We implemented an FDTD code that could simulate the effects of anisotropic dielectric constants and conductivities. To keep the formulation simple we restricted the anisotropy to the x and y axes. The standard two-dimensional FDTD formulation iterates these equations:

$$H^{n+1/2}_{z_{i,j}} = H^{n-1/2}_{z_{i,j}} + \frac{\Delta t}{\mu}\left[\frac{E^{n}_{x_{i,j+1}} - E^{n}_{x_{i,j}}}{\Delta y} - \frac{E^{n}_{y_{i+1,j}} - E^{n}_{y_{i,j}}}{\Delta x}\right]$$

$$E^{n+1}_{x_{i,j}}\left(\frac{\varepsilon}{\Delta t} + \frac{\sigma}{2}\right) = E^{n}_{x_{i,j}}\left(\frac{\varepsilon}{\Delta t} - \frac{\sigma}{2}\right) + \frac{1}{\Delta y}\left(H^{n+1/2}_{z_{i,j}} - H^{n+1/2}_{z_{i,j-1}}\right)$$

$$E^{n+1}_{y_{i,j}}\left(\frac{\varepsilon}{\Delta t} + \frac{\sigma}{2}\right) = E^{n}_{y_{i,j}}\left(\frac{\varepsilon}{\Delta t} - \frac{\sigma}{2}\right) - \frac{1}{\Delta x}\left(H^{n+1/2}_{z_{i,j}} - H^{n+1/2}_{z_{i-1,j}}\right).$$

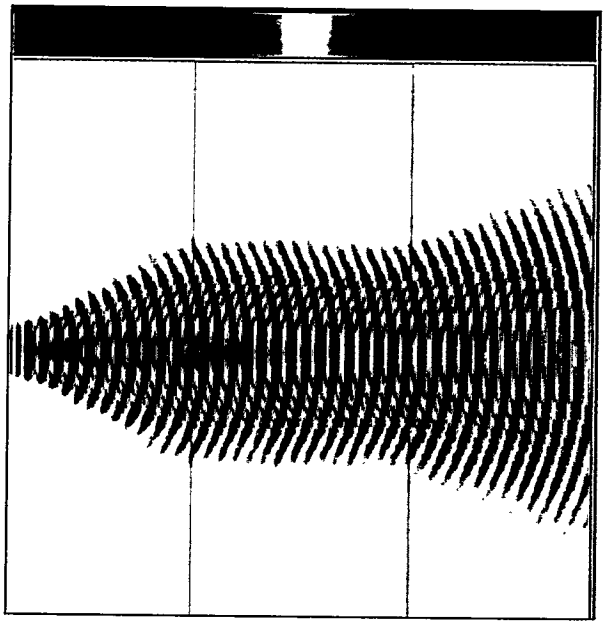


Figure 6. The effect of a strong dielectric anisotropy on an electromagnetic wave. The wave is launched from the left and expands cylindrically until it enters the anisotropic region. The electric field in the x direction is suppressed in this region, so the propagation is mostly in that direction. Upon exiting the anisotropic region, the cylindrical expansion resumes.

The modified FDTD algorithm changes the last two equations to reflect the anisotropy:

$$E^{n+1}_{x_{i,j}}\left(\frac{\varepsilon_x}{\Delta t} + \frac{\sigma_x}{2}\right) = E^{n}_{x_{i,j}}\left(\frac{\varepsilon_x}{\Delta t} - \frac{\sigma_x}{2}\right) + \frac{1}{\Delta y}\left(H^{n+1/2}_{z_{i,j}} - H^{n+1/2}_{z_{i,j-1}}\right)$$

$$E^{n+1}_{y_{i,j}}\left(\frac{\varepsilon_y}{\Delta t} + \frac{\sigma_y}{2}\right) = E^{n}_{y_{i,j}}\left(\frac{\varepsilon_y}{\Delta t} - \frac{\sigma_y}{2}\right) - \frac{1}{\Delta x}\left(H^{n+1/2}_{z_{i,j}} - H^{n+1/2}_{z_{i-1,j}}\right).$$

To verify the resultant code we simulated an anisotropic antireflection coating[6]. **Figure 6** shows what happens when a wave interacts with a strongly anisotropic dielectric.

Future Work

Next year we intend to continue examining the QWIP detector. In addition, we have received funding to undertake the generation of a major new piece of modeling software which will be a framework into which all of the codes we have developed can be integrated and hybridized.

References

1. Hadley, G. (1992), "Transparent boundary condition for the beam propagation method," *IEEE J. Quant. Electron.*, Vol. **28**, pp. 363–370.

2. Khaled, E. E. M., S. C. Hill, and P. W. Barber (1994), "Light scattering by a coated sphere illuminated with a Gaussian beam," *Applied Optics*, Vol. **33**, (15), pp. 3308–3314.

3. Van de Hulst, H. C. (1957), *Light Scattering by Small Particles*, John Wiley & Sons, New York, N.Y.

4. Ratowsky, R. P. (1995), "Ball lens reflections by direct solution of Maxwell's equations," *Opt. Lett.* **20**, (1995), in press.

5. Ratowsky, R. P. (1996), "Accurate calculation of ball lens coupling efficiency," *Photonics Technology Letters*, (1996), (in press).

6. Jacquart, P.-M., O. Acher, and P. Gadenne (1994), "Reflection and transmission of an electromagnetic wave in a strongly anisotropic medium: application to polarizers and antireflection layers on a conductive plate," *Optics Communications*, Vol. **108**, (June 1, 1994), pp. 355–366.

Electro-Mechanical Modeling of Nodular Defects in High-Power Multilayer Optical Coatings

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

Francois Y. Genin and Mark R. Kozlowski
Inertial Confinement Fusion
Laser Programs

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

Richard H. Sawicki
Laser Sciences Engineering Division
Mechanical Engineering

In FY-95, we performed the 2-1/2-D electro-mechanical analysis of nodular defects in high-power multilayer optical coatings. In addition, in preparation for the FY-96 defect modeling effort, we are expanding our modeling capabilities to enable us to perform a full 3-D electro-mechanical analysis of nodular defects, using oblique-incidence sources and asymmetric defect geometries.

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 lasers systems, such as the proposed 1.8-MJ National Ignition Facility (NIF) and the prototype 12-kJ Beamlet laser system. 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 are 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 inadequate control over layer thickness.

In multilayer dielectric stack mirrors, the predominant surface defects are micron-scale domes associated with the classic nodule defect.[1,2] These defects are initiated at seed particles that are either present on the substrate, or that are deposited during the film deposition process. Nodular defects, similar to those shown schematically in **Fig. 1**, are produced

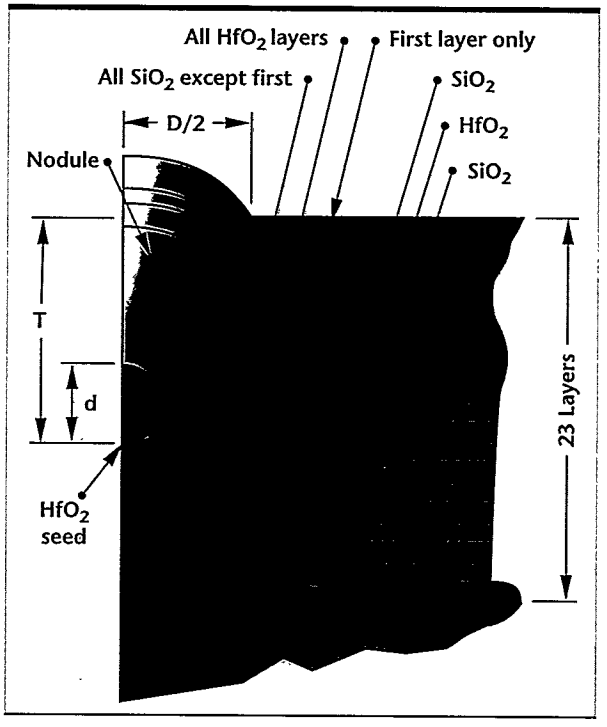


Figure 1. Schematic of a rotationally symmetric nodular defect geometry.

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.

During FY-95, we performed the 2-1/2-D electro-mechanical analysis of rotationally-symmetric nodular defect geometries, assuming normal-incidence illumination. We have studied the influence of defect geometry; e.g., nodule size, depth, and material composition, on the location and magnitude of the electric field enhancements. In addition, we are developing the modeling capabilities required to perform the full 3-D electro-mechanical analysis of nodular defects, in preparation for the FY-96 modeling effort.

Progress

In FY-95, we used the 2-1/2-D finite-difference time-domain (FDTD) electromagnetic (EM) code, AMOS[3,4,5], to study the electric field enhancements in optical coatings containing nodular defects.[6,7] Because AMOS is a 2-1/2-D code, we were limited to analyzing defect geometries that were rotationally symmetric and to using normal incidence illumination. The AMOS FDTD EM code solves the Maxwell equations in the time domain, using the appropriate boundary conditions[8] and material models[9]. During FY-95, we developed a new output diagnostic for the AMOS solver, enabling us to obtain the Fourier coefficients of the electric field components at every node throughout a planar region. By relating the Fourier coefficients of the electric field components to their steady-state values, we were able to calculate the total stead-state electric field everywhere within the defect geometry. A matrix of different defect geometries was studied, varying the seed size, depth, and material composition.

In **Fig. 2**, we show the steady-state electric field distribution throughout a layered mirror geometry with nodular defect, assuming an average input power density of 1 GW/cm^2. The AMOS code actually generates a 2-1/2-D solution, i.e., it generates the 2-D solution in the r-z plane, as shown in **Fig. 2**, and then assumes an explicit ϕ-variation to generate the full 3-D solution. In addition, we show plots of the steady-state electric field along the axis of symmetry of the geometry, as well as along a line far from the axis of symmetry (i.e., far from the defect) where the layered structure behaves like a perfect mirror. Although in this particular case, the electric field enhancement is a factor of 2, we have seen electric field enhancements as high as a factor of 4 for some of the geometries studied. For a rotationally-symmetric defect geometry with normal-incidence illumination, the highest electric field enhancements occur along the axis of symmetry of the nodular defect. Furthermore, it was determined

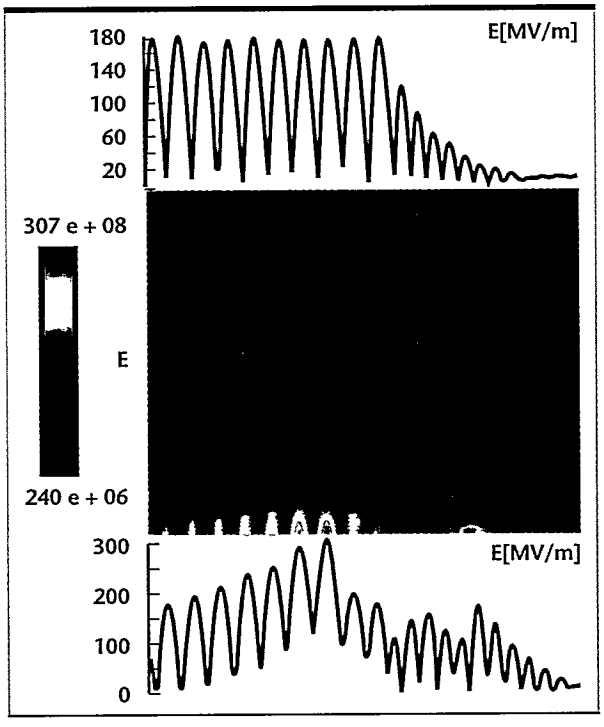


Figure 2. Electric field distribution (in V/m) in a stack mirror geometry with defect nodule, assuming an average input power density of 1 GW/cm^2.

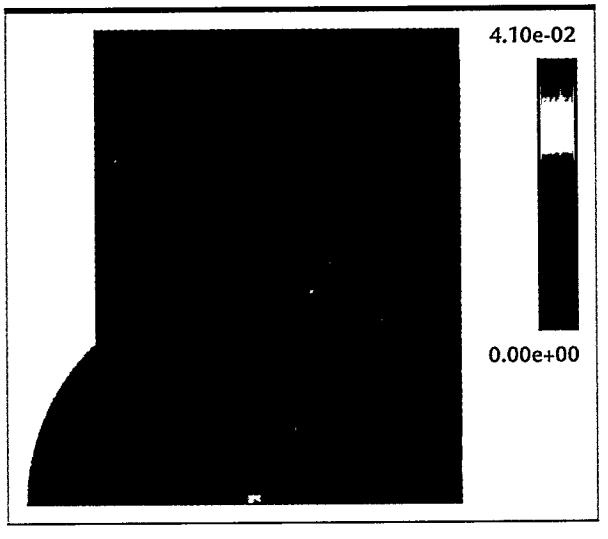


Figure 3. Heat generation distribution (in W/μm^3) in a stack mirror geometry with defect nodule, assuming an average input power density of 1 GW/cm^2.

that the electric-field enhancement decreases with increasing seed depth and that it increases with increasing seed size. However, additional geometries must be studied in order to determine if these are monotonic trends.

Figure 3 shows the corresponding heat generation distribution (σE^2, in units of W/μm^3, with σ, the electrical conductivity and, E, the electric field) for the electric field distribution shown in **Fig. 2**, assuming electrical conductivities of 1.1956 S/m for HfO_2 and 0.3366 S/m for SiO_2, respectively. The "hot spots" in the heat generation distribution typically occur along the axis of symmetry and within the material that is the most lossy, i.e., in the HfO_2. The peak heat generation decreases with increasing seed depth, and increases with increasing seed size.

As shown in the block diagram in **Fig. 4**, the full electro-mechanical analysis procedure for the layered mirror geometry with nodular defect involves four steps: (1) the electromagnetic analysis, (2) coupling between the EM and the thermal-mechanical codes, (3) the thermal analysis, and (4) the structural analysis. As described above, the electromagnetic analysis in the first step is performed using the 2-1/2-D FDTD EM code, AMOS, which calculates the steady-state electric field spatial distribution on the AMOS mesh. In the second step, the steady-state electric field is interpolated from the AMOS mesh to the corresponding thermal-mechanical model mesh, and the heat generation (σE^2) is computed for every element in the thermal-mechanical mesh (**Fig. 3**). In the third step, the heat generation terms are used in the thermal analysis to determine the spatial and temporal distribution of temperature. Finally, the fourth step involves the structural analysis of the nodular defect, determining the spatial and temporal distributions of the stresses and strains in the model.

In the past, the commercially-available 2-D finite-difference thermal-mechanical analysis code, COSMOS/M[10], was used to perform the thermal-mechanical analysis of rotationally-symmetric, normal-incidence nodular defect geometries. These 2-D simulations estimated the transient temperature response of the nodule and the surrounding medium, and predicted the dynamic stresses caused by the thermal impulse. However, while the thermal-mechanical analysis can be performed in 2-D, taking advantage of the rotational symmetry of the nodular defects, the limitations of the 2-D COSMOS/M code become apparent when attempting to increase the accuracy of the 2-D numerical solutions. The current version of COSMOS/M can only handle 3000 nodes to generate the thermal-mechanical model mesh used in the discretization of the various defect geometries, and does not permit nonrotationally-symmetric heat source conditions. These limitations led us to conclude that COSMOS/M would become unsuitable for 3-D problems. A set of thermal-mechanical codes available at Lawrence Livermore National Laboratory (LLNL) (TOPAZ3D and NIKE3D, also known as NITO3D in their coupled form) were therefore identified as suitable for the 3-D thermal-mechanical modeling effort. We are currently involved in making the transition from COSMOS/M to NITO3D and performing preliminary solution convergence tests in preparation for the coupled analysis of the

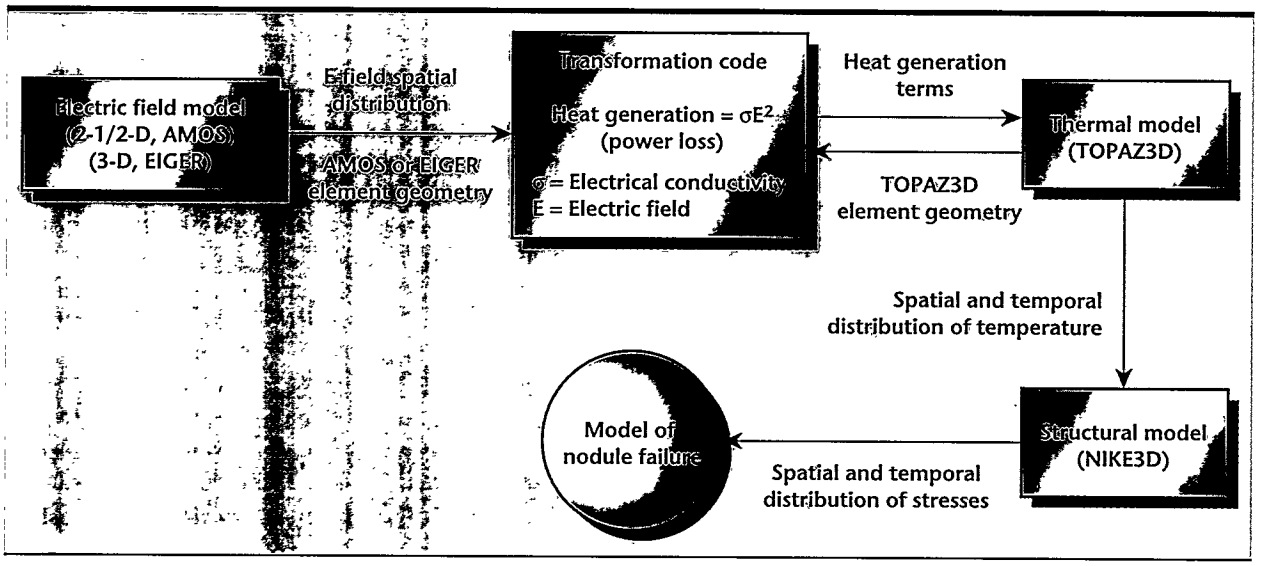


Figure 4. Block diagram of the coupled electromagnetic and thermal-mechanical analysis process.

asymmetric EM-thermal loading of the defect geometry. In addition, towards the end of FY-95 we began the development of the coupling link between the EM code AMOS and the thermal-mechanical code suite, NITO3D. The NITO3D thermal-mechanical model meshes for the various rotationally-symmetric defect geometries are being constructed in preparation for the 2-1/2-D thermal-mechanical analysis of the nodular defects.

Future Work

We anticipate the completion of the 2-1/2-D electro-mechanical analysis of nodular defects in high-power multilayer optical coatings by the end of calendar year 1995. This work will include coupling the 2-1/2-D FDTD EM Code, AMOS, to the 3-D thermal-mechanical code suite, NITO3D, in order to come up with a model of nodule failure.

Future work will extend the electro-mechanical modeling effort to the full 3-D regime, allowing for both generalized sources (arbitrary angle of incidence) and generalized defect geometries (asymmetric geometries). We will use the 3-D frequency-domain EM code, EIGER (currently under development at LLNL), to compute the steady-state electric field distribution throughout these 3-D defect geometries. We will continue to develop our EM-thermal-mechanical coupled code capability to couple the spatial distribution of heat generation terms (σE^2) from EIGER to the 3-D thermal-mechanical code suite, NITO3D. Solution convergence studies will be performed as part of the code-coupling effort. The 3-D thermal-mechanical code suite, NITO3D will then be used to analyze heat diffusion as well as the spatial and temporal distributions of temperature and stresses in order to achieve a fully 3-D electro-mechanical model of nodule failure for various defect geometries.

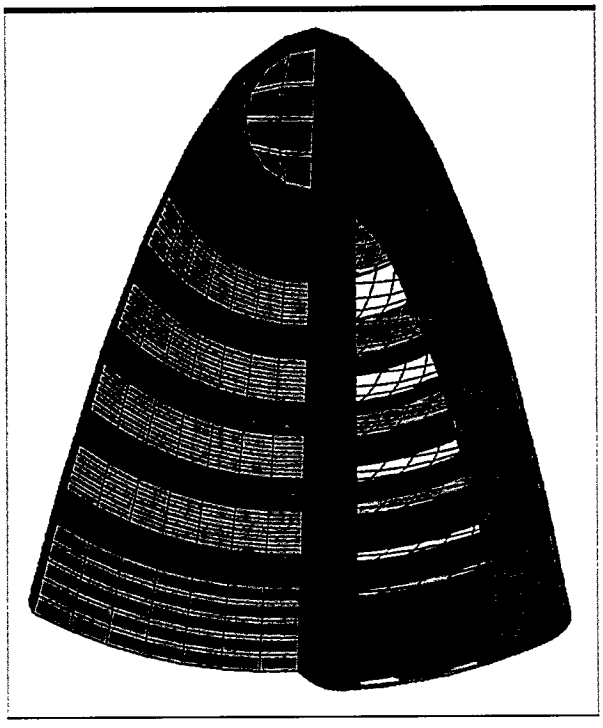


Figure 5. Surface meshes for a 3-D rotationally symmetric nodular defect geometry, to be used with the 3-D frequency-domain EM code, EIGER.

Prior to proceeding with modeling a new matrix of asymmetric, oblique-incidence nodular defect geometries, the 3-D EM and thermal-mechanical codes will be benchmarked using the previously analyzed 2-1/2-D defect geometries (rotationally symmetric, normal incidence illumination). The 3-D surface meshes for these rotationally-symmetric geometries have already been generated in preparation for this effort. **Figure 5** shows a typical 3-D surface mesh for a rotationally-symmetric nodular defect geometry. Because the EIGER code uses stratified media Green's functions to obtain exact solutions for the fields out in the planar region of the mirror, only the parabolic nodular defect region of the geometry needs to be defined using surface meshes. **Figure 5** shows the nodular defect region with one quarter of the geometry removed for easier viewing of the various surface meshes. The entire parabolic surface of the nodular defect is meshed as are the surfaces between the adjacent layers of HfO_2 and SiO_2 within the defect region itself.

Acknowledgments

The authors would like to acknowledge T. L. Swatloski of TGV, Inc., Santa Cruz, California for her contributions in computer science support for this project. We would also like to acknowledge useful discussions with J. F. DeFord of the MacNeil-Schwendler Company, Milwaukee, Wisconsin. In addition, we wish to acknowledge P. P. Weidhaas of LLNL for his contributions in developing the PREAMOS preprocessor. We would also like to acknowledge J. Creighton, a summer employee at LLNL, for his contributions in generating the 3-D surface meshes for the rotationally-symmetric defect geometries. Finally, we would like to acknowledge R. M. Sharpe of LLNL for useful discussions about the EIGER code.

References

1. Johnson, L. F., E. J. Ashley, T. M. Donovan, J. B. Franck, R. W. Woolever, and R. Z. Dalbey (1984), "Scanning Electron Microscopy Studies of Laser Damage Initiating Defects in ZnSe/ThF_4 and SiH/SiO_2 Multilayer Coatings," *Laser Damage in Optical Materials: 1984*, NBS Special Pub. **727**, p. 356.

2. Staggs, M. C., M. Balooch, M. R. Kozlowski, and W. J. Siekhaus (1992), "In Situ Atomic Force Microscopy of Laser-Conditioned and Laser-Damaged HfO_2/SiO_2 Dielectric Mirror Coatings," *Laser-Induced Damage in Optical Materials: 1991*, SPIE Vol. **1624**, p. 375.

3. 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.

4. DeFord, J. F., G. D. Craig, and R. R. McLeod (1990), "The AMOS Wakefield Code," *Proc. Conf. on Computer Codes and the Linear Accelerator Community, Los Alamos, N. Mex.*, January 22-25, p. 265.

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

6. DeFord, J. F., and M. R. Kozlowski (1992), "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.

7. 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.

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

9. Luebbers, R. (1990), *IEEE Trans. Antennas Propag.*, **AP-32**, p. 222.

10. Structural Research and Analysis Corporation (1993), *COSMOS/M Finite Element Analysis computer code, version 1.70*, Santa Monica, Calif.

Development of an Electromagnetic Electron and Ion Fluid Computer Code

David J. Mayhall, Jick H. Yee, and Dale E. Miller
Engineering Research Division
Electronics Engineering

We have made progress in the development of a two-dimensional, cylindrically symmetric, time-dependent, electromagnetic, finite difference electron and ion fluid computer code for modelling the plasma behavior in the LAM Research TCP9400 plasma etching tool in the Microtechnology Center at Lawrence Livermore National Laboratory. We have added convective and flow work terms to the electron conservation equations, implemented a more realistic treatment of the driving electric coil, and added improved boundary conditions for the electron flux at the chamber walls. We have run this improved model out to 28 rf cycles of elapsed time by direct time marching of the solution, with reasonable results, but with the slow buildup of a bipolar numerical noise in the driving azimuthal electric field in the plasma region. We have resited four of the solution components within the computational cell. This change in variable location has resulted in reduction in the azimuthal electric field noise and smoother fluid component distributions in space.

Introduction

In FY-94, we initiated enhancement of an existing, two-dimensional, time-dependent electron fluid code, which solves Maxwell's curl equations with ionization, in order to model plasma behavior in the recently acquired LAM Research TCP 9400 etching machine. We added continuity and momentum conservation equations for positive Argon ions without metastable neutral states. We also added terms for electron and ion transport. For a set of initial conditions, we ran computations out to 10 to 16 rf cycles of elapsed time, with reasonable results.

In FY-95, we have made a number of further improvements to make the computational model more realistic. These improvements include the addition of terms to make the equation set more accurate physically, the addition of a more realistic representation of the electric coil, the addition of a more realistic treatment of the electron flux at the chamber walls, and a resiting of the solution components to reduce numerical noise in the driving electric field E_δ and produce spatially smoother fluid component distributions. A simplified physical geometry for the computational model is shown in **Fig. 1**.

Progress

More Physically Accurate Equations

Part of our developmental work has involved the addition, in the interior solution space, of terms to the governing equation set to make the computations more accurate. We have added the $\vec{v}_e \bullet \nabla \vec{v}_e$ terms to the three electron momentum conservation equations, where $\vec{v}_e$ is the electron fluid velocity vector, and ∇ is the spatial gradient. We have also added the convective $\vec{v}_e \bullet \nabla u_e$ terms and the $\nabla \bullet (\vec{v}_e P_e)$ flow work terms to the electron energy conservation equation, where u_e is the electron energy, and P_e is the electron pressure. The updated electron fluid

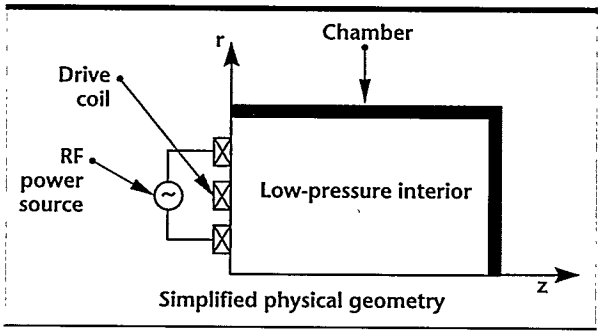


Figure 1. The simplified geometry of the LAM Research TCP 9400 etcher used for the computational model.

equations are presented in **Table 1** with newly added terms underlined. The Maxwell's equations and the positive ion fluid equations are unaltered. They were presented in a previous report.[1]

More Realistic Model of the Electric Coil Structure

We have improved the treatment of the driving electric coil and the dielectric plate which supports it. Previously, a time-programmed radial magnetic field distribution $H_r(r,t)$, where r is the radial coordinate and t is the time, was specified across the radial input plane. Now, in the radial input row of computational cells, we have introduced three perfectly conducting, rectangular electric coil turns with time-programmed current densities as driving inputs. The spaces between these coil turns are assumed to be filled with air, which is above the threshold for electric breakdown. A transverse electromagnetic wave boundary condition is applied to the radial magnetic

Table 1. The enhanced electron fluid equations.

$$\frac{\partial n_e}{\partial t} = \nu_i n_e - \frac{1}{r}\frac{\partial}{\partial r}\left(r n_e v_{er}\right) - \frac{\partial}{\partial z}\left(n_e v_{ez}\right)$$

$$\frac{\partial v_{er}}{\partial t} = -\frac{1}{n_e m_e}\frac{\partial P_e}{\partial r} + \frac{eE_r}{m_e} + \frac{ev_{e\phi}}{m_e}\mu_0\left(H_{z0}+H_z\right) - v_{er}\left(\nu_{eN}+\nu_i\right) - \underline{v_{er}\frac{\partial v_{er}}{\partial r} + \frac{v_{e\phi}^2}{r} - v_{ez}\frac{\partial v_{er}}{\partial z}}$$

$$\frac{\partial v_{e\phi}}{\partial t} = \frac{eE_\phi}{m_e} + \frac{ev_{ez}}{m_e}\mu_0\left(H_{r0}+H_r\right) - \frac{ev_{er}}{m_e}\mu_0\left(H_{z0}+H_z\right) - v_{e\phi}\left(\nu_{eN}+\nu_i\right) - \underline{v_{er}\frac{\partial v_{e\phi}}{\partial r} - \frac{v_{er}v_{e\phi}}{r} - v_{ez}\frac{\partial v_{e\phi}}{\partial z}}$$

$$\frac{\partial v_{ez}}{\partial t} = -\frac{1}{n_e m_e}\frac{\partial P_e}{\partial z} + \frac{eE_z}{m_e} - \frac{ev_{e\phi}}{m_e}\mu_0\left(H_{r0}+H_r\right) - v_{ez}\left(\nu_{eN}+\nu_i\right) - \underline{v_{er}\frac{\partial v_{ez}}{\partial r} - v_{ez}\frac{\partial v_{ez}}{\partial z}}$$

$$\frac{\partial u_e}{\partial t} = e\vec{E}\bullet\vec{v}_e - \underline{\frac{1}{rn_e}\frac{\partial}{\partial_r}\left(rv_{er}P_e\right) - \frac{1}{n_e}\frac{\partial}{\partial z}\left(v_{ez}P_e\right)} - \varepsilon_i\nu_i - \varepsilon_{ex}\nu_{ex} - u_e\nu_i - \underline{v_{er}\frac{\partial u_e}{\partial r} - v_{ez}\frac{\partial u_e}{\partial z}}$$

with

$$\vec{E}\bullet\vec{v}_e = E_r v_{er} + E_\phi v_{e\phi} + E_z v_{ez}$$

$$P_e = \frac{2n_e}{3}\left[u_e - \frac{m_e}{2}\left(v_{er}^{\ 2} + v_{e\phi}^{\ 2} + v_{ez}^{\ 2}\right)\right]$$

Table 2. Initial values for a typical computation in a reduced-size model etcher chamber.

Pressure	20 mTorr
Initial uniform electron density	8.0×10^{12}m^{-3}(8.0×10^{6}cm^{-3})
Initial uniform ion density	8.0×10^{12}m^{-3}(8.0×10^{6}cm^{-3})
Initial uniform electron energy	0.025 eV
Neutral temperature	300 K
RF frequency	13.56 MHz
Peak coil current density	0.851 A/m^2($E_\phi\approx$11.4 V/cm)
Chamber inner radius	5.1 cm
Chamber inner height	2.5 cm
Grid size	(44 points in r) × (26 points in z)
Secondary electron emission coeffiecient	0.05

field component in these coil spaces to allow electromagnetic energy to pass out of the numerical solution space with minimal generation of nonphysical wave reflection from the grid boundary. We have also placed a dielectric layer below the driving electric coil. Such a dielectric layer is used in the actual machine to support the pancake coil and insulate the low-pressure interior of the chamber, where the etching is performed.

More Realistic Chamber Wall Boundary Conditions

We have modified the boundary condition on the electron flux at the chamber walls, to make the treatment in the computer code more realistic. We simulate the outflow of electrons from the plasma to the chamber walls by allowing the electrons to flow to the walls at the velocities calculated by the momentum conservation equations in the code. Secondary emission of electrons from the chamber walls due to positive ion impingement is also included. The secondary emission of electrons back into the plasma is assumed to be proportional to the positive ion flux onto the walls. The proportionality constant, the secondary emission coefficient, is presently taken to be a constant input parameter.

Computations

We have run the improved computer code stably out to 28 rf cycles by time marching from the set of initial conditions shown in **Table 2**. **Figure 2** shows the computed electron density at the end of 28 rf cycles, or 2.06 μs after the initiation of the computation. An rf cycle at 13.56 MHz lasts 73.7 ns. The dielectric layer under the pancake coil begins at the first axial grid line, and ends at the fifth axial grid line from the origin, in the z direction. The dielectric layer is 0.8 cm thick. With the peak electron density at $5.60\times10^{15}m^{-3}$ ($5.60\times10^{9}cm^{-3}$), the peak density has increased by about 700 times beyond its initial value.

Figure 3 shows the electron energy in electron volts at the end of the 28th rf cycle. The peak value at this instant of time is 8.99 eV, which represents a gain of 360 times over the initial value. **Figure 4** similarly shows the ion kinetic energy at the conclusion of the 28th rf cycle. The peak value is 6.31 eV, which shows that, at this time, the peak positive ion kinetic energy is approaching the peak value of the electron energy. The ion kinetic energy distribution is more sharply peaked in space than is the electron energy distribution.

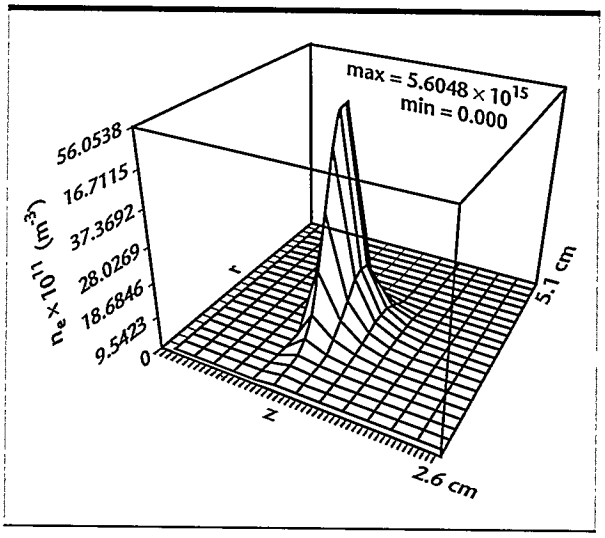


Figure 2. The computed electron density n_e after 28 rf cycles.

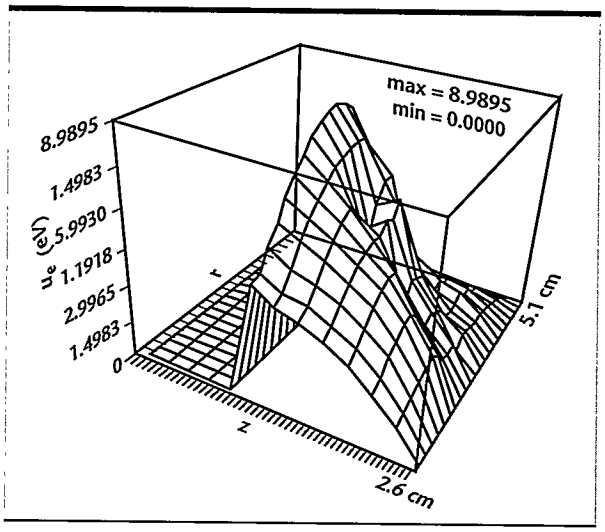


Figure 3. The computed electron energy u_e after 28 rf cycles.

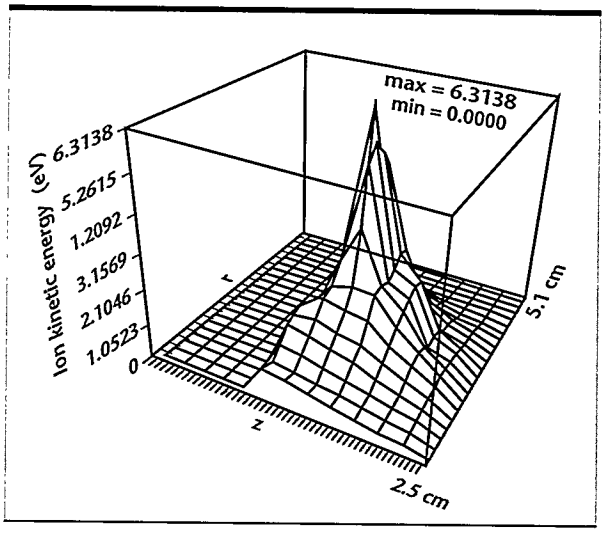


Figure 4. The ion kinetic energy after 28 rf cycles.

Table 3 presents the peak values of the electron and positive ion densities and energies at several times, which are portrayed as the ends of a given number of rf cycles. It can be seen that the peak electron density increases slowly from 15 to 28 rf cycles, and that, by the end of the 28th rf cycle, the peak ion kinetic energy is catching up with the peak electron energy. At the end of each rf cycle, the peak ion density slightly exceeds the peak electron density.

Reduction of Azimuthal Electric Field Numerical Noise

We have noticed that a bipolar numerical noise occurs in the azimuthal electric field in the plasma region. An example of this noise is shown in **Fig. 5(a)**. This noise slowly builds up from rf cycle to rf cycle, unless some measure is taken to suppress it. To reduce this noise, we have moved two of the electron fluid velocity components from the center of the computational cell to the boundaries of the cell. The radial velocity component v_{er} has been moved to the lower boundary parallel to the z axis. The axial velocity component has been moved to the right cell boundary, parallel to the r axis. The E_r and E_z electric field components have also been moved from the center of the compuational cell. The radial electric field is now coincident with the radial electron velocity, and the axial electric field is coincident with the axial electron velocity.

This resiting of solution components has reduced the noise in the azimuthal electric field in the plasma region. This reduction of the numerical noise is illustrated in **Fig. 5**. The azimuthal field in **Fig. 5(a)** is for the original siting of the solution components at 54.0 ns into the first rf cycle of the calculation. This is close to the positive peak of the sinusoidal driving electric field at 55.3 ns. The azimuthal field in **Fig. 5(b)** is for the new siting of the solution components at 54.3 ns into the first rf cycle. The three peaks in the azimuthal electric field along the first axial grid line are the field values at the surfaces of the three coil turns at the interfaces with the dielectric plate. This plate begins at the first axial grid line and extends to the fifth axial grid line. Calculations with resited solution grid have produced smoother energy and density distributions than were obtained with the original placement of solution components. The resiting of the electric field and electron velocity components has improved the computational results.

Table 3. Peak electron and positive ion densities and energies versus number of completed rf cycles.

Number of rf cycles	Peak electron density cm^{-3}	Peak ion density cm^{-3}	Peak electron energy eV	Peak ion kinetic energy eV
15	2.42×10^8	2.46×10^8	8.80	0.48
20	7.25×10^8	7.45×10^8	8.91	1.28
28	5.60×10^9	5.66×10^9	8.99	6.31

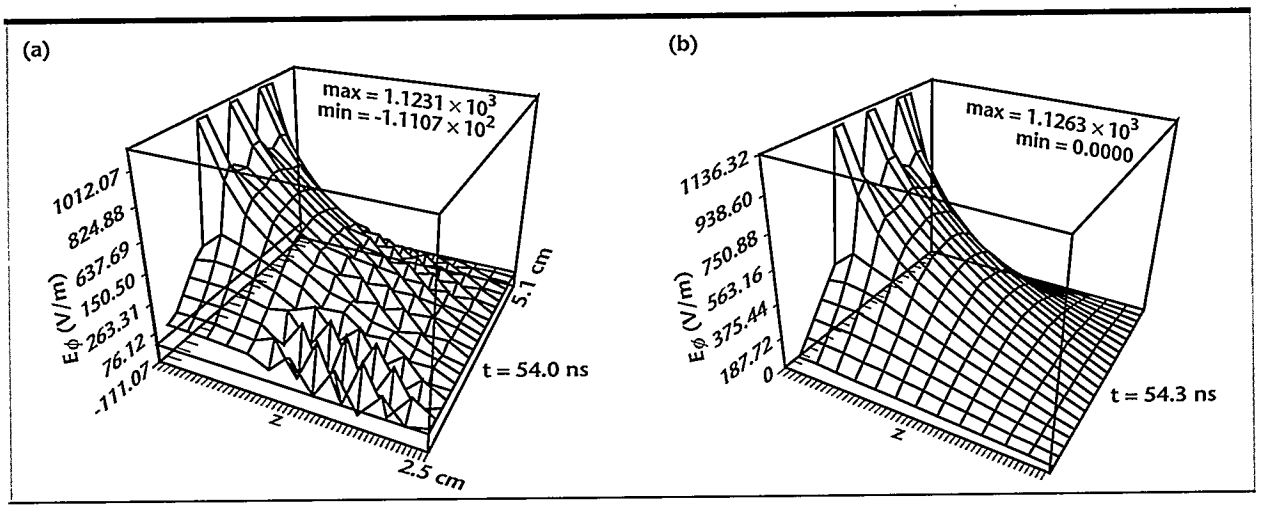


Figure 5. The azimuthal electric field (a) before, and (b) after solution component resiting.

Future Work

In our future work, we will march the resited computer code out in time through a number of rf cycles and compare the results with those from the previous siting. We will also develop solution acceleration techniques, which allow the periodic steady state solution of an equation system to be determined in a more computationally efficient way than marching faithfully through time from the initial state of the system. Such techniques have been successfully used in nonlinear electric circuit analysis and one-dimensional plasma processing calculations in chemical engineering. We will then add the equations for the inclusion of negative ions to our set of electromagnetic-fluid, time-dependent differential equations.

Reference

1. Koo, J. C., K. H. Krause, D. J. Mayhall, and D. E. Miller (1995), *Transformer-Coupled Plasmas for Advanced Etching*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-55868-94).

Single-Pass Surface-to-Borehole Reconstructive Imaging

Scott D. Nelson and Erik Johansson
Defense Sciences Engineering Division
Electronics Engineering

This report describes a series of computational imaging code development efforts, material effects, electromagnetic simulations, experimental data sets, and reconstructive imaging results, which characterize the current state of the surface-to-borehole imaging project. These efforts build on previous work done in the more generic field of ground penetrating imaging radar.

The existing imaging codes were modified to allow for the extremely low frequencies (10 to 10,000 Hz) that are used in this field. These modifications required additional soil information in the kernel of the imaging codes. Results using the modeled data agree to within one down-range resolution cell. The preliminary imaging results for the first experimental data set are inconclusive due to the soil type used.

Introduction

Present procedures for subsurface imaging involve model-based iterative techniques for the estimation of subsurface features (typically the Born iterative method[1]). This approach basically iterates a problem space, based on the results from a forward propagation through the modeled problem space, to determine the difference between the model and the actual results. The model space is then adjusted and the problem is rerun. Although accurate, this requires a considerable amount of CPU time and memory resources. Note, however, that these iterative techniques are a vast improvement over the raw data approach taken by earlier experimental efforts[2,3,4] at Lawrence Livermore National Laboratory (LLNL) in the area of data interpretation.

Imaging techniques developed as part of other LLNL projects use single-pass, coherent wavefield back-propagation to reconstruct the problem space. These techniques[5,6] do not require *a priori* knowledge of target geometries, and make only minimal assumptions regarding the media characteristics, although knowledge of the media involved aids in the reconstruction accuracy (see **Fig. 1**). These techniques are extremely fast and memory-efficient, and are suitable for high frequency applications that require small sample intervals and large amounts of data. Preliminary results from applying these techniques to the borehole data are promising, based on the one-pass nature of the algorithms. Results for problem sets with 400 cross components (20 × 20) can be generated in a few seconds.

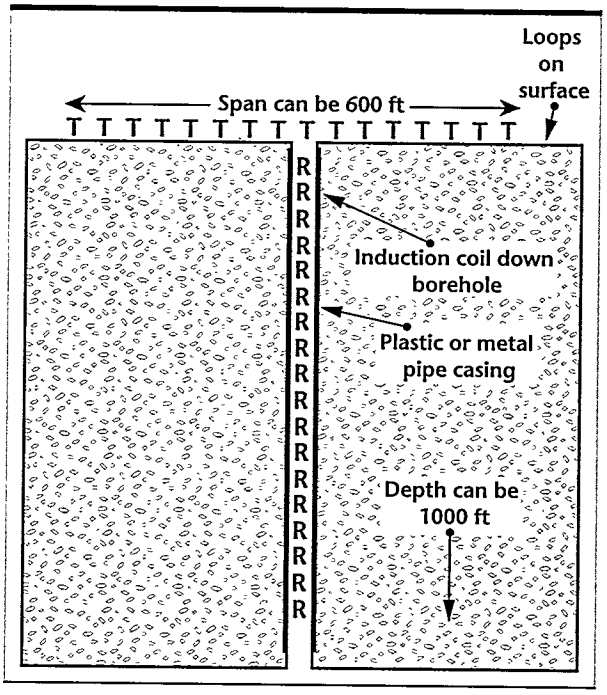


Figure 1. Illustration of surface-to-borehole data acquisition and imaging. This technique uses coils on the surface and down in the borehole. Applications typically use existing boreholes, and are thus constrained by their configuration.

The current techniques used in borehole imaging require that data be acquired, and then that images be generated in an off-line fashion. The motivation for this project was that several 2-D swaths could be taken in the field in a timely fashion; the imaging could be done in the field; and those 2-D images could be used to identify a 3-D region of interest that should be scanned at higher detail. This would decrease the field time involved in interrogating a site. This project's results could also be used as the initial guess for the existing iterative techniques, which should help them to converge faster.

Progress

Computational Imaging Code Development Efforts

The main focus of the imaging code development was to adapt the coherent wavefield back-propagation model used in ground penetrating radar (GPR) applications to the geometry required for surface-to-borehole imaging. The imaging geometry for a typical GPR application is based on a monostatic reflection-mode model, where a transmitter/receiver is repeatedly scanned over a surface to synthesize a 1-D or 2-D aperture and its associated coherent wavefield. The geometry in the surface-to-borehole case is substantially different: a single transmitter is scanned over the surface, and a coherent wavefield is received along the receivers in the borehole for each transmitter position. The transmitter is scanned horizontally over the surface of the ground, but the receivers are located at increasing depths. Hence, the receiver array is orthogonal to the region scanned by the transmitter. This geometry fits neither the reflection-mode imaging model nor the transmission-mode imaging model (where the transmitter and receiver pair are opposite each other, with the object in between), but represents a hybrid of the two: a multi-static geometry.

GPR imaging uses frequency diversity, resulting from the broad-band nature of the transmitted pulse, to estimate the spatial frequency spectrum of the object or medium. This is analogous to using multiple views or angles when acquiring data in x-ray computed tomography (CT). Each frequency present in the GPR-transmitted pulse contributes to a region of the spatial frequency spectrum; and similarly, each angle used in x-ray CT contributes information along a line in the spatial frequency spectrum. Hence, GPR imaging can be interpreted as multi-frequency diffraction tomography. In the surface-to-borehole case, however, there is only a single frequency. Therefore, additional information about the object or medium must be obtained by the coherent combination of multiple images resulting from imaging the object or medium at varying transmitter positions.

The mathematical relationships for surface-to-borehole imaging using coherent wavefield back-propagation are easily derived from the equations relating to the GPR case. In GPR imaging, the spatial frequency spectrum $S(f_x,f_y;z_o,\lambda_o)$ of the object or medium at a plane z_o, resulting from an illumination wavelength λ_o is given by

$$S(f_x,f_y;z_o,\lambda_o) = H^{-1}(f_x,f_y;z_o,\lambda_o)R(f_x,f_y;\lambda_o) \quad (1)$$

where $R(f_x,f_y;\lambda_o)$ is the spatial frequency spectrum of the received wave field, and f_x and f_y are the spatial frequency coordinates corresponding to x and y, respectively.[7] The filter function $H^{-1}(f_x,f_y;z_o,\lambda_o)$ is given by

$$H_p^{-1}(f_x,f_y;z_o,\lambda_o) = \exp\left(-j2\pi z_o\sqrt{1/\lambda^2 - f_x^2 - f_y^2}\right), \quad f_x^2+f_y^2 < 1/\lambda^2 \quad (2)$$

and

$$H_e^{-1}(f_x,f_y;z_o,\lambda_o) = \exp\left(2\pi z_o\sqrt{f_x^2+f_y^2-1/\lambda^2}\right), \quad f_x^2+f_y^2 > 1/\lambda^2 \quad . \quad (3)$$

The condition distinguishing the two cases of the filter function arises from the dispersion relation. The first case (H_p) represents a phase-only filter, which corresponds to an inversion of the propagating mode of the wavefield. The second case (H_e) is a magnitude-only filter, which corresponds to an inversion of the evanescent portion of the wavefield. Since the inversion of the evanescent portion of the wavefield is an exponentially increasing function of z, care must be taken to ensure that it does not swamp out the inversion of the propagating mode. Therefore, the inversion of the evanescent portion is usually neglected, and its filter function set to zero. The final estimate of the object is achieved by direct Fourier inversion of $S(f_x,f_y;z_o,\lambda_o)$.

In the surface-to-borehole geometry, we have an array of receivers along the z direction; the ground surface lies along the x axis; and we are imaging along the x-z plane. Hence, we will perform the

wavefield back-propagation from the borehole along the *x*-*z* plane. Given a received wavefield $R(f_z;\lambda_o)$, the estimate of the object area is given by

$$S(f_z;x,\lambda_o) = H^{-1}(f_z;x,\lambda_o)R(f_z;\lambda_o) \quad (4)$$

where x is the back-propagation distance away from the borehole. The filter functions are given by

$$H_p^{-1}(f_z;x,\lambda_o) = \exp\left(-j2\pi x\sqrt{1/\lambda^2 - f_z^2}\right), \quad f_z^2 < 1/\lambda^2 \quad (5)$$

for the propagating mode, and

$$H_e^{-1}(f_z;x,\lambda_o) = \exp\left(2\pi x\sqrt{f_z^2 - 1/\lambda^2}\right), \quad f_z^2 > 1/\lambda^2 \quad (6)$$

for the evanescent mode.

Prior to the wavefield back-propagation, the data must be preprocessed to compensate for the signal propagation loss, the near-field effects of the transmitter (including direct coupling effects), and the phase error resulting from the multi-static geometry.

Material Effects

In addition to the change of geometry in adapting the GPR algorithms to the surface-to-borehole imaging problem, the effects of conductivity must be considered as well. For a homogeneous conductive medium, the wavelength is given by[7]

$$\lambda = \frac{\sqrt{2}}{f\sqrt{\mu\varepsilon}}\left[\sqrt{1+\left(\frac{\sigma}{2\pi f\varepsilon}\right)^2}+1\right]^{-1/2} \quad (7)$$

where f is the frequency, and μ, ε, and σ are the permeability, permittivity, and conductivity of the medium, respectively.

Furthermore, the filter functions shown above must also be updated to account for the conductivity of the medium. The filter functions now become

$$H_p^{-1}(f_z;x,\lambda_o) = \exp\left(-j2\pi x\sqrt{1/\lambda^2 - f_z^2 - j\frac{f\mu\sigma}{2\pi}}\right), \quad f_z^2 < 1\ \lambda^2 \quad (8)$$

for the propagating mode, and

$$H_e^{-1}(f_z;x,\lambda_o) = \exp\left(2\pi x\sqrt{f_z^2 - 1/\lambda^2 - j\frac{f\mu\sigma}{2\pi}}\right), \quad f_z^2 > 1/\lambda^2 \quad (9)$$

for the evanescent mode.

Electromagnetic Simulations

Several computational electromagnetic simulations were performed to gauge the accuracy and sensitivity of the imaging code (see composite in **Fig. 2**). These simulations included layer boundaries, semi-infinite target bodies, bodies near to, and far from, the borehole, and different soil conductivity states. Three different code sets were used in the course of completing these simulations: *em3d* (a University of Utah code); *NEC*[8] and *Patch* (both LLNL codes). All three of these codes are frequency domain codes that model semi-infinite half spaces with inclusions. This differs from earlier efforts[9] which focused on time domain studies. Flat layer interfaces were

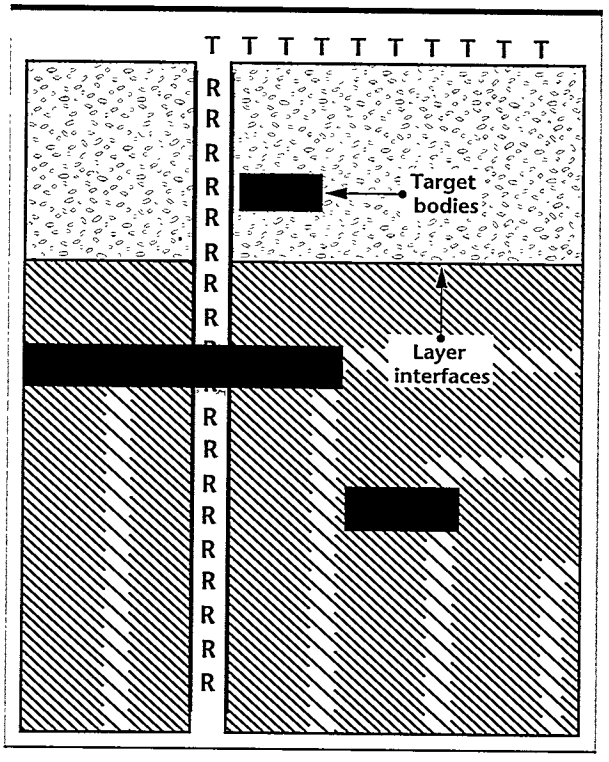


Figure 2. Composite of several of the test scenarios, including the semi-infinite target, the layer interface, and various target locations. Numerous computational electromagnetic simulations were performed in an effort to evaluate the imaging algorithms. In the EM models, only one configuration was used at a time.

introduced to simulate borehole field sites typical of what is found in the experimental data. One of the simulation sets and the resulting reconstructive image is shown in **Fig. 3**. The folds and undulations of the layer structure were not included at this time, nor were the casing effects. The electrical effects of the loops used in the measurement process were included in the later models.

These codes basically track the magnetic field intensities through the problem space. The *NEC* code, in particular, uses a moment method technique which does not require the "meshing" of the entire problem space, but instead drives currents on the various elements to compute the fields. This was the principal advantage of using this code over the *em3d* code, which required a gridded structure,

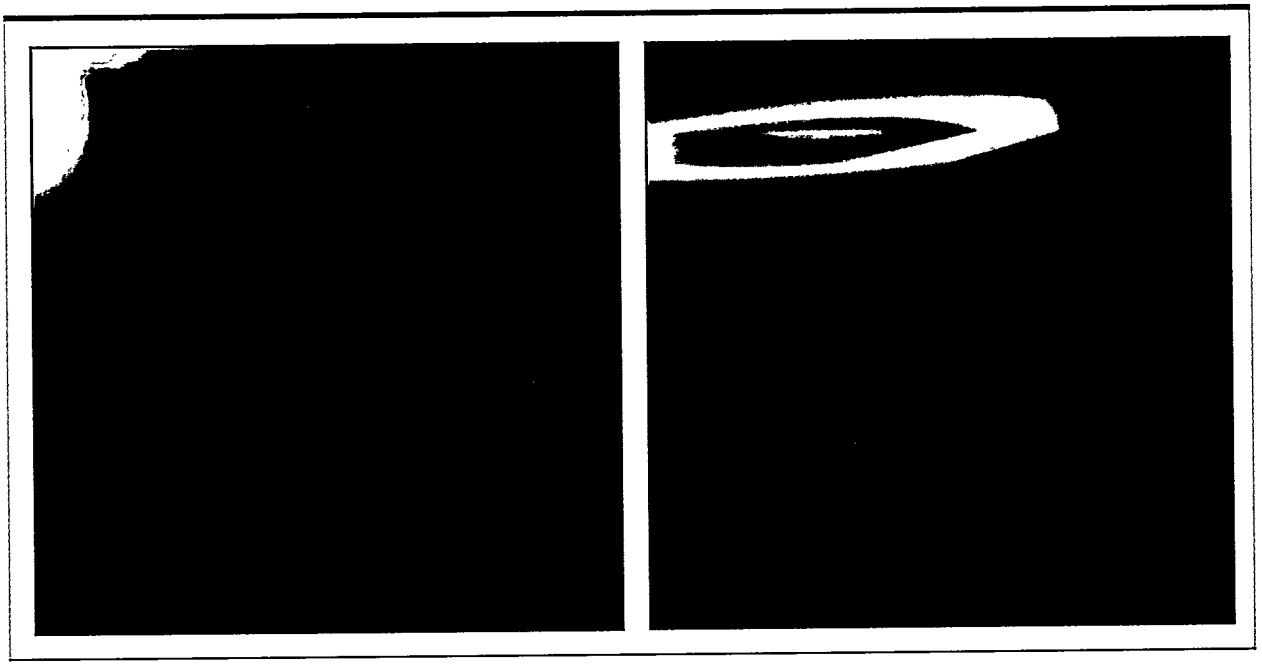
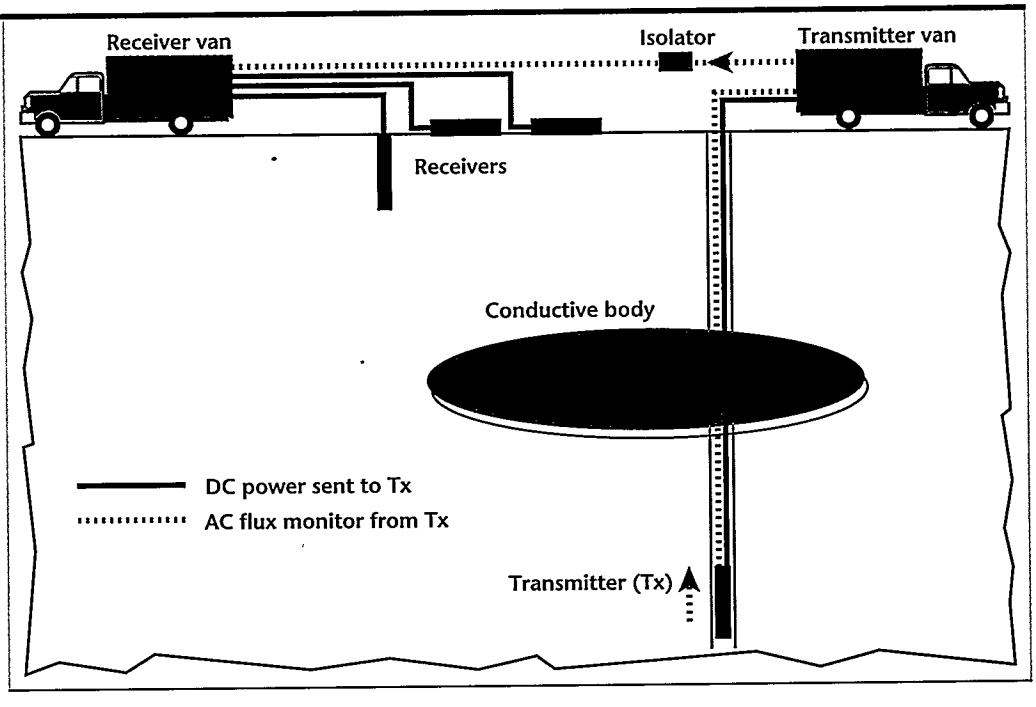


Figure 3. Raw magnetic field data (left image) which results from the simulation. The principal effect is the coupling between the transmitter and receiver. Signal preprocessing is required, as part of the reconstructive imaging code, to get the effects due to the target. The resulting reconstructed image (right) shows the single target clearly in the image. The wavelength for this case is approximately 32 m, which gives a theoretical resolution of 16 m, or 3 resolution cells.

Figure 4. Field layout of the data acquisition system used at the Richmond Field Station, (©1994, M. Wilt). The surface-to-borehole geometry was used for the reconstructive imaging analysis.



and necessitated larger computational resources at the frequency extremes, which resulted in degradation in accuracy, and prompted the change to the NEC code.

Experimental Data Sets

Preliminary experimental data from the Richmond Field Station[10] was used to create a series of images using the reconstructive imaging algorithms. This data set involved a series of transmitter and receiver locations, as shown in **Fig. 4.** The data set was taken at 9.8 kHz, and included a very small salt water body. This body was too small for the spatial resolution that was used, and the lower layer conductivities were out of bounds of the operating parameters of the imaging code. The resulting image was inconclusive (**Fig. 5**). The next section contains a discussion about the conditions required for meaningful reconstructions.

Reconstructive Imaging Results

Reconstructive imaging was used to convert the signals at the receivers into spatial reconstructions of the medium. Although not required, the well

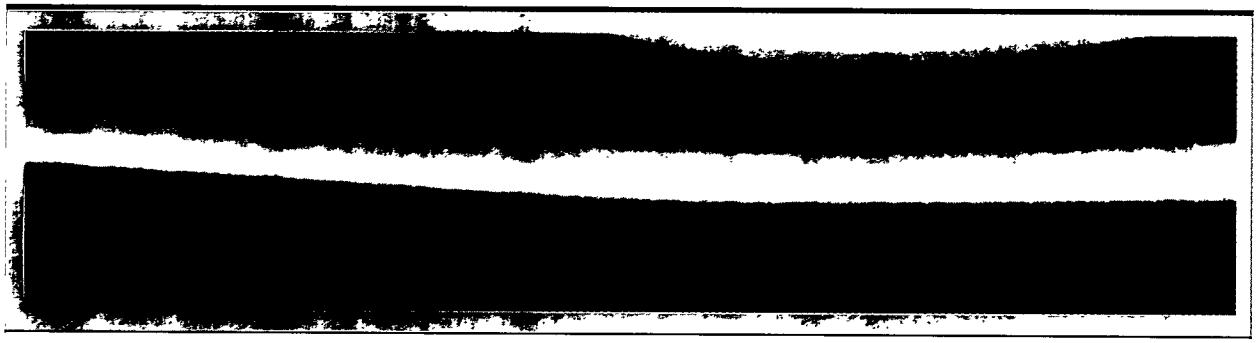


Figure 5. Reconstructed image from the experimental data taken in the surface-to-borehole configuration. The image is not quite coming into focus (spot in upper right) because of the frequency of operation and the in-situ soil conditions. The 1/DRR value in this case was approximately 0.017 m^{-1}, thus generating a poor image.

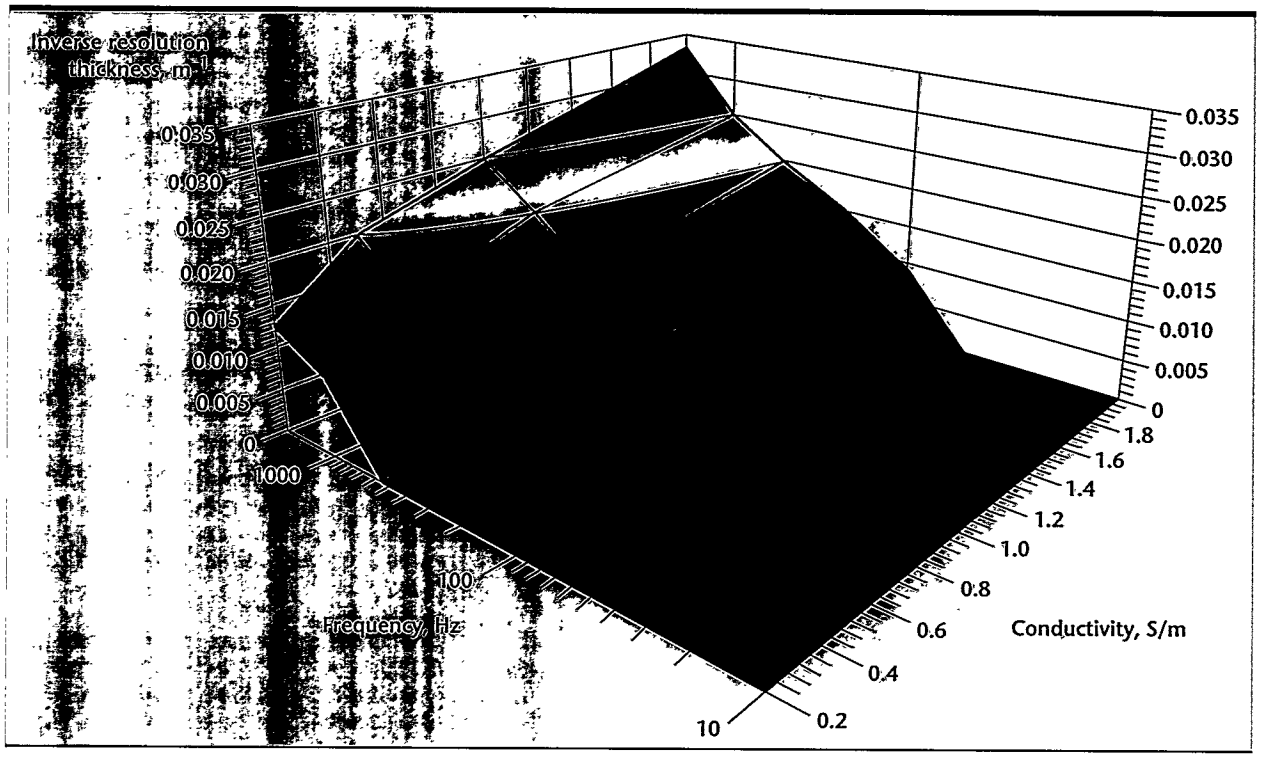


Figure 6. Relationship between the down-range-resolution (DRR) in the reconstructed images and the frequency of operation and soil conductivity. The inverse resolution thickness, 1/DRR, is plotted for clarity. The best resolution in this study occurred at the highest conductivity and frequency that was used (2 S/m, 2 kHz). 1/DRR values above 0.02 generated well-defined target centroids.

logs were used to aid in the processing of the experimental data, and knowledge of the soil was used to aid in that of the modeled cases.

Because of the low frequencies of operation, the reconstructive imaging codes used the magnetic near-field terms of the antenna systems, given by

$$H_r = \frac{\ell I_m \cos(\theta)}{2\pi\eta r^2}\left(1+\frac{1}{jkr}\right)e^{-jkr} \quad (10)$$

$$H_\theta = \frac{jk\ell I_m \sin(\theta)}{4\pi\eta r}\left(1+\frac{1}{jkr}-\frac{1}{(kr)^2}\right)e^{-jkr} \quad (11)$$

$$E_\theta = \frac{-jk\ell I_m \sin(\theta)}{4\pi r}\left(1+\frac{1}{jkr}\right)e^{-jkr} \quad (12)$$

where the propagation constant $k = (\omega^2\mu\varepsilon\text{-}j\omega\mu\sigma)^{1/2}$, and the "magnetic current" $I_m = jk\alpha^2 I_o \pi\eta/l$.

Due to the results from the experimental data set,[10] a parametric study was performed, using the EM modeling codes to examine the frequency band of interest, in connection with the various soil conductivities at these frequencies. The frequency range of interest was 10 Hz to 2 kHz, and soils with conductivities from 0.2 to 2.0 S/m were examined (see **Fig. 6**).

The procedure for the reconstructive imaging algorithms involves creating a single zero-padded copy of the input data, a frequency domain spatial filter, and the output image. Currently, the image corresponding to each transmitter position is stored separately, and a final image is generated by summing coherently over the set. In practice, this can be reduced to the sum image and the current estimate.

Processing and computer memory requirements are limited in scope, and the code runs on numerous platforms. Under 10 s is required to form a coherent image from a 20 × 20 matrix composed of 20 different transmitter positions and 20 different receiver positions.

Future Work

Future work could involve extrapolation of this algorithm to multi-2-D or to true 3-D. The issue of non-planar layers would still need to be addressed for more generic sites. The more general problem of the casing effects would need to be added to the near-field correction part of the expansion terms. The calibration factor to absolute units also needs to be addressed.

Acknowledgments

The authors would like to thank M. Wilt for his experimental efforts, and J. Mast for his feedback on algorithms.

References

1. Wang, Y. M., and W. C. Chew (1989), "An iterative solution of two-dimensional electromagnetic inverse scattering problem," *International Journal of Imaging Systems Tech.* **1**, pp. 100–108.

2. Lytle, R. J., E. F. Laine, D. L. Lager, and D. T. Davis (1979), "Cross-borehole electromagnetic probing to locate high-contrast anomalies," *Geophysics*, Vol. **44**, No. 10, pp. 1667–1676.

3. Laine, E. F. (1980), *Detection of Water-filled and Air-filled Underground Cavities*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-53127).

4. Lytle, R. J., and M. R. Portnoff (1984), "Detecting High-Contrast Seismic Anomalies Using Cross-Borehole Probing," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-**22**, No. 2, pp. 93–98.

5. Johansson, E. M., and J. E. Mast (1994), "Three-dimensional ground penetrating radar imaging using synthetic aperture time domain focusing," *Proceedings on Advanced Microwave and Millimeter Wave Detectors*, (SPIE), Vol. **2275**, July 1994.

6. Mast,J. E., and E. M. Johansson (1994), "Three-dimensional ground penetrating radar imaging using multi-frequency diffraction tomography," *Processing on Advanced Microwave and Millimeter Wave Detectors*, (SPIE), Vol. **2275**, July 1994.

7. Mast, J. (1993), *Microwave pulse-echo radar imaging for the nondestructive evaluation of civil structures*, Ph.D. thesis, University of Illinois—Urbana/Champaign, Urbana, Ill.

8. Burke, G. J., and A. J. Poggio (1981), *Numeric Electric Code (NEC)—Method of Moments*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-18834).

9. Nelson, S. D. (1994), "Electromagnetic modeling for ground penetrating radar using finite difference time domain modeling codes," *Proceedings on Advanced Microwave and Millimeter Wave Detectors*, (SPIE), Vol. **2275**, July 1994.

10. Wilt, M. H., W. Tseng, A. Becker, and M. Descz Pan (1994), *Fluid Injection Monitoring with a Borehole-to-Surface Electromagnetic Method, Field Activities at the UCB Richmond Field Station*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif.

akefield Modeling of Advanced Accelerator Components

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

Clifford C. Shang
Engineering Research Division
Electronics Engineering

George J. Caporaso
Inertial Confinement Fusion Program
Laser Programs

This report describes the development of wakefield post-processing tools to be used in conjunction with finite-difference time-domain (FDTD) electromagnetic codes for the modeling of accelerator components. Presently, wakefield analysis has been restricted to modeling 3-D fields of 2-D cylindrically symmetric structures using the azimuthal mode simulator code, AMOS. Our new post-processing tools are used in conjunction with the 3-D FDTD code, TSAR (for "temporal scattering and response") to model 3-D structures without imposing symmetry conditions. These tools will eventually be coupled to new 3-D time domain codes under development. Preliminary results are shown for a 3-D model of a beam position monitor used in high energy accelerators.

Introduction

In FY-95, we started development of tools to analyze the interaction of relativistic charged particle beams with accelerator components. The work is primarily motivated by the need for advanced modeling techniques to be applied to problems relevant to the design of electron beam accelerators for the advanced radiographic machine, ARM. In the ARM, a linear induction accelerator generates a high-current (~5 kA), relativistic (~20 MeV) electron beam for the generation of x rays for use in pulsed radiography applications. It is desirable to have multiple lines of sight on the target for 3-D tomographic reconstruction of a hydrodynamic event. However, multiple electron beam accelerators for each line of sight increase the cost of such a facility dramatically. Currently we are investigating whether it is possible to generate a long-pulse (200 ns - 1 μs) electron beam in a single accelerator, and split the beam into two beam transport systems using a "kicker" to modulate and spatially separate the beam into two electron beam pulses. Each of the two beams may be split again into additional beams in the downstream beam transport systems using additional "kickers," and then transported in curved beam lines to the target. This concept requires the development of new tools to aid in the design of the accelerator and the beam transport system. Specifically, the operation and design of the "kicker" and the development of curved beam lines in the beam transport system need to be addressed, which requires a full 3-D analysis. In addition, current accelerator codes designed for the particle physics community assume the beam current (space charge) is small, which requires that existing codes be modified to include space charge effects, since our applications involve currents of several kA. In this report, we will concentrate on a discussion of the "kicker" and present results of the modeling of a beam position monitor which is similar in design to the "kicker."

Progress

Kicker

The "kicker" is a structure designed to spatially separate an electron beam pulse into two separate electron beam pulses. The "kicker" consists of two pairs of nonlinearly tapered transmission lines enclosed in a cylindrical beam line. The tapers are designed to steer the relativistic electron beam (REB) into the correct trajectory for transport into a magnetic field region used to bend the beam into the downstream beam lines. 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 separate beam pulses, each beam pulse is directed into the appropriate beam line by the applied magnetic field. The magnetic field has opposite field directions on each side of the midplane separating the downstream beam transport systems, to bend each beam into the appropriate beam transport system. The "kicker" structure is shown schematically in **Fig. 1**.

Figure 2 is a photograph looking into the downstream end of a prototype "kicker," showing the external coaxial lines connecting one pair of transmission lines and the grounded connections on the other pair of transmission lines. The upstream end has a similar appearance, except that the transmission lines taper down to a smaller radius at that end. The strip transmission lines are approximately 120 cm long; the enclosure has an inner radius of 10 cm; the inner radius of the transmission line at the upstream end is 1.75 cm; and the inner radius of the transmission line at the downstream end is 3.28 cm. The analysis of the "kicker" can be broken down into two categories: the response of the REB to the externally applied electromagnetic field, and the wakefield generated in the structure by the passage of the REB. In this report we will discuss only the latter case of wakefield analysis.

Wakefield Analysis

A crucial area in the design of induction linac 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 break-up instabilities, with the resultant loss of beam quality at the target. A substantial amount of work in the analysis of short induction linac accelerator cavities has been done using the 2 1/2 - D FDTD code, AMOS, for the calculation of the wake potentials and wake impedances of such structures[1,2]. We are extending these calculations to more general 3-D accelerator structures that are typically many wavelengths long, in contrast to the localized wakefields generated by short accelerator gaps. While AMOS allows various multipole mode distributions to be launched, the geometry of the structure must be cylindrically symmetric. We use

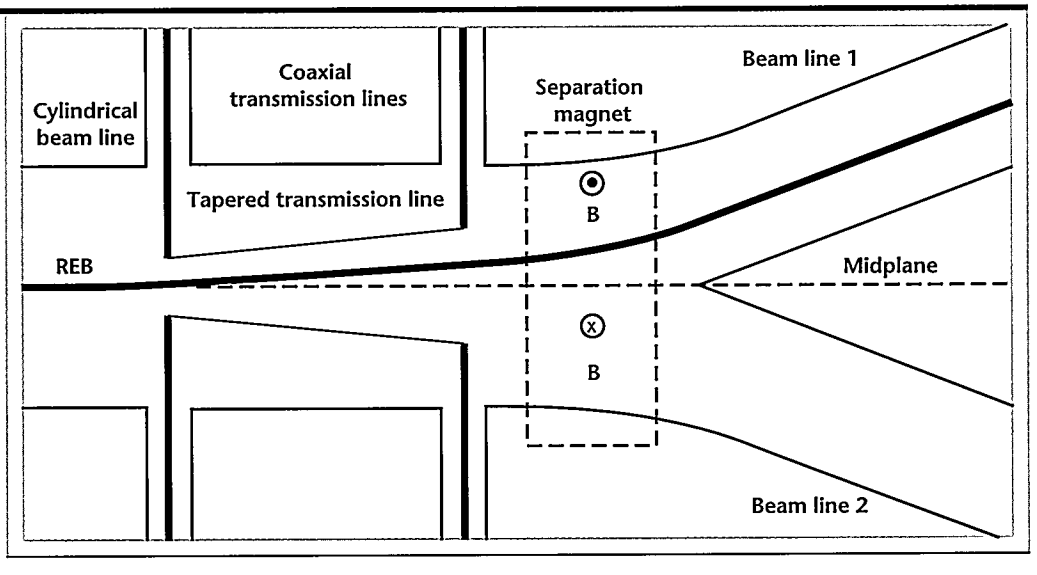


Figure 1. Schematic representation of the "kicker." For clarity, the second pair of grounded transmission lines is not shown. The "kicker" spatially separates the electron beam; and the magnetic field, with opposite polarities on each side of the midplane, bends the individual beams into the appropriate beam line. The dimensions of the structure are defined in the text.

TSAR, a 3-D FDTD code, to model more general 3-D accelerator structures and excitations that do not necessarily possess symmetries in the structures or excitations. The following is a summary of the relevant details of wakefield analysis.[1,2]

Consider a beam transport system consisting of a cylindrical drift tube with some perturbation in its structure, followed again by a cylindrical drift tube,

Figure 2. End-on view of the prototype "kicker," showing the external coaxial transmission lines connecting the top and bottom strip lines.

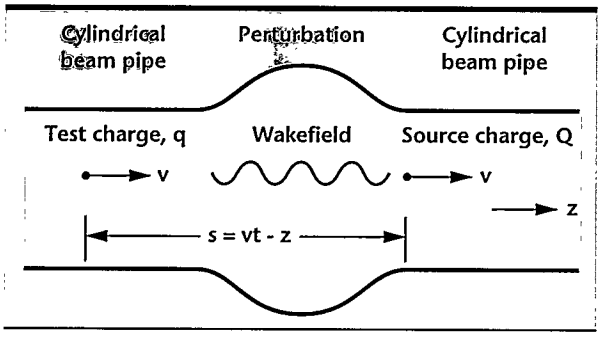


Figure 3. Geometry of a wakefield calculation. The source of the electromagnetic wakefield is the charge, Q. The spectral characteristics of the longitudinal and transverse forces are determined on a test charge, q, subjected to the wakefield.

as shown in **Fig. 3**. The perturbation is not necessarily cylindrically symmetric, and may consist of changes in the cross section, apertures in the wall, different material properties, or other features. To simulate a propagating point charge a series of point axial current sources are distributed appropriately in the transverse plane within the accelerator structure, and phased in the propagation direction, such that the speed of propagation is that of the electron beam ($v \approx c$ for relativistic electron beams). This effectively models a source of the form defined in **Eq. 1**.

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

where 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 equivalent 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. It should be noted that this type of excitation does not involve introducing charge into the system, and therefore, the effects of space charge are not included in the analysis. The purpose of the wakefield analysis is to determine the response of the electromagnetic structure due to the described excitation, on the assumption that the presence of the electron beam does not significantly alter the dispersion characteristics of the structure. This, of course, is not true for high-current, long-pulse electron beams, where for example, the effect of beam space charge waves may be of importance. This is, in fact, an area of current research. 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**.

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

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

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

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

$$\mathbf{W}(s) = \frac{1}{Q}\int_{-\infty}^{\infty}\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right)\Big|_{t=(s+z)/v} dz. \tag{4}$$

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. This vector impedance is given by

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

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.

Modeling

In this section, we discuss modeling techniques for the "kicker," and present modeling results for a similar structure known as a beam position monitor. 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. The "kicker" model is particularly difficult to model with cubical cells, since the structure is a long, narrow, cylindrical structure with small tapers. To model the structure effectively, a cell size of 2 mm was chosen to accurately represent only the gross characteristics of the structure. This cell size was not sufficient to model the external coaxial transmission lines connecting with the strip transmission lines. To incorporate the effect of the external coaxial transmission lines, a wire made of resistive material with a total resistance of 50 Ω was connected from each end of the driven pair of transmission lines to the enclosing cylindrical case. Approximately 20 cm of cylindrical pipe were added to each end of the "kicker" structure to model the entrance and exit beam lines. The beam pipe split was not modeled. In addition, to satisfy stability requirements for the FDTD field solver, the Courant 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. For our model, with a cell size of 2 mm, the time step is 3.33 ps. In determining the wake properties of this structure, it is important to resolve the lowest frequency that the electron beam may see. An estimate of the lowest frequency for this structure, can be found from the length of the strip transmission lines, which are approximately 120 cm. If we assume that the lowest resonance of the structure corresponds to the length of the strip lines being a quarter-wavelength, we find that the lowest frequency to be resolved is about 62.5 MHz, requiring about 10 000 time steps to resolve one RF period. Therefore, obtaining accurate low frequency spectra would typically require a run of several 100 000 time steps. We have successfully modeled the "kicker" for typically 20 000 time steps, but this was not sufficient to resolve the low frequency behavior. In addition, the high frequency behavior was compromised by the poor performance of the MUR radiation boundary conditions applied to the axial boundaries of the problem. Future improvements in the code will help us resolve these issues.

To gain a better understanding of the modeling of a structure similar in design to a "kicker," we modeled a beam position monitor (BPM) before attempting the long run times required for the full scale "kicker." The BPM is similar in design to the "kicker," and is used to monitor the position of beams in high energy accelerators.[3] The TSAR model of a BPM is represented schematically in **Fig. 4**. The radius of the beam pipe and the strip transmission lines is 1.65 cm, and the radius of the enclosure around the BPM is 1.9 cm. The four strip lines are 10 cm long, 1 mm thick, and each strip subtends an angle of 55°. All four strip lines were terminated at both ends with 25 Ω wire resistors.

The 25 Ω resistors were chosen based on a frequency domain calculation of the TEM impedance of a single strip transmission line in the enclosure that resulted in a 25-Ω characteristic impedance. Since the four strip lines are not strongly coupled, due to the large angular sector between them, 25 Ω is good estimate of the impedance. In this case, the cell size was 0.5 mm, corresponding to a time step of .83 ps, but the length of the strip transmission lines was only 10 cm, allowing us to sufficiently resolve the low frequency behavior by running for 16 000 time steps. The quarter-wavelength resonance in this case is about 750 MHz.

To excite the structure and maximize the effect of non-symmetric modes ($m \neq 0$) only one strip transmission line was excited by using point current radiators in the two gaps between the upper ($\phi_0 = \pi/2$) strip transmission line and the waveguide wall. Only the gap regions were excited, since this is the only place where axial image currents experience any discontinuity. About 25 cm of beam line were modeled on each side of the BPM to allow evanescent waves to decay before reaching the radiation boundary conditions imposed on the axial boundaries of the model. The MUR radiation boundary conditions do not behave well near the beam pipe cutoff, and this poses some problems that need to be resolved for the modeling of higher frequency wakefields, as discussed previously. Several diagnostics were used in the TSAR code run.

The primary diagnostic was a volume sensor diagnostic. This diagnostic writes out a single precision binary data file consisting of the vector field components at specific locations within the computational volume. The diagnostic has the capability to interpolate the calculated field values to a local Cartesian, cylindrical, or spherical coordinate system. In our case, we obviously wrote the field variables interpolated to a cylindrical coordinate system, whose axis was coincident with the axis of the BPM. We chose 16 uniformly spaced points in azimuth ϕ, only one point in radius at $r = 1.5$ cm, every two cells in z, and every four time steps, so we effectively have a cylindrical shell of data at a radius of 1.5 cm. These data files are typically about 600 MB in size for each vector field data set (600 MB for three electric field components and 600 MB for three magnetic field components). Each field component is azimuthally decomposed into the time history of a multipole amplitude at every axial slice, so we now have a multipole amplitude matrix $\psi_m^l(z,t)$ for all field components at a specific radius. With only 16 azimuthal locations, we can only extract the $m = 0$ monopole mode and $m = 1$ dipole mode with any confidence. Higher order modes will require higher resolution in azimuth, with the accompanying increase in disk space requirements. The multipole amplitude matrices are used in the computation of the wake impedance functions.

Another diagnostic is the planar slice diagnostic. In this diagnostic, a specific field variable on the requested plane is saved as a function of time. This is useful for making movies of the evolution of the field data as a function of time. It is also possible to take the temporal Fourier transform of the time history of the field variable at every point in the planar matrix to look at the distribution of the fields in the plane at a specific frequency. For the BPM calculation, two planes of data were generated. The first plane was a y-z plane at x=0, and the second plane was a x-y plane at a z location in the middle of the BPM. At every point in each plane a temporal Fourier transform was applied to each field variable to obtain a field pattern at specific frequencies. These field distributions, however, represent the total field components, as opposed to the multipole field component.

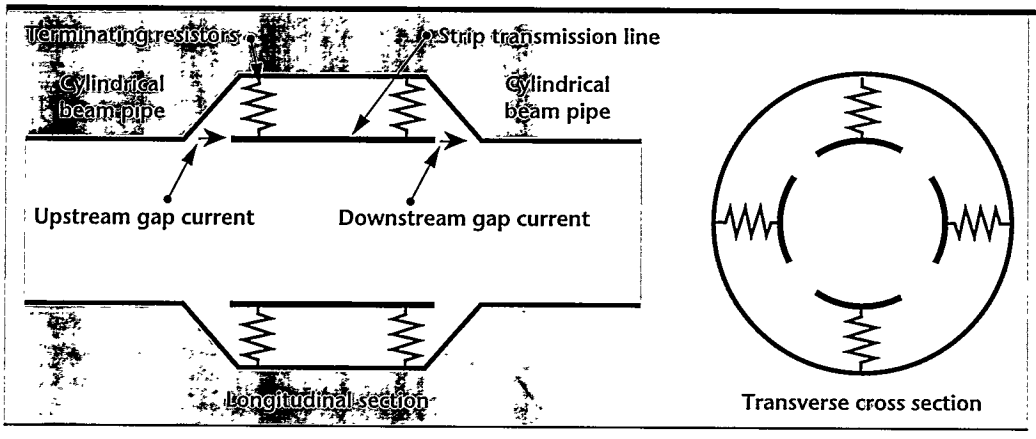


Figure 4. Schematic representation of the beam position monitor. For clarity, the second pair of strip lines is not shown in the longitudinal section. The downstream exciting current is phase-delayed by L/c, to account for the speed of the electron beam. The dimensions of the structure are defined in the text.

Figure 5 shows the distribution of transverse y-component of the wake electric field in the y-z plane at 750 MHz, which corresponds to the quarter-wavelength resonance of the strip transmission line. It is clearly seen that the electric field is maximum at the left gap and near zero at the right gap. Since the BPM can support TEM waves, the contribution to the transverse wakefield is distributed within the entire structure, including the gaps, while the longitudinal wakefield is concentrated in the gap region, since the BPM cannot support low frequency TM waves. **Figure 6** shows the distribution of the transverse y-component of electric field in the x-y plane at the center of the BPM. It is easily seen here that the transverse field is concentrated between the upper strip transmission line and the enclosure. Typical spectra of the longitudinal monopole and dipole impedances are shown in **Fig. 7**, and the transverse dipole impedance is shown in **Fig. 8**. The quarter-wave resonance at 750 MHz is clearly indicated in the longitudinal impedance plot in **Fig. 7**. The transverse dipole impedance shows a broader low frequency characteristic in **Fig. 8**, since this mode is distributed throughout the structure, as previously discussed. Also clearly evident in the plots are the strong contributions from the TM_{01} and TM_{11} cutoff frequencies at 7 GHz and 11 GHz, respectively. These strong contributions at the cutoff frequencies are a consequence of the poor performance of the MUR boundary conditions near cutoff.

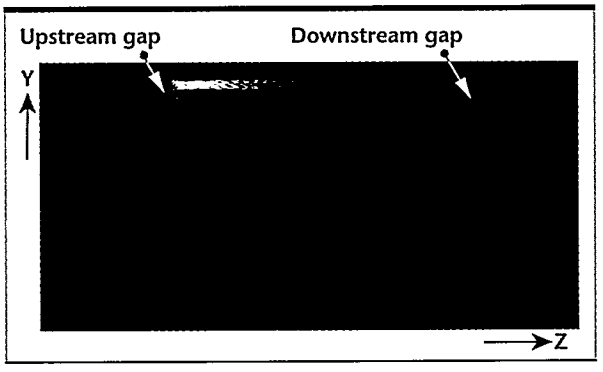


Figure 5. Distribution of the vertical component of wake electric field at 750 MHz on the x=0 plane. The quarter-wave resonance is clearly evident, with a maximum field at the left (upstream) gap and a minimum at the right (downstream) gap. The distributed nature of the transverse field is also evident.

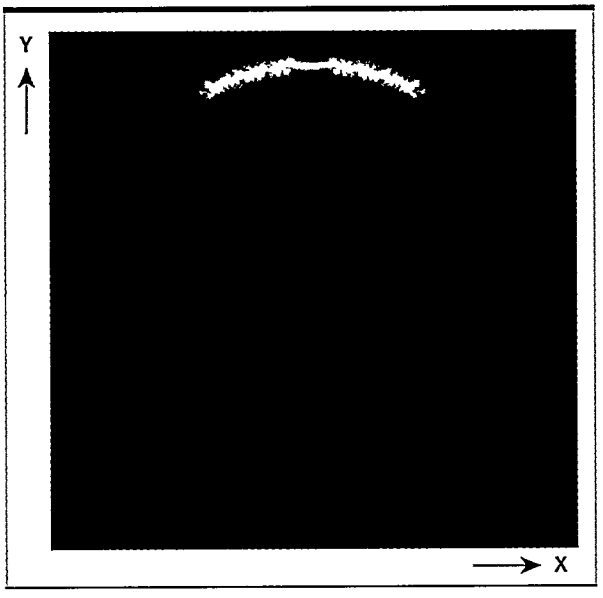


Figure 6. Distribution of the vertical component of the electric field on a plane at the center of the beam position monitor at 750 MHz, showing the contentration of field between the top strip transmission line and the enclosure.

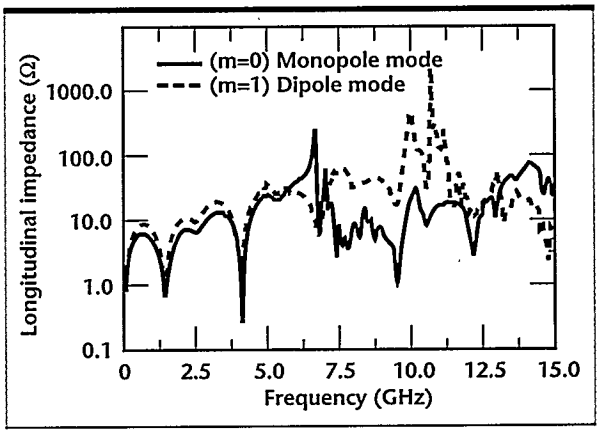


Figure 7. Longitudinal impedance spectra for the monopole and dipole modes. The peak near 750 MHz corresponds to the quarter-wavelength resonance of the strip transmission lines, and the dip near 1.5 GHz is due to the half-wavelength mode along the transmission lines. The peaks near 7 GHz and 11 GHz are due to the TM_{01} and TM_{11} beam line cutoff frequencies, respectively.

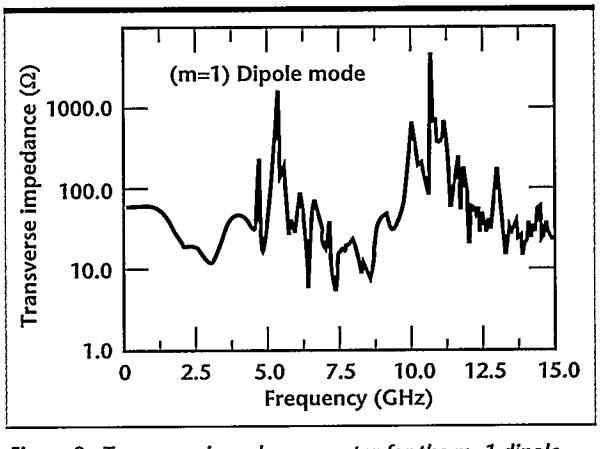


Figure 8. Transverse impedance spectra for the m=1 dipole mode, showing the broader low frequency spectral characteristics due to the support of TEM low frequency waves in the beam poition monitor. The peak near 11 GHz is due to the TM_{11} beam line cutoff.

The modeling of the BPM provides some insight regarding the wakefield properties of the "kicker." However, the effects of the taper need to be addressed. In addition, higher order transmission line modes can be supported by the "kicker," since the gap between the strip transmission lines and the outer enclosure are larger than in the BPM. Also, the effect of the external coaxial lines needs to be included.

Future Work

In this coming year, we will be focusing our efforts in several areas. To better model large, geometrically complex, 3-D volumes, the time domain electromagnetic codes will be modified to handle conforming, unstructured, nonorthogonal grids, and be ported to a massively parallel architecture. In addition, it is extremely important to include better absorbing boundary conditions to absorb out-going waves at the boundaries of the problem space. In the context of accelerator physics, this is important for extending our wakefield computations to higher frequencies. Material models also need to be improved to include dispersive, anisotropic, and nonlinear materials to better model realistic electrodynamic structures. We also plan to incorporate the wakefield post-processing tools into the new codes. Further, the effects of space charge need to be investigated for inclusion into beam transport codes.

Acknowledgments

The authors wish to acknowledge useful discussions with S. Nelson of the Defense Sciences Engineering Division, and S. Pennock of the Engineering Research Division.

References

1. DeFord, J. F., G. D. Craig, and R. R. McLeod (1990), *The AMOS Wakefield Code,* Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-102731).

2. DeFord, J. F., and G. D. Craig (1992), "Wake Potential and Impedance Calculations for Stiff Subrelativistic Beams," *Particle Accelerators* **37-38**, p. 111.

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

IGER: A New Generation of Computational Electromagnetics Tools

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 David R. Jackson
University of Houston
Houston, Texas

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

We have made progress on a multi-institutional project, EIGER, to bring a variety of spectral domain analysis methods into a single, integrated software tool set. In FY-95, the EIGER suite was extended to treat 2- and 3-D problems.

Introduction

The electromagnetic interactions generalized development project (EIGER) is a multi-institutional effort to bring a variety of spectral domain analysis methods into a single, integrated software tool set. Researchers from the Sandia National Laboratories, the U.S. Navy (NCCOSC/NRaD), Lawrence Livermore National Laboratory, and the University of Houston are collaborating to develop this package, which will have an unparalleled ability to take general purpose analysis methods and optimize the approach used for each specific application. 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, as we have already observed.

The object-oriented paradigm enables the abstraction of the basic underlying components of frequency domain electromagnetics, and yields an unprecedented flexibility. The key components and their roles are 1) elements, which are used to describe the geometry; 2) basis (expansion) functions, which interpolate the unknowns (fields) locally; and 3) operators, which are the underlying formulation of the physics used to propagate the energy or enforce fundamental principals.

This is in contrast to standard design procedures, where entire codes are developed around a single element with a specialized basis function for a specific operator. Although such tools can be effectively used 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 this 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 will 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. These 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 preprocessor then reads these elements, and assigns such features as material characteristics and excitation parameters, to generate the actual input for the physics code.

The EIGER suite was extended this year to 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. 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, research is underway to extend these capabilities to higher order elements to more accurately resolve the geometry.

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 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.

A specific basis set was incorporated this year to allow the dynamic form of Maxwell's equations to be solved at extremely low frequencies. These so-called loop and star basis functions are, respectively, divergence-free and nearly curl-free linear combinations of linear basis functions. While the underlying linear basis functions yield ill-conditioned matrices at low frequencies, the loop and star set stabilize this effect. The latter is evidenced in currents circulating about isolated nodes in the mesh, clearly unphysical at low frequencies.

Operators

An electromagnetic operator is a mathematical construct that relates the field at a point to the sources which produced the field. Since this year's effort 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 often known as the method of moments. This year a variety of integral operators was incorporated into EIGER to treat the following boundary conditions: 1) perfect electric conductors; 2) conductors coated with complex impedance materials; and 3) homogeneous dielectrics (penetrable materials). In addition to the dynamic forms of the operators, a rigorous treatment for static fields has been developed. This treatment allows application in both electrostatics and magnetostatics.

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. This year a general layered-media Green's function was developed to treat a variety of problems, including multi-layered electronic circuit boards, high powered optical mirrors, and geophysical problems.

Future Work

We will continue to develop each of the areas outlined above to reach broader classes of problems with a greater diversity of tools. Specific examples of future development are briefly noted here.

The element modules of the code are presently being extended to include higher order elements. These elements can be of arbitrary order, and will more accurately resolve the local variations of the geometry. Also, volume elements will be incorporated for treating inhomogeneous material regions.

Along with the more accurate elements, higher order (smoother) basis functions are being developed to allow better resolution of the local fields with fewer elements. A complement to the higher order or smoother basis functions is an effort to directly incorporate known singularities for certain edge conditions directly into the basis set, rather than numerically solving for the singular behavior. Also, extensions can be made to accurately treat conductor losses directly in the solution.

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 a model with the most appropriate combinations. During FY-96 we will develop a mechanism to automatically adapt these combinations, as appropriate. Our long term adaptive gridding goal for changing both elements and basis functions will significantly impact user interactions. Efforts to ensure converged results will become mostly automated, and modelers will need less intimate knowledge of EIGER's rapidly growing set of algorithms.

Alternative integral operators will continue to be added to the EIGER suite. In addition, differential operators, which are the underlying mathematics associated with finite element modeling, are currently being integrated into the code.

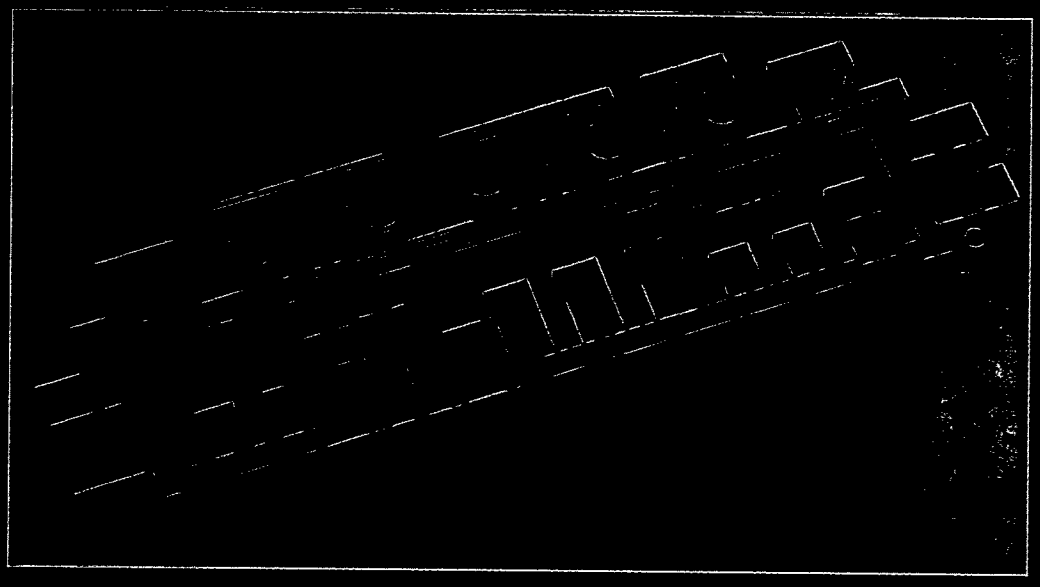
Finally, additional Green's functions will continue to be cast into the EIGER format. Specifically, the following classes of analysis are presently being considered: 1) general periodic Green's functions for cavities and arrays; 2) general symmetry Green's functions for reflection and rotational symmetry; and 3) additional extensions to the layered-media Green's functions.

Gerald L. Goudreau, Thrust Area Leader

The Computational Mechanics thrust area sponsors investigation into the underlying solid, structural, and fluid mechanics and heat transfer necessary for the development of state-of-the-art general purpose computational software. The scale of computational capability spans office workstations, departmental computer servers, and Cray-class supercomputers. Our codes have achieved world fame through our broad collaborators program, in addition to their strong support of ongoing Lawrence Livermore National Laboratory (LLNL) programs. Several technology transfer initiatives have been based on these established codes, teaming LLNL analysts and researchers with counterparts in industry, extending code capability to specific industrial interests of casting, metal forming, and automobile crash dynamics. The next-generation solid/structural mechanics code, ParaDyn, is targeted toward massively parallel computers, which will extend performance from gigaflop to teraflop power.

Our work for FY-95 is described in the following seven articles: 1) Biomechanical Modeling of the Human Knee; 2) A Hyperelastic Shell Model for Elastomers and Biological Soft Tissues; 3) Investigation of the Radiative Properties of Ceramic Casting Materials; 4) Performance of NIKE3D with PVSOLVE on Vector and Parallel Computers; 5) GEMINI Re-Certification and Development; 6) Material Model Development for Thermal Forming of Anisotropic Materials; and 7) Multilength-Scale Modeling: Crystal-Plasticity Models in Implicit Finite Element Codes.

Computational Mechanics



2

2. Computational Mechanics

Overview
Gerald L. Goudreau, Thrust Area Leader

Biomechanical Modeling of the Human Knee
Dale A. Schauer, Jeffrey A. Weiss, Bradley N. Maker, and Michael A. Puso 2-1

A Hyperelastic Shell Model for Elastomers and Biological Soft Tissues
Michael A. Puso 2-7

Investigation of the Radiative Properties of Ceramic Casting Materials
James D. Maltby, Barbara T. Kornblum, Charles S. Landram, and Charles W. Dingus 2-13

Performance of NIKE3D with PVSOLVE on Vector and Parallel Computers
Bradley N. Maker, Qin Jiangning, and Duc T. Nguyen 2-17

GEMINI Re-Certification and Development
Stan Bumpas, Michael A. Gerhard, and Stanley C. Sommer 2-25

Material Model Development for Thermal Forming of Anisotropic Materials
Roger W. Logan 2-31

Multilength-Scale Modeling: Crystal-Plasticity Models in Implicit Finite Element Codes
David D. Sam 2-39

Biomechanical Modeling of the Human Knee

Dale A. Schauer, Jeffrey A. Weiss, Bradley N. Maker, and Michael A. Puso
New Technologies Engineering Division
Mechanical Engineering

A 3-D finite element model of the human knee has been developed for examining the contribution of the main ligamentous structures of the knee to resisting normal and injurious loads. The model includes detailed geometric representations of the knee region, specifically the femur, tibia, fibula, menisci and the four major knee ligaments. New material models were developed to simulate soft tissue behavior and allow for application of initial tension to the soft tissue structures. The model was used to simulate both normal and injurious knee joint kinematics, including flexion-extension, varus-valgus rotation, tibial axial rotation, hyperextension and anterior-posterior displacement. The model has proven to be extremely useful for understanding both normal and pathologic knee joint kinematics as well as the guiding and restraining role of the four major ligaments. Future work will extend this model to study sports injuries and the effects of muscle loading, and to evaluate new total joint replacement designs.

Introduction

The importance of healthy joints cannot be overemphasized. Joints such as the knee and hip are the primary load bearing joints of the body. Their function can be severely compromised by articulation surface damage or ligament injury. Over 14 million joint dislocations and sprains were recorded between 1985 and 1988.[1] The associated cost of these injuries was over $125 billion for 1988. Orthopedic implant figures for the U.S. (1988) indicate that over 220,000 total joint replacement procedures were performed, at a cost of one billion dollars.[2] Of these, 42% were for the knee. To reduce costs, an approach that can improve the diagnosis, treatment and prevention of joint injuries is crucial.

The present implant market is served by an inadequate product, resulting in surgical replacement of implants during the natural life of the recipient. Implant recipients can anticipate anywhere from 1 to 6 revision surgeries by age 75. A 10% extension in implant life would reduce the total annual U.S. health care costs by millions of dollars.

In recent years, advances in the fields of constitutive modeling, computational mechanics, and computer hardware have enabled the construction of finite element models of entire joints. Many of these advances are due to research at Lawrence Livermore National Laboratory (LLNL), and have been incorporated into our own finite element codes for nonlinear solid mechanics analysis.

The 3-D motion of the knee joint is dictated by articular surface geometry, the anatomical arrangement of the ligaments and their insertions to bone, the menisci, soft tissue material properties, and muscle activity. Motion is the result of a move towards an equilibrium of internal and external forces. External forces include those due to muscle action, gravity, and acceleration which are applied to the femur and tibia, while internal forces are a result of reaction forces at the articulating surfaces and resistance from ligamentous and capsular restraints.

Past investigators have developed mathematical models of the knee joint based upon equilibrium of the restraining forces of ligamentous and capsular restraints, joint reaction forces, and external loading of the knee.[3–7] These models have sought to predict the relative motion of the femur and tibia. However, large simplifications have been introduced in these simulations. More recent studies of joint mechanics have used the finite element method.[8–11] These studies have answered important basic science and clinical questions, and have provided some of the necessary methods for 3-D joint modeling. However, all the above attempts have fallen short in

several regards: unrealistic geometry, unrealistic material properties for the ligaments, and the lack of a 3-D representation for the ligaments.

One must be capable of correctly modeling normal healthy anatomy and characterize the total joint behavior before moving on to study implants, evaluate joint mechanics and identify design improvements. The present work has developed a completely 3-D model of the knee, with continuum representations for the ligaments and their material properties, and full contact between all parts of the model. This is an important and crucial initial step in the development of the capability to address joint and implant problems through computational modeling.

Progress

Knee Joint Model

We have succeeded in developing a working finite element model of the knee. Starting with digitized surfaces of the major bones (femur, tibia and fibula) and soft tissues (anterior and posterior cruciate ligaments, tibial and fibular collateral ligaments, and lateral and medial meniscus) of the knee, high-resolution hexahedral finite element meshes were constructed using TrueGrid[12]. The distal femur and proximal tibia were especially challenging because of the multi-directional curvature on the femoral condyles and tibial plateaus. **Figure 1** shows the entire mesh, consisting of 23,000 brick elements. To reduce the computational cost of the analysis, the femur, tibia and fibula were modeled using the rigid material model in NIKE3D[13,14].

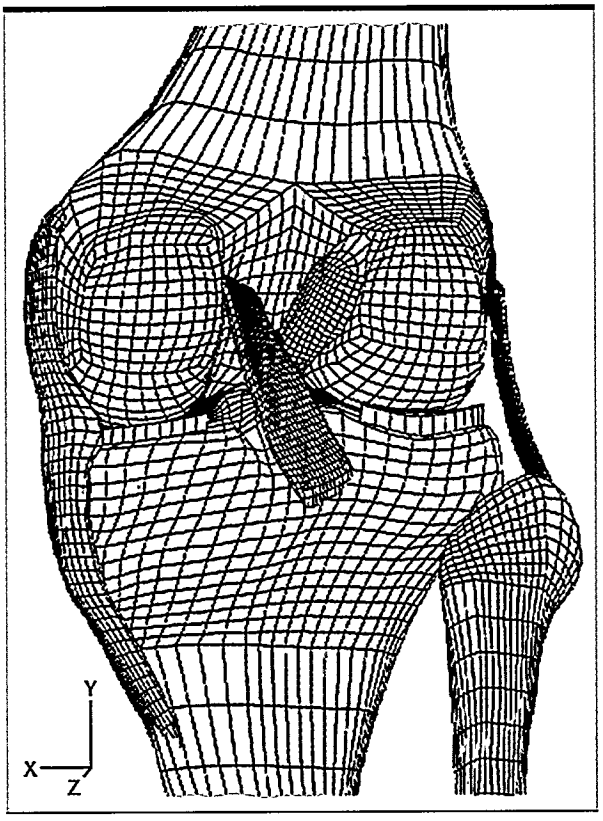


Figure 1. Finite element mesh of knee, including the femur, tibia, fibula, the lateral and medial meniscus, the anterior and posterior cruciate ligaments, and the tibial and fibular collateral ligaments.

Material Model Development

The soft tissues of the knee are anisotropic and inhomogeneous. They undergo large, nearly incompressible deformations, and have some viscoelastic behavior. Most biological soft tissues are under some residual stress *in vivo*. These stresses can be relieved by selective cutting of the tissue or its attachments. The purpose of these residual stresses can vary. For ligaments, it serves to provide initial joint stability in the absence of externally applied loads. Because of the nonlinear, upwardly concave tensile stress-strain behavior of these tissues, failure to include residual stresses in computational mechanics models can result in large errors in stress calculations. The ligaments are under an initial tension *in vivo* as well. Therefore, it was necessary to provide for this behavior. To accommodate these important material characteristics, special constitutive models were developed[15,16] and implemented in NIKE3D. To model the fiber-reinforced nature of the ligaments and menisci, a transversely isotropic hyperelastic strain energy was developed. This allowed for an exponential "toe region," corresponding to the uncrimping of the collagen fibers, as well as a linear region, after which the fibers were assumed to be straightened. The fiber and matrix contributions to the stress were model individually, using a Mooney-Rivlin model to represent the matrix behavior.

To allow for nearly- and fully-incompressible behavior without locking, the model was implemented using a three-field solid element original developed by Simo *et al.*[17] This element has separate interpolation for the displacement, pressure and volume ratio. Because tri-linear shape functions were used for the displacement fields, the pressure and volume ratio were constant over an element.

Tissue structures such as ligaments and tendons can be relatively thin and can undergo large bending, shear and membrane strains *in vivo*. A shell element was also developed to allow the use of the hyperelastic model for thin membrane like structures.[18]

The 3-D shell finite element is a computationally efficient alternative to solid hexahedral elements for thin ligaments that undergo large bending shear and membrane strains. For example, a finite element mesh of the medial collateral ligament would require many elements over the thickness and length to accurately capture its behavior under complex loading. Because shell elements enforce the zero normal stress condition explicitly, and provide Gauss point sampling of the stress over the thickness, far fewer elements are required, providing a computational savings. The implementation in NIKE3D follows the degenerated solid treatment presented by Hughes and Carnoy,[19] with extensions to allow for improved treatment of incompressibility and a four node quadrilateral element formulation which avoids shear lock. This method proved to be very stable.

We have provided a method for tensioning finite element models of biological soft tissues. This allows stabilization of models of joint mechanics such as the knee, by tensioning of the ligamentous structures. We developed constitutive material models that can be used to simulate soft tissue behavior under load. Both of these developments are necessary to enable realistic simulation of joint mechanics, which includes both soft tissue and rigid bone structure.

Simulations of Knee Joint Mechanics

To assess the model's ability to represent normal knee joint kinematics, a series of simulations were carried out. Model inputs, including applied motions and material properties, were taken from the biomechanics literature. The specific motions to simulate were chosen from a knowledge of normal joint kinematics. This knowledge was augmented through two dissection courses using human cadaveric knees at the Orthopedic Biomechanics Institute in Salt Lake City, Utah. In all cases, the ligaments of the knee, as shown in **Fig. 2**, were first pretensioned to the appropriate levels, using a procedure developed during the course of this research project.[20] All degrees of freedom were fixed for the tibia. Various combinations of motions were then simulated by specifying appropriate values for the six kinematic parameters for the femur over time.

To test the model's ability to simulate a normal motion that should generate minimal stress in the ligaments, knee flexion/extension was simulated, and is shown in **Fig. 3**. As the ligaments were first pretensioned, one can observe a slight shift of the femur as the bones move into a stable position. Since each ligament has a variable cross sectional area, the smaller areas produce higher ligament pretension stresses, which show up in the stress color fringe contours. The knee was flexed to ninety degrees and then back to full extension. By examining the effective stress fringe contours of the soft tissue, the role of the collateral ligaments, which are on the sides of the knee, is clearly seen. The model demonstrates a completely understandable stress variation for each ligament as the knee is flexed and extended. For example, the medial collateral ligament shows higher stresses on the side that is stretched during flexion. As the knee is flexed more, the high stresses concentrate in an ever narrower strip along the outer region of tissue that is being stretched the most.

Figure 2. The knee ligaments (a) before and (b) after pretensioning to provide initial joint stability.

To test the model's ability to predict the potential for injurious loadings to cause knee ligament injury, varus-valgus rotation, tibial axial rotation, anterior-posterior translation and hyper-extension were also simulated. **Figure 4** shows the knee during varus-valgus motion. Examination of the knee model response clearly indicates that the collateral ligaments are the primary restraining structures. Reasonable stress distributions were obtained for each of the simulations. The results correlated well with the sparse experimental data that were available in the literature. The model was able to predict injurious loadings of the medial collateral ligament and anterior cruciate ligament for varus-valgus rotation and anterior-posterior tibial translation, respectively. For the hyperextension simulation, injurious stress levels were calculated for the posterior cruciate ligament.

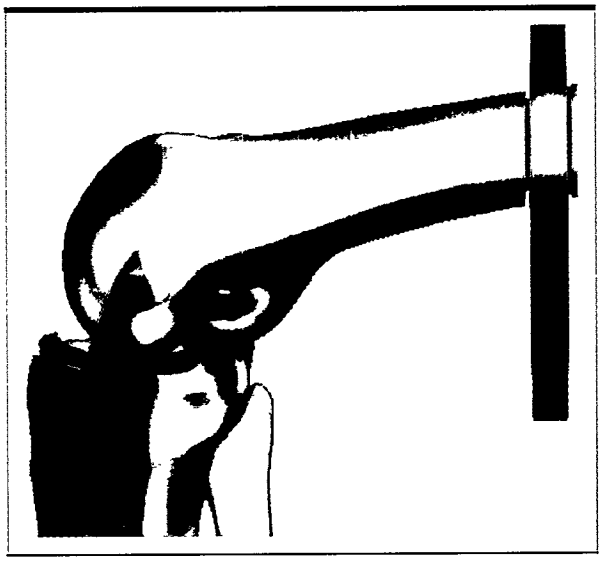


Figure 3. Knee model in flexion/extension, demonstrating the role of the collateral ligaments located on the sides of the knee.

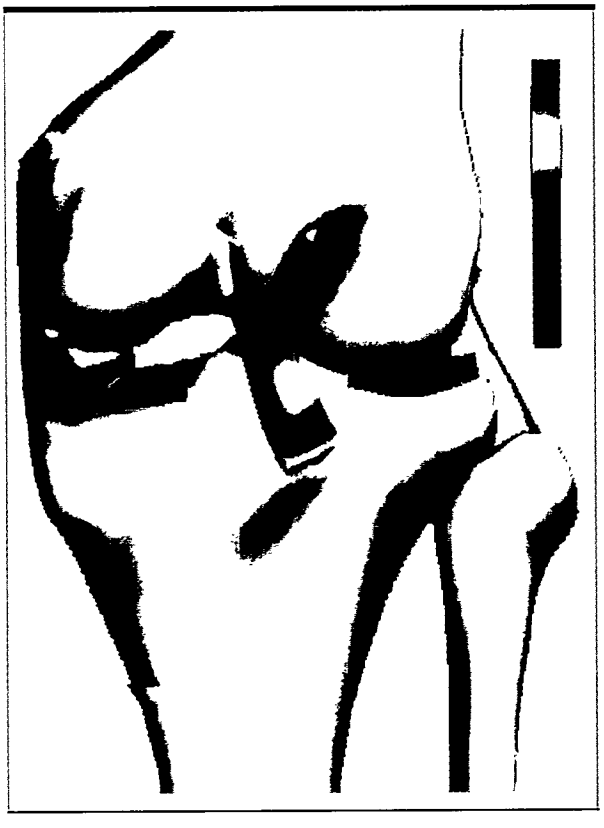


Figure 4. Simulation of knee varus-valgus rotation, verifying that the collateral ligaments are the primary restraining structures.

We have created a detailed 3-D model of the human knee joint which included the four major ligaments and the menisci. For the first time, we have simulated knee motion in which the contribution of the knee ligaments to joint stability and joint kinematics can be assessed directly by the stress distributions within these structures. Application of this approach to specific joints will result in a better understanding of normal joint mechanics, and alterations that occur due to injury or disease.

Future Work

In this past year, we have established LLNL as a leader in biomechanics computational research. However, much work still remains before we will be able to address common joint and orthopedic problem areas in a routine manner. We need to reduce the time required to produce a quality mesh of a joint including the soft tissue. More time and effort are needed to better characterize the computational conditions of a healthy knee. The simulated knee mechanics must be compared with experimental data to validate the knee model. Both the meniscus and cartilage provide shock absorption and smooth joint kinematics. The meniscus has circumferential fiber reinforcement, undergoes large deformation, is poroelastic and nearly incompressible. The meniscus must be more closely approximated before we can fully characterize the knee mechanics of a normal knee. We need to develop a material model for the cartilage in the knee that has poroelastic behavior to provide a representative transfer and distribution of loads across the joint. To better simulate dynamic motion, NIKE3D needs a dynamic extension of the rigid body capability. Several total knee replacement designs need to be modeled and their mechanics evaluated.

Ultimately, if this method is properly applied, it offers the possibility to develop implant design improvements that achieve a close match to a normal healthy joint.

References

1. Knowledge Enterprises, Inc., (1994-5), *Bare Bones*, Vol. 3(4), Vol. 4(4,5).

2. Pramer, A., S. Furner, and D. P. Rice (1992), *Musculoskeletal conditions in the United States*, American Academy of Orthopedic Surgeons, Park Ridge, IL.

3. Andriacchi, T. P., R. P. Mikosz, S. J. Hampton, and J. O. Galante (1983), "Model studies of the stiffness of the human knee joint," *J Biomech* **16**, pp. 23–29.

4. Blacharski , P. A., J. H. Somerset, and D. G. Murray (1975), "A three-dimensional study of the kinematics of the human knee," *J Biomech* **8**, pp. 375–384.

5. Blankenvoort, L., R. Huiskes, and A. de Lange (1988), "The envelope of passive knee joint motion," *J Biomech* **21**, pp. 705–720.

6. Crowninshield, R., M. H. Pope, and R. J. Johnson (1976), "An analytical model of the knee," *J Biomech* **9**, pp. 397–405.

7. Wismans, J., F. Veldpaus, and J. Janssen (1980), "A three-dimensional mathematical model of the knee joint," *J Biomech* **13**, pp. 677-686.

8. Chand, R., E. Haug, and K. Rim (1976), "Stresses in the human knee joint," *J Biomech* **9**, pp. 417–422.

9. Shirazi-Adl, A. (1989), "Nonlinear finite element formulation of wrapping ligaments," *ASME Adv in Bioeng*, BED **15**, pp. 93–94.

10. Tissakht, M., R. Farinaccro, and A. M. Ahmed (1989), "Effect of ligament attachments on the response of the knee menisci in compression and torsion: A finite element study," *Trans 35th Annual ORS*, p. 203.

11. Spilker, R. L. (1985), "Finite element models in composite materials," *ASME Adv in Bioeng* BED, pp. 37–38.

12. XYZ Scientific Applications, Inc., TrueGrid™, Livermore, California.

13. Maker, B. N., R. M. Ferencz, J. O. Hallquist (1990), *NIKE3D: A nonlinear, implicit, three-dimensional finite element code for solid and structural mechanics*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-MA-105268).

14. Maker, B. N. (1994), "Rigid bodies for metal forming analysis with NIKE3D," *Int J Num Methods Eng* **37**, pp. 1–18.

15. Weiss, J. A., B. N. Maker, and S. Govindjee (1995), *Finite element implementation of incompressible, transversely isotropic hyperelasticity*, U.C. Berkeley Report #UCB/SEMM-95/07 (in review, Computer Methods in Applied Mechanics and Engineering).

16. Weiss, J. A., B. N. Maker, and D. A. Schauer (1995), "Treatment of initial stress in hyperelastic finite element models of soft tissues," *ASME Bioengineering Conference* BED-**29**, pp. 105–106.

17. Simo, J. C., R. L. Taylor, and K. S. Pister (1985), "Variational and projection methods for the volume constraint in finite deformation elastoplasticity," *CMAME*; **51**, pp. 177–208.

18. Puso, M. A., J. A. Weiss, B. N. Maker, and D. A. Schauer (1995), "A transversely isotropic hyperelastic shell finite element," *Proc. ASME Bioengineering Conference* BED-**29**, pp. 103–104.

19. Hughes, T. J. R., and E. Carnoy (1983), "Nonlinear finite element shell formulation accounting for large membrane strains," *Computer Meth Appl Mech Engng* **39**, pp. 69–82.

20. Weiss, J. A. (1995), "A model for the material behavior of ligaments and tendons," in review, *ASME J Biomech Eng.*

Hyperelastic Shell Model for Elastomers and Biological Soft Tissues

Michael A. Puso
Defense Technologies Engineering Division
Mechanical Engineering

We have implemented a constitutive law developed specifically for biological soft tissues. We report progress made using a shell model for the medial collateral ligament in the high fidelity knee model.

Introduction

The material behavior of biological soft tissues is typically modeled using incompressible transversely isotropic hyperelastic constitutive laws. Furthermore, tissue structures such as ligaments and tendons can be relatively thin and can undergo large bending, shear and membrane strains *in vivo*. Elastomers such as hevea rubber, polyisobutylene and neoprene are also incompressible and hyperelastic. Typically, analysts use transversely isotropic hyperelastic composite models for analyzing tires,[1] since the treads provide a stiffening effect in the circumferential direction. Three dimensional shell finite elements are a computationally efficient alternative to solid elements for these types of problems. Problems due to parasitic shear lock, incompressibility and through thickness stress gradients are handled more efficiently by shell elements, allowing much coarser meshes to achieve accurate solutions. Because most of the research on shell element technology has been devoted to elasto-plastic materials, a new theory was developed in this work for handling specific applications to large strain hyperelastic materials. The constitutive law[2] developed specifically for biological soft tissues has been implemented in this shell model and used to model the medial collateral ligament in the high fidelity knee model. Other strain energy formulations have also been implemented for modeling elastomers.

Progress

Shell Kinematics

The usual shell element kinematics are based on the following interpolation of element geometry using a four-node bilinear shell element,

$$\mathbf{x} = \sum_{i=1}^{4} N_i(\xi,\eta)\left(\bar{\mathbf{x}}_i + z(\zeta)\bar{\mathbf{x}}_i\right) \tag{1}$$

where $\mathbf{x} = \mathbf{x}(\xi, \eta, \zeta)$ is the global position vector of a point in the shell in terms of the element natural (or local) coordinate system (**Fig. 1**); $\bar{\mathbf{x}}_i$ is the position of the reference surface for the i^{th} node; and $\hat{\mathbf{x}}_i$ is the fiber normal unit vector, which locates points through the thickness at node i. Also in **Eq. 1**, $N_i(\xi,\eta)$ are the bilinear shape functions and $z(\xi)$ is a linear function of ξ through the thickness. An isoparametric formulation is used, such that the displacement vector is of the same form as in **Eq. 1**. The incremental nodal displacement vector $\Delta\boldsymbol{d}$ is composed of nodal translations and nodal rotations at the i^{th} node, such that

$$\Delta\boldsymbol{d} = \begin{Bmatrix} \Delta\bar{\boldsymbol{u}} \\ \Delta\hat{\boldsymbol{u}} \end{Bmatrix} . \tag{2}$$

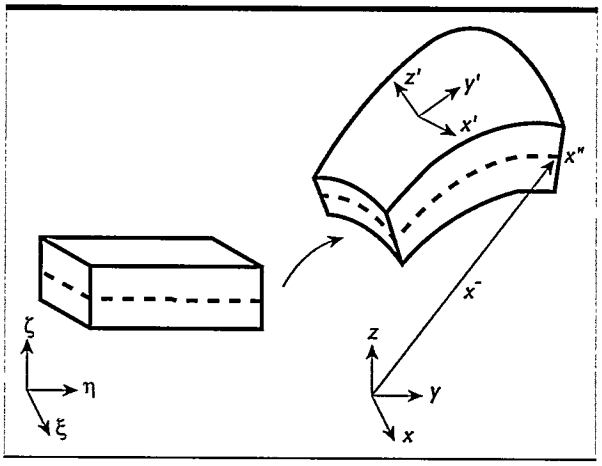


Figure 1. Local element coordinate system ξ,η,ζ; global coordinate system x,y,z; and convected coordinate system x', y', z'.

In the following development, stress and strain tensors are referred to the orthonormal convected basis parallel to the shell lamina (**Fig. 1**).

Hu-Washizu Assumed Strain Method

Most shell element formulations are co-rotational, and are based on small strain, large displacement kinematics; i.e., the YASE shell[3]. Available large strain shell models[4,5] are formulated for incremental constitutive laws. Because hyperelastic materials often experience large strains, and the stress can be calculated from the gradient of a strain energy function W, a novel shell formulation was sought.

In order to alleviate the shear locking experienced by low order finite element interpolations, a mixed field or assumed strain method is necessary. The following potential functional is used to develop the Euler-Lagrange equations

$$\int\left(\frac{1}{\bar{J}}W(\bar{\boldsymbol{F}})+\boldsymbol{\sigma}\bullet\left(\bar{\boldsymbol{F}}^{-1}-\boldsymbol{F}^{-1}\right)\right)dA\,dz+\Pi_{\text{ext}}=0 \qquad (3)$$

where $\boldsymbol{F}=\nabla\boldsymbol{x}$ is the gradient of deformation of **Eq. 1** with respect to the undeformed reference configuration; $\bar{\boldsymbol{F}}$ is the assumed deformation gradient, $\bar{J}=\det(\bar{\boldsymbol{F}})$; dA is the element of area along the surface convected frame; dz is the element of thickness perpendicular to the surface convected frame; σ is the independent stress field; and Π_{ext} is the potential energy of the applied loads. **Eq. 3** is a large strain version of the Hu-Washizu potential function[6]. Using the current configuration as the fixed reference frame, and taking variations of **Eq. 3** with respect to the displacement field, the assumed strain field and the stress field provide the Euler-Lagrange equations, which weakly enforce strain compatibility, the constitutive law and momentum balance. Taking the appropriate steps, the following weak form of the equilibrium equations results:

$$\int\left(\bar{\boldsymbol{F}}\frac{\partial W}{\partial\bar{\boldsymbol{E}}}\bar{\boldsymbol{F}}^{T}+p\,\boldsymbol{1}\right):\delta\bar{\boldsymbol{L}}\,dA\,dz+\delta\Pi_{ext}=0 \qquad (4)$$

where $\bar{\boldsymbol{L}}$ and $\bar{\boldsymbol{E}}$ are the assumed forms of the velocity gradient and Green strain; p is the undetermined Lagrange multiplier due to the incompressibility constraint on $\bar{\boldsymbol{F}}$; and $\boldsymbol{1}$ is the identity tensor. The assumed velocity gradient is an interpolation of the gradient of the displacement field $\boldsymbol{L}$[7] such that,

$$\bar{\boldsymbol{L}}=\bar{\boldsymbol{B}}\Delta\boldsymbol{d} \qquad (5)$$

where $\bar{\boldsymbol{B}}$ is a modification of the usual strain displacement matrix $\boldsymbol{B}$. The assumed strain field is chosen such that it allows a constant shear strain field in bending without shear lock. From the identity $\bar{\boldsymbol{F}}=\int\bar{\boldsymbol{L}}\,\bar{\boldsymbol{F}}dt$, the following recursion relation provides exact updates of the assumed deformation gradient from time t_n to t_{n+1}, for constant nodal velocities,

$$\bar{\boldsymbol{F}}_{n+1}=\bar{\boldsymbol{F}}_{n}+\bar{\boldsymbol{L}}_{n}\bar{\boldsymbol{F}}_{n}\quad . \qquad (6)$$

Because the shell thickness is not a nodal degree of freedom from **Eq. 2**, the through thickness strain $(\bar{\boldsymbol{L}}_n)_{33}$ is undetermined from **Eq. 5**. Calculation of this strain is made to exactly satisfy the incompressibility constraint such that,

$$\det\left(\bar{\boldsymbol{F}}_{n+1}\right)=1\quad . \qquad (7)$$

Taking the variation of **Eq. 7** produces the result trace $(\delta\bar{\boldsymbol{L}})=0$. It is then seen that $p\boldsymbol{1}:\delta\bar{\boldsymbol{L}}=p$ trace $(\delta\bar{\boldsymbol{L}})=0$, and the pressure term in **Eq. 4** vanishes leaving a generalized displacement form for the equilibrium equation,

$$\int\bar{\boldsymbol{F}}\frac{\partial W}{\partial\bar{\boldsymbol{E}}}\bar{\boldsymbol{F}}^{T}:\delta\bar{\boldsymbol{L}}\,dA\,dz+\delta\Pi_{ext}=0\quad . \qquad (8)$$

The pressure can then be calculated by assuming a zero normal stress through the thickness. **Equation 8** can be exactly linearized to provide a consistent tangent formulation.

Because large membrane strains typically occur, the shell will often experience large amounts of thinning. Shell thickness cannot be used as a degree of freedom since ill conditioning would occur for thin shells. The displacements are linearly interpolated over the thickness by the function $z(\zeta)=h/2(1+\zeta/2)$, where h is the shell thickness. By varying the quantity h, a constant through thickness strain $(\boldsymbol{L})_{33}$ results. The strain at time t_n, $(\boldsymbol{L}_n)_{33}$, is calculated from the average of the quantity $(\bar{\boldsymbol{L}}_n)_{33}$ over the current thickness. The thickness is then recursively updated from time t_n to t_{n+1} as follows,

$$h_{n+1}=h_n\left(1+\left\|\boldsymbol{L}_n\hat{\boldsymbol{x}}_n\right\|\right) \qquad (9)$$

where $\hat{\boldsymbol{x}}_n$ is the fiber normal at time t_n.

Membrane Behavior

The shell element developed here provides significant savings in computational effort for modeling membrane behavior over solids, because the incompressibility constraint is maintained kinematically by **Eq. 7**. Methods for achieving near incompressibility

in solids, such as the penalty method,[8] and Lagrange multiplier methods,[9] are far more intensive, since they require many more load steps and equilibrium iterations to converge. Furthermore the penalty method causes problems due to ill conditioning.

Figure 2 shows a membrane (l = 10 in., w = 10 in., t = 0.5 in.) stretched to four times its original length. Fringes of shell thickness are also shown on the deformed mesh. The model with 100 shell elements was compared with models of 100 solid elements. For the solid elements, both the penalty and augmented Lagrangian methods were used. The following CPU times were achieved using 20 time steps on an SGI Indigo 2 R4400 for the three cases:

Shells	33 CPUs
Solids with penalty	45 CPUs
Solids with augmented Lagrangian	75 CPUs

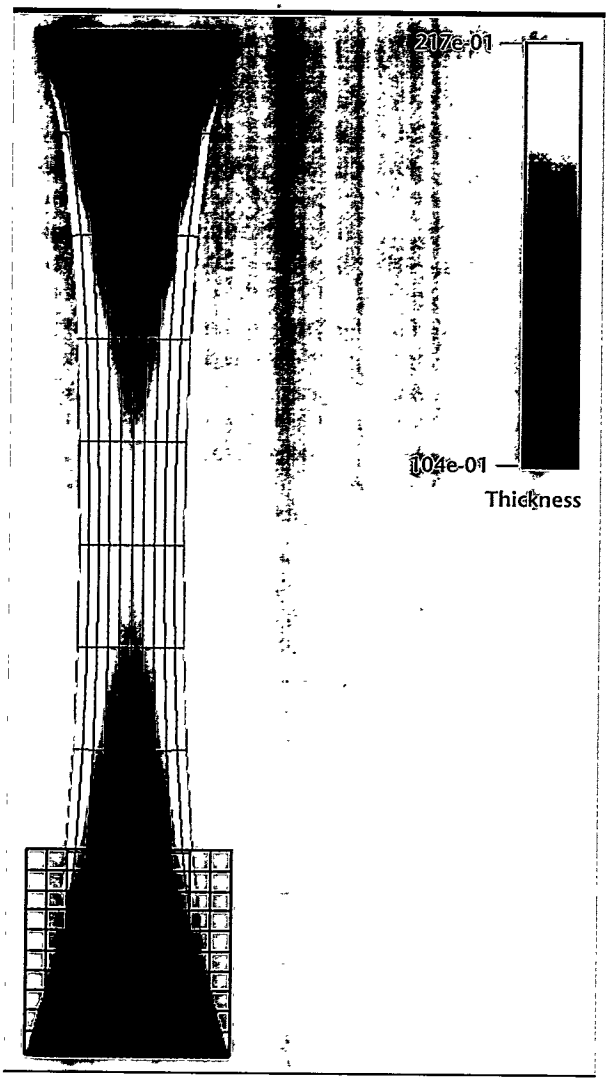


Figure 2. Undeformed and deformed mesh for membrane stretched to four times its original length. Shell thickness is shown in fringe.

In **Fig. 3**, the average uniaxial stress is shown for the three cases. As verification of the shell formulation, the shell model compares well with the augmented Lagrangian solids model. The penalty method grossly under estimates the stress at large strain since it can be shown that the incompressibility constraint is not adequately maintained. Attempts to increase the penalty factor will cause problems with convergence due to ill conditioning and element inversion.

Bending Behavior

The bending behavior of the shell elements is enhanced by the assumed shear strain fields. This allows far fewer elements in the plane of the membrane. Furthermore, the through thickness

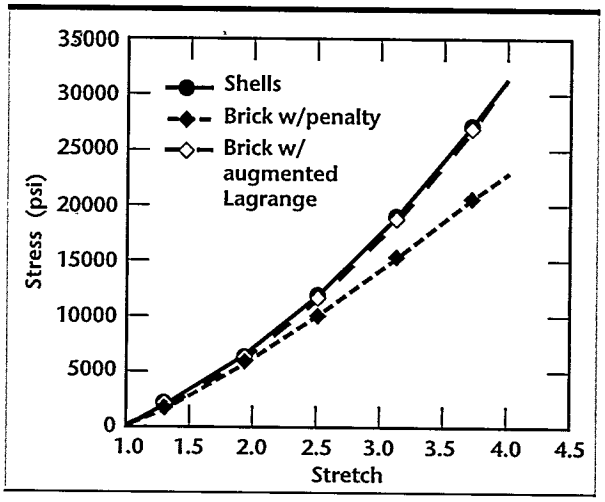


Figure 3. Average stress vs uniaxial stretch.

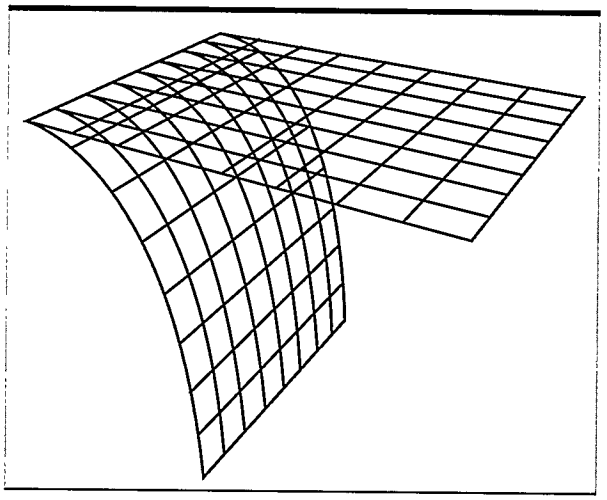


Figure 4. Undeformed and deformed mesh of rubber sheet with applied shear tip load.

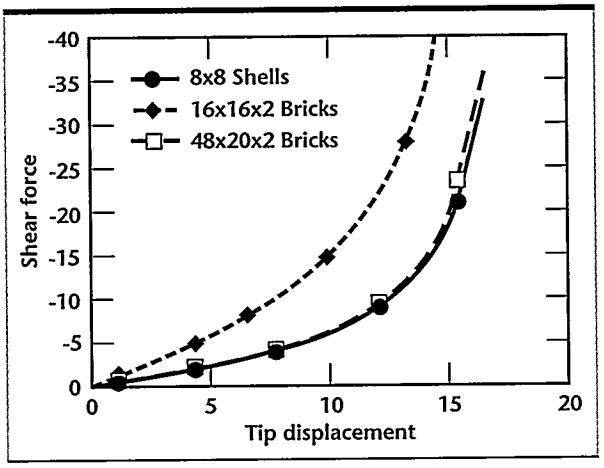


Figure 5. Shear tip force vs tip displacement.

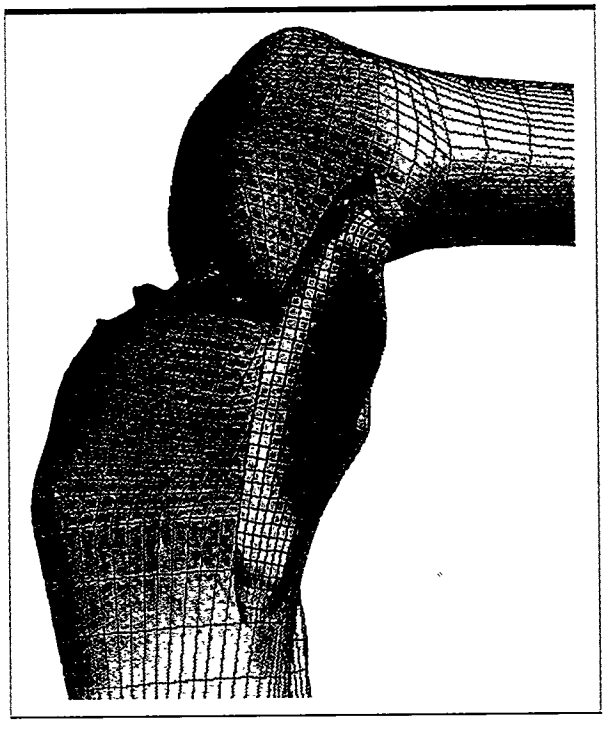


Figure 6. High fidelity knee model with shell elements modeling MCL. Effective stress is shown for 90° flexure.

stress gradients for non-linear material models can be captured without adding elements through the thickness.

Figure 4 shows an initially flat rubber sheet (l = 20 in., w = 20 in., t = 1 in.) clamped at one end and loaded in shear at the tip. Three different meshes were used: 8×8 shell mesh, a $16 \times 16 \times 2$ solid element (2 elements through thickness) and a $48 \times 20 \times 2$ solid element mesh. Results of the force vs displacement are shown in **Fig. 5**. It is seen that the $16 \times 16 \times 2$ solid element mesh is far too stiff. The shell model compared well with the fine $48 \times 20 \times 2$ mesh. The following CPU times for 20 time steps occurred for the three analyses:

8×8 shells	34 CPUs
$16 \times 16 \times 2$ solid	438 CPUs
$48 \times 20 \times 2$ solid	2100 CPUs

Future Work

The hyperelastic shell is currently implemented and working in the high fidelity knee model. Using a transversely isotropic constitutive law[2], the shell is used to model the medial collateral ligament (**Fig. 6**) due to its membrane structure. Comparison to solid element modeling of the MCL has yet to be made.

Currently all the shell models in NIKE3D possess three displacement and three rotational degrees of freedom (see **Eq. 2**). Because the fiber normal vector $\hat{x}_I$ is a unit vector, only two rotational degrees of freedom are necessary. By reducing the degrees of freedom from six to five, CPU times can be reduced up to 30% for large shell problems, further increasing the efficiency of the shell elements. In addition, this shell formulation and the large strain Hughes-Liu shell element contain two kinematic spurious modes. These modes can cause problems under certain conditions but can be easily eliminated by using a Bathe Dvorkin[10] form of the assumed shear strain field in the current formulation.

References

1. Noor, A. K., K. O. Kim, and J. A. Tanner (1990), "Analysis of Aircraft Tires via Semianalytic Finite Elements," *Finite Elements in Analysis and Design* **6** (no. 3), pp. 217–233.

2. Weiss, J., B. Maker, and S. Govindjee (1995), *Finite Element Implementation of Incompressible, Transversely Isotropic Hyperelasticity*, U.C. Berkeley Report, Berkeley, Calif. (UCB/SEMM-95/07).

3. Engelmann, B. E., R. G. Whirley, and G. L. Goudreau (1989), "A Simple Shell Element Formulation For Large-Scale Elastoplastic Analysis," *Proceeding of the Symposium on Analytical and Computational Models for Shells*, ASME Winter Annual Meeting.

4. Hughes, T. J. R., and W. K. Liu (1981), "Nonlinear Finite Element Analysis of Shells: Part I-Three-dimensional Shell," *Computer Methods in Applied Mechanics and Engineering* **26**, pp. 331–362.

5. Hughes, T. J. R., and E. Carnoy (1983), "Nonlinear Finite Element Analysis of Shell Formulation Accounting for Large Membrane Strains," *Computer Methods in Applied Mechanics and Engineering* **39**, pp. 69–82.

6. Simo, J. C., and T. J. R. Hughes (1986), "On the Variational Foundations of Assumed Strain Methods," *Journal of Applied Mechanics and Engineering* **53** (no. 1), pp. 51–54.

7. Hughes, T. J. R. (1980), "Generalization of Selective Integration Procedures to Anisotropic and Nonlinear Media," *International Journal for Numerical Methods in Engineering* **15**, (no. 9), pp. 1413–1418.

8. Malkus, D. S., and T. J. R. Hughes (1978), "Mixed Finite Element Methods, Reduced and Selective Integration Techniques: A Unification of Concepts," *Computer Methods in Applied Mechanics and Engineering* **15**, pp. 63–81.

9. Glowinski, R., and P. Le Tallec (1988), *Augmented Lagrangian Methods for the Solution of Variational Problems*, Springer-Verlag, Berlin, Germany.

10. Bathe, K. J., and E. N. Dvorkin (1984), "A Continuum Mechanics Based Four Node Shell Element for General Non-Linear Analysis," *International Journal for Computer-Aided Engineering and Software* **1**.

Investigation of the Radiative Properties of Ceramic Casting Materials

James D. Maltby, Barbara T. Kornblum, Charles S. Landram, and Charles W. Dingus
New Technologies Engineering Division
Mechanical Engineering

The radiative properties of ceramic casting materials at elevated temperatures are of great interest to the casting industry, but are very difficult to measure. We have developed a technique for predicting these properties based on a combination of experimental measurements and computer simulation using the Monte Carlo participating media radiative heat transfer code, TRIM3D. Initial results show the materials to be highly scattering, with optical depths near 10 for typical thicknesses.

Introduction

Ceramic materials are commonly used for investment casting molds, as they have been for thousands of years. With the advent of near-net-shape casting, where precision parts are produced in a single casting operation, detailed knowledge of heat transfer during solidification becomes important. In the past, it has been assumed that the major heat transfer modes are conduction through the ceramic, and convection and radiation from the outer surface. However, visual inspection of molds at very high temperatures shows them to be translucent, with subsurface details becoming visible. If these materials show significant transmission at high temperatures and long (infrared) wavelengths, radiant transport from the metal through the ceramic could be an important heat loss mechanism.

Unfortunately, the transmissive properties of these materials are not available in the literature, making it impossible to estimate the magnitude of this effect. We used a three pronged approach for the measurement of these properties. First, we did experiments to measure the amount and angular distribution of light transmitted through samples of the material. Second, we performed a theoretical analysis of the experiments to produce first estimates of the material properties. Third, we used a Monte Carlo radiative transport code to model the experiment in a detailed manner and predict accurate values for the important material parameters.

The properties required for prediction of thermal radiation transport in a participating medium are the absorption coefficient, scattering coefficient and scattering phase function. The absorption coefficient determines how rapidly radiation is absorbed by the medium; the scattering coefficient tells how often photons are scattered; and the scattering phase function describes the distribution of scattered photons. Unfortunately, all three parameters interact to produce the final distribution of radiation, so the experimental results do not read out the properties directly. An iterative approach was used instead, using TRIM3D to predict the transmitted fluxes and a heuristic procedure to refine the property guesses. This problem is one of a class of inverse property prediction problems currently of great interest in the research community.

This coupled experimental and computational technique was also used for quartz flashlamp material of the type used in the National Ignition Facility (NIF) beamlet. This information will be useful for future coupled radiation/conduction/thermal stress analyses of flashlamp failures.

Progress

The experimental apparatus is described in detail by M. Havstad, in the Materials Science and Engineering section of this report. Basically, a collimated beam of radiation is incident on a thin circular wafer of ceramic material 2 in. in diameter, and a few millimeters thick. The angular distribution of the transmitted radiation is measured using a detector on a swing arm, as diagrammed in **Fig. 1**. A profile of the radiation incident on the back of the sample is shown in **Fig. 2**, and the

angular intensity profile of the transmitted radiation is shown in **Fig. 3**, for a 0.19-cm-thick sample in the wavelength range 0.4 μm to 1.0 μm.

To model this process using TRIM3D, a simplified mesh was built showing just the ceramic sample and an array of detecting surfaces. The profile of **Fig. 2** was used as a photon source input for the back side of the ceramic disk, and the intensity on each of the detector surfaces was predicted by the code. The detector surfaces are 5° wide to match the experimental measurement interval. A set of material properties was assumed, and the difference between the predicted intensity profile and the experimental profile of **Fig. 3** was used to improve the estimate of the material properties. Iteration continued until the two profiles matched reasonably closely.

To come up with a reasonable initial guess, we adopted a method first used by Glicksman[1] to analyze a similar experiment. The procedure uses the ratio of the intensities of the two peaks to calculate an initial extinction coefficient:

$$\frac{I}{I_0} = e^{-(a+\sigma)x} = e^{-kx}$$

where the extinction coefficient is defined as:

$$k = (a + \sigma).$$

The scattered distribution from 5° outward is then used to estimate the scattering coefficient, σ. The initial guess for our material properties was k = 62 cm^{-1}, a = 22 cm^{-1} and σ = 40 cm^{-1}. However, when these numbers were used for a TRIM3D simulation, the estimated peak height was too small by several orders of magnitude. Further investigation of the derivation showed that an implicit 1-D assumption was used that was not valid for this type of experiment. The use of three approaches to cross check a single prediction proved useful in preventing errors.

Examination of our initial computational runs showed that the shape of the intensity profile was strongly influenced by the scattering albedo, defined as:

$$\Omega = \frac{\sigma}{a+\sigma} = \frac{\sigma}{k}$$

Several predicted intensity profiles at different scattering albedos are shown in **Fig. 4**. It can be seen that as the scattering albedo decreases, and

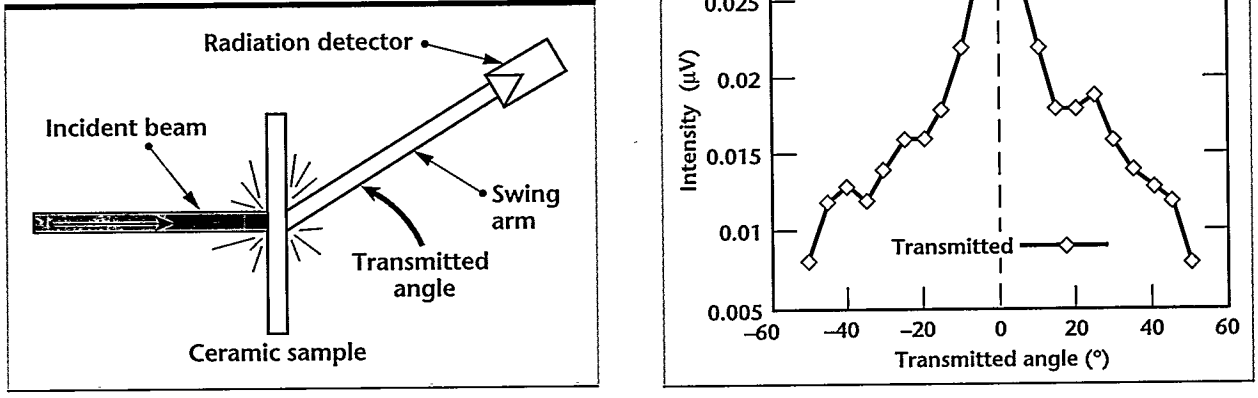


Figure 1. Illustration of detector on a swing arm, used to measure the angular distribution of transmitted radiation.

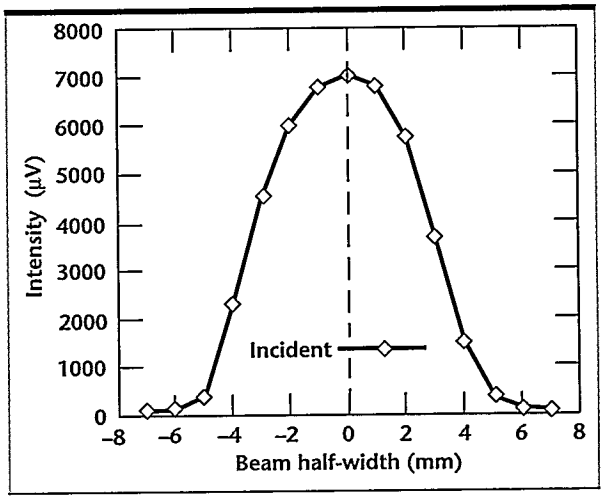


Figure 2. Profile of radiation incident on the back of the sample.

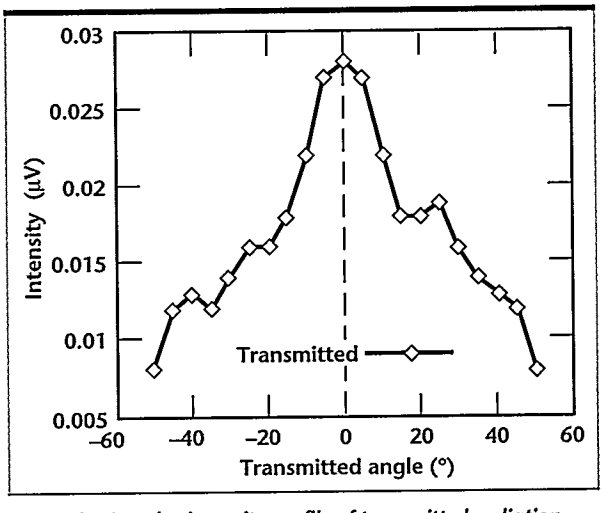


Figure 3. Angular intensity profile of transmitted radiation.

the material becomes less scattering and more absorbing, the distribution becomes more peaked and drops off far more rapidly out near 50°. We developed a new iteration procedure that involves first picking an albedo that matches the shape of the experimental curve, then iterating over the absorption coefficient, a, to match the peak height. Usually, this only took a few iterations.

Several typical material property sets were tested in order to quantify the effect of anisotropy in the scattering phase function. Extremely small changes in the predicted profile were seen, less than the error in the experimental data. This effect has also been seen in analyses of similar

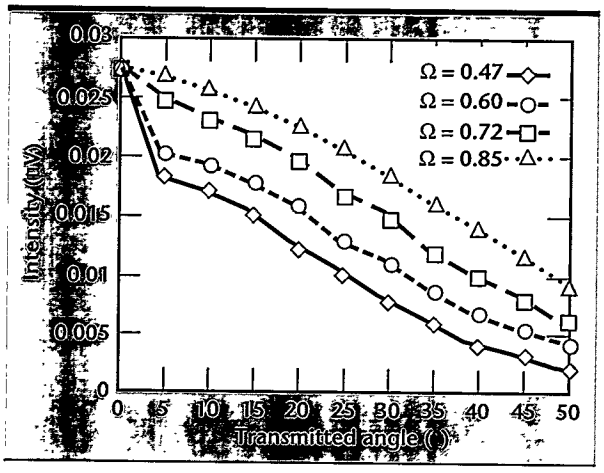


Figure 4. Four predicted intensity profiles, at different scattering albedos.

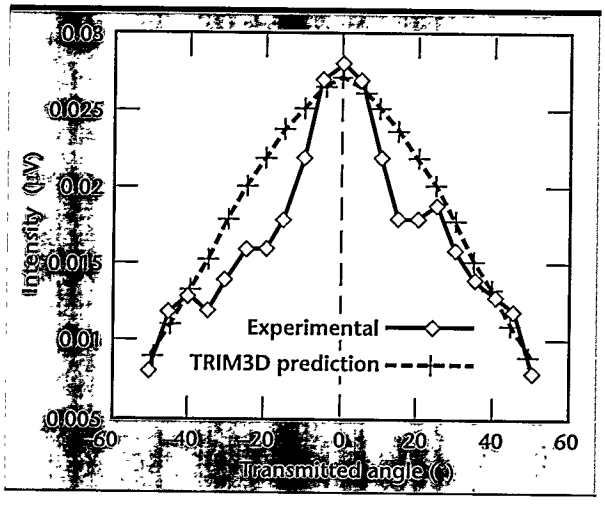


Figure 5. Plot of best fit for sample. Estimated properties are $a = 8\ cm^{-1}$, $\sigma = 45\ cm^{-1}$, $k = 53\ cm^{-1}$, and $\Omega = 0.85$.

benchmark problems[2], so it was felt that the experiment was insufficiently sensitive to deduce a phase function. The simulation also shows that an isotropic scattering phase function is a good approximation for problems of this type.

The best fit for the 0.19-cm sample is shown in **Fig. 5**, and the estimated properties are $a = 8\ cm^{-1}$, $\sigma = 45\ cm^{-1}$, $k = 53\ cm^{-1}$, and $\Omega = 0.85$. This shows that the material is highly scattering, but that radiative transport through the layer is a significant heat loss mechanism. The results also indicate that the material may be considered optically thick in typical casting shells, so a temperature-dependent thermal conductivity may be used as a reasonable approximation instead of a more detailed (and expensive) participating media transport simulation. Most importantly, an experimental and computational method has been developed to measure the properties of scattering participating media.

Future Work

In FY-96 we will need to analyze more sample thicknesses in order to correlate these initial parameter estimates. A least squares curve fitting algorithm will need to be developed to generate a best fit. The measurements will be extended over a broader range of wavelengths and for other types of ceramic material.

The experimental apparatus will also be used to measure the spectral transmissivity of doped quartz and borosilicate glass used in the large laser amplifiers intended for NIF. These measurements will then be used in a TRIM3D analysis of flashlamp failures due to rapid thermal stress.

References

1. Glicksman, L., M. Scheutz, and M. Sinofsky (1987), *JHMT* **30**, p. 187.

2. Maltby, J. D. (1993), *HTD*-vol. **276** (*ASME*), p. 161.

Performance of NIKE3D with PVSOLVE on Vector and Parallel Computers

Bradley N. Maker
Defense Technologies Engineering Division
Mechanical Engineering

Qin Jiangning and Duc T. Nguyen
Multidisciplinary Parallel-Vector Computation Center
Old Dominion University, Norfolk, Virginia

The cost of large, implicit, finite element calculations is dominated by the linear equation solver. In this work, several versions of the direct linear equation solver PVSOLVE were implemented in the general purpose, nonlinear finite element code NIKE3D. Timing studies were performed on Sun and IBM workstations, on Cray Y/MP and C90 supercomputers, and on the distributed memory MEIKO CS-2 parallel computer. Cost-breakdowns and efficiency curves are presented for a benchmark problem with several mesh densities. The distributed memory implementation is shown to be most efficient for large problems, where MEIKO performance is comparable to that of a single Cray processor.

Introduction

From the earliest days with NASTRAN and SAP, finite element calculations have been performed using implicit methods. The characteristic of these methods, typically applied to static, low-rate dynamic, and eigenvalue problems, is the generation and solution of a large system of coupled algebraic equations. The kernel of the implicit computer algorithm is then the linear equation solver. As Ferencz reports,[1] for the n x n x n-element Boussinesq problem, storage requirements for the stiffness matrix rise proportional to n^5, while CPU costs for its factorization rise as n^7. These numbers justify a great deal of effort in developing an efficient linear equation solving algorithm.

The particular computer code of interest in this paper is NIKE3D, a general purpose, nonlinear, implicit finite element code[2]. The nonlinear equation solving algorithms in NIKE3D are based on quasi-Newton schemes, such as the BFGS method[3]. These rely heavily on the relative economy of the forward elimination/back substitution phase of the linear solver, for which Ferencz reports a CPU cost of n^5 on the aforementioned problem. Even so, the linear equation solver dominates NIKE3D simulation costs for even moderate size (5000 element) problems.

A great volume of research has addressed this difficulty by considering iterative solution methods. The element-by-element (EBE) solver, Ferencz's landmark contribution to NIKE3D, exemplifies these methods. The element level implementation of EBE avoids assembly of the full stiffness matrix, storing instead the dense element stiffness matrices, and allowing for full vectorization. For large problems, the EBE method can yield staggering savings over direct methods. On a recent model with 292,000 degrees of freedom, solved on a Cray C90, EBE reduced storage requirements by over 97%, and CPU cost by over 98%[4].

However, the convergence rate of iterative solvers is dependent on the condition number (ratio of largest to smallest eigenvalues) of the numerical system, which can vary widely from model to model. Experience with NIKE3D has demonstrated the existence of several problems which are so poorly conditioned that a solution cannot be obtained with EBE at any cost. Therefore, in recognition of the general purpose nature of NIKE3D, it is prudent to develop an efficient direct linear equation solver as an alternative to EBE. The following sections describe a family of direct solvers which are optimized for vector and parallel computing architectures.

Progress

Vector Linear Equation Solver

The first in the family of PVSOLVE linear equation solvers was developed for Cray-style vector computers[5]. This solver achieves enhanced performance over scalar algorithms by using vectorization and loop unrolling[6]. The key to enhanced vector performance is the row-wise storage of the variably banded, symmetric stiffness matrix. This provides a contiguous data structure for vector operations. In the NIKE3D implementation, a loop unrolling level of eight was chosen. This requires storage to be allocated such that each block of eight contiguous equations has a uniform row length. The factorization operation is then unrolled, or explicitly coded, for the block of eight equations. Vectorization may then proceed as a uniform reduction operation along the block of rows. The reduction operation takes the form of a SAXPY ($z = \alpha * x + y$), using the chaining capability of the Cray hardware to double floating point performance. This implementation generates vector lengths equal to the row lengths of each block, or roughly proportional to the bandwidth of the system. Performance is therefore optimal on large bandwidth systems.

Benchmark Problem Description

The classical Boussinesq problem of a point load in an elastic half-space was chosen as a benchmark problem[7]. The domain was discretized as a quarter symmetric cube of n x n x n elements. Loading was scaled to produce infinitesimal deformations, so that results could be verified using linear elasticity solutions. The full nonlinear element formulation of NIKE3D was used, allowing timings for this linear problem to be extrapolated for use in anticipating performance on nonlinear problems. Characteristics of the linear system are given in **Table 1** for several levels of discretization.

Performance on Scalar and Vector Machines

The vectorized PVSOLVE code was implemented in NIKE3D, and compared to the existing FISSLE solver[8]. **Table 2** presents results for a range of mesh densities executed on several computer platforms. Timings were measured using the system elapsed timer, and are reported in seconds. On the scalar SUN and IBM machines, performance of the two solvers was comparable, but on the CRAY machine vectorization of PVSOLVE produced a significant speed-up, in spite of the slightly larger operation count compared to FISSLE.

Parallel-Vector Linear Equation Solver

The next generation of PVSOLVE code includes parallel processing capability.[5,9,10] The target architecture is the Cray multitasking environment. To facilitate machine independence, the FORCE language[11] was used to direct multitasked operations within PVSOLVE. In the NIKE3D implementation, only the PVSOLVE routines were allowed to operate under multitasking. The stiffness matrix generation and assembly phases of the computation executed sequentially.

Table 1. Characteristics of the linear systems chosen as benchmark problems. Storage is given in units of words. Notice that PVSOLVE's row-wise storage scheme requires slightly more storage that FISSLE's column-based method.

Mesh Density, n	4	8	16	24
Number of Elements	64	512	4,096	13,824
Number of Equations	300	1,944	13,872	45,000
Total Storage (FISSLE)	20,181	446,909	11,683,773	(n/a)
Total Storage (PVSOLVE)	(n/a)	(n/a)	11,866,552	82,727,421
Column Height, max (FISSLE)	85	257	889	(n/a)
Column Height, max (PVSOLVE)	(n/a)	(n/a)	896	1,912
Column Height, avg (FISSLE)	67	229	842	(n/a)
Column Height, avg (PVSOLVE)	(n/a)	(n/a)	854	1,846

Table 3 lists timing results for the 16 × 16 × 16 Boussinesq problem executed on a Cray-C90 16/128 machine. The average elapsed time per processor is reported. For comparison, the single processor vectorized PVSOLVE timing on the C90 is also included. The parallel-vector PVSOLVE includes several algorithm enhancements over the vector version, yielding improved single processor performance.

Distributed Parallel-Vector Linear Equation Solver

The latest version of PVSOLVE for NIKE3D is designed to operate in a distributed parallel architecture.[12,13] To efficiently implement this solver in NIKE3D, several fundamental code changes were necessary. The distributed memory nature of the target computers requires that communication between processors be kept to an absolute minimum. It is therefore necessary to compute and assemble the stiffness matrix "in place." This requires a strategy for anticipating which portion of the global stiffness matrix will be stored on each processor, resulting in a domain decomposition algorithm.

Distributed Memory PVSOLVE. Like the previous versions, the massively parallel PVSOLVE creates blocks of contiguous equations, and performs loop unrolling within each block. In the version of PVSOLVE used with NIKE3D, the block size is set to eight. Ownership of the blocks is then distributed over the processors. The first block is allocated to the first processor, the second block to the second processor, and so on. If the number of blocks exceeds the number of processors, which is usually the case, allocation continues in the same sequential manner, beginning again with the first processor. When the allocation process is completed, each processor owns approximately the same number of blocks, distributed evenly throughout the matrix.

The factorization process begins with the first processor factorizing the first block, while the other processors wait. The motivation behind choosing a small block size is to minimize the waiting time for each processor before it can begin factorizing one of its blocks. As each block is factorized, messages are passed, containing the diagonal entries required for factorization of successive blocks.

Stiffness matrix data is stored columnwise, and padded with zeroes as necessary, to ensure that all columns of a block have the same height, and that this height is an integer multiple of the block size. This enables efficient unrolling of the reduction operation, and provides an opportunity for vectorization. During initialization, each processor computes its own diagonal pointer, based on the portions of the global matrix which it owns. This results in an efficient, scalable storage scheme.

Table 2. Timing results, in CPU s, comparing vectorized PVSOLVE with NIKE3D's existing solver FISSLE. Timings represent cost of one factorization, forward elimination, and back substitution, for Sun Sparc 10-51 (SUN), IBM RS/6000-550 (IBM), and Cray Y-MP 8/64 (YMP) machines.

Mesh Density, n	4	8	16
SUN (PVSOLVE)	0.120	8.300	(n/a)
SUN (FISSLE)	0.130	8.630	(n/a)
IBM (PVSOLVE)	0.060	4.010	346.380
IBM (FISSLE)	0.050	3.590	306.980
YMP (PVSOLVE)	0.015	0.605	42.879
YMP (FISSLE)	0.060	1.663	66.565

Table 3. Timing results for one linear equation solution of the 16× 16× 16-element Boussinesq problem, using vectorized and parallel-vector versions of PVSOLVE running on a Cray C90 machine. Timings represent average elapsed CPU time per processor, in seconds.

Number of Processors	1	2	3	4	5	6	7
Vectorized PVSOLVE	16.48						
Parallel-vector PVSOLVE	14.68	7.34	4.89	3.68	2.93	2.45	2.11

Domain Decomposition. Given the blocking and allocation strategy of PVSOLVE, creation of a domain decomposition for NIKE3D becomes straightforward. At initialization, each processor performs a loop over the list of all elements, checking each degree of freedom for possible inclusion in one of the blocks of equations owned by the processor. A local element list is generated, containing all of the elements required to assemble the portion of the global stiffness matrix which the processor owns. As the element stiffness matricies are generated, an assembler routine builds the local stiffness matrix, discarding the portions of each element stiffness matrix which do not contribute to the local stiffness matrix.

Two remarks are appropriate here:

1) This decomposition procedure sacrifices some redundant computation to avoid communication cost. Most elements are processed by more than one processor, since it is unlikely that the 24 element degrees of freedom will be allocated into blocks corresponding to the same processor. This approach is in contrast to decomposition schemes commonly applied to explicit finite element codes, where elements are uniquely distributed over processors, and nodal data is computed redundantly.
2) Using this decomposition scheme, the bandwidth minimizer effectively determines spatial decomposition of the mesh. This happens because PVSOLVE distributes equations across processors. However, the bandwidth minimizer is responsible for assignment of equation numbers to actual nodal degrees of freedom. This results in a mapping of nodes and elements to processors. This raises interesting issues regarding optimal bandwidth minimization schemes for distributed memory solvers.

Implementational Considerations. This initialization procedure fits well into the existing architecture of NIKE3D. Since the global connectivity of the mesh can change as contact surfaces evolve, NIKE3D allows for re-initialization of the diagonal pointer array before each stiffness matrix reformation. Connectivity data for the static portions of the mesh (brick, beam, and shell elements) are used to build an initial column height array. A copy of this array is augmented with current connectivity information from the contact surfaces before computing the current diagonal pointer array. In the parallel implementation, PSOLVE uses NIKE3D's current column height array to build local diagonal pointer arrays on each processor. This compatability, combined with a general assembly routine which accepts element stiffness matrices of variable dimension, yields a modular and manageable code.

The external force contribution to the load vector is presently computed redundantly to avoid communication. The internal force (stress divergence) contribution is zero for the linear analyses presented herein. To perform nonlinear or multi-step analyses, the stress divergence terms must be computed in parallel, followed by a global sum communication of results. This communication is not expected to present a significant cost penalty, since a global sum operation is currently used to communicate the solution vector after the backward substitution phase of PVSOLVE, and costs for the entire forward/backward phase are relatively small.

Performance on MEIKO Machine. The parallel NIKE3D/PVSOLVE code was developed on the MEIKO CS-2 computer. The CS-2 processors are 66 MHz Sparc chips, with 1 Mb internal cache. Each processor also accesses two 66 MHz-Fujitsu vector units, with 8 Kb vector registers and cache coherency with the Sparc chip. The vector units were not used in the timing studies conducted for this paper, due to their developmental immaturity. The CS-2 communication network is a multi-stage fat-tree architecture, with peak bandwidth of 40 Mb/s. At the time of this writing, the machine is undergoing acceptance testing for 256 nodes. The compiler used was PGI-Fortran

Table 4. Performance data for the 16× 16× 16-element Boussinesq problem using NIKE3D with plarallel PVSOLVE on the MEIKO CS-2 computer. All timings are reported in seconds.

Number of processors	1	4	8	16	32	64	96
Elements per Processor	4096	3525	1830	1620	820	415	270
Storage (K-words)	11794	2979	1502	770	403	220	162
Generation/Assembly	71.6	58.2	30.0	26.7	13.4	6.73	4.49
Factorization	2326	588.8	305.0	185.9	134.8	115.8	113.1
Forward/Backward	11.70	3.94	2.64	2.30	2.55	5.97	19.19
Communication	0	16.0	31.5	48.9	63.2	79.9	101.0

v1.0.3, developed by The Portland Group. Unfortunately, compiler optimization had to be disabled, and the de-bug option activated to obtain a reliable executable code.

A series of calculations were performed on the MEIKO machine for the Boussinesq problem with mesh densities of 8 × 8 × 8, 16 × 16 × 16, and 24 × 24 × 24. Timings were measured for the element stiffness generation and assembly phase, and the solution phase. During the solution phase, separate accounting was taken of the time spent during communication. The system elapsed time clock was used in all cases.

Table 4 presents data for the 16 × 16 × 16-element mesh. Timings for factorization and forward/backward solution include communication. The total communication cost during the solution phase is also reported. No communication is required during the element stiffness generation and assembly phase, however redundant computation results in non-optimal parallel efficiency. This efficiency loss is most apparent when few processors are used. Increasing from one to four processors reduces the average element per processor count from 4096 to only 3525, a mere 29% efficiency. However, increasing from 32 to 64 processors decreased element counts from 820 to 415, for a local efficiency of nearly 99%. Similar trends are observable for other costs associated with local element count, such as storage and generation and assembly.

Figures 1, 2, and **3** show breakdowns of cost associated with a single factorization, forward elimination, and back substitution using parallel PVSOLVE. The lower shaded portion of each histogram bar indicates the portion of each calculation consumed by communication. This communication includes synchronization at the beginning of the solve, messaging during the solve, and a global sum communication at the end of the solve, to pass the entire solution vector to each processor. These figures demonstrate that as the number of processors is increased, communication costs dominate the cost of the solve. As the problem size is increased, more processors may be effectively used since more parallel work is available.

Figures 4, 5, and **6** show parallel efficiency for the three benchmark problems. For each problem, efficiency curves are shown for the storage, element stiffness generation and assembly, and linear solution phases of the calculation. Efficiencies are computed as $e = N*T_N/T_1$, where T_1 is the time or storage required on a single processor, and T_N is the effort required by N processors. In the case of the 24 × 24 × 24-element problem, a single processor does not have the capacity to solve the problem. Single processor timings were extrapolated from the four-processor data, assuming linear speed-up, for purposes of computing efficiency.

Efficiency curves are presented for three components of the analysis: the storage required for the stiffness matrix, the time required to generate and

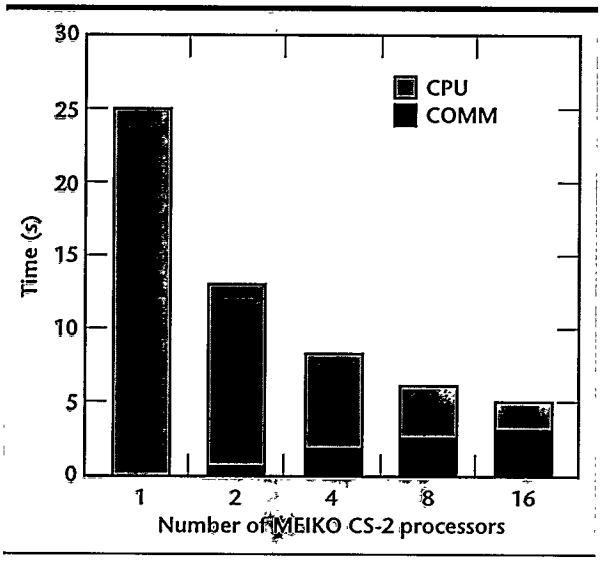


Figure 1. Cost breakdown for linear solver on 8 × 8 × 8-element mesh using MEIKO machine. Lower portion of each bar indicates time consumed by communication. For this small problem, 68% of cost is due to communication when only 16 processors are used.

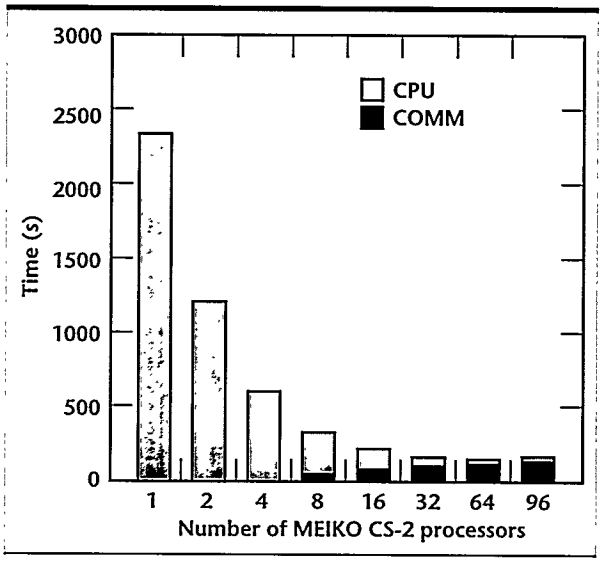


Figure 2. Cost breakdown for linear solver on 16 × 16 × 16-element mesh using MEIKO machine. Lower portion of each bar indicates time consumed by communication. For this medium-sized problem, more processors may be used before communication dominates cost. The 96-processor timing is 76% communication, and 9% slower overall than the 64-processor result.

assemble the local stiffness matrix on each processor, and the solution time. The storage requirements are by far the most efficient of these three operations, exceeding 90% efficiency in most cases. The generation and assembly phase approaches 50% efficiency for the largest problem. However, after the number of processors reaches a threshhold value, doubling the number of processors cuts cost nearly in half, implying good local efficiency (scalable performance). The seemingly constant efficiency of the generation and assembly curves for large numbers of processors is, in part, an artifact of the uniform bandwidth structure of the stiffness matrices for the cubic meshes used in this study. In previous studies using this decomposition scheme, nearly ideal efficiencies were obtained for 2- and 3-node elements[9,13]. NIKE3D's 8-node, 24-degree of freedom brick element naturally spans more processors, resulting in increased redundancy.

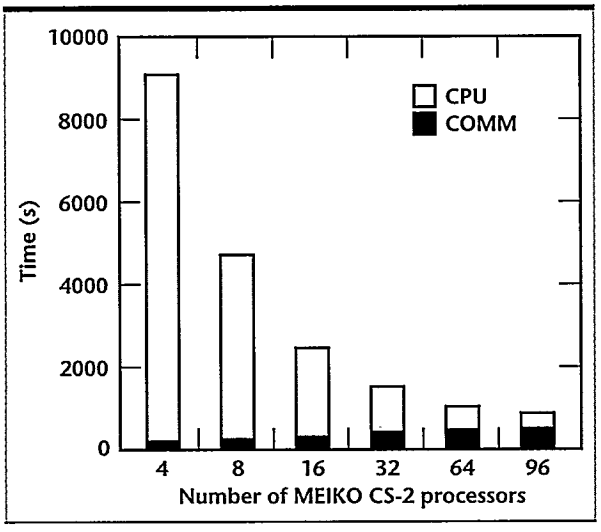


Figure 3. Cost breakdown for linear solver on 24 × 24 × 24-element mesh using MEIKO machine. Lower portion of each bar indicates time consumed by communication. For this large problem, communication consumed only 54% of the 96-processor timing.

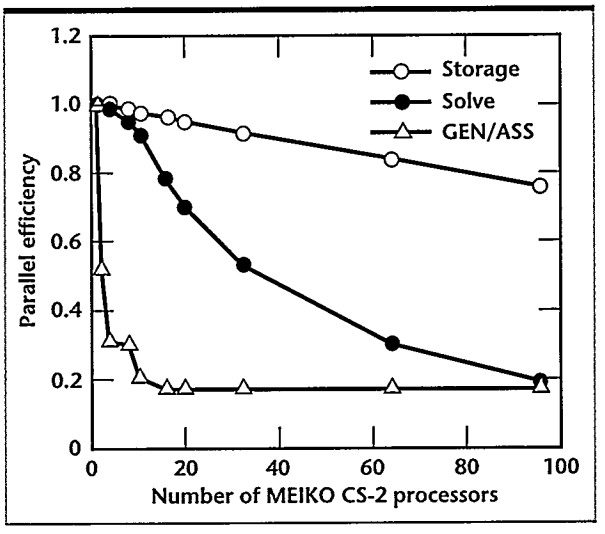


Figure 5. Efficiency data for three phases of the 16 × 16 × 16-element analysis using MEIKO machine. Stiffness matrix storage is highly efficient. Good scalability of the generation/assembly phase is obtained when more than ten processors are used. Solution phase efficiency is greater than 50% for up to 38 processors.

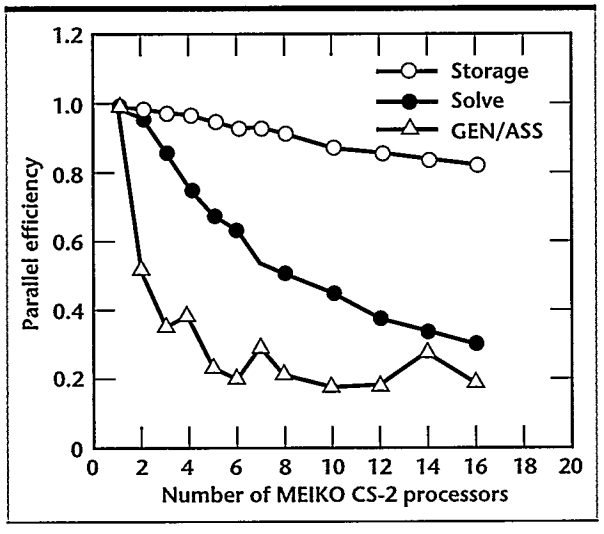


Figure 4. Efficiency data for three phases of the 8 × 8 × 8-element analysis using MEIKO machine. Stiffness matrix storage is highly efficient. Element stiffness generation and assembly show low efficiency, relative to the single processor result, but scalability is good when five or more processors are used. Solution phase efficiency is greater than 50% for up to eight processors.

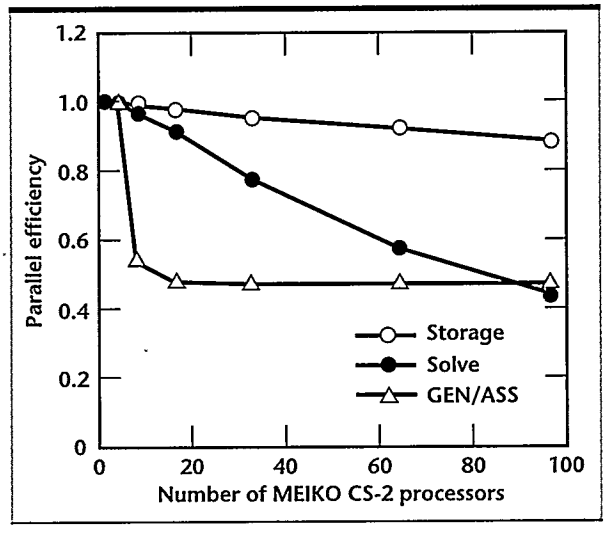


Figure 6. Efficiency data for three phases of the 24 × 24 × 24-element analysis using MEIKO machine. Storage requirements exceed the capacity of a single processor, so timing data was extrapolated assuming ideal efficiency of the four-processor result. Good scalability of the generation/assembly phase is obtained when 16 or more processors are used. Solution phase efficiency is greater than 50% for up to 80 processors.

Not surprisingly, the efficiency of PVSOLVE is highly dependent on problem size. In the case of the small $8 \times 8 \times 8$ mesh, efficiency dropped below 50% using nine or more processors. For the $16 \times 16 \times 16$-element mesh, enough work was available to sustain an efficiency of 50% with 38 processors, and for the $24 \times 24 \times 24$-element problem, PVSOLVE maintained efficiency of 50% or more using up to 80 processors.

Conclusions

Vectorization of PVSOLVE produced significant savings on Cray machines, compared with NIKE3D's non-vectorized solvers, without significantly reducing performance on scalar machines. The multitasking PVSOLVE provided significant additional savings on the Cray C90 platform, but extensive code modifications are required for parallel execution of the entire NIKE3D code. In contrast, the distributed memory, message passing implementation used for the MEIKO machine resulted in a manageable, modular code.

Performance of the MEIKO code was limited, since compiler optimization could not be used. Single processor performance, therefore, fell far short of the SUN and IBM workstations. While multi-processor timings will improve with compiler optimization, communication costs will be unchanged. Efficiency curves presented here are, therefore, upper bounds for the current algorithm.

With respect to the linear solver, this study shows that communication costs increase with number of processors, eventually dominating the problem. In the case of the $16 \times 16 \times 16$-element problem, performance on the MEIKO machine cannot exceed the Cray Y/MP timing, no matter how many processors are used. But as problem size increases, more processors may be used efficiently. For the $24 \times 24 \times 24$-element problem, 96 MEIKO processors required 818 s, soundly routing the 2002-s Cray X/MP timing that Ferencz obtained using FISSLE.

Direct methods are very often bound by storage limitations rather than by CPU speed. However, with 128 Mb of memory per processor, and nearly ideal local efficiency of the storage algorithm for large numbers of processors, the MEIKO machine can easily exceed the capacity of many Cray machines. This is perhaps the most significant advantage of parallel NIKE3D/PVSOLVE.

References

1. Ferencz, R. M. (1989), *Element-By-Element Preconditioning Techniques for Large-Scale, Vectorized Finite Element Analysis in Nonlinear Solid and Structural Mechanics*, Ph.D. thesis, Stanford University, Stanford, Calif.

2. Maker, B. N., R. M. Ferencz, and J. O. Hallquist (1991), *NIKE3D: A Nonlinear, Implicit, Three Dimensional Finite Element Code for Solid and Structural Mechanics—User's Manual*, University of California, Lawrence Livermore National Laboratory, Livermore, Calif. (UCID-105268).

3. Matthies, H., and G. Strang (1979), "The Solution of Nonlinear Finite Element Equations," *Int. J. Num. Meth. Eng.*, **14**,(11), pp. 1613–1626.

4. Hokanson, K. E., University of Minnesota, private communication.

5. Storassli, O. O., D. T. Nguyen, and T. K. Agarwal (1990), "A Parallel-Vector Algorithm for Rapid Structural Analysis on High-Performance Computers," *Proceedings of the AIAA/ASME/ASCE/AHS/ASC 31-st SDM Conference, Long Beach, CA*, pp. 1149–1159.

6. Poole, E. L., and A. L. Overman (1988), *The Solution of Linear Systems of Equations With A Structural Analysis Code On The NAS Cray 2*, NASA (CR 4159).

7. Love, A. E. H. (1944), *A Treatise on the Mathematical Theory of Elasticity*, Dover Publications, New York, N.Y.

8. Taylor, R. L., E. L. Wilson, and S. J. Sackett (1980), "Direct Solution of Equations By Frontal and Variable Band Active Column Methods," *Europe-U.S. Workshop: Nonlinear Finite Element Analysis in Structural Mechanics*, Bathe, Stein, and Wunderlich, Eds., (Springer-Verlag, Berlin, Germany), pp. 63–89.

9. Storaasli, O. O., D. T. Nguyen, M. A. Baddourah, and J. Qin (1993), "Computational Mechanics Analysis Tools For Parallel-Vector Supercomputers," *Computing Systems In Engineering*, Vol. **4**,(4–6), pp. 349–354.

10. Storaasli, O. O., D. T. Nguyen, and T. K. Agarwal (1990), "The Parallel Solution of Large-Scale Structural Analysis Problems On Supercomputers," *AIAA Journal*, Vol. **28**, No. 7, pp. 1211–1216.

11. Jordan, H. F., M. S. Benton, and N. S. Arenstorf (1986), "FORCE User's Manual," *ECE Tech. Rept. 86-1–4*, Dept. of Electrical and Computer Engineering, University of Colorado, Boulder, Colo

12. Qin, J., and D. T. Nguyen (1994), "A Parallel-Vector Equation Solver For Distributed Memory Computers," *Computing Systems In Engineering Journal*, Vol. **5**, No. 1.

13. Qin, J., and D. T. Nguyen (1993), "A New Parallel-Vector Finite Element Analysis Software On Distributed Memory Computers," *Proceedings of the AIAA/ASME/ASCE/AHS/ASC 34-th SDM Conference, La Jolla, Calif.*

EMINI Re-Certification and Development

Stan Bumpas, Michael A. Gerhard, and Stanley C. Sommer
New Technologies Engineering Division
Mechanical Engineering

This report discusses four tasks established for benchmarking and developing the GEMINI computer code: (1) update the users' manual and validate the code; (2) convert GEMINI to the SUN, SGI and IBM workstations; (3) incorporate composite modal damping; and (4) calculate mode shapes using the block Lanczos method.

Introduction

The GEMINI computer code performs linear structural and stress analyses of buildings and mechanical components using the finite element method. It evolved from the SAP4 linear structural analysis program from the University of California, Berkeley. GEMINI has many improvements over SAP4, including additional and refined finite element types, faster equation solution techniques, more efficient modal extraction methods, and random vibration analysis options. GEMINI interfaces directly with the pre-processors SLIC and INGRID for mesh generation, and with the thermal analysis code TOPAZ3D, and the post-processors TAURUS and GRIZ for graphical presentation of deflection and stresses. The code has many modal analysis options, such as modal time history integration, response spectrum analysis, and random vibration analysis using a specific post-processor GEMPSD. Since 1984, GEMINI has been used on the main frame computers at Lawrence Livermore National Laboratory (LLNL).

Progress

GEMINI was used extensively to obtain estimates of the rms deflections due to random vibrations, in LLNL's National Ignition Facility (NIF) preliminary conceptual design. During the analysis of the NIF conceptual design, a thorough QA evaluation was performed, and showed that GEMINI performed the random vibration calculations correctly. **Table 1** shows rms deflection and rotation results from the NIF analyses compared with NASTRAN results.

GEMINI has also been used to perform response spectrum and time history integration analyses of structures under seismic excitation, most recently for the seismic evaluation of the Device Assembly Facility (DAF) underground structure. **Figure 1** shows typical stress results for the DAF structure subjected to earthquake loading.

GEMINI has many other element types than the comparable analysis code NIKE3D, and has linear analysis options that do not exist in NIKE3D. Because it would be expensive to add these elements and options to NIKE3D, effort was expended on the benchmarking and development of GEMINI. This effort was aimed at making analyses of the engineering design of NIF more precise and cost effective. Specifically, the following four tasks were established:

(1) update the users' manual and validate the code;
(2) convert GEMINI to the SUN, SGI and IBM workstations;
(3) incorporate composite modal damping; and
(4) calculate mode shapes using the block Lanczos method.

These tasks will now be discussed in order.

Update of Users' Manual and Validation of the Code

The GEMINI users' manual was originally distributed in October, 1984. Since then it has been updated with only letter additions and with header text in the output file. A new users' manual has been completed in draft form and incorporates the letter additions and header information. This new

Table 1. Random vibration results for target area—1% critical damping, GEMPSD analysis.

	Node	Location	X Disp. (microns)	Y Disp. (microns)	Z Disp. (microns)	X Rotation (microrad.)	Y Rotation (microrad.)	Z Rotation (microrad.)
NASTRAN	906	podium floor	1.13	1.18	0.60	0.02	0.10	0.06
GEMPSD	906	podium floor	1.07	1.13	0.58	0.02	0.10	0.05
NASTRAN	1358	diagnostic floor	1.22	1.31	0.61	0.09	0.04	0.02
GEMPSD	1358	diagnostic floor	1.17	1.26	0.58	0.09	0.04	0.02
NASTRAN	1407	upper floor	1.54	1.64	0.53	0.02	0.10	0.03
GEMPSD	1407	upper floor	1.47	1.57	0.52	0.02	0.09	0.03
NASTRAN	1811	bleacher floor	1.63	1.64	0.53	0.02	0.10	0.03
GEMPSD	1811	bleacher floor	1.56	1.57	0.52	0.02	0.09	0.03
NASTRAN	2130	dome-rib	2.31	2.32	1.28	0.03	0.03	0.00
GEMPSD	2130	dome-rib	2.21	2.23	1.24	0.03	0.03	0.00
NASTRAN	3102	chamber waist	1.27	1.31	0.80	0.14	0.15	0.02
GEMPSD	3102	chamber waist	1.22	1.26	0.76	0.13	0.13	0.02
NASTRAN	4759	final optic	1.88	1.90	1.02	0.16	0.15	0.06
GEMPSD	4759	final optic	1.77	1.80	0.96	0.13	0.14	0.03

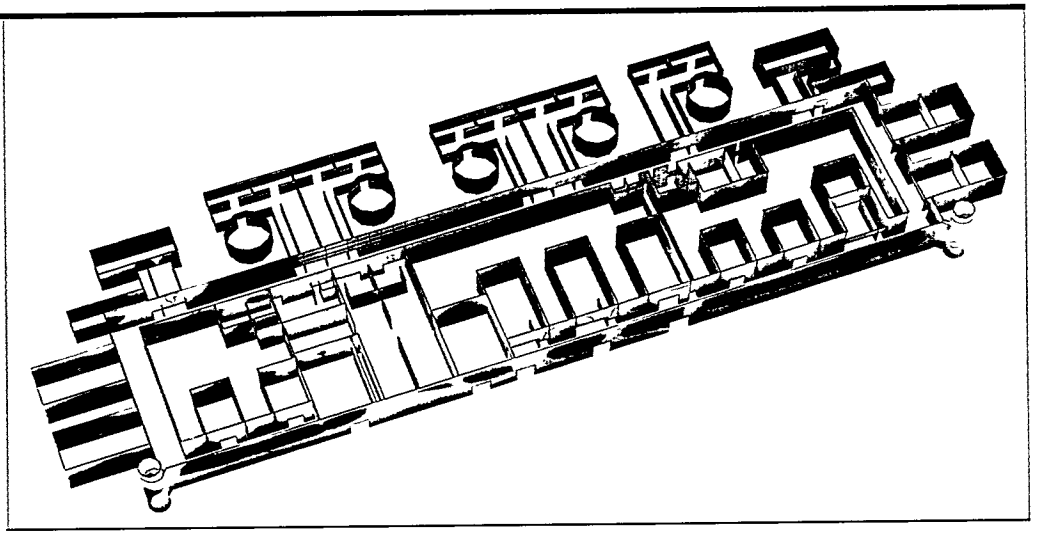


Figure 1. Effective surface stresses for the shell elements in the DAF structure under seismic loading.

manual is one-third larger than the original manual, compatible with on-line documentation, and more explicit. The SAP4 and USC suites of check problems were obtained for the validation studies. These suites of problems were augmented with some problems used in a comparable MODSAP validation. These problems have theoretical or widely accepted solutions. Because their input format was different from that used in GEMINI, the problems were converted in format and analyzed on the workstation and the main frame computers. **Table 2** shows the types of problems that were checked for the validation studies. In addition, a few additional problems were analyzed to show the workstation effects of

Table 2. GEMINI check analyses.

Analysis Name	Analysis Title	Analysis Type	Loading	Element Types	GEMINI on SUN
Problems from LLNL Users Manual for SAP 4 (subset of Problems in USC SAP IV Sample Problems)					
Problem 1.1	plane truss	static	temperature and concentrated forces	truss and boundary spring	compares well
Problem 2.1	6 story, 2 bay frame	static	distributed forces	beam	compares well
Problem 2.4a	plane frame earthquake response	dynamic	eigenvalues	beam	compares close
Problem 2.4b	plane frame earthquake response	dynamic	modal time history	beam	comparison differences
Problem 2.4c	plane frame earthquake response	dynamic	response spectrum	beam	compares well
Problem 3.1	torsion of square box beam	static	concentrated forces	2-D plane stress	compares well
Problem 4.1	thick walled	static cylinder	pressure and	2-D solid temperature	compares well
Problem 4.3a	cylindrical tube forced response	dynamic	modal time history	2-D solid	comparison differences
Problem 4.3b	cylindrical tube forced response	dynamic	direct integration	2-D solid	comparison differences
Problem 4.5	short beam with uniform load	static	pressure	2-D solid	compares well
Problem 5.1	thick walled cylinder	static	pressure and concentrated forces	3-D solid and boundary spring	compares well
Problem 6.1	spherical cap	static	pressure	shell and boundary spring	comparison differences
Problem 12.1	pipe network	static	temperature and concentrated forces	pipe and boundary spring	compares close
Problems from previous GEMINI analyses on CRAY					
FREEEI	plane frame	dynamic	eigenvalues	beam	comparison differences
FRMG5I	optics frame	dynamic	eigenvalues	beam and shell	compares close
UFRAME8I	EBIT experimental stand	dynamic	response spectrum	beam and shell	compares close
UFRAME9I	EBIT experimental stand	static	concentrated forces	beam and shell	compares close

lower bit precision, which is approximately half the precision on the main frame. Almost all of these check problems agreed with the expected solution, as also shown in **Table 2**. The differences discovered in the few problems that did not check will be resolved and corrected. The benchmark suite of problems will be checked before any new version of GEMINI is put into public use.

Conversion of GEMINI to SUN, SGI, and IBM Workstations

The conversion of GEMINI source code from the 64-bit CRAY main frame to the 32-bit SUN, SGI and IBM workstation computers has been completed with single precision arithmetic used throughout for the workstation. The converted source was generalized, so that it should easily transport to other workstations, such as the HP or DEC workstations. The conversion removed any remaining 64-bit special coding that was specific to the CRAY main frame version. During conversion, additional coding was installed to permit a user to have complete comment lines anywhere in the input. In addition, the bandwidth minimization routines in GEMINI were made more flexible and efficient for the workstation. The workstation version will allow faster analysis of small-and medium-sized problems, as well as easy check runs of larger problems. We will combine the sources for the main frame and workstation versions of GEMINI into one easily transportable source.

Incorporation of Composite Modal Damping

A version of GEMINI has been completed which calculates composite modal damping for small structures. Composite modal damping is important for structures like those in the NIF conceptual design, because these structures are usually composed of materials with widely varying damping values.

Composite modal damping is the damping which may be realized for each mode of a structure that is composed of several materials. The composite damping for a particular mode of a structure is computed in GEMINI using **Eq. 1** below.

$$\beta_j = \phi_j^T \mathbf{MK}(el)\,\phi_j / \omega_j^2 \qquad (1)$$

where

β_j	is the composite modal damping ratio for the j-th mode,
ϕ_j	is the j-th mass normalized mode shape of the undamped structure,
$\mathbf{MK}(el)$	is the modified mass (or stiffness) matrix equal to the sum over all elements of the damping ratio for an element and the diagonal mass (or full stiffness) matrix for an element, that is, $\mathbf{MK}(el) = \sum_{i=1}^{nel} \beta_i \mathbf{MK}_i$

where

nel	is the number of elements,
β_i	is the damping for the i-th element,
$\mathbf{MK}_i$	is the mass (or stiffness) matrix for the i-th element, and
ω_j	is 1.0 if using mass matrices, or is the radian frequency associated with the j-th mode if using stiffness matrices.

These damping values are called either the mass-proportioned or the stiffness-proportioned composite modal damping values. The values for composite modal damping that GEMINI calculated were checked against the values that the MODSAP code determined using the same equation. In addition, GEMINI allows a user to blend the mass- and stiffness-proportioned damping values.

Calculation of Mode Shapes

The algorithm to compute approximate modal coordinates, as derived from block Lanczos coordinates, has been adapted from the single Lanczos coordinate algorithm that is already installed in GEMINI. It is important to compute modal coordinates using the Lanczos algorithm, because only the significant modes of a structure are efficiently found using the associated spatially non-varying dynamic load shape that starts the Lanczos algorithm. It is important to use the block Lanczos algorithm because typical base-driven loadings such as the seismic or random vibrations in NIF, have three directions of uncorrelated dynamic input, constituting a block. An algorithm starting with a single Lanczos coordinate does not efficiently capture three-directional loading. Once the initial block of Lanczos coordinates ($\mathbf{G}_1$) is obtained from the partial dynamic loading, a recursive relation is used to compute additional blocks according to the algorithm given below.

Given Lanczos blocks $\mathbf{G}_j$ and $\mathbf{G}_{j-1}$ compute

$$\overline{\mathbf{R}}_j = \mathbf{K}^{-1}\,\mathbf{M}\,\mathbf{G}_j$$

$$\mathbf{A}_j = \mathbf{G}_j^T\,\mathbf{M}\,\overline{\mathbf{R}}_j$$

then, factor

$$\mathbf{R}_{j-1} = \mathbf{G}_j \mathbf{B}_j$$

$$\mathbf{R}_j = \overline{\mathbf{R}}_j - \mathbf{G}_j \mathbf{A}_j - \mathbf{G}_{j-1} \mathbf{B}_j$$

then, factor

$$\mathbf{R}_j = \mathbf{G}_{j+1} \mathbf{B}_{j+1}$$

where **K** is the stiffness matrix of the structure, **M** is the mass matrix of the structure, and $\mathbf{G}_0$ is understood to be zero. If it is anything like the single Lanczos algorithm, block Lanczos algorithm may save on the order of one order of magnitude of time for modal extraction for NIF.

The block Lanczos algorithm has been installed in a version of GEMINI that computes Lanczos block vectors for a specialized structure.

Future Work

The new users' manual will be issued, and differences in the check problems will be resolved. The specialized structure restriction will be removed from the Lanczos algorithm. Approximate eigenvalue extraction will be incorporated into the block Lanczos algorithm. Inclusion of comments, calculation of composite modal damping, and the complete block Lanczos algorithm will be incorporated into the main frame code.

Material Model Development for Thermal Forming of Anisotropic Materials

Roger W. Logan
Defense Technologies Engineering Division
Mechanical Engineering

During the latter portion of FY-95, the goal of this work was to provide a basis for follow-on work in the finite-element modeling of the thermal forming process, and of the laser and electron beam welding processes. Preliminary simulations showed that combinations of DYNA3D, ALE3D, or NIKE3D with TOPAZ3D have the potential to solve this problem, if mass, capacity, and conduction scaling are applied to the latter explicit variations. However, extensive remeshing capability will be needed for beam paths of any significant distance. Concurrently, since key applications envisioned include uranium and other anisotropic metals, an important precursor was considered to be the development of an anisotropic constitutive model that would provide an acceptable fit to the uranium data available. This model was implemented in the DYNA3D code during FY-95, with plans for implementations of extended versions into companion codes in FY-96 and beyond. Applications of the isothermal version are shown, and extension to the thermal isotropic case is discussed as the next step.

Introduction

This work was motivated by impending applications to fabrication and joining of uranium and its alloys for the weapons program at Lawrence Livermore National Laboratory (LLNL), and also for efforts involving technologies as applied to new industrial joining processes. Modeling of beam-forming and welding processes involves consideration of heat flow, fluid flow, materials response, and mechanical response of the base material. With over 15 years of numerical weld modeling effort worldwide, no models have yet been developed that couple all of the necessary heat, fluid and mechanical codes together into a single code for the prediction of all relevant welding information, such as the weld pool shape, heat affected zone properties, fusion zone properties, weld distortion, and residual stress. The state of the art in finite element analysis of beam-forming is far less mature, and quantitative results to date are almost totally lacking in the community.

However, many finite element method (FEM) codes have been developed to solve specific aspects of the overall thermal-mechanical deformation problem. These codes generally take one of three forms: heat transfer only codes for predicting thermal profiles; coupled thermal and fluid codes for predicting weld pool shape; and coupled thermal and mechanical (CTM) codes for predicting thermal-mechanical distortion and residual stress. From a practical standpoint, CTM codes have the highest payoff for solving welding and related problems at LLNL and in industry. Attempts have been made at LLNL to develop a CTM model for welding by combining the NIKE and TOPAZ codes to produce the PALM code in 2-D and the NITO code from NIKE3D[1] and TOPAZ3D.[2] However, these attempts have only been partially successful, lacking from either the proper materials model, or specific details about the welding heat-source/materials interaction, or lack of rezoning and advection capability, as present in the ALE3D (Arbitrary Lagrangian-Eulerian) code.[3]

Progress

Tools Needed

Thermal and Mechanical Codes. Three prominent CTM weld modeling efforts are: 1) the work at Carleton University[4–7]; 2) the work at Sandia National Laboratory[8–10]; and 3) the work at the University of Illinois.[11,12] At Carleton

University, the emphasis has been on the application of a fast, nonlinear, transient 3-D FEM code to solve thick-section welding problems with complex geometries. The meshes may contain 50,000 nodes, and can be analyzed on workstations, using preconditioned conjugate gradient solvers.[13] At Sandia National Laboratory, the emphasis to date has been on the application of a 2-D FEM code, PASTA, with plane stress assumptions to solve thin, planar, axisymmetric, autogeneous arc welding problems. This code was originally developed to solve mechanical forming problems, and is a fully explicit code that takes advantage of vector processors.[8]

Coupling the LLNL TOPAZ and NIKE codes has already been performed at LLNL,[14] and they can be applied qualitatively to the beam-forming and welding problem, as shown in **Fig. 1**. The example shows beam-shaping (thermal forming), which is not only a first step toward quantitative weld-distortion analysis, but an emerging manufacturing technology in its own right. However, additional work needs to be performed to optimize a 3-D code to handle realistic problems, and to incorporate the proper materials and heat source models in a generalized fashion to make the code provide accurate 3-D quantitative distortions on a reasonable computational timescale. As shown in **Table 1** below, none of our existing LLNL codes will truly handle this class of problem. Even the versions being developed [including ALE with Thermal effects (TALE) and ALE with a NIKE-like implicit solver (NALE)] fall short in one capability or another. Future options include NALET (from ALE/TOPAZ), DYNAMITE (from DYNA with MESHING, including TOPAZ Explicit) and NITRO (from NIKE/TOPAZ). Eventually, one or more of this set will fill the need, especially with the addition of an h-method (subdivision of elements) rezoner to the best candidate.

Each of the three proposed thermal-mechanical code systems in **Table 1** has its strengths regarding thermal forming and welding in 3-D. The DYNA-NIKE-TOPAZ systems are more developed in terms of material models and implicit solvers (NIKE) due to previous work, while the ALE-TOPAZ system has more strength in what is essentially the r-method rezoning (relocation of nodes) methodology with careful advection of state variables. Cross-pollination of the two has long been one strategic objective, and at the same time provides us not one, but two, sources of day-to-day production analysis. A necessary feature to make long path lengths tractable is the availability of at least r-method, and preferably h-method, rezoning. Enhancement of this capability is addressed in a separate work. The r-method is

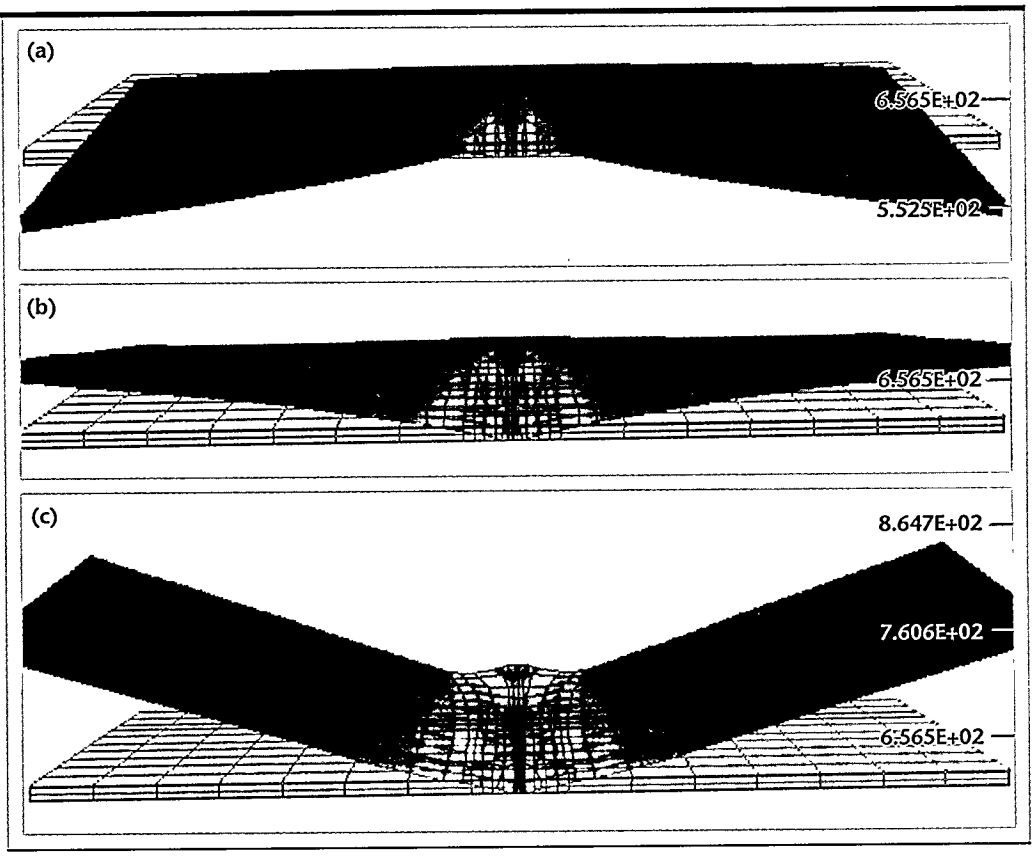


Figure 1. Behavior of flat plate sample during electron beam pass for thermal forming: a) initial heating and expansion of upper surface; b) cooling behind beam, and contraction of heated surface; and c) end of travel, more contraction leading to final first-pass shape. Multiple passes of beam lead to final formed shape.

analogous to ALE in most ways. Thus, the ALE3D code is in this respect closer to the problem at hand than the NIKE-TOPAZ system. Still, extensive work is required (and underway) in coupling to TOPAZ and a solver.

Material Models. In order to attain widespread usage (and thus reliability based on many users' experience), a material model must also be tractable in both formulation and logic of material parameter input. It is for this reason that we are considering, as a first step, development of a thermal version of the most basic or 'total method' rate and pressure-dependent Forming Limit Diagram model,[15] as described currently in the DYNA3D[16] and NIKE3D codes. The rate, pressure, structure, and failure limits in this model are simple in form, additive or multiplicative; thus, any number of features may be toggled to add simplicity or complexity. Although input is largely in tabular form, the relations can be input to correspond to closed-form expressions. The thermal version (Model 35T) is envisioned to take the following constitutive form, using tabular relations for most of the constitutive inputs:

$$\bar{\sigma} = C_{ex}(\bar{\varepsilon}^p)C_{edm}(\dot{\varepsilon})C_{pxm}\left(\frac{P}{X}\right)\times S(T,t,\bar{\varepsilon}^p)\exp\left(\frac{-Q}{RT}\right) \quad (1a)$$

where we begin in the isotropic case with the von Mises yield function:

$$\bar{\sigma}^2 = \frac{1}{2}\left[(\sigma_b-\sigma_c)^2+(\sigma_c-\sigma_a)^2+(\sigma_a-\sigma_b)^2+6\sigma_{bc}^2+6\sigma_{ca}^2+6\sigma_{ab}^2\right] \quad (1b)$$

A major feature of all metals, by definition, is anisotropy of both elastic and plastic properties. In many cases, this part of the constitutive behavior can be, or is of necessity, ignored in numerical calculations. Since one of our major materials of interest (alpha uranium) is highly anisotropic, it will be necessary for good quantitative results to implement anisotropy into the beam-forming/welding systems. Due to this, and other related considerations, the development of the thermal-forming and joining system began on a limited basis, with the implementation during FY-95 of an enhanced anisotropic capability for the isothermal case. We find it is often efficient to implement a methodology in the explicit DYNA3D first, bearing in mind that the formulation must be appropriate for subsequent implementation into NIKE3D and ALE3D as well. Such is the case in this work. The isothermal example shown below uses the implementations of anisotropic yield surfaces in DYNA3D to show that alternate, non-quadratic yield criteria can provide more accurate predictions of the earing observed in

Table 1. Characteristics of LLNL codes.

	Code Name	Implicit Solver	Elastic/Plastic Anisotropy	Thermal/Structure Rate/Failure	Heat: Source Conduction Convection	2.5-D/3-D Rezone
Existing Codes:	DYNA	No	Yes	No	No	No
	NIKE	Yes	No	No	No	No
	TOPAZ	Yes	No	No	Yes	No
	ALE	No	No	No	No	Yes-r
Being Developed:	NITO	Yes	No	No	Yes	No
	(T)ALE	No	No	No	Yes	Yes-r
	(N)ALE	Yes	No	No	No	Yes-r
Proposed Options:	DYNAMITE	No	Yes	Yes	Yes	Yes-r,h
	NALET	Yes	Yes	Yes	Yes	Yes-r,h
	NITRO	Yes	Yes	Yes	Yes	Yes-h,h

isothermal sheet forming processes, thus providing the groundwork for the subsequent thermal version. One of the most straightforward and commonly used is Hill's quadratic equation[17] for anisotropic plastic flow:

$$\bar{\sigma}^2 = \frac{1}{(R+1)}\left[F(\sigma_b - \sigma_c)^2 + G(\sigma_c - \sigma_a)^2 + H(\sigma_a - \sigma_b)^2 + D\right] \quad (2)$$

Equation 2 relates the effective stress to the three normal components of Cauchy stress, with the term D containing the shear stress terms:

$$D = 2L\sigma_{bc}^2 + 2M\sigma_{ca}^2 + 2N\sigma_{ab}^2 \quad (3)$$

The values for the constants in **Eqs. 2** and **3** can be expressed in terms of the strain ratios R, Q, and P as described above, with the following additional relations needed:

$$F = R/P \quad (4)$$

$$G = 1 \quad (5)$$

$$H = R \quad (6)$$

$$L = (Q_{bc} + \tfrac{1}{2})(R+1) \quad (7)$$

$$M = (Q_{ca} + \tfrac{1}{2})(R+Z) \quad (8)$$

$$N = (Q_{ab} + \tfrac{1}{2})(1+Z) \quad (9)$$

In addition to being comparatively straightforward to implement, the quadratic Hill criterion permits the relatively simple calculation of the ratio, $X(\theta)/X$, of the yield stress, in a direction at an angle in the plane of the sheet to the rolling direction, as well as the calculated R-value in that direction:

$$\frac{X(\theta)}{X} = \left[\frac{R+1}{2Ns^2c^2 + R(c^4 - s^4) + c^4 + Zs^4}\right]^{1/2} \quad (10)$$

$$R(\theta) = \frac{1}{(Zs^2 + c^2)}\left[R + ((2Q_{ab}+1)(1+Z) - Z - 1 - 4R)s^2c^2\right] \quad (11)$$

where:

$$c = \cos(\theta)$$
$$s = \sin(\theta) \quad (12)$$

Equations 10 and **11** are relatively easy to comprehend, but show several trends that are not usually borne out by experimental data. For example, at 45° to the rolling direction, Hill predicts (if $R = P$):

$$\frac{X(45^\circ)}{X} = \left[\frac{R+1}{Q_{ab}+1}\right]^{1/2} \quad (13)$$

Equation 13 shows a high dependence of yield stress on orientation in the plane of the sheet. This dependence is a likely factor in overprediction of earing as in **Fig. 2**, as the strong material in the 45°-direction tends to pull in, to form the wall of the punch (forming a deep trough), while compressing the 0°- and 90°-walls (forming high ears). One yield criterion which in principle should show better agreement with experiment is **Eq. 14**, extending Hill to a non-quadratic form, with values of the exponent a in the range of $a = 8$ for fcc[18] (the Hosford criterion), and $a = 6$ for bcc[19] metals:

$$\bar{\sigma}^a = \frac{1}{(R+1)}\left[F(\sigma_2 - \sigma_3)^a + G(\sigma_3 - \sigma_1)^a + H(\sigma_1 - \sigma_2)^a\right] \quad (14)$$

This equation predicts much milder dependencies of strength ratios in stress states and directions other than tension in the rolling direction. For example, in biaxial tension, where $\sigma_a = \sigma_b = B$, forms predict the relatively simple value:

$$\frac{B}{X} = \left(\frac{R+1}{Z+1}\right)^{1/a} \quad (15)$$

Clearly, values of $a > 2$ will give a much milder dependence, often observed experimentally, in unalloyed uranium and other metals of interest. Further, if we assume the R-value orientation dependence as in **Eq. 11**, the orientation dependence of uniaxial flow stress is much milder. This ratio is important in the flange of the cup during draw-in, and thus indicates promise for greater accuracy in earing calculations. However, difficulties arise in the implementation of the Hosford criterion for cases in other than principal stress/strain space. This stems

from the lack of shear terms in the criterion. **Equation 13** must remain in principal stress space,

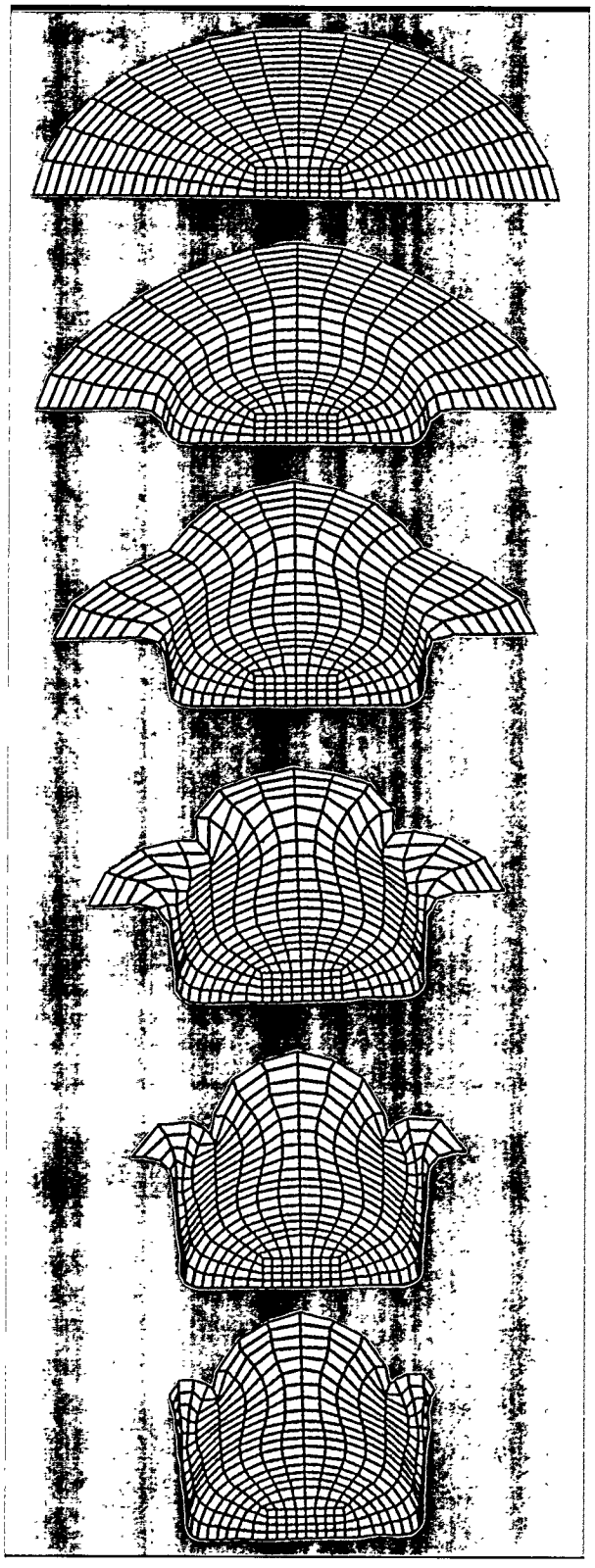


Figure 2. DYNA3D simulation of earing development during deep drawing, R, Q_{ab}, P = 4, 1, 4, and a = 2.

to be used without spurious results and non-convexity problems. To do so, we must make several approximations in updating the Cauchy stress tensor (expressed as a vector of six) when using **Eq. 13.** First and foremost is a rotation to the principal stress coordinate space. This does not normally coincide with either the material (rolling and transverse direction) coordinate system, nor with the axes of principal strain. Our first assumption is that the axes of principal stress and strain coincide, although for planar isotropy we know that they normally will not. However, this assumption, which leads us to ignore cross-terms in the constitutive matrix, is believed to lead to only small errors for the degree of anisotropy observed in most sheet metals. In the following section, we will demonstrate the effect of the chosen yield surface (**Eq. 2** or **13**) on the extent of earing in cupping.

Experimental Verification: Comparison to the Earing Phenomenon. To demonstrate the correlation of DYNA earing simulations with experiments, we compare the extensive data obtained by Wilson and Butler[20] with that obtained using simulations with either $a = 2$ or $a = 8$ in **Eq. 14.** Numerous runs with $R = P$ were made as outlined in **Table 2.** These results are plotted as $\Delta h/h$ vs $\Delta R/R$ in **Fig. 3,** along with the original data from Wilson and Butler. It is strikingly clear that the use of $a = 2$ (identically Hill) overpredicts earing by about a factor of two. In contrast, the use of $a = 8$ matches almost exactly the observed earing data; in fact $a = 8$ seems to under predict earing slightly. However, recall that we were forced to ignore cross-terms in the matrix C_{ij}^{ep} above when using $a = 8$. We do not need to do this when $a = 2$, but we can choose to do so to examine the effect. There is indeed some difference, as shown by the data labeled '$a = 2x$' in **Table 2** and in **Fig. 3.** Thus, if cross-terms existed for $a = 8$, we might expect to see slightly more calculated earing, and perhaps slightly over-predict the data instead of slightly under-predicting it. Either way, the value of $a = 8$ is much more accurate than $a = 2$.

Several variations of runs were made in the process of compiling the DYNA comparisons in **Fig. 3.** These included variations of the SVE (Selective Velocity Enhancement) factor from 20 x to 50 x, with very little difference noticed in earing. The runs presented here were run at SVE = 20, and otherwise the parameters were as used previously.[21,22] Meshing consisted of 24 nodes radially along the blank, and 9 nodes around the quarter-symmetry hoop direction from 0 to 90°. Meshing of 12 × 9 or even 12 × 5 nodes gave reasonable ear development early on, but 5 nodes

in the hoop direction showed difficulties in completing the cupping operation. A much more important fact is that in the DYNA3D codes there is obviously a degree of freedom in the hoop direction in the flange. Early simulations by the author[23] considerably under-predicted earing, even though $a = 2$ was used in the MARMOT code of that work. However, a key constraint of the MARMOT sheet-forming code was the assumption of no displacement in the hoop direction. When sliding boundary planes were used in DYNA to preclude motion in the hoop direction at $\theta = 0$, 22.5, 45, 67.5, and 90°, earing is cut drastically (see the data for '2m' in **Table 2** and 'u2 = 0' in **Fig. 3**). This explains why the early MARMOT code under-predicted earing even with Hill, and shows the importance of hoop displacement and shear in the earing process.

Future Work

We have discussed the issues of thermal anisotropic elastic/plastic material models necessary for a basis for a thermal-forming and joining finite-element capability, keyed toward uranium and other anisotropic metals. As a first step, we have implemented and verified yield surfaces (quadratic and non-quadratic) that will enable us to closely approximate the known trends of behavior in anisotropic plasticity of uranium and other metals for which von Mises or Hill is not an adequate fit.

Future work will involve implementation of the elastic/plastic anisotropy in a thermal context as discussed above, and likely in combination with models due to Bammann.[24] This, combined with a coupled thermal/mechanical/rezoning capability, will enable us to address this technology in a quantitative manner leading to improved designs for forming and joining.

Acknowledgments

The author wishes to thank J. W. Elmer for many helpful discussions, for experimental data, and for a thorough written review of the state of the art in weld modeling, some of which has been excerpted here. G. J. Leake performed the initial simulations of welding with his NIKE/TOPAZ model. Discussions with R. G. Couch were helpful in determining the future course to be taken regarding ALE3D and rezoning methods.

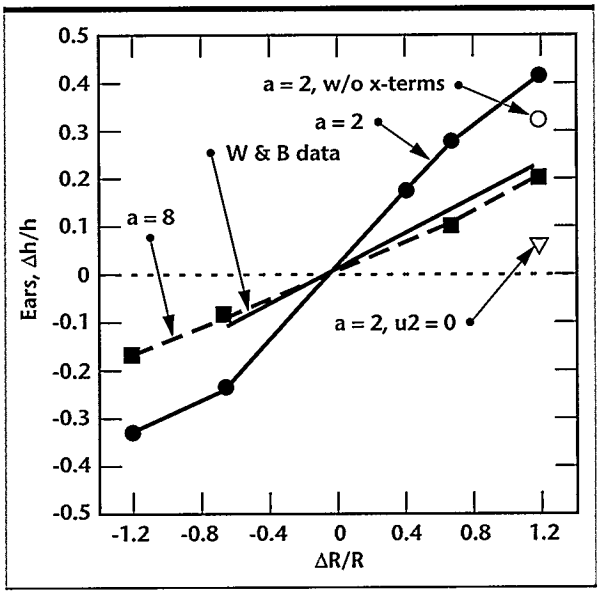


Figure 3. Plot of earing (Δh/h vs ΔR/R) including Wilson and Butler data, compared to DYNA results with a = 2 and a = 8.

Table 2. Data used in DYNA simulations of earing.

a	R	Q	P	h	Δh	ΔR/R	Δh/h
2	4.0	1.0	4.0	7.32	2.98	1.20	.407
2	2.0	1.0	2.0	7.20	1.92	0.67	.267
2	1.5	1.0	1.5	7.22	1.20	0.40	.166
2	0.5	1.0	0.5	7.08	-1.50	-0.67	-.212
2	1.0	2.0	1.0	7.18	-1.72	-0.67	-.240
2	0.5	2.0	0.5	7.26	-2.40	-1.20	-.330
2x	4.0	1.0	4.0	7.62	2.34	1.20	.312
8	4.0	1.0	4.0	7.84	1.54	1.20	.196
8	2.0	1.0	2.0	7.64	0.73	0.67	.096
8	1.0	2.0	1.0	7.78	-0.79	0.67	-.102
8	0.5	2.0	0.5	7.66	-1.27	-1.20	-.166
2m	2.0	1.0	2.0	7.48	0.36	0.67	.048

References

1. Maker, B. N., R. M. Ferencz, and J. O. Hallquist (1994), *NIKE3D: A Nonlinear, Implicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics—User's Manual,* University of California, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-MA-105268).

2. Shapiro, A. B. (1985), *TOPAZ3D: A Three-Dimensional Finite Element Heat Transfer Code—User's Manual,* University of California, Lawrence Livermore National Laboratory, Livermore, Calif. (UCID-20484).

3. Anderson, S., E. Dube, S. Futral, I. Otero, and R. Sharp (1995), *ALE3D: An Arbitrary Lagrange/Eulerian 3D Code System—User Manual,"* University of California, Lawrence Livermore National Laboratory, Livermore, Calif. (in press).

4. Goldak, J., M. McDill, A. Oddy, R. House, X. Chi, M. Bibby (1986),"Computational Heat Transfer for Weld Mechanics," *Advances in Welding Science and Technology,* S. A. David, Ed. (ASM International, Metals Park, Ohio), p. 15.

5. McDill, J. M., A. S. Oddy, and J. A. Goldak (1993),"Comparing 2-D Plane Strain and 3-D Analyses of Residual Stresses In Welds," *International Trends in Welding Science and Technology,* (ASM International, Metals Park, Ohio), p. 105.

6. Goldak, J., and M. Bibby (1988), "Computational Thermal Analysis of Welds: Current Status and Future Directions," *Modeling of Casting and Welding Processes IV,* A. F. Giamei and G. J. Abbaschian, Eds. (TMS, Warrendale, Pa.), p. 153.

7. Oddy, A. S., J. A. Goldak, and J. M. J. McDill (1990), "Transformation Effects in the 3-D Finite Element Analysis of Welds," *Recent Trends in Welding Science and Technology,* S. A. David and J. M. Vitek, Eds. (ASM International, Metals Park, Ohio), p. 97.

8. Winters, W. S., and W. E. Mason (1988), "The Computer Code PASTA and its Application to the Traveling Tungsten Arc Welding Problem," *Modeling of Casting and Welding Processes IV,* A. F. Giamei and G. J. Abbaschian, Eds. (TMS, Warrendale, Pa.), p. 197.

9. Winters, W. S., and K. W. Mahin (1991), "PASTA2D Modeling of an Axisymmetric Stationary Gas Tungsten Arc Weld," *Modeling of Casting, Welding, and Advanced Solidification Processes V,* M. Rappaz, M. R. Ozgu, and K. W. Mahin, Eds. (TMS, Warrendale, Pa.), p. 97.

10. Mahin, K. W., W. S. Winters, J. Krafcik, T. Holden, R. Hosbons, and S. MacEwen (1990), "Residual Strain Distributions in Gas Tungsten Arc Welds: A Comparison Between Experimental Results and Weld Model Predictions," *Recent Trends in Welding Science and Technology,* S. A. David and J. M. Vitek, Eds. (ASM International, Metals Park, Ohio), p. 83.

11. Terkiwal, P., and J. Mazumder (1988), "Three-Dimensional Finite Element Analysis of Multi-Pass GMA Welding," *Modeling of Casting and Welding Processes IV,* A. F. Giamei and G. J. Abbaschian, Eds. (TMS, Warrendale, Pa.), p. 167.

12. Terkiwal, P., and J. Mazumder (1990), "Thermomechanical Analysis of Residual Strains and Stresses in a GMA Weld," *Recent Trends in Welding Science and Technology,* S. A. David and J. M. Vitek, Eds. (ASM International, Metals Park, Ohio), p. 91.

13. Goldak, J. (1991), "Computational Weld Mechanics as a Coupled Problem," *Modeling of Casting, Welding, and Advanced Solidification Processes V,* M. Rappaz, M. R. Ozgu, and K. W. Mahin, Eds. (TMS, Warrendale, Pa.), p. 85.

14. Mahin, K. W., A. B. Shapiro, and J. O. Hallquist (1986), "Assessment of Boundary Condition Limitations of the Development of a General Computer Model for Fusion Welding," *Advances in Welding Science and Technology,* S. A. David, Ed. (ASM International, Metals Park, Ohio), p. 215.

15. Logan, R. W., D. B. Thomas, and G. K. Young (1995), *Implementation of a Pressure and Rate Dependent Forming Limit Diagram Model into NIKE and DYNA,* Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-JC-105760), prepared for ASME-WAM 95, San Francisco, Calif.

16. Whirley, R. G., and J. O. Hallquist (1991), *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).

17. Hill, R. (1948), *The Mathematical Theory of Plasticity,* Clarendon Press, Oxford, England.

18. Hosford, W. F. (1979), "On Yield Loci of Anisotropic Cubic Metals," *Proc. 7th N. Amer. Metal Working Research Conf.* (SME, Dearborn, Mich.), p. 191.

19. Logan, R. W., and W. F. Hosford (1980), "Upper-Bound Anisotropic Yield Locus Calculations Assuming <111> -Pencil Glide," *Int. J. Mech. Sci.* **22**, p. 419.

20. Wilson, D. V., and Butler, R. D. (1962), "The Role of Cup-Drawing Tests in Measuring Drawability," *J. Inst. Metals* **90**, p. 473.

21. Whirley, R. G., B. E. Engelmann, and R. W. Logan (1992), "Some Aspects of Sheet Forming Simulation Using Explicit Finite Element Techniques," *CED*-Vol. **5**, M. J. Saran, A. B. Pifko, N. Kikuchi, and K. K. Tamma, Eds. (ASME).

22. Whirley, R. G., B. E. Engelmann, and R. W. Logan (1992), "Numerical Methods for Simulation of Industrial Metal Forming Processes", *AMD* Vol. **156**, M. J. Saran, A. B. Pifko, N. Kikuchi, and K. K. Tamma, Eds. (ASME, N.Y., N.Y.).

23. Logan, R. W. (1984), *Sheet Formability: Simulation and Experiment*, Ph.D, thesis, University of Michigan—Ann Arbor, Mich.

24. Bammann, D. J. (1990), "Modeling Temperature and Strain Rate Dependent Large Deformations of Metals," *Applied Mechanics Reviews*, **43**(5), part 2, p. s312.

ultilength-Scale Modeling: Crystal-Plasticity Models in Implicit Finite Element Codes

David D. Sam
New Technologies Engineering Division
Mechanical Engineering

The material modeling capabilities of NIKE3D have recently been expanded to include the anisotropic plastic behavior of single and polycrystals that flow by slip on specified crystallographic planes, using an elastic-viscoplastic constitutive relation. The implementation of the model is based on a forward gradient integration procedure which was previously implemented in the commercial nonlinear code, ABAQUS™. Details of the kinematic description for the single crystal constitutive model are included in this report. The material model is demonstrated by simulating a 3-D channel-die compression test of an aluminum bicrystal.

Introduction

A multidisciplinary team of researchers at Lawrence Livermore National Laboratory has recently completed the first year of an initiative that involves the development of physics-based multi-length-scale models to optimize casting and deformation processing of polycrystalline materials. The purpose of this project is to develop models to simulate the material's mechanical response on the meso-length scale. The meso-length scale refers to the realm where the material microstructure (e.g., grain size and shape, crystallite orientation) is resolved, but dislocations are homogenized. This phase of the initiative was to build a bicrystal consisting of two crystallites joined by a single, planar grain boundary. The bicrystal will be deformed by channel-die compression while recording the deformation response. Comparisons will be made between the measured responses with predictions using a 3-D finite element simulation that uses a crystal plastic material model.

In this study, we are primarily interested in the deformation response of crystalline solids composed of face-centered cubic (fcc) crystallites. For an fcc system, the deformation is assumed to result solely from shear slipping along close-packed crystallographic planes. **Figure 1** illustrates three possible deformation modes that may operate on one of the slip planes in an fcc unit cell. From the figure we see that the slip planes, which are defined by their unit normal vectors, are of the type given by the set

$$\hat{\mathbf{m}} = \left\{ \frac{1}{\sqrt{3}}(1,1,1) \right\} \quad . \tag{1}$$

Each of these planes has three slip directions (excluding their negatives), and is of the type given by the set

$$\hat{\mathbf{s}} = \left\{ \frac{1}{\sqrt{2}}(1,\bar{1},0) \right\} \quad . \tag{2}$$

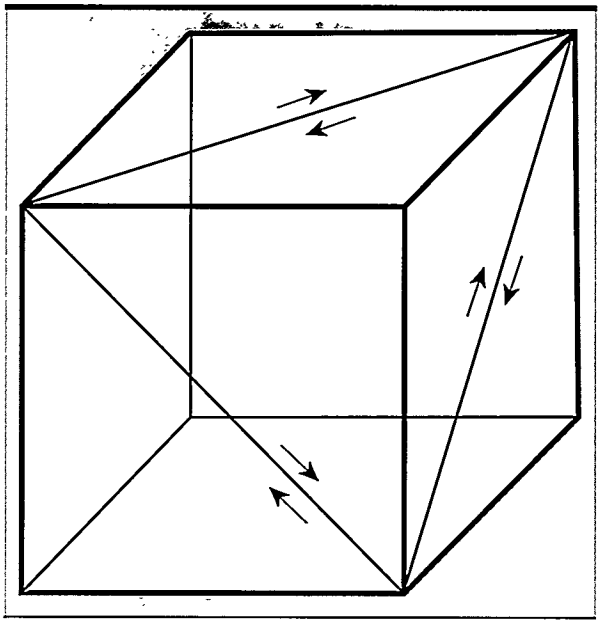


Figure 1. Shear deformation illustrated for a cubic crystal unit cell.

By inspection of the fcc unit cell, it is seen that it has four unique slip planes, each with three unique slip directions (not counting their negatives). A slip plane, together with an associated slip direction, constitutes a slip system. Note that for each slip system, $\hat{\mathbf{s}}$ and $\hat{\mathbf{m}}$ are orthogonal. There are 12 slip systems that may operate in the fcc single crystal to achieve a deformation. An arbitrary deformation in a single crystal is achieved by the composite effect of all contributing (active) slip systems.

Our objective is to develop a tensor-based single crystal constitutive model that includes the kinematics of the crystal's deformation response, coupled with a scalar law that relates the shear stress, resolved onto the slip plane and slip direction, with the associated slip rate.

Progress

From the description of the crystal's deformation response, to achieve plastic slipping on a given slip system, k (k = 1,...,12), the yield criterion must be satisfied, according to

$$\hat{\mathbf{s}}^{(k)} \cdot \sigma \cdot \hat{\mathbf{m}}^{(k)} = s_i^{(k)} \sigma_{ij} b_j^{(k)} = \tau^c \quad , \tag{3}$$

where σ is the symmetric Cauchy stress tensor. In the above expression and hereafter, summation is implied over the repeated tensorial subscripts. **Equation 3** states that plastic slipping initiates when the shear stress, resolved onto a slip system, reaches a critical value. **Equation 3** is known as Schmid's Law.

The development of the constitutive model that follows falls within the general framework of the finite deformation theory of crystals.[1] After Asaro and Rice,[2] we assume that the elastic response of the crystallite is undisturbed by plastic slip. The plastic deformation of a crystal may then be envisioned as a decomposition into two subsequent deformations (see **Fig. 2**): a homogeneous shearing on a given slip system, followed by the simultaneous rigid body rotation and elastic distortion of the lattice. The deformation gradient tensor $\mathbf{F}$ is defined as the gradient of the time-varying spatial coordinates, $\mathbf{x}$, with respect to the fixed material reference coordinates, $\mathbf{x}^0$. Mathematically, the decomposition of $\mathbf{F}$ is expressed as

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{x}^0} = \mathbf{F}^* \cdot \mathbf{F}^p \tag{4}$$

where the superscripts * and p refer to the elastic distortion part, and the plastic slip part, respectively.

The kinematic description of the crystal mechanics now follows that presented by Peirce, Asaro and Needleman.[3] Thus, according to the above decomposition of $\mathbf{F}$, the velocity gradient is decomposed additively as follows:

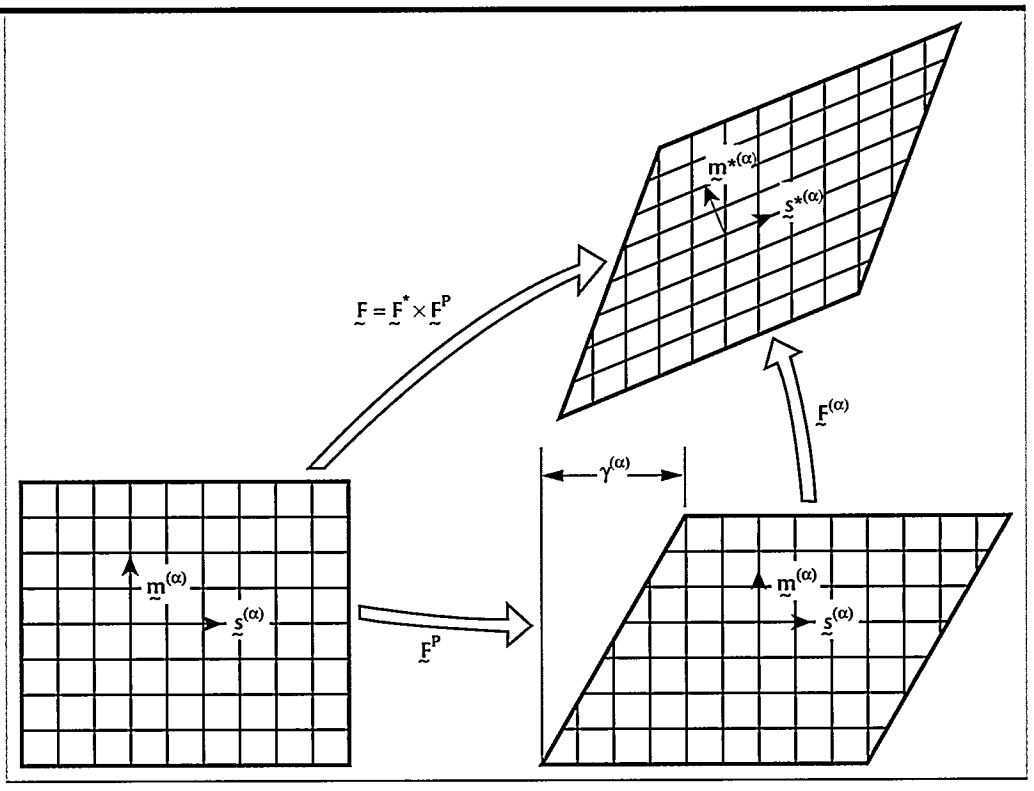


Figure 2. Decomposition of deformation into plastic slip and lattice distortion/rotation components.

$$\mathbf{L} = \frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{x}} = \dot{\mathbf{F}} \cdot \mathbf{F}^{-1} = \dot{\mathbf{F}}^* \cdot \mathbf{F}^{*-1} + \mathbf{F}^* \cdot \dot{\mathbf{F}}^p \cdot \mathbf{F}^{p-1} \cdot \mathbf{F}^{*-1} \quad , \tag{5}$$

or equivalently, $\mathbf{L} = \mathbf{L}^* + \mathbf{L}^p$. Now, from the geometry of a single crystal that is flowing by a slip rate, $\dot{\gamma}$, on a predetermined slip system, α,

$$\dot{\mathbf{F}}^p \cdot \mathbf{F}^{p-1} = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \hat{\mathbf{s}}^{(\alpha)} \otimes \hat{\mathbf{m}}^{(\alpha)} \quad , \tag{6}$$

where, for generality, the summation is over N potentially active slip systems. Substituting **Eq. 6** into **Eq. 5** we have

$$\mathbf{L}^p = \mathbf{F}^* \dot{\mathbf{F}}^p \cdot \mathbf{F}^{p-1} \mathbf{F}^{*-1} = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \mathbf{s}^{*(\alpha)} \otimes \mathbf{m}^{*(\alpha)} \quad , \tag{7}$$

where

$$\mathbf{s}^{*(\alpha)} = \mathbf{F}^* \cdot \hat{\mathbf{s}}^{(\alpha)} \qquad \mathbf{m}^{*(\alpha)} = \hat{\mathbf{m}}^{(\alpha)} \cdot \mathbf{F}^{*-1} \quad . \tag{8}$$

Because the unit vector $\hat{\mathbf{s}}$ convects with the lattice during the deformation $\mathbf{F}^*$, we see from **Eq. 8** that $\hat{\mathbf{s}}^*$ and $\hat{\mathbf{m}}^*$ are orthogonal; but because of the stretching of the lattice, they are not, in general, unit vectors. Now, decomposing the velocity gradient tensor into its symmetric and asymmetric parts, **D** and **W**, respectively, we find their respective plastic slip components as

$$\mathbf{D}^p = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \frac{1}{2}\left(\mathbf{s}^{*(\alpha)} \otimes \mathbf{m}^{*(\alpha)} + \mathbf{m}^{*(\alpha)} \otimes \mathbf{s}^{*(\alpha)}\right) = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \mathbf{P}^{(\alpha)} \tag{9}$$

$$\mathbf{W}^p = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \frac{1}{2}\left(\mathbf{s}^{*(\alpha)} \otimes \mathbf{m}^{*(\alpha)} - \mathbf{m}^{*(\alpha)} \otimes \mathbf{s}^{*(\alpha)}\right) = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \mathbf{W}^{(\alpha)} \quad . \tag{10}$$

To formulate a constitutive relation between stress and deformation rate, we begin by selecting a resolved shear stress on each slip system α, $\tau^{(\alpha)}$, such that it is work-conjugate to the associated slip rate, $\dot{\gamma}^{(\alpha)}$, and write the plastic dissipation energy as

$$\tau : \mathbf{D}^p = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \tau : \mathbf{P}^{(\alpha)} = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \mathbf{m}^{*(\alpha)} \cdot \tau \cdot \mathbf{s}^{*(\alpha)} = \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)} \tau^{(\alpha)} \tag{11}$$

with respect to the reference coordinate system. Here τ is the Kirchoff stress tensor defined according to $\tau = \det(\mathbf{F})\sigma$, where σ is the Cauchy stress tensor.

The existence of a relationship between elastic (lattice) strain and stress leads naturally to the notion of a strain energy function, ϕ, where, if $\mathbf{E}^*$ is the Lagrange strain tensor that is associated to the lattice deformation which is work-conjugate to the 2nd Piola-Kirchoff stress tensor, $\mathbf{P}^*$:

$$\mathbf{E}^* = \frac{1}{2}\left(\mathbf{F}^{*T} \cdot \mathbf{F}^* - \mathbf{I}\right) \quad , \tag{12}$$

$$\mathbf{P}^* = \mathbf{F}^{*-1} \cdot \tau \cdot \mathbf{F}^{*-T} \quad , \tag{13}$$

then

$$\mathbf{P}^* = \frac{\partial \phi}{\partial \mathbf{E}^*} \tag{14}$$

and

$$\dot{\mathbf{P}}^* = \frac{D\mathbf{P}^*}{Dt} = \frac{D}{Dt}\left(\frac{\partial \phi}{\partial \mathbf{E}^*}\right) = \frac{\partial^2 \phi}{\partial \mathbf{E}^* \cdot \partial \mathbf{E}^*} : \frac{D\mathbf{E}^*}{Dt} = \frac{\partial^2 \phi}{\partial \mathbf{E}^* \cdot \partial \mathbf{E}^*} : \mathbf{E}^* = \mathbf{C} : \dot{\mathbf{E}}^* \quad . \tag{15}$$

The time rate of change of the Kirchoff stress is obtained from **Eqs. 13** and **15** and is given by

$$\dot{\tau} = \mathbf{F}^* \cdot \left(\mathbf{C} : \dot{\mathbf{E}}^*\right) \cdot \mathbf{F}^{*T} + \dot{\mathbf{F}}^* \cdot \mathbf{F}^{*-1} \cdot \tau + \tau \cdot \mathbf{F}^{*-T} \cdot \dot{\mathbf{F}}^{*T} \quad . \tag{16}$$

Similarly, as stated in **Eq. 5**, the lattice part of the velocity gradient may be decomposed into its symmetric and skew-symmetric parts, according to

$$\dot{\mathbf{F}}^* \cdot \mathbf{F}^{*-1} = \mathbf{D}^* + \mathbf{W}^* \tag{17}$$

and, from **Eq. 12** the lattice Lagrange strain rate may be written as

$$\dot{\mathbf{E}}^* \cdot \mathbf{F}^{*\mathrm{T}} = \mathbf{D}^* \cdot \mathbf{F}^* \quad . \tag{18}$$

Substitution of **Eqs. 17** and **18** into **Eq. 16** simplifies the expression for Kirchoff stress rate as

$$\dot{\tau} = \boldsymbol{L}:\mathbf{D}^* + \mathbf{W}^* \cdot \tau - \tau \cdot \mathbf{W}^* \tag{19}$$

where the generalized convected modulus, $\boldsymbol{L}$, is defined (in Cartesian coordinates) as

$$L^{ijkl} = F^{*i}{}_m F^{*j}{}_n F^{*k}{}_o F^{*l}{}_p C^{mnop} + \frac{1}{2}\left(\delta^{ik}\tau^{jl} + \delta^{il}\tau^{jk} + \tau^{il}\delta^{jk} + \tau^{ik}\delta^{jl}\right). \tag{20}$$

Here, δ^{ij} is the Krönecker delta, and C^{mnop} are the components of the local lattice elastic modulus for materials with cubic symmetry (which have three independent components).

Substituting the lattice elastic deformation and spin rates in terms of their total and plastic slip components, $\mathbf{D}^* = \mathbf{D} - \mathbf{D}^{\mathrm{p}}$ and $\mathbf{W}^* = \mathbf{W} - \mathbf{D}^{\mathrm{p}}$, and using the definition of the Jaumann rate of a rank 2 tensor, from **Eq. 19**, the Jaumann rate of Kirchoff stress is given by

$$\overset{\nabla}{\tau} = \dot{\tau} - \mathbf{W}\cdot\tau + \tau\cdot\mathbf{W} = \boldsymbol{L}:\mathbf{D} - \boldsymbol{L}:\mathbf{D}^p - \mathbf{W}^p\cdot\tau + \tau\cdot\mathbf{W}^p \tag{21}$$

or, expressing it in terms of slip system geometry, with the definitions of **Eqs. 9** and **10**, we obtain

$$\overset{\nabla}{\tau} = \boldsymbol{L}:\mathbf{D} - \sum_{\alpha=1}^{N} \dot{\gamma}^{(\alpha)}\mathbf{R}^{(\alpha)} \quad , \tag{22}$$

which gives a constitutive expression relating stress rate to strain rate and slip rate. Here the tensor $\mathbf{R}^{(\alpha)}$ is defined by

$$\mathbf{R}^{(\alpha)} = \boldsymbol{L}:\mathbf{P}^{(\alpha)} + \mathbf{W}^{(\alpha)}\cdot\tau - \tau\cdot\mathbf{W}^{(\alpha)} \tag{23}$$

We now note that the constitutive description given in **Eq. 22** is completed only when an additional (scalar) constitutive relation is introduced which relates the rate of plastic slipping in the lattice to the resolved shear stress and slip system hardness. Because of problems with nonuniqueness of predicted active slip systems, inherent in the above formulation for rate-independent crystals, a viscoplastic approximation has been used.[4,5] For a viscoplastic crystal, following the formulation of Peirce, Asaro and Needleman,[5] we select the power-law relation used by Hutchinson,[6] in a self-consistent estimate of the creep behavior of polycrystals:

$$\dot{\gamma}^{(\alpha)} = \dot{\gamma}_0 \mathrm{sign}\left(\tau^{(\alpha)}\right)\left|\frac{\tau^{(\alpha)}}{g^{(\alpha)}}\right|^{1/m} \tag{24}$$

where the function sign(α) returns +1 or –1, if α is positive or negative, respectively; $\dot{\gamma}_0$ is an arbitrary reference slip rate; $g^{(\alpha)}$ is the slip system hardness; and m is the hardening exponent. Note that strain insensitivity is approached as $m \to 0$. The evolution relation for $g^{(\alpha)}$ is given by assuming that latent hardening of inactive slip systems increases more rapidly than does the self-hardening of the active slip systems:[7,8]

$$g^{(\alpha)} = \int_0^t \sum_{\beta=1}^{N} h_{\alpha\beta}\left|\dot{\gamma}^{(\beta)}\right|\frac{\partial q(\gamma)}{\partial \gamma}dt = \int_0^t \sum_{\beta=1}^{N} H_{\alpha\beta}\left|\dot{\gamma}^{(\beta)}\right|dt \tag{25}$$

where the coefficients $h_{\alpha\beta}$ have the following interpretation: they specify the rate of hardening of slip system α resulting from slip on system β. Here $q(\gamma)$ represents some material-specific smooth function of total accumulated slip, giving the hardening modulus of the self-hardening systems. The total accumulated slip is obtained by summing the magnitudes of the slip rates on each contributing system, and integrating over time:

$$\gamma = \int_0^t \sum_{\alpha=1}^{N}\left|\dot{\gamma}^{(\alpha)}\right|dt \quad . \tag{26}$$

The above formulation completes a constitutive description for the stress/deformation response of a crystalline solid that deforms by slip on select, geometrically defined slip systems. The numerical implementation of such relations follows a tangent modulus method[5] that gives a forward gradient approximation of the slip rates, based on a Taylor series expansion about the current slip rate. The details of the numerical implementation and results reported for shear localization studies in a finite set

of grains, is given by Becker,[9,10] using a user-defined material routine in the commercial finite deformation code, ABAQUS™. The lengthy details of this will not be repeated in this report.

To illustrate the behavior of this material model, a finite element model was produced for an assumed bicrystal, with each crystallite having the shape of a cube. A flat planar surface was modeled as the interface between the two crystallites. Such a model is illustrated in **Fig. 3**, which shows the orientation for each crystallite, with respect to the laboratory reference frame. The deformation was then modeled as a channel-die compression operation by compressing the bicrystal in the 3-direction, constraining it in the 2-direction and allowing free plastic flow in the 1-direction. The respective crystal orientations were selected such that the top crystallite would be more resistant to plastic slipping than would the bottom crystallite, thus illustrating the inhomgeniety of strain due to the misorientation across the interface. The model was compressed by 30% in the 3-direction. The predicted deformed mesh is shown in **Fig. 4**. The inhomogeneity of strain, as predicted by our constitutive law, is clearly visible by inspection of the surface, which was initially normal to the direction of free plastic flow, in the deformed mesh.

Future Work

Of necessity, the validation of the constitutive law, described above, requires correlation with deformation response measured from experiments such as the channel-die compression test. Several such experiments are to be completed in the early part of FY-96.

This single crystal constitutive formulation includes no information about the length scale of the material, its geometry or deformation process. Recent work has shown strong correlations between plasticity phenomena (e.g., hardening) and material length scale, and has suggested the need for more sophisticated plasticity laws.[11] Such a formulation, based on strain gradient theory, has been developed and is currently being demonstrated for microindentation predictions in isotropic materials.[12] The size effects that are demonstrated in microindentation tests can be predicted by the strain gradient theory.

In the near future, work will include development of a new crystal plasticity model in the context of strain-gradient plasticity. Such a model will then be implemented in an implicit finite element code and demonstrated with microindentation calculations on a single crystal. Such calculations will be compared with the results from the conventional crystal-plastic model, as described in this report.

Finally, the results from these calculations will be directly compared with microindentation experiments on single crystals of copper or aluminum. Correlated with these experiments, the analyses can be used to either confirm or rebut the strain gradient theory.

Acknowledgments

The author gratefully acknowledges the helpful discussions and consultations with R. W. Logan, M. Puso, P. J. Raboin, J. Y. Shu, and E. Zywicz. The motivation for this work was the result of an initiative, lead by W. E. King. Finally, valuable contributions concerning the implementation of the material model were provided by R. C. Becker of the ALCOA Technical Center.

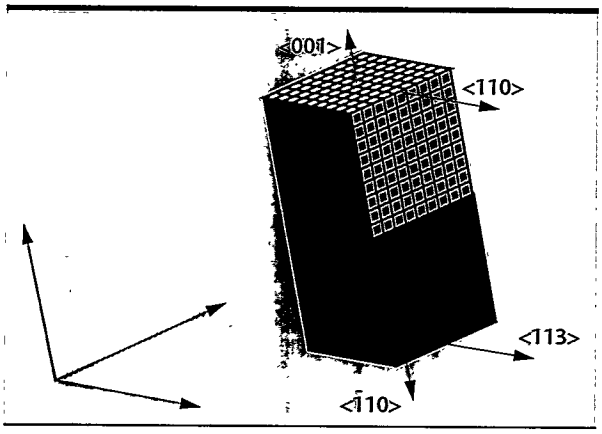


Figure 3. Finite element model of the bicrystal system, illustrating the misorientation between the top crystallite (light shading) and the bottom crystallite (dark shading).

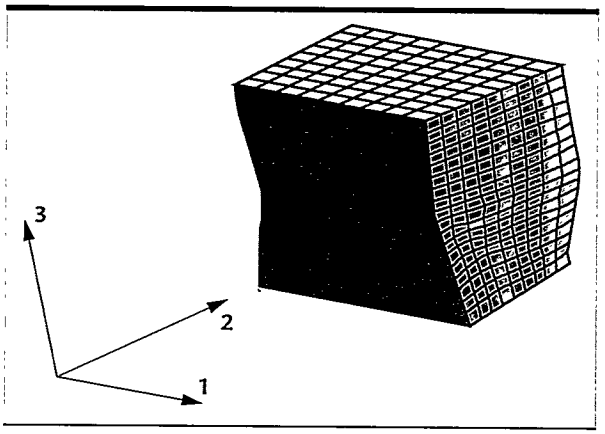


Figure 4. Finite element mesh deformed by channel-die compression, showing localized deformation in the bottom crystallite.

References

1. Hill, R., and J. R. Rice (1972), "Constitutive Analysis of Elastic-Plastic Crystals at Arbitrary Strain," *J. Mech. Phys. Solids*, **20**, p. 401.

2. Asaro, R. J., and J. R. Rice (1977), "Strain Localization in Ductile Single Crystals," *J. Mech. Phys. Solids*, **25**, p. 309.

3. Peirce, D., R. J. Asaro, and A. Needleman (1982), "An Analysis of Nonuniform and Localized Deformation in Ductile Single Crystals," *Acta Metall.*, **30**, p. 1087.

4. Pan, J., and J. R. Rice (1983), "Rate Sensitivity of Plastic Flow and Implications for Yield -Surface Vertices," *Int. J. Solids Structures*, **19**, p. 973.

5. Peirce, D., R. J. Asaro, and A. Needleman (1983), "Material Rate Dependence and Localized Deformation in Crystalline Solids," *Acta Metall.*, **31**, p. 1951.

6. Hutchinson, J. W. (1976), "Bounds and Self-Consistent Estimates for Creep of Polycrystalline Materials," *Proc. R. Soc. Lond. A*, **348**, p. 101.

7. Kocks, U. F., and T. J. Brown (1966), "Latent Hardening in Aluminum," *Acta Metall.*, **14**, p. 87.

8. Franciosi, F., M. Berveiller, and A. Zaoui (1980), "Latent Hardening in Copper and Aluminum Single Crystals," *Acta Metall.*, **28**, p. 273.

9. Becker, R. C. (1988), *Implementation of Crystal Plasticity Constitutive Relations in ABAQUS*, internal report, ALCOA Technical Center, May 31.

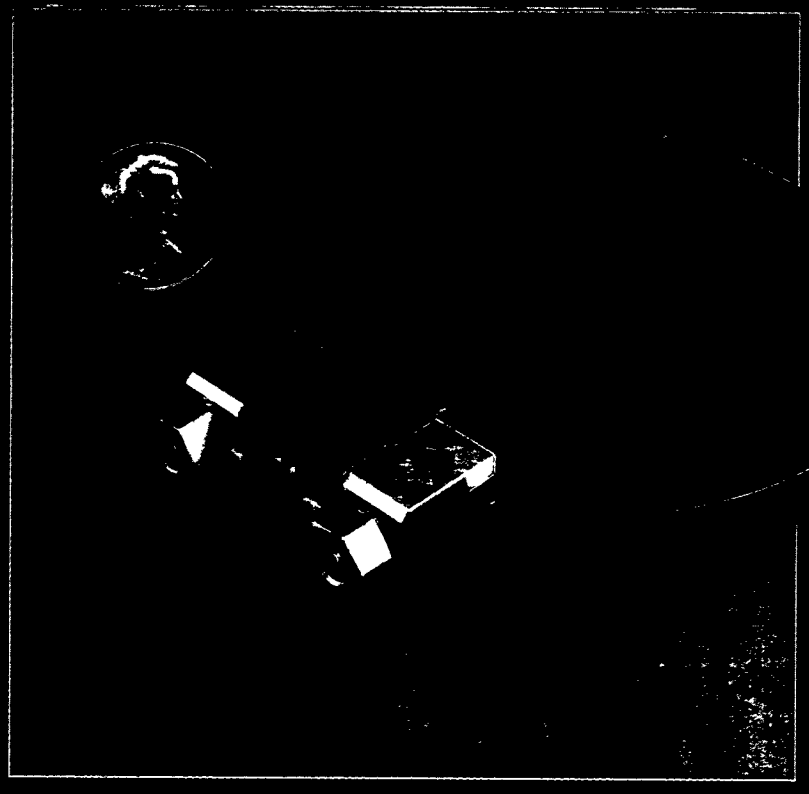
10. Becker, R. (1991), "Analysis of Texture Evolution in Channel Die Compression—I. Effects of Grain Interaction," *Acta Metall. Mater.*, **39**, p. 1211.

11. Fleck, N. A., G. M. Muller, M. F. Ashby and J. W. Hutchinson (1994), "Strain Gradient Plasticity: Theory and Experiment," *Acta Metall. Mater.*, **42**, p. 475.

12. Shu, J. Y., and N. A. Fleck (1995), work in progress.

Raymond P. Mariella, Jr., Thrust Area Leader

Microtechnology



3

3. Microtechnology

Overview
Raymond P. Mariella, Jr., Thrust Area Leader

Optical Coating, Testing, and Etching for Advanced Semiconductor Optical Amplifier Technology
Sol P. DiJaili, Frank G. Patterson, and Robert J. Deri 3-1

Micromachined Sample Injector for a Portable Hand-Held Gas Chromatograph
Conrad M. Yu 3-3

Fiber-Optic DNA Detector
Laurence R. Brewer and Donald A. Masquelier 3-5

Silicon Dioxide Film Deposition and Characterization
Norman F. Raley and Dino R. Ciarlo 3-7

Interconnections for Microflow Systems
James A. Folta and William J. Benett 3-11

Shape Memory Film Microactuators
Peter A. Krulevitch, Abraham P. Lee, Philip B. Ramsey, and M. Allen Northrup 3-15

Polyimide Electrostatic Microvalve Development
Abraham P. Lee, Julie Hamilton, and Jimmy Trevino 3-19

Microinstrumentation for Analysis of Biological Samples
Raymond P. Mariella Jr., Gerald G. Eveleth, Richard G. Langlois, Donald A. Masquelier, Dino R. Ciarlo, Mukund C. Venkatesh, and Shadi Shakeri 3-21

SMA-Actuated Silicon Microgripper
Abraham P. Lee, Peter A. Krulevitch, Jimmy Trevino, Stacy Lehew, Dino R. Ciarlo, and M. Allen Northrup 3-23

Optical Coating, Testing, and Etching for Advanced Semiconductor Optical Amplifier Technology

Sol P. DiJaili
Engineering Research Division
Electronics Engineering

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

The three main objectives for this project are: 1) to develop an antireflection (AR) and high reflectivity coating set-up for III-V integrated optics; 2) to refurbish our main chamber on the chemically-assisted ion beam etcher (CAIBE); and 3) to establish a performance characterization set-up for the semiconductor optical amplifiers (SOA) chips that we fabricate. SOAs provide optical gain in optical systems, and thus can compensate for losses due to long distance transmission or splitting losses in future multi-user all-optical networks.

Introduction

The basic goal of this project is to build three systems that support SOA activities for our programmatic customers. The role of support that each of these systems plays is briefly stated as follows. First, the AR coating reduces Fresnel losses at the two facets of the chip and improves fiber-to-fiber gain by about 3 dB. Second, the CAIBE is used to fabricate laser mirrors monolithically on the amplifier chip. These mirrors are critical in the fabrication of the crosstalk-free SOA, for which we have recently received a U.S. patent. Third, the characterization set-up is essential to determine the gain and other important performance parameters of our SOA devices.

Progress

Shown in **Fig. 1** is a picture of the AR coating system, which essentially involves applying a dielectric coating to both optical facets of the SOA. The critical parameters are the index of refraction and optical thickness. The desired index of refraction is the geometric mean between the waveguide index of refraction and the index of refraction of air. Since the exact waveguide index of refraction is a difficult parameter to measure, the bulk index of refraction is used. The burden of optimization then falls on controlling the thickness of the dielectric film to the optimum level, which is very nearly one-quarter wavelength.

Our system includes an *in-situ* monitor to determine the optimum thickness. This monitor involves detecting the amplified spontaneous emission from the facet of the SOA, and maximizing it. The dielectric material used is SiOx (x ~ 1.2), so that the index of refraction is nearly -1.85. The index of refraction

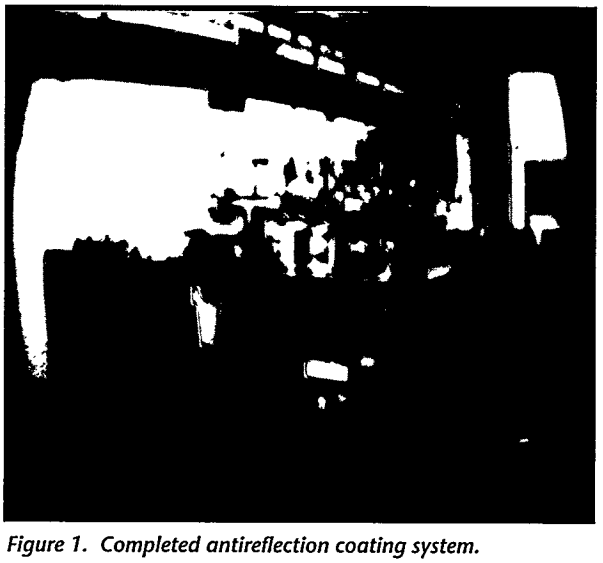


Figure 1. Completed antireflection coating system.

is controlled by magnetron-sputtering silicon and backfilling with oxygen. Water cooling of the *in-situ* monitor was added, due to the temperature sensitivity of our SOAs operating at 1.3 μm. The vacuum system features a small volume cross with standard conflat flanges. A turbo-pump backed by a roughing pump is used to evacuate the chamber to the 10^{-7} Torr range in approximately 20 min. Various subsystems have been tested, and the final assembly is nearly completed.

Shown in **Fig. 2** is a picture of the new CAIBE chamber. Chemically-assisted ion beam etching involves ion milling of the chip, along with the introduction of a dry etching gas. Our system uses a xenon ion beam and chlorine etching gas for the etching of III-V semiconductors. The etch rate of the system is around 1500 Å/min. The proper operation of the system depends critically on reducing the partial pressure of any residual oxygen-containing species to UHV levels of around 10×10^{-10} Torr levels. The old chamber had many o-rings and was difficult to troubleshoot, especially when a low-level leak occurred. One of the key reasons for switching over to the new chamber was to reduce down-time when leaks occur, and to provide a solid platform to improve the overall vacuum integrity of the system. The old chamber has been removed and the new chamber has been mounted on the frame. All the necessary mating flanges have been designed, and fabrication has been completed. What remains at this writing is the final assembly of all the subsystems onto the main chamber, which we expect to be complete soon.

Shown in **Fig. 3** is a picture of the SOA testing station. The station measures optical gain, spectrum, transparency current, and residual facet reflectivity. The station is currently used to screen-test the batches of SOA chips that we fabricate. The station consists of a rotation table with a vacuum, to hold the small (1 mm × 0.5 mm × 0.2 mm) SOA chips. A pulsed current source supplies current through a needle probe that contacts the top of the chip. A conically polished fiber is then used to couple the amplified spontaneous emission from SOA into an optical spectrum analyzer to check the spectral integrity of the SOA. The device is then turned around and the power levels of the two facets are compared. Often, if there is a break or bad facet on a SOA, the two power levels are different. To measure the gain, laser light is coupled into the opposite end via another conically polished fiber. The transparency current is then measured by an optical saturation technique and lock-in detection. This optical method allows for accurate determination of the material transparency point. This parameter is then used to determine the optical gain of the device. This station has been set up for characterizing both our 0.8 μm and our 1.3 μm wavelength SOAs.

Acknowledgments

We would like to thank W. Goward, M. Lucas, R. Hicks, L. Ngyuen, P. Keiser, and H. Petersen for their diligent efforts in completing these projects.

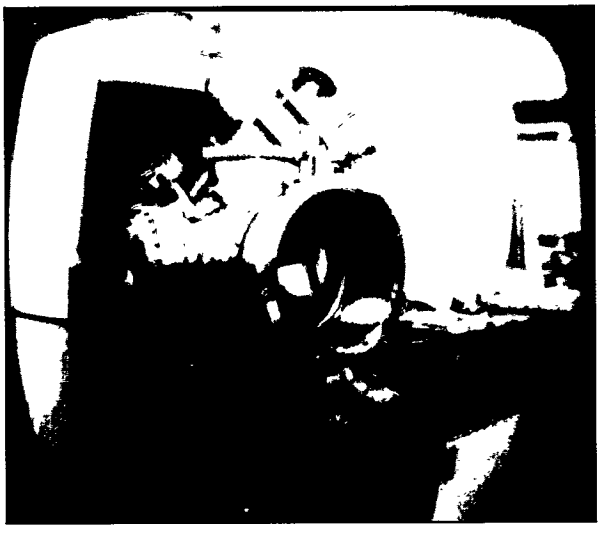


Figure 2. New chemically-assisted ion beam etcher vacuum chamber.

Figure 3. Testing station for semiconductor optical amplifiers.

Micromachined Sample Injector for a Portable Hand-Held Gas Chromatograph

Conrad M. Yu
Engineering Research Division
Electronics Engineering

In this project, a miniature silicon sample injector, and a polysilicon thermal-conductivity detector were designed, built, and tested.

Introduction

Gas chromatography has proven to be a reliable method for identification of liquid and gas species, with detection sensitivity approaching as high as several parts per million for thermal-conductivity cell detection, or several parts per billion for ion cell detection. However, the conventional laboratory gas chromatographs (GC) are bulky units (many cubic feet in size) that take a long time to analyze a gas sample (typically 20 min).

For the past several years we have been pursuing research directed towards the eventual development of a miniature, portable, low power, hand-held GC analytical instrument. This challenging goal is to be achieved by combining many innovative miniature components fabricated through micromachining processes in silicon.

A standard technique for the introduction of gases and liquids into the conventional GC system is to insert a hypodermic syringe needle through a self-sealing septum, and inject measured volumes from an attached syringe into the GC. Lawrence Livermore National Laboratory's miniature portable GC follows the same injection principles, but with a different design. Our basic injector design involves silicon and silicon nitride, which are compatible with the rest of the system, to allow high temperature operation.

Progress

Design and Fabrication

In our design, the carrier gas in the injection chamber passes through the top of the silicon nitride membrane valves, preventing the outside gases from intruding into the GC column. A pneumatic syringe moves the sample gases in and out of the chamber between two valves, one normally open, and the other normally closed. During sampling, we apply enough pressure on the syringe

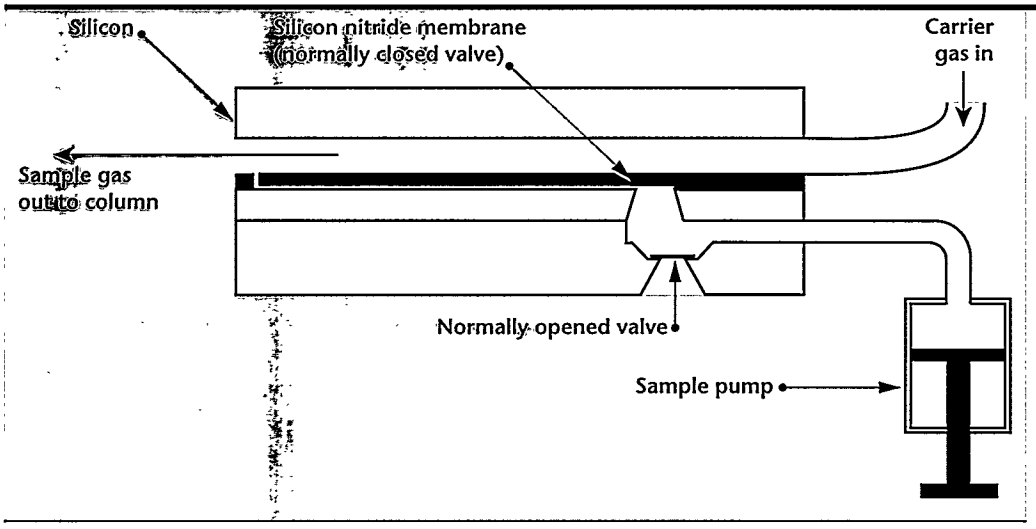


Figure 1. Schematic illustration of gas chromatograph sample injector.

to cause the valves to react. Thus, a plug of sample gas is injected into the column in the stream of the carrier gas (see **Fig. 1**). Together, they enter the GC column.

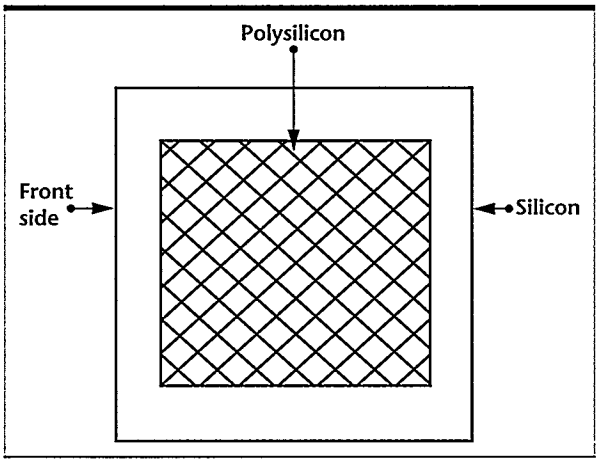


Figure 2. Schematic illustration of thermal conductivity cell.

A thermal-conductivity detector was designed and fabricated. Boron-doped polysilicon was used to sense the thermal resistivity change during detection. The temperature dependence of its thermal resistance was about 1 Ω/° from room temperature up to 150 °C (see **Fig. 2**).

Future Work

In the past year, we tested the principle functions of individual components of the GC. In the coming year, we plan to assemble and test the components as a system. We will use these testing results to determine modifications of these components, so that we can use them as a unit for a particular application.

Acknowledgments

The author would like to thank N. Raley for technical help and discussions; J. Koo for technical discussions; T. Rodriguez for the fabrication of thermal conductivity cells; and J. Trevino for helping with polysilicon LPCVD and boron-doping processes.

Fiber-Optic DNA Detector

Laurence R. Brewer and Donald A. Masquelier
Engineering Research Division
Electronics Engineering

We have successfully designed, built, and tested a compact fiber-optic-based DNA detector with a sensitivity adequate for use in DNA sequencers.

A photograph of our compact fiber-optic-based detector is shown in **Fig. 1**. The detector used an optical fiber to transmit laser light through a lens, where it was focused onto the sample containing dye-tagged DNA. A second optical fiber was used to collect fluorescent light through the same lens. A dichroic mirror, placed at 45° to the optic axis, was used to separate these two light paths. The detector can be the unit module basis of an array detector, so that many electrophoresis channels can be monitored simultaneously, eliminating mechanically scanned detection, and increasing the signal-to-noise ratio of the detected DNA.

The detector was successfully tested during the course of four different electrophoresis runs using microchannel plates filled with a 6% polyacrylamide gel, and an applied field of 65 V/cm. A four color DNA sequencing standard (Applied Biosystems Taq) was used, and fluorescence corresponding to the C (Cytosine nucleic acid) base was monitored from a single microchannel. The data for a typical experimental run is shown in **Fig. 2**. The standard sequence of C fragments is shown out to approximately 100 base pairs.

The detector will be tested in an array format in the near future, and may be incorporated into advanced, high throughput DNA sequencers for the Human Genome Program at Lawrence Livermore National Laboratory.

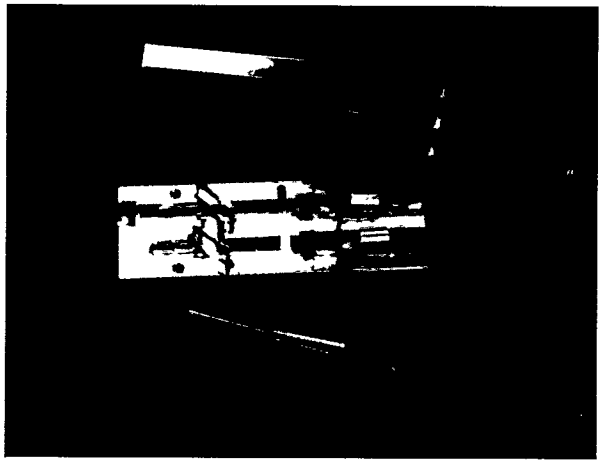


Figure 1. Photograph of fiber-optic-based detector.

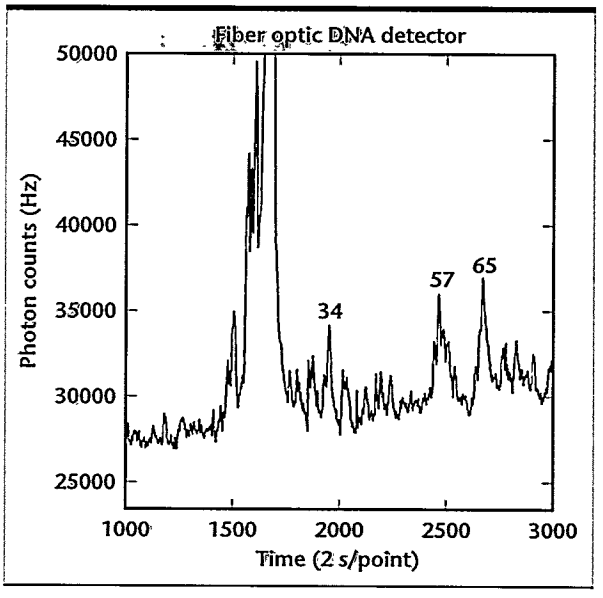


Figure 2. Data for experimental run.

Reference

Davidson, J.C., L.R. Brewer, J.C. Koo, J.W. Balch, A.V. Carrano and J.E. Lamerdin, Microinstrumentation for Large DNA Fragment Analysis and Collection, Lawrence Livermore National Laboratory, Livermore,California, UCRL 53868-94, p. 3-11.

Silicon Dioxide Film Deposition and Characterization

Norman F. Raley and Dino R. Ciarlo
Engineering Research Division
Electronics Engineering

We have modified a low-pressure chemical vapor deposition system (LPCVD) to allow use of silane, oxygen and phosphine gas for deposition of both undoped and phosphorus-doped thin films of silicon dioxide. We have characterized film deposition rate and phosphorus content, as well as the etch rate in HF, buffered HF, KOH and H_3PO_4. Tensile doped-oxide films were obtained, and oxide membranes were formed, with optical transmission extending to 200 nm. Expected applications include mirror supports for a micromachined optical interferometer, thin transmission ports for detection probing in microinstruments, an etch mask and a sacrificial layer.

Introduction

A continuing thrust in the Microtechnology Center at Lawrence Livermore National Laboratory relates to the fabrication of microelectromechanical system (MEMS) devices. The capability for LPCVD of silicon dioxide thin films at low temperature (LTO) is expected to impact a variety of such devices. Specific examples include use as 1) a membrane that is transmissive to ultraviolet, for use in an optical interferometer or to facilitate probing of reactions in chemical or biological microinstruments; 2) an etch mask for fabrication of multiple deep recesses in silicon or application to wafer bonding; or 3) a sacrificial layer for an interferometer or micropositioner.

Progress

Equipment

A horizontal three-zone resistance-heated reactor was modified to provide for deposition of undoped and doped silicon dioxide thin films. Silane and phosphine gas were fed to the reactor independently for maximum flexibility in the control of silane-to-phosphine gas flow ratio, as opposed to use of a combined silane/phosphine gas source. The source gas for silane was 100%, whereas that for phosphine was diluted in nitrogen. The phosphine percentage was a compromise between minimization of overpressure in the reactor, due to the excessive nitrogen balance in the phosphine source gas, and current facility safety analysis to limit phosphine levels. The temperature profile variation obtained at 450 °C was typically less than 1 °C over 15 cm length at the reactor tube center, with little adjustment.

Various safety features were implemented in system hardware regarding toxic gases, including an overpressure limit of 700 mTorr when toxic gases are in use; a mechanical pump exhaust sensor to monitor the connection to the exhaust system, with interlock to the phosphine mass flow controller; and a controlled decomposition oxidation furnace to eliminate silane and phosphine gases in the exhaust line, in order to prevent potential fires in the exhaust system. Safety features in system operation include primarily appropriate evacuation of the phosphine gas line, to avoid unnecessary potential exposure.

Film Deposition

Deposition processes were developed for undoped and doped silicon dioxide films. Undoped films were deposited using oxygen and silane at 450 °C. The deposition rate was about 325 Å/min. It is important to use less silane flow than oxygen in order to prevent spontaneous combustion in the reactor, evidenced by light emission and a fluctuation in reactor pressure. Doped films were deposited using oxygen, silane and phosphine at 450 °C. The deposition rate was lower at about 142 Å/min, compared to the undoped process, due to

the use of lower gas flows. The deposition rate is fairly independent of reactor pressure in this parameter regime. Lower gas flows for the doped film process were used, to minimize the rise in reactor pressure which occurred with the use of phosphine. The cause is presently unknown, but may be related to generation of phosphoric acid in the mechanical pump. Even with the use of lower gas flows, however, significant rise was possible for long deposition times required for thick (2 μm) films. A throttle valve was used to keep the pressure constant, to eliminate any process irreproducibility due to reactor pressure variation. Finally, the effect of phosphine on the deposition rate was minimal, since the rate for doped films was about 87% of that for undoped films for the same oxygen and silane gas flows.

Film Characterization

Film thicknesses up to 2 μm were deposited for both undoped and doped processes. The index of refraction for both undoped and doped films was about 1.45, and thickness uniformity was typically less than 10% for wafer diameters up to 4 in. The wt% of phosphorus in the doped films was determined by published data,[1] which relate the phosphorus concentration to the sheet resistance of p-type substrates after diffusion using the oxide film as dopant source, as well as wet etch rate in HF-HNO_3 solutions. The wt% phosphorus in the LTO films is linear with phosphorus flow rate, with about 7 wt% phosphorus obtained using the doped film process described above. Etch rates in various wet etchants such as 49% HF, BHF (40% NH_4F:49% HF = 5:1), 42% KOH and H_3PO_4 were also measured and tabulated. The etch rates in HF and BHF were 5.8 and 0.30 μm/min, respectively, for undoped films as-deposited, whereas the rates increased roughly linearly with phosphorus concentration, to values of 60.6 and 0.85 μm/min, respectively, for 7 wt% films. The etch rate in KOH at 85 °C was about 215 Å/min for undoped films and somewhat higher at 560 Å/min for 7 wt% doped films. In general, the etch rates decreased by about a factor of 2.5 after a 900 °C, 20-min nitrogen anneal. The etch rate for annealed undoped LTO film in H_3PO_4 at 150 °C was 4 Å/min.

The stress of the LTO films was measured by determination of the change in radius of curvature of the film and silicon substrate using a laser reflectance tool. The results are shown in **Fig. 1** for as-deposited and annealed films, as a function of wt% phosphorus in the film. Note that tensile stress was achieved for wt% phosphorus greater than 8 for as-deposited films. The stress becomes more compressive after a 900 °C, 20-min nitrogen anneal, in which case we extrapolate that greater than about 24 wt% phosphorus is required to maintain a tensile film. It should be noted that the stress values of as-deposited films were found to decrease over time after deposition, presumably due to absorption of moisture.[2] This effect was more pronounced for undoped films whereas films with about 8 to 10 wt% phosphorus were fairly stable. Any stress instabilities were removed by performing the anneal step.

It should be noted that the anneal step is not necessary to ensure good electrical integrity of the LTO films. Initial measurements indicate a breakdown voltage of at least 250 V over a 1.0 mm × 2.5 mm contact area for as-deposited 10 wt% phosphorus doped films with 0.5μm thickness. This value is quite comparable to published data for high quality thermal oxide films.

Applications

The measured data for LTO films are relevant for various applications. Surface micromachining using etching of doped oxide sacrificial layer in HF can be done using typical polysilicon, but also with nitride, with a differential etch rate ratio (ERR) of 1700 for the latter, with a nitride etch rate in HF of 142 Å/min. Annealed undoped LTO films can be used as masks for silicon etching in KOH above (ERR of 116) and nitride etching in H_3PO_4 at 150 °C (ERR of 11), where the etch rate in the latter case for nitride is 45 Å/min. Finally, oxide membranes

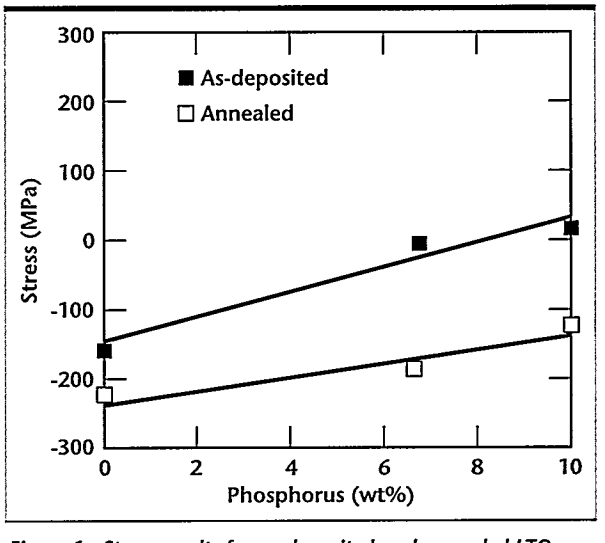


Figure 1. Stress results for as-deposited and annealed LTO films as a function of wt% phosphorus.

suitable for ultraviolet transmission uses have been obtained with 8 wt% phosphorus, by depositing the oxide film onto pre-etched silicon substrates without anneal followed by a short post-deposition silicon etch. The oxide thickness was 1.4 μm and the membrane area was 1.3 mm × 1.3 mm. The optical transmission of these membranes is shown in **Fig. 2.** The oscillation is due to interference of multiple reflections within the membrane. The transmission is significant at wavelengths as low as 200 nm. It is estimated[3] that for the same thickness and transmission, the cutoff wavelength for silicon nitride is higher at about 450 nm.

Future Work

LTO films deposited using phosphine gas source diluted in nitrogen are equivalent in terms of deposition rate and wet (HF) etch rate, compared to known results using a combined silane-phosphine gas source. Tensile stress has been obtained for as-deposited LTO films with greater than 8 wt% phosphorus, and an oxide membrane has been demonstrated with significant optical transmission to 200 nm. The equipment modifications are complete, safety features are in place, and a baseline deposition process with corresponding film characteristics has been documented.

More experiments are needed to assess the stability of stress in the LTO films, as well as the extrapolation that a tensile oxide film can be achieved after post-deposition anneal. Use of the films in hermetically-sealed packages may be advantageous. These developments will be important to ensure that potential moisture absorption in the films and resultant stress variability are minimal.

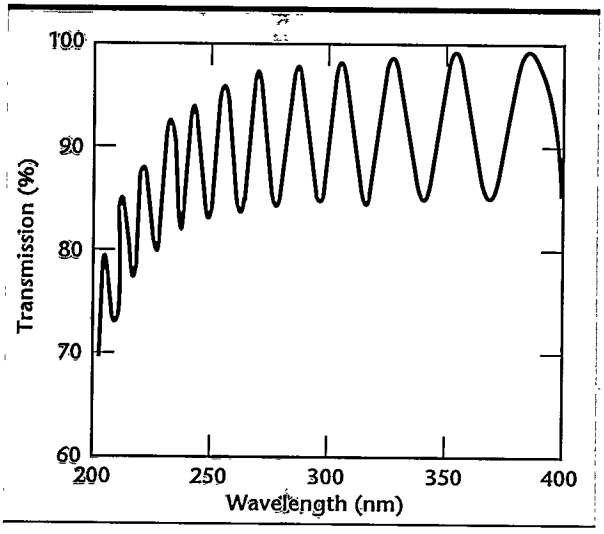


Figure 2. Optical transmission as a function of wavelength for 1.4 μm-thick oxide membranes.

Acknowledgments

The authors wish to acknowledge D. Applegate for gas line welding, R. Hicks for parts machining assistance, P. Krulevitch for use of the laser reflectance tool, and G. Loomis for oxide membrane optical transmission measurement.

References

1. Adams, A. C., and S. P. Murarka (1979), "Measuring the Phosphorus Concentration in Deposited Phosphosilicate Films," *J. Electrochem. Soc.*, **126**(2), p. 334.

2. Sunami, H., Y. Itoh, and K. Sato (1970), "Stress and Thermal-Expansion Coefficient of Chemical-Vapor-Deposited Glass Films," *J. Applied Physics*, **41**(13), p. 5115.

3. Raley, N. F., D. R. Ciarlo, J. C. Koo, B. Beiriger, J. Trujillo, C. Yu, G. Loomis, and R. Chow (1992), "A Fabry-Perot Microinterferometer for Visible Wavelengths," *Tech. Digest IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head, S.C., p. 170.

nterconnections for Microflow Systems

James A. Folta
Engineering Research Division
Electronics Engineering

William J. Benett
Defense Technologies Engineering Division
Mechanical Engineering

At Lawrence Livermore National Laboratory (LLNL), we have developed three methods to allow connection of microfabricated fluidic components to one another and to macroscopic flow systems: 1) a gas chromatography (GC) separation column, microfabricated on bonded silicon wafers, was edge-port connected to a pair of fused silica capillaries to allow connection to a sample injector and detector; 2) a surface-port connection method was used in the development of a chemically-resistant package for microfabricated valve prototypes; and 3) another surface port connector was used for connection to a miniaturized flow cytometer. Additionally, we developed a plastic liner for microfabricated reaction chambers for the polymerase chain reaction (PCR). The liners prevent contamination of the reactants and allow easy re-use of the chambers.

Introduction

While the concept of complete microflow systems integrated on a single silicon wafer offers tremendous appeal, three factors greatly hinder progress toward that goal: 1) compromises in performance for different components, dictated by differing fabrication requirements; 2) lower fabrication yield for highly integrated and complex systems; and 3) the inability to replace individual components that have failed. Since many of the components which comprise a microfabricated chemical analysis system have widely disparate fabrication requirements, they are much more difficult to integrate into a single wafer than electronic devices. We believe the best approach is to develop improved packaging technologies that allow microflow components to interface to macroscopic systems and to one another. This approach allows each component to be individually optimized and replaced. Higher levels of integration should be pursued only when needed to meet compelling performance requirements of the system.

There has been a great deal of progress made in recent years in the development of micro-fluidic devices with micro-electro-mechanical systems (MEMS) technology. However, it is common for microflow devices to be developed with little regard to packaging and interconnection issues. A strong need exists for improved interconnections to take advantage of the low dead spaces, small internal volumes and low cost of microflow devices. There is considerable demand outside LLNL for a zero-dead volume interconnection technology between sample preparation systems, separation capillaries, and detectors; i.e., a generic package itself can be a useful product, independent of the microfabricated devices.

Depending on the application, the common design specifications for interconnections include:

1) chemical resistance,
2) non-contamination of sample and reagent flow streams,
3) no adhesion of sample components and reagents to the package surfaces,
4) reasonable cost,
5) single-use packages (cost considerations may dictate re-usable packaging for disposable microfabricated devices),
6) ability to clean a packaged device between uses,
7) low dead volume interconnections between different types of microflow devices (of particular interest is the mating of liquid microvalves to detectors, capillaries, and microreactors),
8) high temperature, and
9) high pressure.

Progress

Five LLNL projects are focused on development of micro-fluidic devices for chemical analysis: 1) micro-reactors for PCR; 2) capillary electrophoresis for DNA sequencing; 3) microfabricated separation capillaries for GC; 4) miniaturized flow cytometry; and 5) microfabricated valve development. Improvements in existing device packaging and interfacing methods are needed to allow widespread acceptance of the technologies. We initially collected the performance requirements for interconnections for each of the applications. Results are shown in **Table 1**.

Surface Connections

To date most connections to micro-fluidic devices have been made to the wafer surface of the device. The most common type of interconnection is small tubing which is made of plastic, metal or glass. The most common method for creating the surface ports is anisotropic etching through the top wafer of a bonded wafer pair. In this process, the shapes of the etched ports are dictated by the direction of the slower etching crystal planes of the silicon wafer. In [100] silicon (the most common wafer orientation), this creates either square or rectangular ports with sloped sidewalls. The shape mismatch between circular tubing and square ports makes a low-dead volume connection very difficult to accomplish. It is also difficult to make the connections between tubing and the etched ports leak-tight and mechanically robust.

We have developed a packaging and interconnection method in which flow channels that match the shape of the surface ports were machined into a small plastic block with conventional precision machine tools. A thin gasket was laser-cut to a similar shape into a sheet of thin elastomer, and used to form a seal between the silicon chip and the block. A top plate was used to compress the three-layer sandwich. The edges or bottom of the block were machined to match conventional tubing hardware. This method was used to successfully connect to chemically-resistant microvalve prototype components for an extended series of flow tests. It was also used to provide connection to a micromachined component in a miniature flow cytometer, as shown in **Fig. 1**, where the package is opened to expose the silicon chip and gasket. The cytometer package allows liquid to be introduced into ports etched into the surface of the chip. The front edge of the chip remains exposed, to allow fluid to freely exit through edge ports into an optical detection system.

While the "surface port with gasket" method works for some applications, it has three main disadvantages. First, it is difficult to machine the block and distribute the flow so that no dead spaces or flow stagnation points exist. Second, the ports can take up significant chip area and raise the cost of the devices. A typical wafer thickness of 0.5 mm, combined with the anisotropic etch behavior of [100] silicon, provides a minimum port size of about 0.7 mm. Third, precision machining of the block and gasket can easily exceed the cost of the chip.

Table 1. Interconnection requirements of selected applications.

	Application				
	GC columns	**Microvalves**	**Electrophoresis**	**PCR reactors**	**Flow cytometer**
Fluid	gas	liquid	liquid	liquid	liquid
Temperature range (°C)	20-280	20-50	20-40	20-100	20-40
Port size (mm)	0.1	0.3-0.6	0.1-1	0.3-1	0.3
Location of port*	edge	edge or surface	edge or surface	edge or surface	surface
Zero dead space required?	yes	yes	yes	depends	no
Resistant to aggressive chemicals?	no	yes	no	no	no

*Refers to connection to the edge or the surface of the silicon chip.

Edge Connection for GC Columns

One potential solution to these problems is to create ports on the edges of the micro-fluidic devices. These ports can be made with isotropic etching to yield near-circular ports, or by anisotropic etching, which can approximate circular ports. In both cases, the sidewalls of these ports are perpendicular to the wafer edge.

The microfabricated separation capillaries for GC were selected as a testbed for the edge connection approach. The application requires a connection that has near-zero dead volume, is helium leak-tight, and must maintain a seal at elevated temperatures (~280 °C). The first step towards a test of the seal concept was to redesign the GC wafers to enable an edge connection. This required a new mask layout that had the GC column spiral into the center of the wafer, then turn and spiral back out to the edge, creating two ports on the edge of the wafer. Once the columns were etched into the wafers, the wafers were bonded together with standard silicon-to-silicon bonding techniques. Glass rails were then anodically bonded to the edge of the wafer pair. The glass and wafer sandwich was then precisely sawed to create a high quality sealing surface. The glass rails provided a mechanically robust "handle" for attaching the sealing hardware to the wafer assembly.

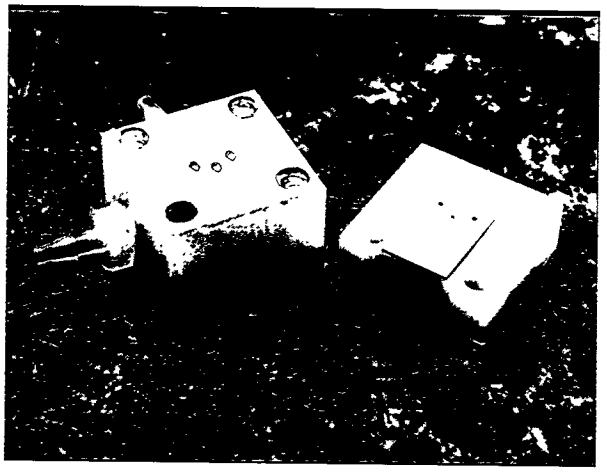


Figure 1. Interconnection for the miniature flow cytometer. The top plate and silicon chip (right) are removed to expose the base (left). Liquid is ported from the base into the surface of the chip and exits through edge ports (too small to see) on the front edge of the chip which remains free of the package.

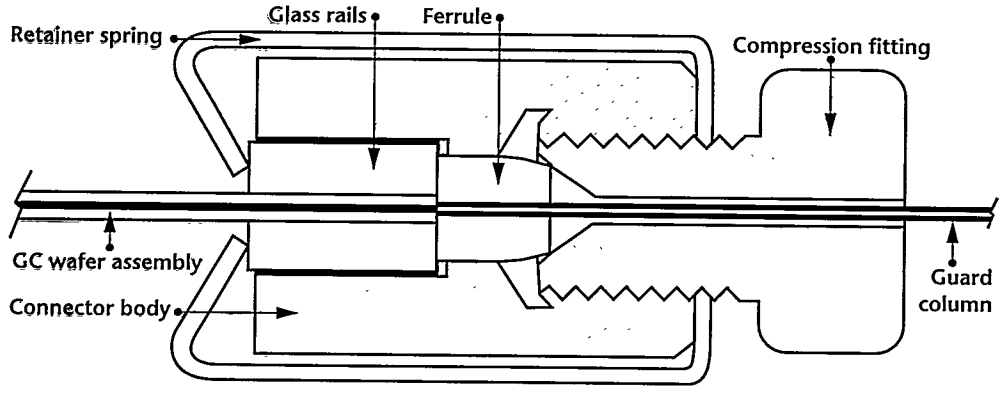


Figure 2. Cross sectional view of the internal sealing mechanism of the edge connector used to interconnect silica capillaries to a GC column micromachined in a silicon wafer pair.

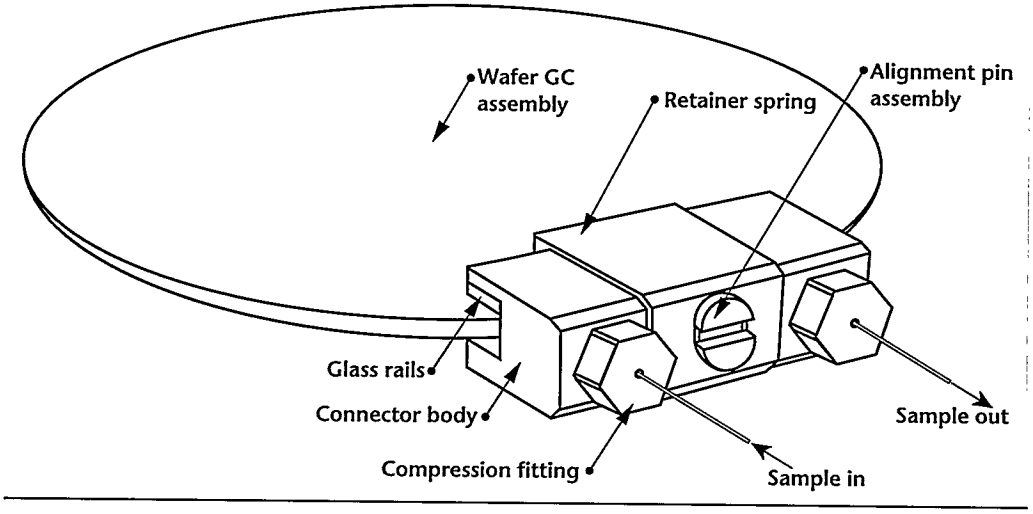


Figure 3. Three dimensional view of the hardware which aligns and grips the GC column wafers.

The seal consists of a standard GC ferrule, which allows a silica guard column to slip through. A cross section of the connecting mechanism is shown in **Fig. 2**. A standard GC fitting was used to compress the back face of the ferrule and the silica column against the wafer assembly. At the same time, the ferrule was compressed around the silica column to complete the seal. A 3-D view of the hardware used to accomplish the seal, shown in **Fig. 3**, consisted of a body, two standard GC compression fittings with ferrules, an alignment pin assembly, and a spring retainer. The alignment pin engages an etched feature in the edge of the silicon wafer assembly. This feature is created during the same photolithographic and etch process that created the GC column itself. The alignment pin aligned the sample input and output fittings in the body with the GC wafer assembly.

Figure 4 is a photograph of the complete assembly, including the GC column formed between 3-in. silicon wafers. We have just begun measuring the performance of the described connection. However, a prototype of this interconnection approach with a single capillary sealed against a flat surface showed the method to be leak-tight when tested with a helium/vacuum leak detector.

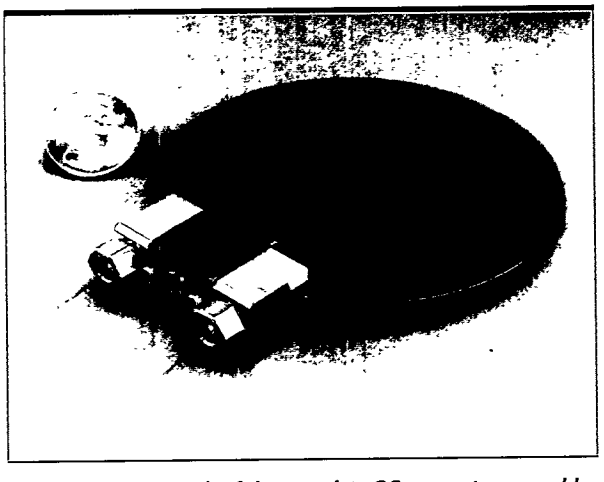


Figure 4. Photograph of the complete GC connector assembly. Two silica capillaries are attached to a GC column formed between etched and bonded silicon wafers of 3 in. diameter.

PCR Reaction Chamber Liner

LLNL has developed a PCR reaction chamber which is fabricated using silicon-based micromachining technology. These devices have several advantages over conventional PCR thermal cyclers. The micromachined PCR chambers are smaller, faster, portable, and more power efficient. While the devices can be cleaned for re-use, doing so is tedious and can allow contamination of subsequent analyses. A possible solution to this problem is to create a liner that can be slipped into the PCR chamber. This is also a logical progression toward automating the process of sample loading and unloading.

The injection molding process was chosen to fabricate the liners, owing to the low cost of mass production. Polypropylene was selected as the liner material because of ease in injection molding and proven compatibility with conventional PCR systems. The LLNL Plastics Shop fabricated a prototype mold and developed an injection molding procedure to mold liners to the design specifications. This was a very challenging component to mold because of its small size and high aspect ratio.

At present more than 100 liners have been fabricated and successful PCR reactions have been demonstrated. The testing of the liners is continuing and further development is being planned.

Future Work

We are continuing to characterize the performance of the interconnection techniques described here, with our main focus on the GC capillary connector. We are also developing improved surface interconnection techniques for chemically-resistant microvalve applications.

Acknowledgments

The authors wish to thank P. Krulevitch for technical discussions and assistance in the design of the PCR chamber liners. We also thank D. Schumann, W. Jensen, and R. Thigpen of the LLNL Plastics Shop for developing the injection molding process for the PCR liners.

Shape Memory Film Microactuators

Peter Krulevitch, Abraham P. Lee, Philip B. Ramsey, and M. Allen Northrup
Engineering Research Division
Electronics Engineering

We have developed a process for depositing Ni-Ti-Cu shape memory films which have been used to actuate a 900-μm-long silicon gripper. We have also designed and fabricated a shape memory film microvalve actuator. In this report, we present the critical processing techniques for depositing shape memory alloy (SMA) films, and discuss results from the microvalve actuator work.

Introduction

The shape memory effect, exhibited by a number of binary and ternary alloys, involves a thermally-induced crystalline transformation between a ductile phase and a high strength phase. While in the low temperature martensitic phase, shape memory alloys (SMAs) are easily deformed, but when heated, they transform to the austenitic phase (also known as the parent phase), and recover their initial shape. This reversible transformation forms the basis for shape memory actuators, in which a biasing force, produced by a spring, for example, deforms the SMA element at low temperature, and the SMA element overcomes the bias when heated. In bulk form, Ni-Ti-based SMAs are capable of recovering up to 8% strain, or generating as much as 400 MPa if constrained during recovery.[1]

Bulk SMA actuators have experienced limited success in specialized applications, such as temperature sensitive switches and robotic manipulators.[1,2] For micron- to millimeter-scale applications, including micro-electromechanical systems[3,4] and miniature robotics,[5,6] the high power-to-weight ratio, large stroke, and low driving voltage of SMA films and thin wires suggest a potential for superior microactuators. Shape memory actuators can be heated resistively, and the small thermal mass and large surface-to-volume ratios associated with thin film structures allow for rapid heat transfer, increasing the maximum switching rate to the 100 Hz range. Furthermore, SMA thin film devices have the potential for batch fabrication and integration with control electronics. Despite these apparent advantages, successful implementation of SMA film microactuators requires that several challenges must be met, including deposition and heat treatment of compositionally sensitive alloys, integration with micromachining processes, and mechanical design that takes advantage of the complex thermo-mechanical behavior of these films.

At Lawrence Livermore National Laboratory, the primary applications for SMA films include micromedical tools, such as the silicon gripper for catheter-based interventions,[7] and fluid control in microfluidic systems, such as micromachined biochemical reaction chambers and microchannel electrophoresis. We have developed a process for

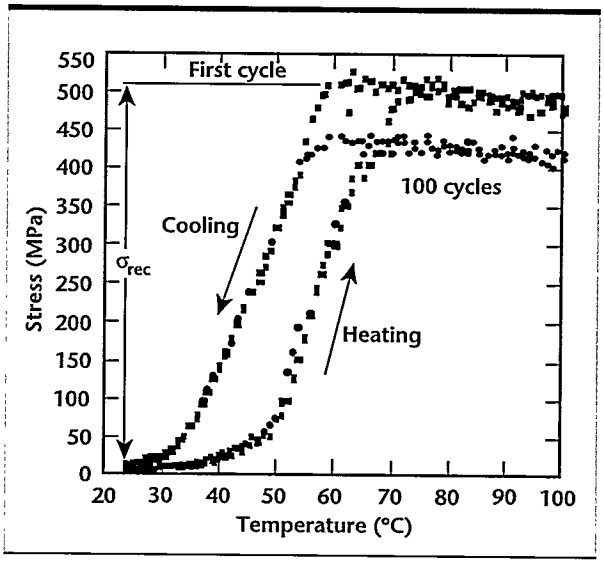


Figure 1. Stress vs temperature curves for a Ni-Ti-Cu film. Approximately 75 MPa recoverable stress is lost after 100 thermal cycles.

depositing Ni-Ti-Cu shape memory films which have been used to actuate a 900-μm-long silicon gripper. We have also designed and fabricated a shape memory film microvalve actuator. In this report, we present the critical processing techniques for depositing SMA films, and discuss results from the microvalve actuator work.

Progress

The primary challenge in creating SMA microactuators is establishing a deposition process that allows for precise compositional control. Several investigations into the properties of Ni-Ti SMA films have been conducted, including sputtering[4] and laser ablation[8] of single alloy targets, multilayer sputtering,[9] and multiple source sputtering.[10] Binary Ni-Ti is extremely sensitive to composition. An increase in the atomic percentage of Ni by 1% from an equi-atomic ratio of Ni:Ti lowers the martensitic start temperature, M_S, from 50 to –100 °C.[2] Even if a film has the proper composition overall, local inhomogeneities result in the nucleation and growth of precipitates which do not exhibit shape memory behavior, as demonstrated by Chang and Grummon[9] for Ni-Ti-based shape memory films. We have shown that the addition of Cu to the Ni-Ti system makes SMA films less sensitive to variations in composition,[10] consistent with bulk Ni-Ti-Cu alloys. For optimum shape memory properties, it is also important to limit impurities, especially oxygen and carbon.[1] Unbound Ti readily oxidizes, and less than 1 at.% oxygen has been shown to inhibit the shape memory effect in Ni-Ti alloys,[1] making target purity and high vacuum deposition conditions essential for SMA films.

Thermal history also plays a critical role in determining SMA film characteristics. When the substrate temperature is near ambient during deposition, sputtering[11] and laser ablation[12] produce amorphous films with low surface roughness, which subsequently must be crystallized by

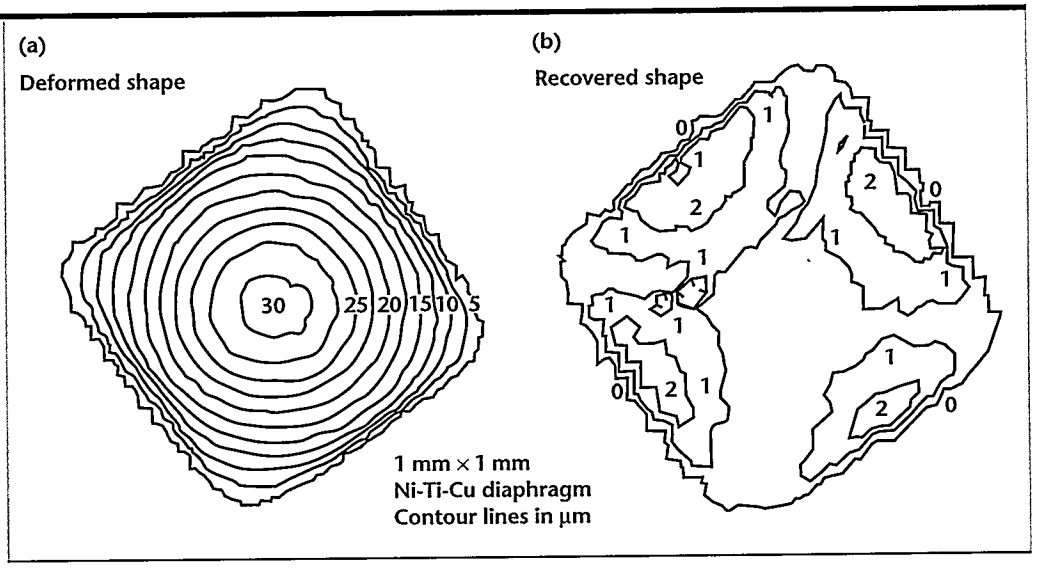


Figure 2. Deformed and recovered shape of a 1 mm × 1 mm Ni-Ti-Cu diaphragm.

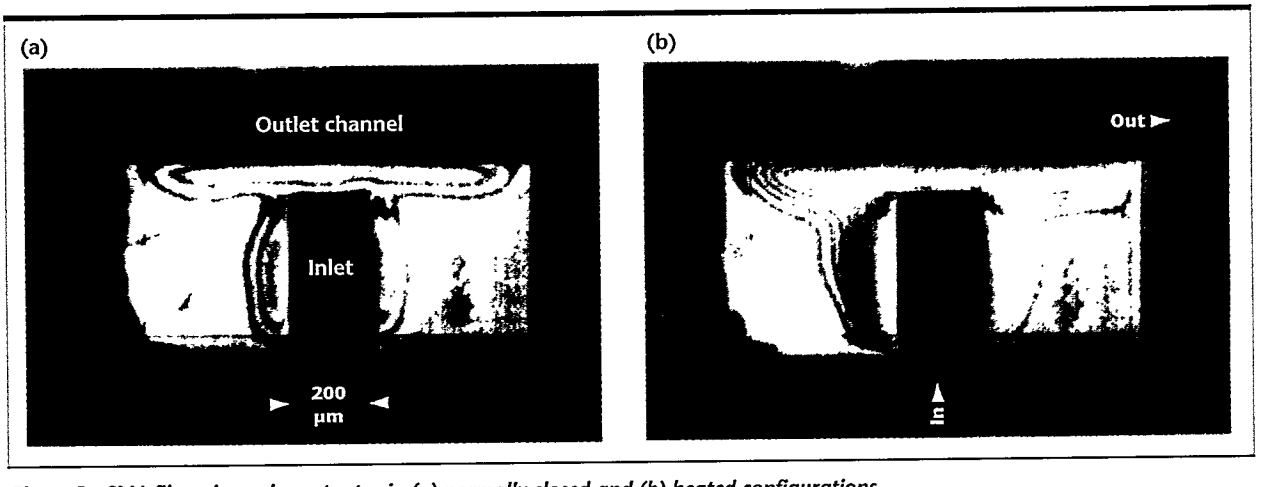


Figure 3. SMA film microvalve actuator in (a) normally closed and (b) heated configurations.

an anneal near 500 °C.[11] Films processed in this manner retain their smooth surfaces after undergoing the crystallizing anneal. If the SMA film remains attached to a substrate, the large tensile intrinsic stress associated with densification during the amorphous to crystalline transformation, combined with a large tensile thermal stress, can destroy the film. To avoid the intrinsic residual stress due to film densification and to lower the crystallization temperature, the substrate may be heated during deposition,[13] producing as-deposited crystalline films, at the expense of enhanced surface roughness.

Figure 1 shows the stress-temperature behavior of a Ni-Ti-Cu film on silicon, measured by substrate curvature. This film was deposited by sputtering from an alloy Ni-Ti-Cu target with a Ti-rich composition, to compensate for preferential sputtering of nickel. Upon heating, the film recovers more than 500 MPa of the residual stress, and has a hysteresis width of 10 °C. After 100 thermal cycles, the recoverable stress drops by 75 MPa, but is expected to stabilize between 300 and 400 MPa. The shape memory behavior of a Ni-Ti-Cu diaphragm, fabricated by sputtering onto a 1000 Å silicon nitride membrane, is shown in **Fig. 2**. A deflection of 30 µm at the membrane center is recovered by heating, and there is a small permanent deformation at the four edges, in the region of high stress.

The SMA film microvalve actuator is shown in **Fig. 3** in the normally closed **(Fig. 3)** and heated **(Fig. 3)** configurations. A shape memory/silicon composite diaphragm deflects downwards upon heating, pulling away from the glass cover layer and linking the inlet and outlet flow channels. In **Fig. 3**, each fringe corresponds to a change in gap distance of approximately 0.25 µm between the diaphragm and the glass. Due to residual tensile stress in the silicon/SMA diaphragm, the valve is slightly open in the unheated state, and opens further with heating.

Future Work

We are currently working on the development of high-force, large-displacement SMA microactuators for medical and microfluidic control applications.

References

1. Funakubo, H., Ed. (1984), *Shape Memory Alloys*, J. B. Kennedy, Translator, Gordon and Breach Science Publishers, New York, N.Y.
2. Duerig, T. W., K. N. Melton, D. Stockel, and C. M. Wayman, Eds. (1990), *Engineering Aspects of Shape Memory Alloys*, Butterworth-Heinemann Ltd., London, England.
3. Walker, J. A., K. J. Gabriel, and M. Mehregany (1990), "Thin Film Processing of TiNi Shape Memory Alloy," *Sensors and Actuators*, **A21-A23**, p. 243.
4. Johnson, A. D. (1991), "Vacuum-Deposited TiNi Shape Memory Film: Characterization and Application in Microdevices," *J. Micromech. Microeng.* **1**, p. 34.
5. Kuribayashi, K., T. Taniguchi, M. Yositake, and S. Ogawa (1992), "Micron Sized Arm Using Reversible TiNi Alloy Thin Film Actuators," *Mat. Res. Soc. Symp. Proc.* **276**, p. 167.
6. Kuribayashi, K., S. Shimizu, T. Nishinohara, T. Taniguchi, M. Yoshitake, and S. Ogawa (1993), "Trial Fabrication of Micron Sized Arm Using Reversible TiNi Alloy Thin Film Actuators," *Proc. 1993 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Yokohama, Japan*, (IEEE 0-7803-0823-9), p. 1697.
7. Lee, A. P., D. R. Ciarlo, P. Krulevitch, S. Lehew, J. Trevino, and M. A. Northrup (1995), "A Practical Microgripper by Fine Alignment, Eutectic Bonding, and SMA Actuation," *Proc. 8th Int. Conf. Solid State Sens. and Act., Transducers '95, Stockholm, Sweden*, p. 368.
8. Ikuta, K., M. Hayashi, T. Matsuura, and H. Fujishiro (1994), "Shape Memory Alloy Thin Film Fabricated by Laser Ablation," *Proc. IEEE Workshop on Micro Electro Mechanical Systems, Oiso, Japan*, p. 355.
9. Chang, L., and D. S. Grummon (1991), "Elevation of Premartensitic Transformation Temperature by $(Ni+Cu)_2$ Ti Precipitation in Sputtered $Ni_{(1-X)}$ $TiCu_{(X)}$ Thin Films," *Scripta Met.* **25**, p. 2079.
10. Krulevitch, P., P. B. Ramsey, D. M. Makowiecki, A. P. Lee, and M. A. Northrup (1995), "Mixed-Sputter Deposition of Ni-Ti-Cu Shape Memory Films," *Thin Solid Films*, (in press).
11. Busch, J. D., A. D. Johnson, C. H. Lee, and D. A. Stevenson (1990), "Shape Memory Properties in Ni-Ti Sputter-Deposited Film," *J. Appl. Phys.* **68**, p. 6224.
12. Ikuta, K., M. Hayashi, T. Matsuura, and H. Fujishiro (1994), "Shape Memory Alloy Thin Film Fabricated by Laser Ablation," *Proc. IEEE Workshop on Micro Electro Mechanical Systems, Oiso, Japan*, p. 355.
13. Gisser, K. R. C., J. D. Busch, A. D. Johnson, and A. B. Ellis (1992), "Oriented Nickel-Titanium Shape Memory Alloy Films Prepared by Annealing During Deposition," *Appl. Phys. Lett.* **61**, p. 1632.

Polyimide Electrostatic Microvalve Development

Abraham P. Lee, Julie Hamilton, and Jimmy Trevino
Engineering Research Division
Electronics Engineering

During FY-95 we made progress towards our objectives to 1) design a low power (< 1 mW), low leakage, high efficiency microvalve; 2) develop the fabrication process of polyimide multilayer cantilever actuators; and 3) test microvalves, including characteristics such as the sustaining pressure vs the applied voltage, the response time of the actuators, and testing in liquid and gas.

Introduction

The concept of our microvalve is shown in **Fig. 1.** A design process flow has been developed, integrating polymer surface micromachining with bulk silicon micromachining **(Fig. 2).**

Progress

Ground Plane Electrode. Processes for generating the ground plane electrode have been evaluated. It has been verified that by using *in situ* phosphorous-doped low-pressure chemical vapor deposition (LPCVD) (developed by N. Raley and reported elsewhere in this section), and then driving the phosphor into the top layer by annealing at 1000 °C for 30 min, the resistivity is the most controllable and moreover, the smoothest in surface topography. Boron-doped surfaces generate an undesired etch stop (ρ is not readily controllable), and gaseous phosphine diffusion generates rough surfaces full of craters of concentrated phosphorous sites.

Etching. Single-sided etching is critical in our project. The high temperature (> 65 °C) induces high stress in most of the single-sided fixtures, due to the thermal expansion of the rubber O-ring. Also, the teflon screws used to clamp the fixtures fracture at higher temperatures. Therefore, chemical etching of silicon, which requires 105 °C, is not feasible.

Al as a sacrificial layer is difficult to etch since the etchant phosphoric acid is too viscous to etch underneath the cantilevers (<1 μm space). We also used

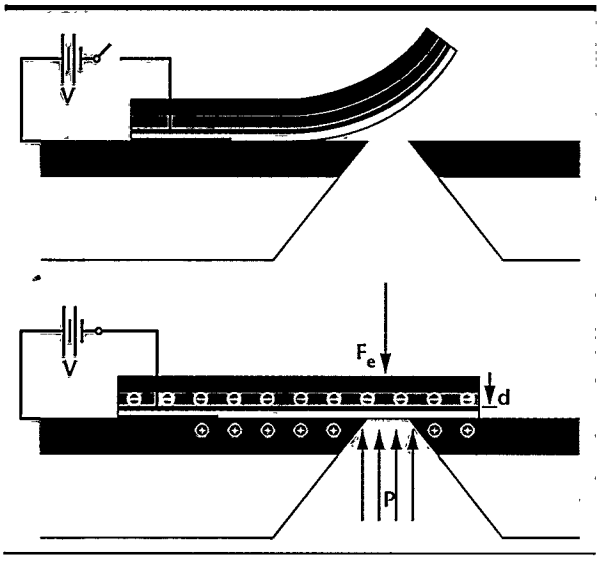


Figure 1. Schematic of operation principle of polyimide electrostatic microvalve.

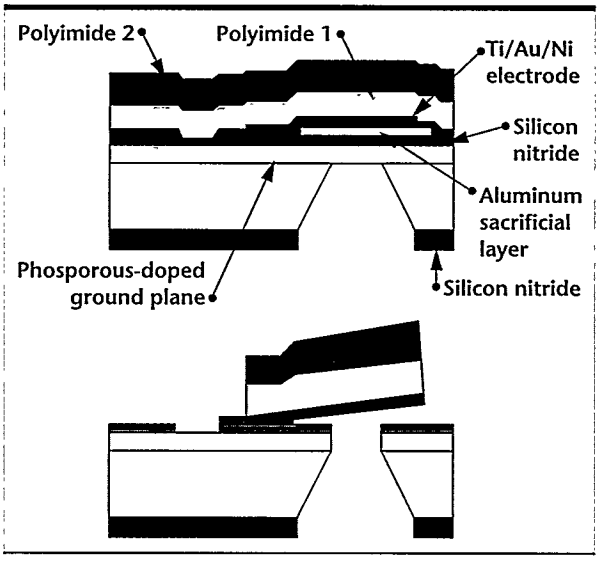


Figure 2. Process flow for polyimide electrostatic microvalve.

plasma-enhanced chemical vapor deposition (PECVD)-deposited oxide as a sacrificial layer.

Adhesion of polyimide to nitride is sufficient with 400 °C curing for 30 min.

Figure 3 shows a finite element analysis of the released polyimide cantilever, of width 100 µm and length 500 µm.

Results. We successfully fabricated a series of microvalves with a variety of hole sizes, cantilever geometries, and layer thicknesses **(Fig. 4a and 4b)**. We also developed a successful process flow that can be used by Lawrence Livermore National Laboratory (LLNL) designers in the future.

Due to adhesion problems during the release process, many cantilevers lifted off. Initial results have demonstrated the microvalve completely sealing with an applied bias voltage of 20 V. The total vertical displacement generated was 300 µm. In the future, it is believed that the required voltage could be lowered by reducing the gap between the electrodes. Also, an improvement in the ground pad contact resistance would render to the speed of the valve.

Furthermore, the nitride in the field was etched away during plasma etching, the polyimide causing breakdown at the rim of the electrodes. These two problems could be fixed by returning to the original process with a thinner first polyimide layer, and by changing the gas composition of the reactive ion etching (RIE) process (less CF_4 and more O_2).

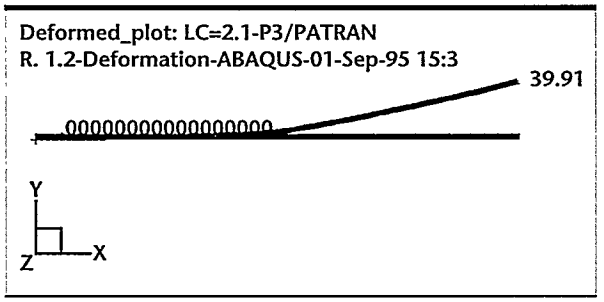


Figure 3. Finite element analysis of the released polyimide cantilever. The width of the cantilever is 100 µm and the length is 500 µm.

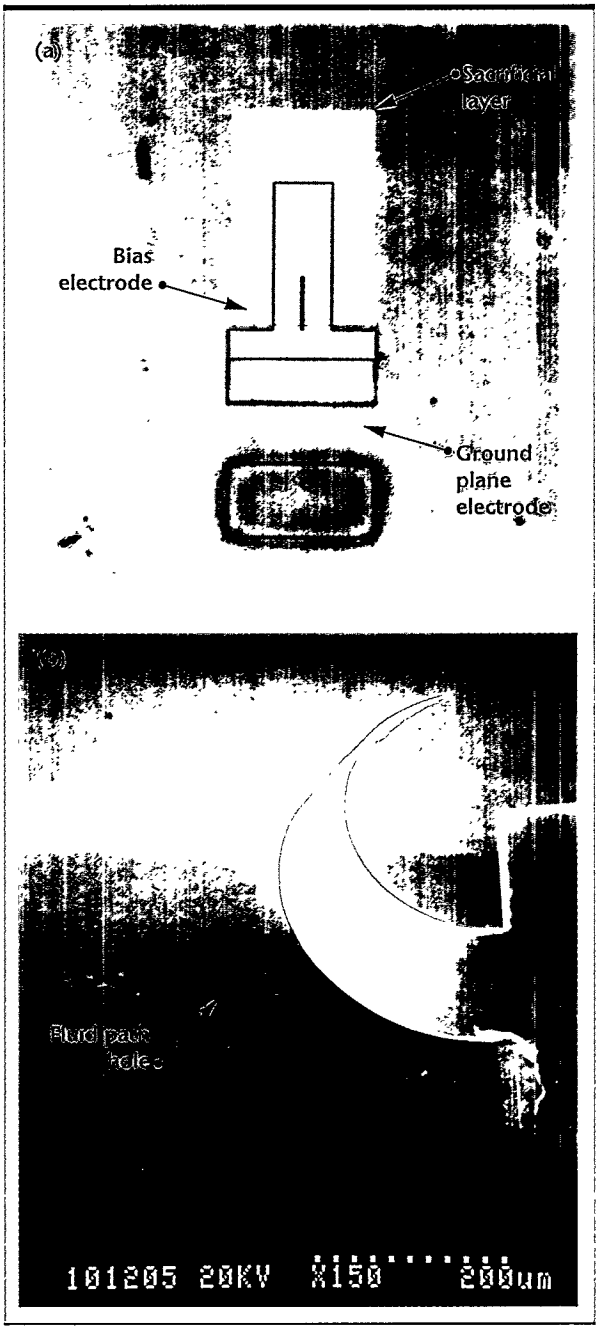


Figure 4. (a) Photograph of microvalve before release; (b) SEM micrograph of microvalve after release, showing the etched hole beneath it.

Microinstrumentation for Analysis of Biological Samples

Raymond P. Mariella, Jr., Gerald G. Eveleth, Richard G. Langlois, Donald A. Masquelier, Dino R. Ciarlo, Mukund C. Venkatesh, and Shadi Shakeri
Engineering Research Division
Electronics Engineering Department

We have developed a new way of performing large-fragment DNA analysis, using flow cytometry. We have also completed a microfabricated sheath-flow sample injector; designed and fabricated the components to test a new concept for combining very slow flow rates in flow cytometry with total internal reflection, to collect scattered light; and collaborated with the Department of Laboratory Medicine at the University of California, San Francisco (UCSF) to develop an instrument dedicated to an assay for HIV activity in blood samples.

Introduction

The goal of this project is to use microfabrication for the development of less expensive instrumentation for the analysis of biological samples. Both electrophoresis and flow cytometry for large-fragment DNA analysis were initially targeted. However, we conceived of a new way to perform large-fragment DNA analysis using flow cytometry instead of electrophoresis, and, therefore, devoted all of our efforts this year to flow cytometry. We also completed our first microfabricated sheath-flow/sample injector, which performed well.

We have evaluated a variety of optical configurations for the detection of fluorescence in our flow cytometer, with the intent of maximizing the low-light-level sensitivity. We have designed and had fabricated the components to test our new patented concept (IL-9444), which will combine the use of very slow flow rates in flow cytometry with our total-internal-reflection technique to collect scattered light. We began a collaboration with a blood-bank authority in the Department of Laboratory Medicine at UCSF, to develop an instrument dedicated to an assay for HIV activity in blood samples. In the process, an employee of Lawrence Livermore National Laboratory participated in a significant advance in blood typing.

Progress

Flow cytometry is a powerful diagnostic tool for monitoring microscopic particles in solution. We have already made significant progress, which has given us an order of magnitude larger signals and easier assembly, reduced noise levels, increased experimental accuracy, and reduced costs. Our new, patented idea (IL-9106) has been to use the total internal reflection of the flow stream itself to capture elastically and inelastically (e.g., fluorescent) scattered light.[1] We use this flow stream as an optical waveguide to transmit this scattered light efficiently to an optical detector, rather than using the traditional collector for perpendicular light scatter (PLS), which is an expensive, high-numerical-aperture lens.[2] **Figure 1** is a schematic drawing of the experimental set up. One important point we have already demonstrated is that our new technique allows the positioning, and, thus, the automatic alignment, of our PLS collector with no laser on, and with no sample being consumed. We just ensure that the liquid stream is directed onto the fiber optic tip.

We devised a technique to microfabricate the sheath-flow/sample injector so that the entire apparatus can also be inexpensive, using advanced etching and wafer-to-wafer bonding techniques. Although we found this first model to be more difficult to use than the highly-engineered commercial version, a success of our first unit is gratifying. The data collected using the microfabricated source is shown in **Fig. 2.**

Shortly after the success of our microfabricated sheath-flow/sample injector in a miniature flow cytometer, we began a collaboration with Prof. Girish Vyas at UCSF, to develop an instrument dedicated to a bead-based assay for viral pathogens in blood-bank samples. Since the demanding requirements

for the best single-color low-light-level sensitivity for his assay strongly overlapped the technical requirements for a two-color technique we had conceived for the sizing of DNA fragments after digestion with restriction fragments, we focused our efforts on this collaboration. This concept is described in the next paragraph.

During the course of our collaboration, we also participated in discovering a method for rapid, simultaneous typing of blood for ABO and Rh factors,

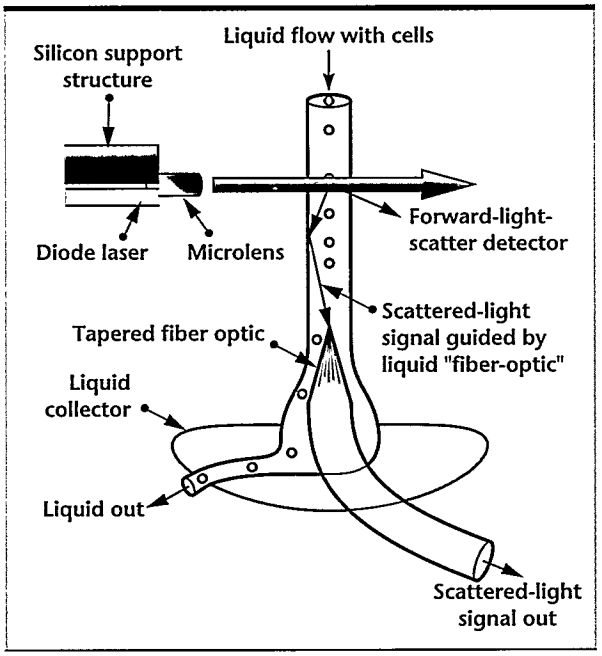


Figure 1. Schematic drawing of diode-laser-based flow cytometer. Cells in an aqueous carrier flow down through an orifice and form a free stream. When any cell passes through the laser beam, it will simultaneously scatter (reflect) the laser light into the forward-light-scatter and the right-angle-scatter detectors. This reveals the cell's size and internal structure.

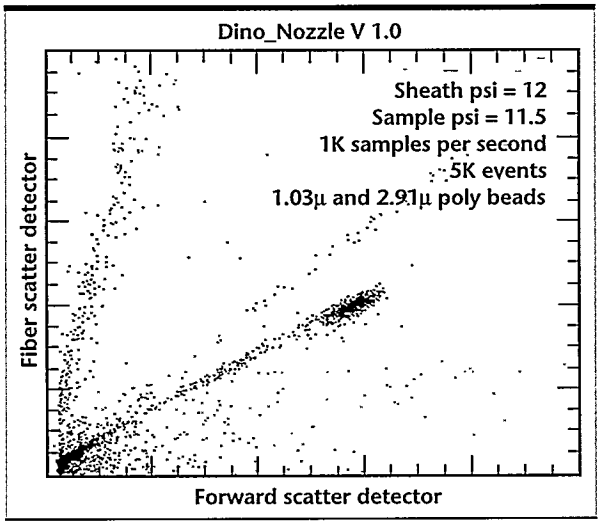


Figure 2. Data collected using microfabricated source.

using non-fixed erythrocytes in a three-color assay with the flow cytometer.[3] Although this was not planned when we started our project, we have achieved the first advance in blood typing in decades.

The previous flow-cytometric method for sizing DNA fragments after digestion by restriction enzymes was to saturate the fragments with intercalating dyes, such as "TOTO," a commercially-available homodimer of thiazol orange, and to size the fragments according to the laser-induced fluorescence of each fragment. The limitation has been that the system electronics uses this fluorescence itself as the trigger for detection of a fragment's passage through the laser beam.[4] This has meant that fragments smaller than 2000 bases are effectively unseen, since fluorescence from such a small fragment is equal to the background electronics noise in the trigger systems. We are pursuing a suggestion to ligate a different fluorescent label to each DNA strand before the restriction enzyme does its digestion (carried out in varying degrees of completion, this is known as "partial fragment analysis"); or to ligate this second label to each fragment after the digestion. Thus, the trigger signal would be independent of the fragment's length.

In a recent demonstration of the high optical throughput of our system, along with its very low background noise levels, we have observed weak green fluorescence from 0.14-µm and even 0.044-µm diameter beads, manufactured by Duke Scientific Corp. We were unable to observe either of these on the industry's leading commercial flow cytometers.

References

1. Mariella Jr., R. P., G. van den Engh, M. A. Northrup (1995), *Aqueous Carrier Waveguide in a Flow Cytometer*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (U.S. Patent 5,475,487).

2. Ormerod, M. G., Ed. (1990), *Flow Cytometry. A Practical Approach*, (2 volumes) IRL Press, (Oxford University Press), Oxford, Great Britain.

3. Shakeri, S. A., K. S. Vekateswaran, R. Mariella Jr., and G. N. Vyas (1995), "Simultaneous ABO and Rh Blood Typing by Flow Cytometry Using a Three-Color Fluorescence Detection System," *presented at the 37th annual meeting of the American Society of Hematology, Seattle, Wash.*

4. Goodwin, P. M., M. E. Johnson, J. C. Martin, and W. P. Ambrose (1993), "Rapid Sizing Of Individual Fluorescently Stained DNA Fragments By Flow Cytometry," *Nucleic Acids Research* **V21**, N4:803-806.

SMA-Actuated Silicon Microgripper

Abraham P. Lee, Peter A. Krulevitch, Jimmy Trevino, Stacy Lehew, Dino R. Ciarlo, and M. Allen Northrup
Engineering Research Division
Electronics Engineering

The objective of this project is to develop a microgripper that can operate in liquid as well as in air, and can access confined areas through small holes or tubes. This report describes the fabrication process for a practical silicon microgripper.

Introduction

Although microgrippers have been demonstrated in the past, they have been, for the most part, either small in gripping force, low in structural rigidity, relatively small in opening displacements, or inflexible in the design parameters. Another issue is that these microgrippers are usually attached to a large base for signal transduction and packaging purposes (**Fig. 1**). Furthermore, many of the designs could not function well in a liquid environment, due either to the high temperature required for activation, or to the high surface tension forces of the liquid, limiting the gripping capability. These issues have prevented the generation of a practical microgripper.

Progress

We have developed a microgripper with a large gripping force, a relatively rigid structural body, and flexibility in functional design. The actuation is generated by NiTiCu shape memory alloy (SMA) thin films, and the stress induced can deflect each side of the microgripper up to 55 μm, for a total gripping motion of 110 μm. This opening motion corresponds to a 13-mN opening force on the tip of the gripper.

The microfabrication process can be categorized into bulk micromachining, fine alignment, and NiTiCu SMA thin film deposition. **Figure 2** shows a schematic of a Lawrence Livermore National Laboratory silicon microgripper; **Fig. 3** shows SEM micrographs of a completed device. **Figure 4** illustrates the attachment of the microgripper to a small guide wire with the same cross-section as the microgripper itself.

The microgripper is 900 × 200 × 380 μm^3 in dimension. Each silicon cantilever is 12.5 μm thick, and 5-μm NiTiCu SMA thin films are deposited on both sides of the microgripper for actuation. The pushing pad is designed to assist in releasing the gripped object by pushing forward as the gripper is opened. This pad is 30 μm wide, while the gripping jaws are 60 × 110 × 100 μm^3. The hollow channel is 110 μm wide.

The gripping jaws, pushing pads, and hollow channel are shaped by a combination of precision sawing and bulk-machining of silicon. The connection of the microgripper to external leads and milli-end effectors, however, requires assembly and, therefore, does not allow batch fabrication. There is

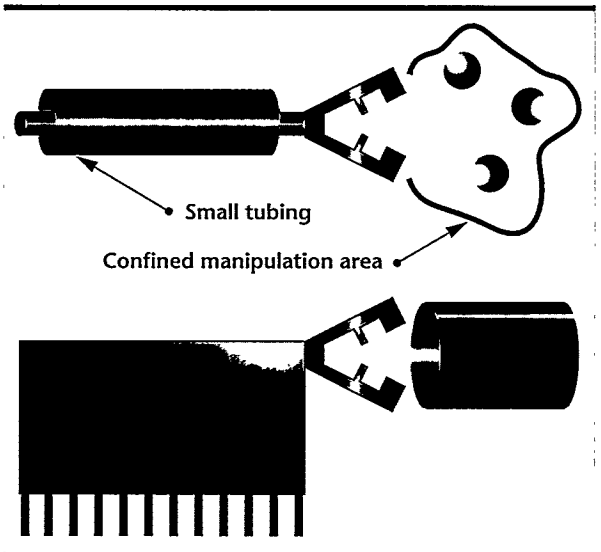
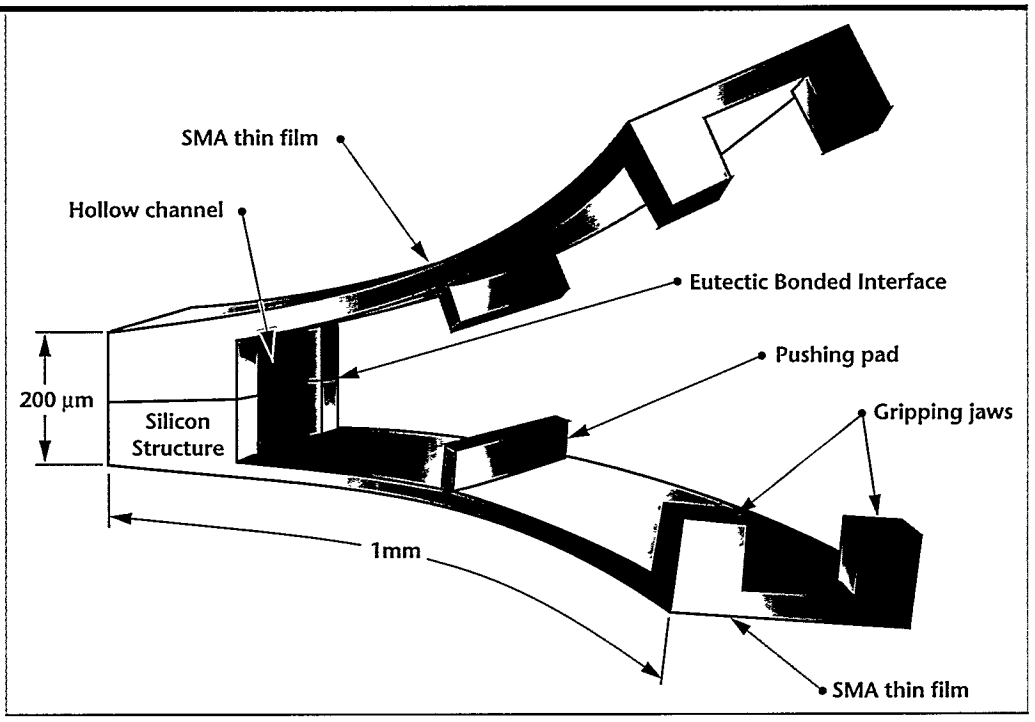


Figure 1. Schematic of microgripper, illustrating the capability to manipulate through small tubing to access confined or hazardous environments.

on-going effort to develop assembly techniques for packaging. To determine the power requirements for microactuators attached to the end of catheters/guide wires for endovascular stroke treatments, we developed a micromachined silicon microcantilever (with exact dimensions of a microgripper), integrated with insulated serpentine heaters and SMA thin film. This provides a structure available for testing in water/blood, with extended feedthrough wires to simulate the long electrical passages in catheter-based procedures.

The test wafer that has been completed was thermoelectrically heated using the NiTiCu film as the heater. Tests in air (power consumption 0.1 W), as well as in water (power consumption 0.2 to 0.3 W), demonstrated an order of magnitude less power required. The deflection of the cantilevers is 35 to 45 µm each in water (total of 70 µm with two). The cantilevers are 12 µm thick and the completed cantilevers are flat through stress compensation. The test cantilevers will be thermally insulated from the environment by a thin film oxide layer, and retested for comparison of the power consumption. More wafers are being processed with polysilicon as the heater material, instead of directly heating the NiTiCu. The stress induced by the SMA is about 400 MPa. (See "Shape Memory Film Microactuators," also in the Microtechnology thrust area section.)

Figure 2. Detailed schematic of silicon microgripper.



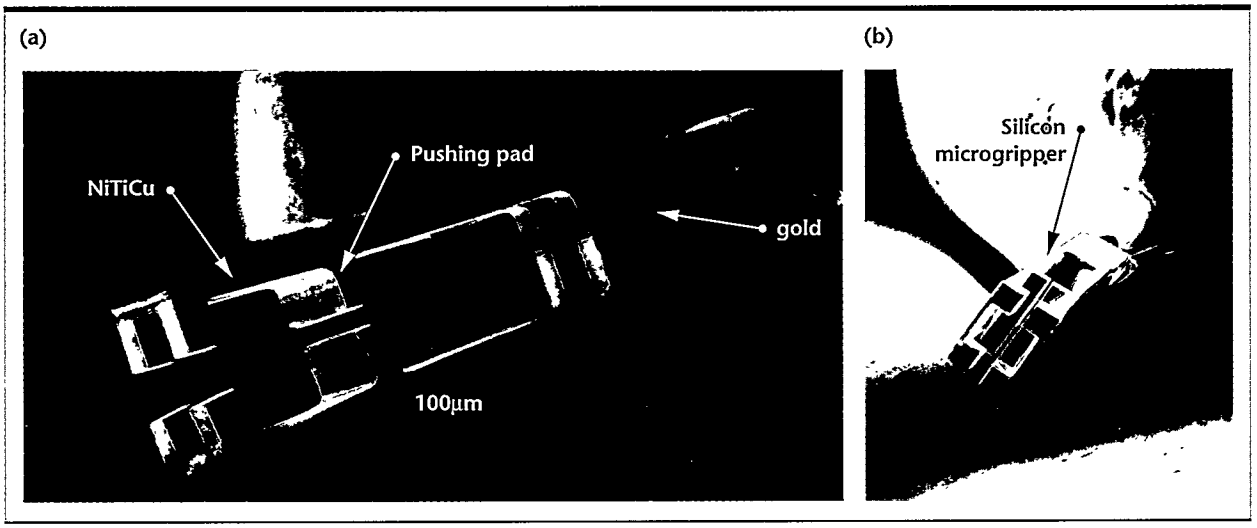


Figure 3. SEM micrographs of (a) side-view, and (b) front-view of microgripper.

The temperature of the surrounding water was measured by a thermocouple to be slightly below 60 °C without insulation. With insulation, it is possible that a 50° to 60 °C range is achievable.

For biomedical applications in the body, and possibly in the blood vessels, it is important to know that the SMA must reach 70 °C initially, but the temperature can immediately (within milliseconds) be lowered to 55 °C, and the gripper will remain open, indicating the possibility of lowering heat transfer into the surrounding medium. Optimization of the thickness of the thin film insulation reduces power consumption and enhances biocompatibility.

Figure 4. SEM micrograph illustrating microgripper attached to a guiding wire.

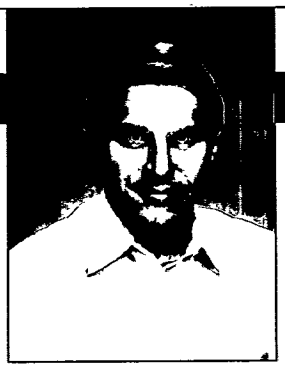
Summary

In this report, a process of fabricating a practical silicon microgripper is described. This microgripper has large gripping force (13 mN) and is actuated at relatively low actuation temperatures (70 °C). Fine alignment, AuSi eutectic bonding, precision sawing, NiTiCu SMA deposition, as well as bulk silicon machining are the core of the process. The fabrication processes of this microgripper, although not standard, do not involve exotic equipment. All the individual processes could be streamlined for batch processing and, therefore, fabrication costs (excluding the packaging with small guidewires) could be reduced.

Future Work

For future developments, the outer surfaces of the gripper are being integrated with heaters and/or strain sensors for remote active heating and, possibly, feedback control. The hollow channel has the potential for either wire connection or injection of liquids and therapeutic medicine. Another important advantage is the possibility to apply alternative actuation mechanisms on the microgripper structure, either hydraulic, or simply thermal bimorphic. Many creative designs of practical microgrippers can be conceived using this basic process.

Applications for the microgripper include assembling small parts for manufacturing; minimally-invasive *in vivo* biopsy tissue sampling; catheter-based endovascular therapeutic procedures; and remote handling of small particles in extreme environments, such as high or low pressures, and hazardous fluids.



Kenneth L. Blaedel, Thrust Area Leader

During FY-95, the Manufacturing Technology thrust area focused on precision manufacturing, particularly trying to increase the precision-to-cost ratio of processes and devices that are important to our future. Our motivation is that demand for the utmost in precision is giving way to demand for less expensive precision. This is an extension of our traditional domain, since we still must confront the frontiers of absolute precision.

We moved some of our activity from the ultimate in precision, usually applied at the last stages of a manufacturing process, to the roughing or intermediate finishing stages. In some cases, cost can be saved by applying precision earlier in a manufacturing process.

Our environment at Lawrence Livermore National Laboratory (LLNL) includes Laboratory-directed research and development (LDRD), CRADA projects, and development work that is done and paid for within a program. In general, LDRD incorporates the higher risk research associated with manufacturing technology, while CRADAs embody the lower risk reduction to practice associated with industry. Ideas developed in the thrust area fall somewhere in between.

In addition to the projects whose progress is discussed in the two following articles, a number of small efforts were conducted within the thrust area, and therefore warrant mention here.

A project was initiated to explore the manufacture of an astronomic telescope for hard x-ray imaging. Assessment revealed that the critical issues facing such a telescope were, first, the fabrication of the nested, cylindrical optical mirrors; and second, the alignment and assembly of the mirrors. The deliverable was a mission concept proposal to NASA.

The challenge of fabricating the optics boils down to making about 2000 reflective surfaces with adequate accuracy, light in weight, and reasonable in cost. One method for making such optics is to replicate the surfaces of a specially made mandrel by electroplating nickel, separating the nickel from the mandrel, and then depositing a multilayer on the nickel shell to increase the reflectivity of the optical surface.

One main problem associated with this technology is that the stiffness-to-weight ratio for nickel dictates that almost half the telescope weight resides in the optical surfaces. Therefore, a higher performance material is sought with a factor of two improvement over nickel in specific stiffness, and with marginal improvement in thermal stability. The relevant characteristics are 1) the expense of depositing the material; 2) the residual stress of the deposited material; 3) the long-term dimensional stability of the structure; and 4) the coefficient of expansion, which matches the integrating structure, probably beryllium. The material must also have a reasonable fracture toughness and yield strength, to easily endure the launch.

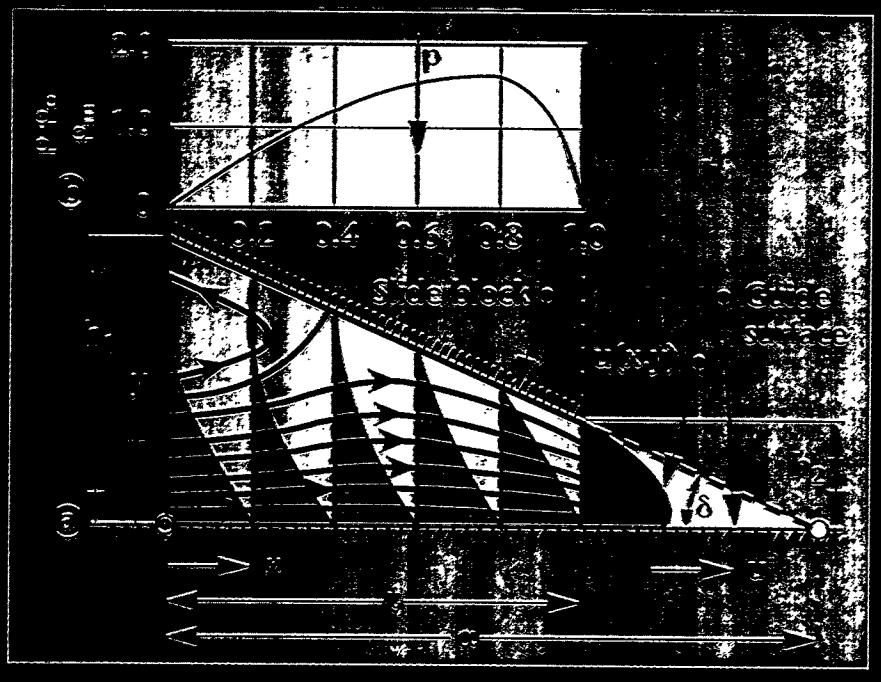
The second challenge is to ensure the precision alignment and assembly of the integrated telescope. We have selected the idea of precision-machined integrating end-plates that allow the shells to assemble without further alignment directly into position, vastly reducing the labor of assembly. The required tolerances of assembly fall well within that which can be precision-turned.

We are considering ways of bringing precision manufacturing to new customers, at minimum cost. One new customer is the Human Genome Project. A limiting issue will be the mechanical precision with which automated wet chemistry can be conducted. The tests are now conducted by a mechanical robot, but if the test requirements are scaled out a few years, then an order-of-magnitude increase in precision will be required for very little escalation in cost. With regard to the time-scale for this, LLNL appears to be ahead of the ability or motivation of industry to respond.

In a cooperative effort with LLNL's Energy, Manufacturing, and Transportation Technologies Program, we participated in a symposium with the University of California, at Berkeley to understand the critical role of precision manufacturing in selected industries. Representatives of industry, academe, national laboratories, and government heard industrial representatives give their perspective on precision manufacturing. Speakers from the automotive, airframe, engine, semiconductor, and machine tool industries expressed their need for a foundation in precision.

We continue to seek a local ally to help grow and preserve precision manufacturing as a core technology available to LLNL.

Manufacturing Technology



4

4. Manufacturing Technology

Overview
Kenneth L. Blaedel, Thrust Area Leader

Precision Saw for Slicing Wafers of Crystalline Materials
Bradley L. Fix *4-1*

Reactive Sputter-Deposition of Boron Nitride
Daniel M. Makowiecki, Alan F. Jankowski, Jeffrey P. Hayes, and Mark A. McKernan *4-5*

Precision Saw for Slicing Wafers of Crystalline Materials

Bradley L. Fix
Manufacturing and Materials Engineering Division
Mechanical Engineering

We have developed a method of reducing the cost and the waste involved in the manufacture of thin wafers from bulk optical and electronic materials. We report on progress in three areas: hydrodynamic bearing design and modeling, testbed development, and the conceptual design of a prototype precision bandsaw.

Introduction

The goal of this project is to reduce the cost and the waste involved in manufacturing thin wafers from bulk optical and electronic materials, and therefore, to increase the yield of a boule. The eventual result of this work will be a saw that incorporates both new blade technology and precision machine tool design. The advantage of this saw is that, as a system, it changes the very first machining operation, slicing, from a roughing process to a semi-finishing process.

Although we are generally interested in many crystalline materials, we are particularly interested in KDP and silicon wafers. KDP is critical to Lawrence Livermore National Laboratory's (LLNL's) Laser Program, and the slicing of silicon is of interest to LLNL's EMATT's work with the semiconductor industry. The first step in processing both of these materials is to slice wafers from boules. Current slicing techniques create wafers with significant figure errors and subsurface damage. The difficulty in manufacturing both of these crystals is being increased, since the National Ignition Facility (NIF) will require 40- to 42-cm crystals (greater than the NOVA laser's 30 cm), and the semiconductor industry is currently moving from 200-mm to 300-mm silicon wafers.

One component of our approach to reduce the cost of finishing these crystals is to institute precision manufacturing techniques upstream from final finishing operations. This will be achieved by reducing figure errors and subsurface damage during the roughing (sawing) operation, which reduces or eliminates a number of the follow-on operations.

Advances in two critical technologies allow this. The first advance is a new method for accurately attaching small diamond grit to the leading edge of the blade. This allows cutting the wafer with minimal subsurface damage, and stems from our experience in grinding brittle materials. The second advance is a new method to guide the blade during sawing, to maintain a straight cut, and thereby, minimize the non-flatness of the wafer. The second advance is the subject of this report.

The concept is to develop a hydrodynamic bearing in the saw kerf using the cutting fluid and the relative velocity of the blade with respect to the workpiece. This bearing will create a restoring force orthogonal to the plane of the blade, to resist deflection and bending, and maintain a straight cut. Although embodied here as a bandsaw, this concept is applicable to other slicing configurations including OD and ID saws.

During this project year, we have concentrated our work in three areas: hydrodynamic bearing design and modeling, bearing testbed development, and conceptual design of a prototype precision bandsaw.

Progress

Hydrodynamic Bearing Design and Modeling

Hydrodynamic principles of bearing lubrication theory were used to design and model "features" which, when added to the surface of a bandsaw blade, would create a restoring force to maintain blade straightness. The intent is to use the cutting

Table 1. Bearing designs and characteristics.

Hydro Band Saw, Force on Plane Wedge with Water, 7/9/95

	Input Parameters								
								Finite Width Wedge	
	Visc. vis Pa s	Sliding Velocity U ft/min	Chan. Thk. Ratio k=h1/h2	Outlet Chan. Thk. h2 in.	Slider Length l in.	Slider Width b in.	Reduced Re Re*(H/l)^2	Avg.Pres. Np*p psi	Avg.Shr. Nf*f psi
U = 500, l = 5 B-1, h2 = 00010, k = 1.1	0.001	500.0	1.10	0.0010	0.50	1.00	1.27E-01	2.12E-01	1.60E-02
U = 1000, l = 5 B-1, h2 = 00010, k = 1.1	0.001	1000.0	1.10	0.0010	0.50	1.00	2.55E-01	4.23E-01	3.20E-02
U = 1500, l = 5 B-1, h2 = 00010, k = 1.1	0.001	1500.0	1.10	0.0010	0.50	1.00	3.82E-01	6.35E-01	4.80E-02
U = 2000, l = 5 B-1, h2 = 00010, k = 1.1	0.001	2000.0	1.10	0.0010	0.50	1.00	5.10E-01	8.46E-01	6.40E-02
U = 2500, l = 5 B-1, h2 = 00010, k = 1.1	0.001	2500.0	1.10	0.0010	0.50	1.00	6.37E-01	1.06E+00	8.00E-02
U = 3000, l = 5 B-1, h2 = 00010, k = 1.1	0.001	3000.0	1.10	0.0010	0.50	1.00	7.65E-01	1.27E+00	9.61E-02

Hydro Band Saw, Force on Plane Wedge with Light Oil, 7/9/95

	Input Parameters								
								Finite Width Wedge	
	Visc. vis Pa s	Sliding Velocity U ft/min	Chan. Thk. Ratio k=h1/h2	Outlet Chan. Thk. h2 in.	Slider Length l in.	Slider Width b in.	Reduced Re Re*(H/l)^2	Avg.Pres. Np*p psi	Avg.Shr. Nf*f psi
U = 500, l = 5 B-1, h2 = 00010, k = 1.1	0.087	500.0	1.10	0.0045	0.50	1.00	2.98E-02	9.05E-01	3.08E-01
U = 1000, l = 5 B-1, h2 = 00010, k = 1.1	0.087	1000.0	1.10	0.0045	0.50	1.00	5.96E-02	1.81E+00	6.17E-01
U = 1500, l = 5 B-1, h2 = 00010, k = 1.1	0.087	1500.0	1.10	0.0045	0.50	1.00	8.93E-02	2.72E+00	9.25E-01
U = 2000, l = 5 B-1, h2 = 00010, k = 1.1	0.087	2000.0	1.10	0.0045	0.50	1.00	1.19E-01	3.62E+00	1.23E+00
U = 2500, l = 5 B-1, h2 = 00010, k = 1.1	0.087	2500.0	1.10	0.0045	0.50	1.00	1.49E-01	4.53E+00	1.54E+00
U = 3000, l = 5 B-1, h2 = 00010, k = 1.1	0.087	3000.0	1.10	0.0045	0.50	1.00	1.79E-01	5.43E+00	1.85E+00

fluids available during typical sawing operations to generate the necessary guiding hydrodynamic forces, in the same way that lubricating oil generates forces in sliding mechanical bearings.

If the saw features are in the form of wedges, analytic solutions are available. **Figure 1**[1] shows a typical flow pattern and pressure distribution for a plane guide surface sliding past an infinitely wide inclined wedge slide block.

This model was used to design two prototype bearings, one using water for slicing silicon, and the other using light oil for slicing KDP. **Table 1** details the bearing designs and characteristics.

Hydrodynamic Bearing Testbed

A testbed was developed to evaluate the performance of the hydrodynamic bearing design and to test the accuracy of the simulations. The testbed consisted of a large rotating disk to simulate the bandsaw blade. In fact, this set-up more closely simulates the bearing as it would be used in an OD saw. Wedge features were machined into the face of a bronze pad. The pad was then positioned in close proximity to the rotating disk. A dynamometer was mounted beneath the bronze pad to measure both the normal and shear forces created by the hydrodynamic bearing. Water was initially chosen as the bearing fluid.

Figure 2 illustrates the wedge design used on the testbed. The wedge features were fabricated on the testbed by mounting a diamond tool onto the face of the rotating disk and flycutting the features into the bronze pad.

The experimental tests were conducted by rotating the disk at a fixed speed (300 rpm, or 1260 sfpm), and changing the gap between the bearing pad and rotating disk. The gap was varied from 0.017 in. to 0.0008 in. The results of the test are presented in **Fig. 3**. The nominal operating point for the bearing, as designed, is also shown for a 0.001-in. gap.

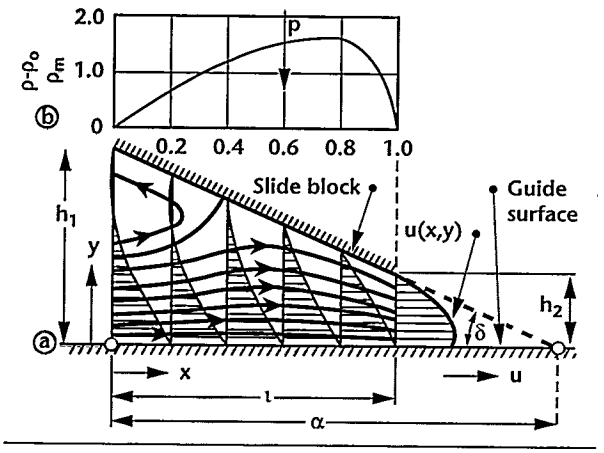


Figure 1. Typical flow pattern and pressure distribution for plane guide surface sliding past infinitely wide inclined wedge slide block.

Conceptual Design of a Bandsaw

During the first year of this project, preliminary work began on the conceptual design of a precision bandsaw to incorporate the blade technology. This concept, illustrated in **Fig 4**, consists of two precision spindles to rotate the blade, and a precision slide to hold and translate the boule for slicing. There are many peripheral issues that must still be addressed.

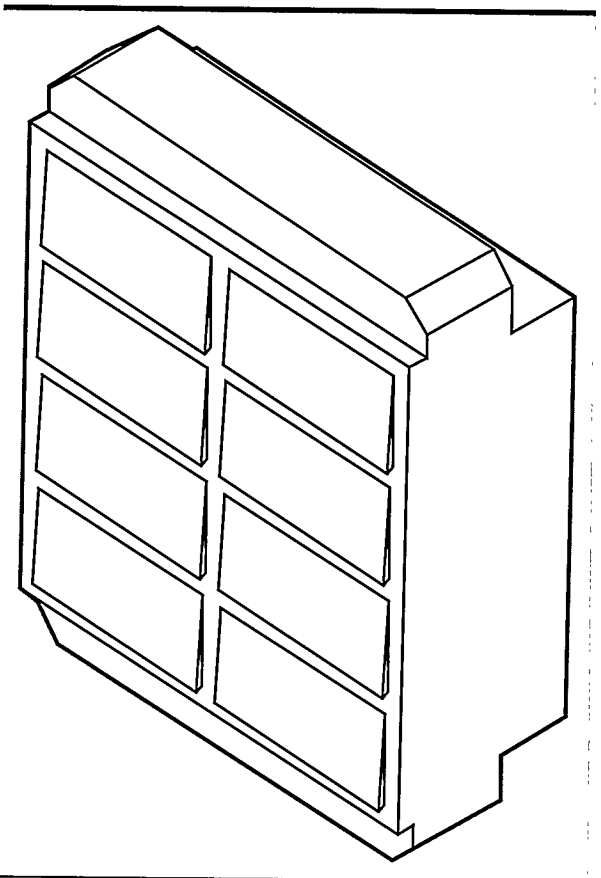


Figure 2. Illustration of wedge design used on testbed.

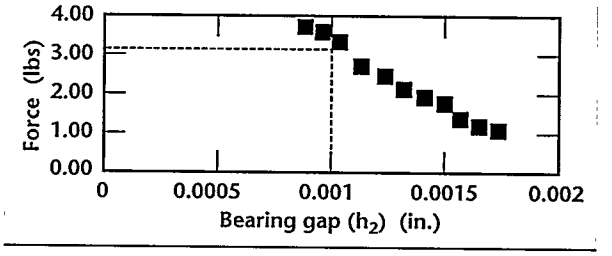


Figure 3. Results of experimental test.

Conclusions

We have found a very good correlation between the modeled forces and experimental results. During the second project year, additional tests will be conducted to corroborate initial results and to test an oil fluid bearing.

The primary issue with this bearing design is that, at high surface speeds, which are necessary to maintain productivity, the Reynolds number becomes large. Our current analytical method does not apply at $Re^* > 0.1$. To minimize the Reynolds number, smaller gaps, < 0.001 in. for water, must be maintained. Maintaining small gaps is possible. However, it may complicate and increase the cost of blade fabrication. In addition, shear forces generated by the bearing increase as the gap decreases, and as the surface speed increases. We would like to minimize the shear force, and thereby, reduce the power requirements for the spindle drive.

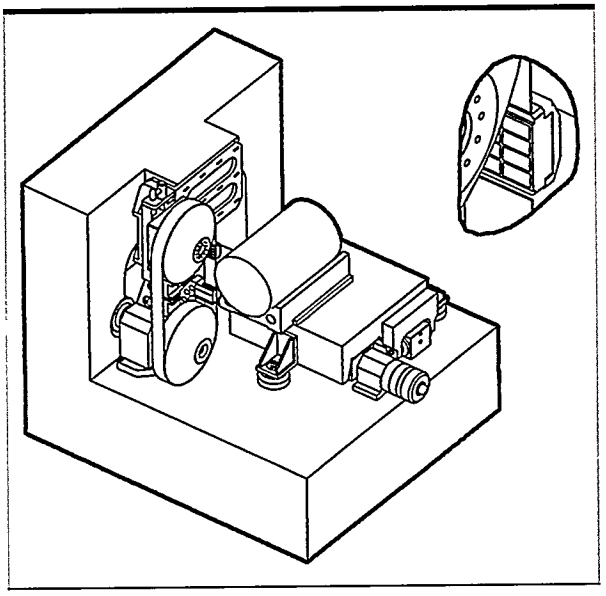


Figure 4. Drawing of conceptual design of precision bandsaw.

Future Work

During FY-96, we plan to investigate other hydrodynamic bearing designs, with the intent of overcoming the issues mentioned above. We will also begin to solve the problems involved in taking the testbed to a useable saw. These include, for example: obtaining a commercial source for fabricating endless bands with integral bearing features; constraining the wafer, so as not to propagate a crack in front of the blade as it is being sliced; determining the best balance between the two restorative forces, the tension in the band, and the hydrodynamic bearing force; and characterizing the effect of the high damping of the hydrodynamic bearing on the machine dynamics.

Reference

1. Schlicting, H. (1968), *Boundary-Layer Theory, 6th ed.*, McGraw-Hill, New York, N.Y., pp. 108–114.

Reactive Sputter-Deposition of Boron Nitride

Daniel M. Makowiecki and Alan F. Jankowski
Chemistry, Materials Science and Technology Division
Chemistry and Materials Science Department

Jeffrey P. Hayes and Mark A. McKernan
Applied Research Engineering Division
Mechanical Engineering

The preparation of fully dense boron targets for use in planar magnetron sources has lead to the synthesis of boron nitride (BN) films by reactive rf sputtering. The deposition parameters of gas pressure, flow and composition are varied, along with substrate temperature, and applied bias. The films are characterized for composition using Auger electron spectroscopy, for chemical bonding using Raman spectroscopy, and for crystalline structure using transmission electron microscopy. The deposition conditions are established which lead to the growth of crystalline BN phases. In particular, the growth of an adherent cubic BN coating requires 400 to 500 °C substrate heating and an applied –300 V dc bias.

Introduction

One of the strengths of Lawrence Livermore National Laboratory (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 rms 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.

Progress

Thin film processing techniques, such as laser ablation and ion beam sputtering have been developed to deposit BN films. A long-standing objective has been to stabilize the growth of the cubic phase of BN for the applications mentioned above. Typically, the source material for these physical vapor deposition methods has been BN in its hexagonal phase.

To date, the stabilization of the cubic phase in thin films has typically been achieved through the use of very energetic deposition processes, e.g., bias-sputtering or ion beam bombardment of the substrate.[1–4] Consequently, the cubic boron nitride (cBN) films have large intrinsic stresses. Therefore, in addition to preferential sputtering effects,[5] a mechanism for stabilization of cBN in films has been attributed to the presence of large stresses.[6] However, such highly stressed films are mechanically unstable and often have poor adherence. A less energetic physical vapor deposition process could have advantages in forming cBN films.

The recent development of fully dense, pure boron targets has made it possible to initiate an investigation of BN deposition without the hexagonal phase precursors.[7,8] The ablation or sputtering of a target material often results in the transport of molecular clusters to the substrate.[9] This is clearly the case for a target consisting of an equilibrium phase, such as hexagonal boron nitride (hBN). The deposition of even small amounts of the target material could promote the epitaxial formation of the hexagonal phase. An energetic bombardment is therefore required to transform the hexagonal phase into the less stable or metastable cubic phase. In this study, we assess the deposition conditions needed to rf-sputter pure boron targets, in order to stabilize crystalline growth in BN films. A comparison of the free energies of formation indicates that the thermodynamic equilibrium for the transformation of hBN to cBN is less favorable than the formation of cBN from pure boron.[10]

We initiate the study by examining the influence of working gas composition and substrate temperature variation on the formation of crystalline BN phases. Next, we explore the energetic conditions needed to stabilize cBN growth in BN films as synthesized with the application of substrate bias. Characterization of the BN films includes chemical evaluation through Auger electron spectroscopy (AES), structural analysis through the use of both light and electron microscopy, as well as an assessment of chemical bonding through Raman and Fourier transform infrared (FTIR) spectroscopy.

Experimental Method

The synthesis of the BN films proceeds from the rf sputter-deposition of fully dense, pure boron targets, using unbalanced planar magnetrons. The patented process used to create these unique boron targets has been described in detail elsewhere.[11] In brief, boron powders are compacted in an evacuated tantalum container which is then hot-isostatically pressed to yield a right circular cylinder of fully dense (i.e., less than 0.15% porosity) boron. The manufacture of boron sputtering targets then proceeds by a patented process involving the brazing of a compatible backing plate to prevent catastrophic target failure at initiation of the sputtering process.[11] X-ray diffraction analysis reveals the targets to be the rhombohedral crystalline phase of boron.

Two deposition chambers are used to evaluate the synthesis of BN using a reactive gas mixture of Ar-N_2. The first deposition chamber is used to establish the baseline effects of gas pressure, flow, and substrate temperature on the composition and phase formation of BN films. This deposition system is cryogenically pumped from atmospheric pressure to a base pressure of 5.3×10^{-6} Pa in 12 h, including a 4-hr, 100 °C bake-out. Some reports have indicated that epitaxial growth of cBN can be enhanced through the use of nickel epilayers.[12,13] As such, 20 to 50 nm of nickel are initially sputter-deposited onto silicon wafers for use as substrates.

The substrate table is positioned 9 cm horizontally away from the center of the 6.4-cm-diameter boron target. The substrate temperature is controlled using a Boralectric™ heater. The wafers are heated up to 600 °C, since above this temperature Cu$K\alpha$ x-ray diffraction reveals the formation of an orthorhombic phase of nickel silicide. The Ar sputter gas pressure is nominally selected as 1 Pa (7 mTorr), with a constant flow rate of 28 cc min^{-1}. A partial flow of N_2 is incorporated to yield a 0 to 55% composition range. The deposition rate is monitored with a calibrated 6 MHz-Au-coated quartz crystal. An increase in applied forward power from 100 to 300 W produces a linear increase in deposition rate from to 0.007 to 0.021 nm s^{-1}. The BN films are grown to a total thickness of 100 to 300nm, as verified using contact profilometry.

The second deposition chamber is used to examine the effects of applied substrate bias on the formation of BN. Some differences from the first deposition chamber are as follows: 1) the second chamber is fixtured to apply a BN hard coating to the surface of cutting tools; 2) provision is made for the initial sputter-deposition of 20 to 50 nm of either titanium or chromium to promote adhesion of the coating; and 3) the substrate table is positioned 5 cm vertically below an array of three magnetron sources, each of 3.3 cm diameter. Also, to assess the effects of film stress on adhesion, the sputter pressure is varied from 0.2 to 3 Pa. Boron nitride deposition rates of 0.02 to 0.04 nm s^{-1} are produced by forward target powers of 50 to 80 W. The bias applied to the forming BN coating is produced by drawing off approximately 5% of the applied rf

power to the target. The reported 0 to –400 V dc bias is the (negative) component of the rf power applied to the substrate. A small percentage of the ionized sputter gas will then strike the substrate and modify the structure of the coating.

The BN thin films are characterized for composition, morphology, and crystalline structure, as well as for chemical bonding. The film composition is measured using AES, coupled with depth profiling by Ar-sputtering. The intensities of the B, C, N, and O *KLL* Auger electrons (as data accumulated in the derivative mode) are used to compute the atomic concentrations. The detection limit for carbon and oxygen impurities is approximately 2 to 3 at.% in these BN films.

Transmission electron microscopy (TEM) is used for the assessment of film morphology and crystallinity. Bright field imaging of BN films prepared in plan view reveals the grain structure. Dark field imaging coupled with electron diffraction patterns, reveals the crystalline state of the films. Raman spectroscopy is used to examine the vibrational states of chemical bonds within the films. Although there are many crystalline phases of BN, including rhombohedral and wurtzite, the distinctive structures and chemical bonding are hexagonal and cubic. The Raman phonon of hBN appears at 1367 cm^{-1} as representative of sp^2 hybridized, planar bonding. The transverse optical (TO) and longitudinal optical (LO) modes of cBN appear at 1057 and 1306 cm^{-1}, respectively, as representative of sp^3 tetrahedral bonding. The relative intensity of the TO and LO modes for cBN can vary, since BN is noted to have strong Raman anisotropy.[14] In addition, the presence of a weak and broad feature at 1245 cm^{-1} represents strong plasmon-phonon coupling in cBN. FTIR is used to verify the bonding trends observed in these BN films using Raman spectroscopy. Whereas the Raman spectra are generated in reflection, the FTIR spectra are generated in transmission. Simulations of the IR spectra have also confirmed secondary hBN features as a peak at 783 cm^{-1}, and the increased asymmetry of the 1367 cm^{-1} peak with increasing film thickness towards a satellite peak at 1560 cm^{-1}.[2] The primary 1367 cm^{-1} hBN feature is associated with in-plane B-N bond stretch while the secondary feature at 783 cm^{-1} is associated with the out-of-plane B-N-B bond bend.[4,15] Finally, conventional light microscopy is used to compare the effects of intrinsic film stress on macroscopic film morphology.

Experimental Results and Analysis

The sputter gas composition and substrate temperature are initially varied to determine effects on BN film formation. AES analysis indicates that the coatings have a constant composition through the depth of the film. As an example, **Fig. 1** shows the depth profile of a film deposited at a substrate temperature of 600 °C with a Ar-55% N_2 gas mixture. The BN film composition is found to be independent over the gas composition range of 25 to 55% N_2 for the limited range of deposition rates used in this study. Heating the substrate above 200 °C during deposition is seen (**Fig. 2**) as advantageous to eliminate contamination of the BN films by oxygen. The average BN film composition is

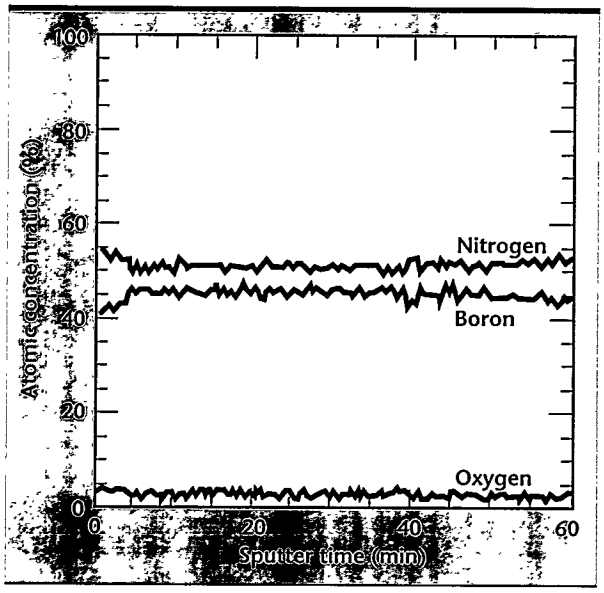


Figure 1. Atomic concentration depth profile of a BN film deposited at 600 °C.

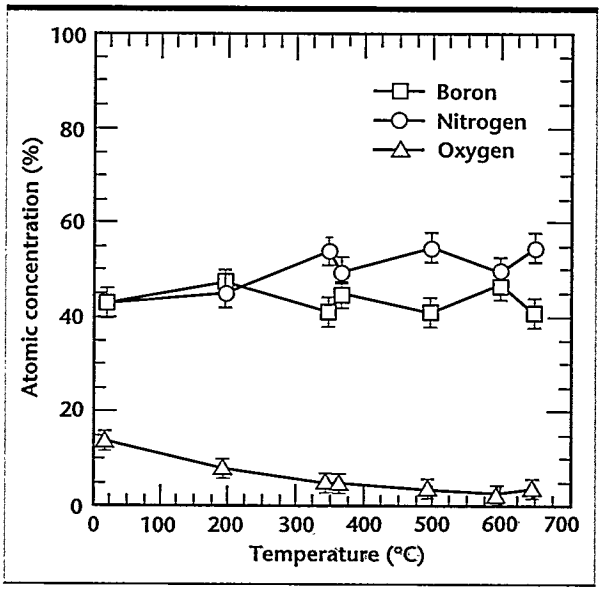


Figure 2. Variation of atomic concentration of BN films with substrate temperature.

found to be 45 at.% boron from the AES analysis. The Raman spectra of 150 to 200 nm thick BN films deposited over the 50 to 600 °C temperature range are shown in **Fig. 3**. Note that the individual curves are offset on the intensity axis to accentuate the Raman spectral features.

As the substrate temperature increases from 50 to 350 °C, the intensity of the primary hBN peak increases. Above 350 °C, the increase in temperature produces an asymmetry in the hBN peak towards the wavenumber corresponding to the LO mode of cBN. The inference from these Raman spectra is threefold: 1) at "room" temperature, an sp^2-bonded structure such as hBN is formed with some disorder; 2) as temperature increases, disordering of the BN phase is reduced, resulting in a refined hBN growth; and 3) a further increase in substrate temperature, i.e., a further increase of energy during film formation, disorders the well ordered sp^2-bonded structure to possibly include some small fraction of sp^3 bonding.

Confirmation of these substrate temperature effects on BN phase formation is given in the FTIR spectra of **Fig. 4** for three of the samples displayed in **Fig. 3**. The individual curves of **Fig. 4** are offset on the intensity axis to accentuate the Raman spectral features. Peak positions for hBN are indicated on the lower wavenumber axis, and cBN peak positions on the top axis. For a 50 °C substrate temperature, diffuse peaks appear above 1350 and 1550 cm^{-1}, corresponding to a disordered hBN structure, as the turbostratic phase (tBN) or an amorphous structure.[1,2] For a 200 °C deposition, a well ordered structure is formed, with primary and secondary peaks matching the simulated IR spectrum[2] of hBN. For a 600 °C deposition, the IR spectrum shows diminished intensity for the hBN peaks, but includes additional broad peaks about 1100 cm^{-1}, possibly indicative of sp^3 bonding. These deposition related trends in chemical bonding are clearly qualitative. A quantitative assessment of chemical bonding is beyond the scope of this initial study. Further TEM characterization for crystalline structure is not yet warranted.

The application of substrate bias is to enhance stabilization of cBN growth. Given the perspective of applying substrate temperature and varying the sputter gas parameters, the deposition conditions

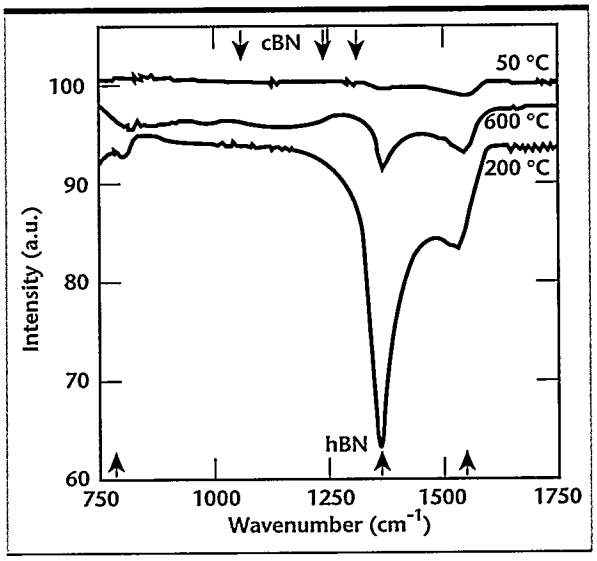


Figure 4. Fourier transform infrared spectra of BN films. Variation with substrate temperature confirms the Raman results of Fig. 3.

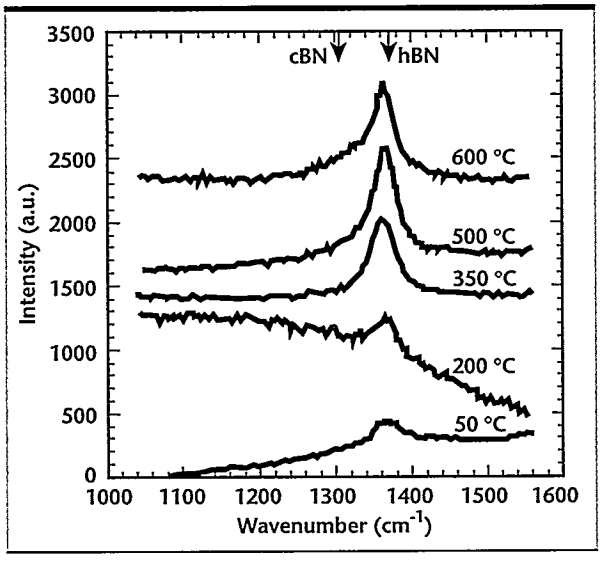


Figure 3. Raman spectra of BN films. Variation with substrate temperature indicates changes in chemical bonding.

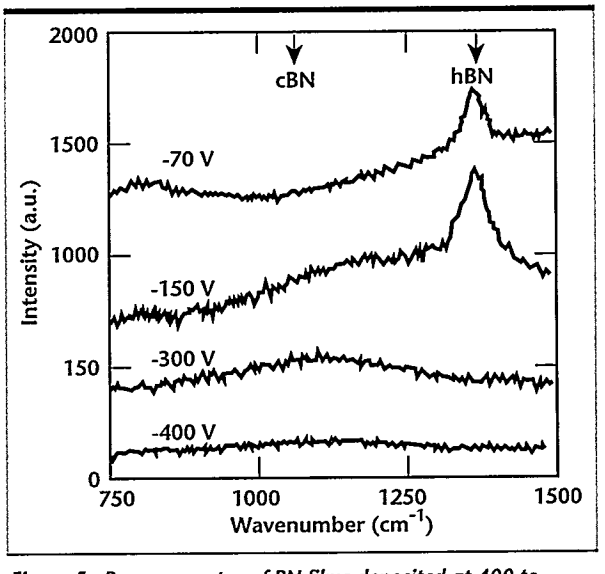


Figure 5. Raman spectra of BN films deposited at 400 to 500 °C. Variation with applied substrate bias indicates changes in the chemical bonding.

established for applying substrate bias are as follows: 1) vary sputter pressure using an Ar-25% N_2 gas mixture at a nominal flow of 25 to 30 cc min^{-1}; and 2) heat the substrate to 400 to 500 °C to promote crystalline formation, yet destabilize the growth of hBN. The effects of substrate bias on BN structure are shown in the Raman spectra of **Fig. 5** for 150 to 200-nm-thick films deposited using a 1.1 Pa sputter gas pressure, noting that individual curves are offset on the intensity scale to accentuate bonding features.

For a –70-V bias, a structure with some characteristic sp^2 bonding is formed in the BN deposit. As the bias is increased to –150 V, a diffuse peak appears at about 1100 cm^{-1}, characteristic of sp^3 bonding as seen in the TO mode for cBN. As the bias is increased further to –300 V, the characteristic sp^2 bonding at 1367 cm^{-1} nearly vanishes while the broad peak at 1100 cm^{-1} increases in intensity. Further increase in the bias to –400V begins to reduce the intensity of the broad peak at 1100 cm^{-1}. The application of substrate bias is to destabilize sp^2 bonding as found in hBN and promote sp^3 bonding as found in cBN. Further microstructural characterization of these bias-applied BN structures is warranted with TEM to confirm the change in crystalline structure.

To facilitate TEM plan view examination, samples deposited without a Cr adhesion layer are used, since these films are removable from the substrates. The bright field (BF) and dark field (DF) images are shown in **Fig. 6** for a BN film deposited on a "room" temperature Si substrate with a –70-V bias at 1.1 Pa Ar-40%N_2 gas pressure. The BF image reveals a fine-grained, continuous film with no distinguishing features. An electron diffraction ring pattern (**Fig. 6** insert) indicates a polycrystalline structure. The interplanar d-spacings are computed from the camera constant as calibrated to a Au standard. The d-spacings are 0.350 ± 0.005, 0.2101 ± 0.0008, and 0.1205 ± 0.0002 nm. The DF image generated from the broad, intense inner ring of the diffraction pattern shows a continuous, fine grained diffracting volume. The structure is probably single phase and nanocrystalline, as the dark field illumination of the diffracting 2 to 5 nm sized grains can be easily translated with small sample tilting. The d-spacings for this sample best fit the turbostratic boron nitride (tBN) phase with previously reported d-spacings of 0.356 and 0.212 nm.[16]

It should be noted that the d_{cBN}(111)of 0.2088 nm or the d_{hBN}(10.0/10.1) of 0.217/0.206 nm and the d_{hBN}(11.0/11.2) of 0.125/0.117 nm are not probable reflections.[17,18] The tBN characterization is quite consistent with the results for bonding trends found in the Raman spectra of **Figs. 3** and **5**. The BN film deposited at low temperature and low bias appears to be a disordered sp^2-bonded structure.

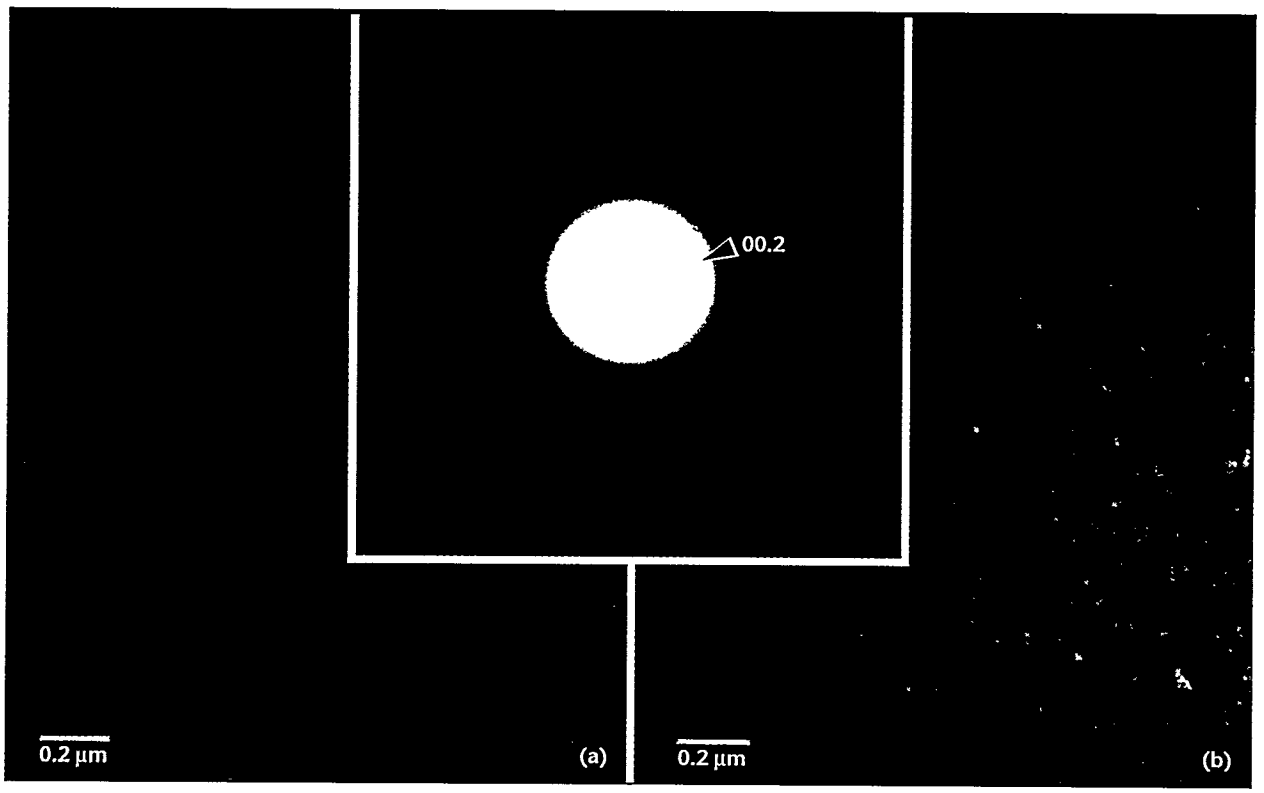


Figure 6. Boron nitride film deposited at room temperature with a –70-V substrate bias. Transmission electron microscopy reveals (a) the microstructure in bright field, and (b) the diffracting grains in dark field, which contribute to the most intense inner ring (insert) of the electron diffraction pattern.

Raman spectra have indicated the presence of sp^3 bonding, possibly for cBN in the –300-V deposit of **Fig. 5**. Since the tBN structure is found to be a consistent identification for the Raman and TEM analysis, verification of a cBN phase is now pursued using TEM. The BF and DF images are shown in **Fig. 7** for a BN film deposited on a 500 °C substrate with a –300-V bias at 1.1 Pa Ar-25%N_2 gas pressure. The BF image reveals a continuous film structure containing a mosaic appearance, attributable to some extent of film curvature. The electron diffraction ring pattern (**Fig. 7** insert) indicates a polycrystalline deposit. The DF image generated from the narrow, most intense inner ring of the diffraction pattern shows a uniformly dispersed, fine grained diffracting volume. Dark field illumination of the diffracting 2- to 5-nm nanocrystals can also be easily translated with small sample tilting.

From the innermost ring outward, the computed d-spacings are 0.2096 ± 0.0005, 0.1283 ± 0.0002, 0.1093 ± 0.0001, 0.0907 ± 0.0001, 0.0830 ± 0.0001, and 0.0741 ± 0.0001 nm. These d-spacings can be indexed to (111), (220), (311), (400), (331) and (422), which fit a cubic structure with a lattice parameter of 0.3625 ± 0.0003 nm. The cubic structure and lattice parameter value match the cBN powder diffraction value of 0.3616 nm. The increased lattice parameter may be a result of the nitrogen-rich (53 at.%) composition in this sample.

The use of optical microscopy further reveals the effect of intrinsic stress on the adhesion of the BN films to the substrate. The formation of cBN has been shown for films deposited on a substrate heated to 400 to 500 °C with a –300-V applied bias. The effect of sputter gas pressure further effects the intrinsic film stress. BN films deposited using Ar-25%N_2 sputter gas pressures of 1.1 and 2.1 Pa are shown in the optical photos of **Fig. 8**. The film deposited at 1.1 Pa is releasing, as seen in the crazed pattern, whereas the 2.1 Pa deposit is adherent, showing no such pattern. The presence of cBN in each of these films is verified using Raman spectroscopy, producing spectra similar to the –300-V bias curve of **Fig. 5**.

Discussion and Summary

It is established that the deposition of a cBN film from a hBN target requires an energetic process.[1–4] For example, only tBN and hBN growth is found for pulsed laser ablation of B or hBN targets onto substrates at elevated (600 °C) temperature without the energetic component of a coincident ion

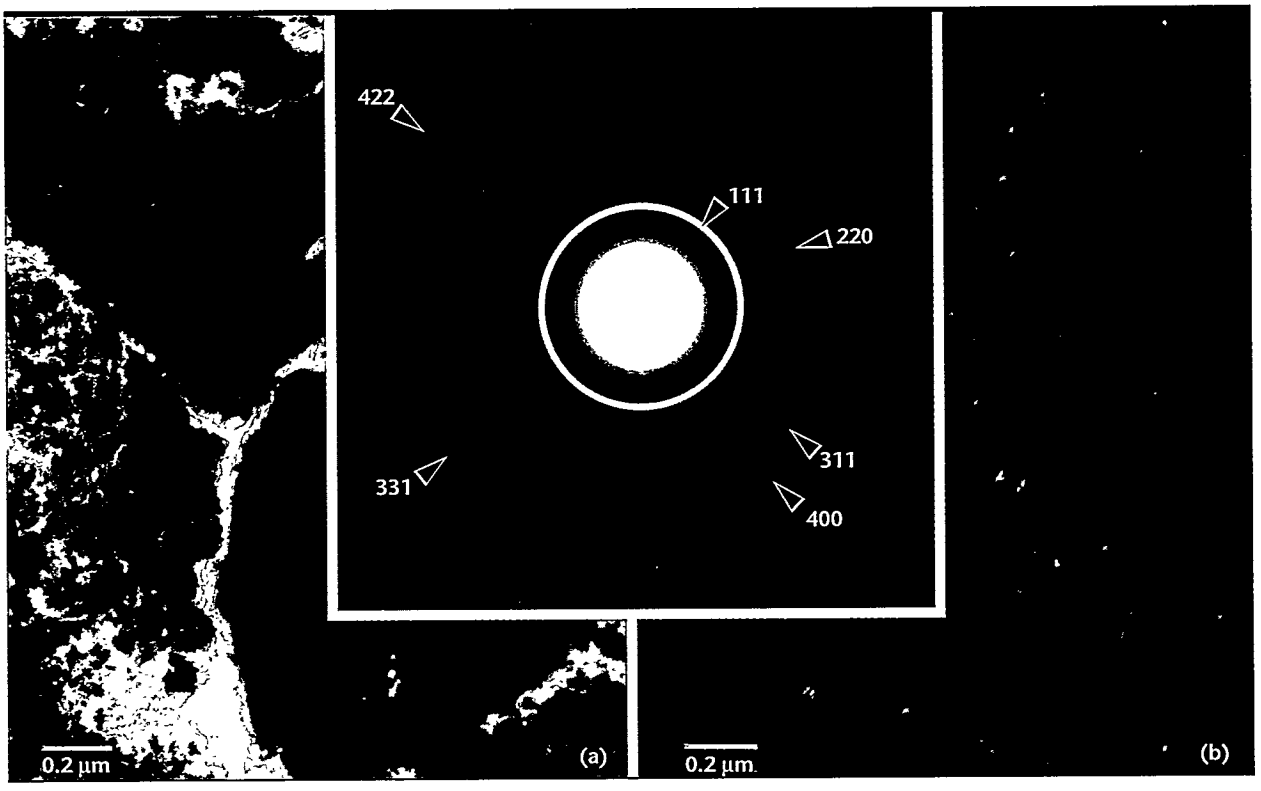


Figure 7. BN film deposited at 500 °C with a –300-V substrate bias. Transmission electron microscopy reveals (a) the microstructure in bright field, and (b) the diffracting grains in dark field, which contribute to the most intense inner ring (insert) of the electron diffraction pattern.

irradiation of the substrate.[19] The apparent stabilization of the cBN phase in the growing film by increasing substrate bias (when sputtering a boron target with Ar-N_2) is a clear indication that an energetic process is needed in comparison to growth of hBN. Although substrate potentials approaching the target discharge voltage are reached, the current density of in-coming Ar^+ ions is clearly much less than that required to noticeably resputter the growing film, due to the applied power distribution and resulting plasma confinement at the planar magnetron target surface. Although unable to measure flux, we provide a quantitative correlation to energy.

The Raman and FTIR spectra are representative of the BN films, and not the substrate or epilayers. Raman and FTIR characterization of the epilayer/substrate combinations did not produce any of the spectral features observed in **Figs. 3–5**. An assessment of enhanced cBN growth using a metal (such as Ni) epilayer will require higher substrate temperatures and the use of Ni single crystals. Results reported for cBN growth using hot filament assisted, rf plasma chemical vapor deposition required substrate temperatures of 900 to 1100 °C.[12,13] The Si substrates of this study are not suitable as silicide formation, and unknown artifacts result above 600 °C.

In summary, thin-film fabrication of more than 80 BN samples has been accomplished by reactive sputter-deposition from a pure, fully dense boron target. The near stoichiometric BN film composition is found to be relatively insensitive to gas composition (25 to 55% N_2) and substrate temperature (200 to 600 °C) over the range of deposition rates (<0.04 nm s^{-1}) studied. This is a consequence of a near equilibrium kinetic process between nitrogen and sputtered boron for the range of low rate deposition parameters examined. The control of substrate temperature and bias allow for the stabilization of BN films with sp^2 and sp^3 bonding. Although not all combinations have been explored, the following trends are confirmed for nearly stoichiometric films with a N:B ratio of 1.2, as characterized for chemical bonding using Raman spectroscopy and crystallinity using TEM: low temperature (<200 °C) and low negative bias (<150 V) produce films with structures characteristic of amorphous or turbostratic BN; intermediate substrate temperatures (200 to 350 °C) stabilize the growth of films with structures characteristic of hexagonal BN; and the application of intermediate negative bias (150 to 400 V) for substrates heated to 400 to 500 °C stabilizes the formation of films with structures characteristic of cubic BN.

Ion-assisted processes have been successful in producing cBN films, but typically these coatings are highly stressed and poorly adherent. An advantage for using pure boron targets in the reactive sputter deposition of cBN is the tractable nature of

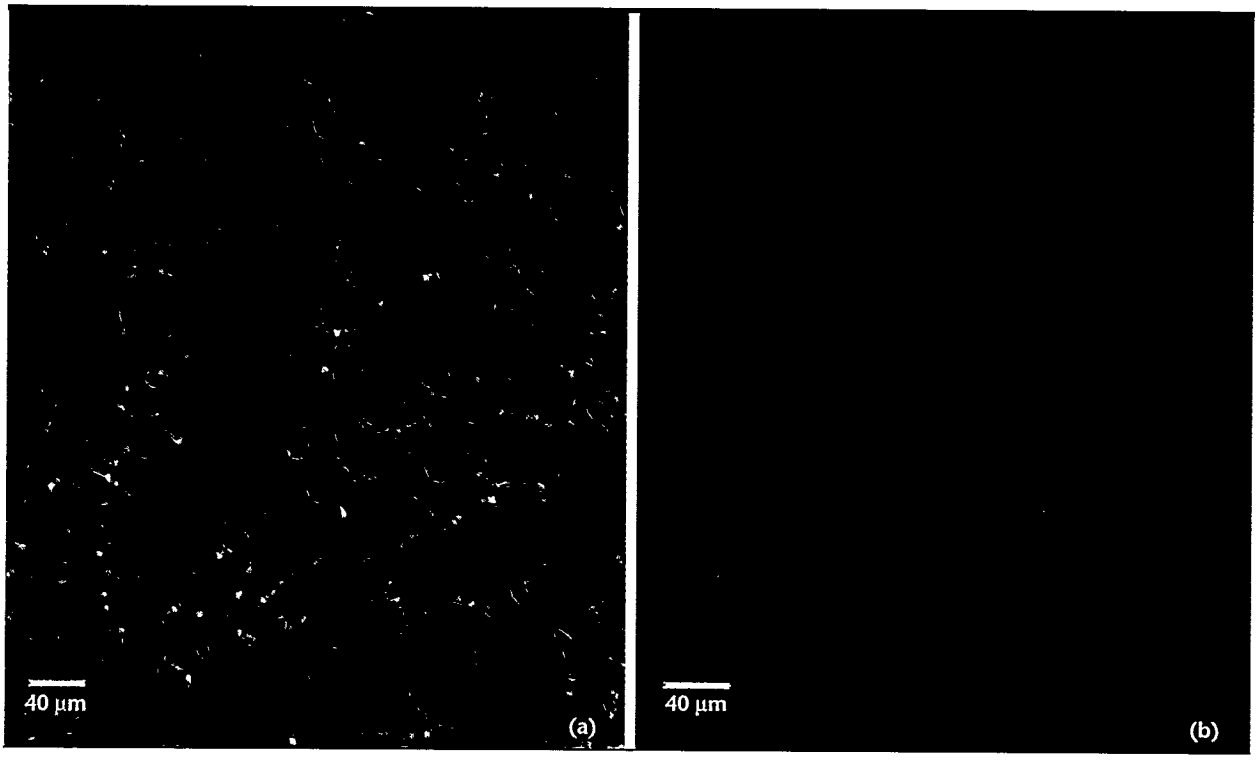


Figure 8. Optical micrographs of BN films deposited at (a) 1.1 Pa and (b) 2.1 Pa Ar-25%N_2 gas pressure.

producing adherent films. The use of sputter gas pressure as a deposition parameter that stabilizes the growth of adherent cBN coatings on flat Si substrates is, therefore, promising for cutting tool applications.

Acknowledgments

The authors acknowledge funding from DOE's Basic Energy Sciences Synthesis and Processing Center for Surface Hardness. We also thank M. Wall of LLNL for the TEM images, and K. McCarty of SNL for the FTIR measurements.

References

1. Kester, D., K. Alley, D. Lichtenwainer, and R. Davis (1994), "Growth and characterization of cubic boron nitride thin films," *J. Vac. Sci. Technol.* **A12** , p. 3074.

2. Friedmann, T., P. Mirkarimi, D. Medlin, K. McCarty, E. Klaus, D. Boehme, H. Johnsen, M. Mills, D. Ottesen, and J. Barbour (1994), "Ion-assisted pulsed laser deposition of cubic boron nitride films," *J. Appl. Phys.* **76**, p. 3088.

3. Wada, T., and N. Yamashita (1992), "Formation of cBN films by ion beam assisted deposition," *J. Vac. Sci. Technol.* **A10**, p. 515.

4. Tanabe, N., T. Hayashi, and M. Iwaki (1992), "Deposition of cubic boron nitride thin films by ion-beam-enhanced deposition," *Diamond Relat. Mater.* **1**, p. 883.

5. Reinke, S., M. Kuhr, and W. Kulisch (1994), "Mechanisms in ion induced c-BN growth," *Diamond Relat. Mater.* **3**, p. 341.

6. McKenzie, D., W. McFall, W. Sainty, C. Davis, and R. Collins (1993), "Compressive stress induced formation of cubic born nitride," *Diamond Relat. Mater.* **2**, p. 970.

7. Makowiecki, M., A. Jankowski, M. McKernan, and R. Foreman (1990), "Magnetron sputtered boron films and Ti/B multilayer structures," *J. Vac. Sci. Technol.* **A11**, p. 3910.

8. McKernan, M., D. Makowiecki, P. Ramsey, and A. Jankowski (1991), "Magnetron sputter deposition of boron and boron carbide," *Surf. Coatings Technol.* **49**, p. 411.

9. Thornton, J. (1982), *Deposition Technologies for Films and Coatings*, R. Bunshah, Ed., Noyes Publ., Park Ridge, N.J., p. 187.

10. Wagman, D., W. Evans, V. Parker, R. Schumm, S. Bailey, I. Halow, K. Churney, and R. Nuttall (1985), *CRC Handbook of Chemistry and Physics, 65th ed.*, R. Weast, M. Astle, and W. Beyer, Eds., Chemical Rubber, Boca Raton, Fla., p. D-55.

11. Makowiecki, D., and M. McKernan (1995), *Fabrication of Boron Sputter Targets*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif., U.S. Patent No. 5,392,981.

12. Zhang, F., Y. Guo, Z. Song, and G. Chen (1994), "Deposition of high quality cubic boron nitride films on nickel substrates," *Appl. Phys. Lett.* **65**, p. 971.

13. Song, Z., F. Zhang, Y. Guo, and G. Chen (1994), "Textured growth of cubic boron nitride film on nickel substrates," *Appl. Phys. Lett.* **65**, p. 2669.

14. Huong, P. (1991), "Structural Studies of diamond films and ultrahard materials by Raman and micro-Raman spectroscopies," *Diamond Relat. Mater.* **1**, p. 33.

15. Geick, R., C. Penny, and G. Rupprecht (1966), "Normal Modes in Hexagonal Boron Nitride," *Phys. Rev.* **146**, p. 543.

16. Thomas Jr., J., N. Weston, and T. O'Connor (1963), "Turbostratic Boron Nitride, Thermal Transformation to Ordered-layer-lattice Boron Nitride," *J. Am. Chem Soc.* **84**, p. 4619.

17. JCPDS-ICDD (1992), cBN File, No. 25-1033.

18. JCPDS-ICDD (1992), hBN File, No. 34-0421.

19. Friedmann, T., K. McCarty, E. Klaus, J. Barbour, W. Clift, H. Johnsen, D. Medlin, M. Mills, and D. Ottesen (1994), "Pulsed laser deposition of BN onto silicon (100) substrates at 600°C," *Thin Solid Films* **237**, p. 48.

Donald R. Lesuer, Thrust Area Leader

The program in Materials Science and Engineering has [illegible] areas: [illegible] understanding of the mechanical behavior [illegible] processing, structure and properties of structural materials that are of interest to the Livermore National Laboratory (LLNL) programs and future funded programs. Many of these programs involve collaboration with U.S. defense and other government agencies. We work in 1) composite materials including metal matrix and polymer matrix composites, 2) [illegible] modeling in which LLNL-developed computer codes are used to [illegible] processing as well as [illegible] structural response of materials.

Our activities [illegible] are focused in both general areas. Projects provide the following reports: 1) Strength and Fracture Toughness of Materials, 2) Dynamic Mechanical Behavior of [illegible], 3) Participating Media Radiative Transport Program, 4) Materials Issues in Investment Casting Molds, 4) Synthesis and Processing of Nanocrystalline Aluminum Hydrocarbons, 5) Fatigue Behavior of Metal Matrix Composites at Elevated Temperatures, and 6) Pearlite in Ultrahigh Carbon Steels. Heat Treatments and Mechanical Properties.

Materials Science and Engineering

5

5. Materials Science and Engineering

Overview
Donald R. Lesuer, Thrust Area Leader

Strength and Fracture Toughness of Material Interfaces
Robert Riddle, David S. Hiromoto, Nhan Nguyen, Wayne King, Geoffrey Campbell, and Walter Wien 5-1

Dynamic Mechanical Behavior of Fiber Composites
Steve J. DeTeresa and Scott E. Groves 5-7

Participating Media Radiant Transport at High Temperature: NIF Flashlamps and Investment Casting Molds
Mark A. Havstad and Charles W. Dingus 5-15

Synthesis and Processing of Nanocrystalline Aluminum and Hydroxyapatite
T. G. Nieh and Donald R. Lesuer 5-21

Fatigue Behavior of Al Metal Matrix Composite at Elevated Temperatures
Chol K. Syn, Donald R. Lesuer, and David S. Hiromoto 5-27

Pearlite in Ultrahigh Carbon Steels: Heat Treatments and Mechanical Properties
Eric M. Taleff, Chol K. Syn, Donald R. Lesuer, and Oleg D. Sherby 5-35

Strength and Fracture Toughness of Material Interfaces

Robert Riddle, David S. Hiromoto, and Nhan Nguyen
Manufacturing and Materials Engineering Division
Mechanical Engineering Department

Wayne King, Geoffrey Campbell, and Walter Wien
Condensed Matter Division
Chemistry and Materials Science Department

Dissimilar materials are often combined to obtain composite properties of strength, fracture toughness, stiffness and creep resistance that neither material alone could provide. The strength and fracture toughness of the material interfaces resulting from the bond of dissimilar materials often determines whether the desired material properties are attainable. This report describes progress in producing experimental and computational tools and methods to fabricate, measure, and predict the strength and fracture toughness of material interfaces.

Introduction

A major application for the combining of dissimilar materials to obtain superior composite material properties is in metal matrix composites (MMC's). Here the interface strength and interface fracture toughness between the embedded ceramic particles and the ductile metal phase play a major role in determining the macroscopic strength and fracture toughness properties of the MMC. The ductile phase plasticity plays an important part in increasing the fracture toughness of MMC's, while the ceramic particles increase the stiffness and strength. The MMC of particular study here consists of an aluminum alloy reinforced with Al_2O_3 particles. We studied the influence of the metal/ceramic interfaces within the MMC by experimentally measuring the interface strength, plastic deformation, and fracture toughness of specimens prepared by ultrahigh vacuum diffusion bonding. We also characterized the deformation process in the ductile interlayer of the specimen by x-ray tomography.

A major thrust of the transportation industries in the U.S. is the incorporation of advanced structural materials in airplanes and automobiles. These advanced materials include MMC's, where particulate or whisker reinforcements stiffen and strengthen a ductile phase matrix. The composite materials offer the potential of considerable weight reduction, leading to increased fuel efficiency and cost savings.

Lawrence Livermore National Laboratory (LLNL) has CRADA's with automobile and aerospace manufacturers, where the fabrication, measurement, and prediction of metal matrix component strengths are important.

Many other LLNL programs are concerned with the strength and fracture toughness of interface strengths in components that they design, fabricate and test. Examples of such programs include weapons development, micro-and opto-electronics systems packaging, and laser optical coatings.

In all these potential applications of advanced MMC materials, there exists a defined need for structural material properties suitable for implementation in finite element computer codes to predict the component stiffness, and tensile, fatigue and fracture strengths, as a function of applied loads and thermal cycles. These structural material properties must describe macro-structural properties in terms of the micro-structural constituents and processing history, to take advantage of our ability to tailor the material properties to the particular application.

Progress

Characterization of Ductile Rupture Using X-Ray Tomographic Microscopy

Mechanical tests have been performed on samples fabricated from cylinders joined in the LLNL ultrahigh vacuum diffusion bonding machine. The material system of pure aluminum bonded to sapphire was chosen for its low x-ray absorption, chemical compatibility, and relevance to potential MMC materials. The use of sapphire in the sample avoids the problems of grain pull-out and porosity present in the use of polycrystalline alumina[1]. The bonded samples were cut into bend beams with nominal dimensions of 3 × 3 × 40 mm. The bonded samples and resulting bend beams are shown in **Fig. 1**.[2] The sandwich configuration samples were fractured in four-point bend, as shown in **Fig. 2**.

Some samples were notched to a depth of 30% of the sample height, to move the failure region away from the edge of the ceramic, and to increase the stress concentration in the ductile aluminum foil.

Notched samples were examined by x-ray tomographic microscopy (XTM) to look for any defects which might exist prior to sample loading.

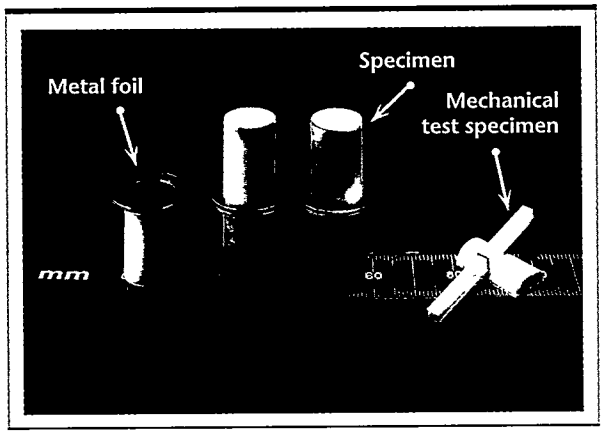


Figure 1. Bonded samples, cut into bend beams with nominal dimensions 3 × 3 × 40 mm.

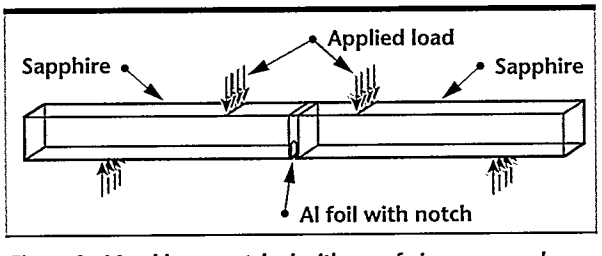


Figure 2. Metal layer, notched with a wafering saw, producing a notch root diameter of only 40 μm.

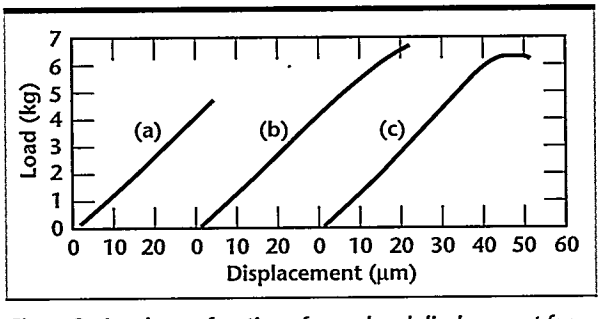


Figure 3. Load, as a function of cross-head displacement for (a) first loading; (b) second loading; and (c) third loading.

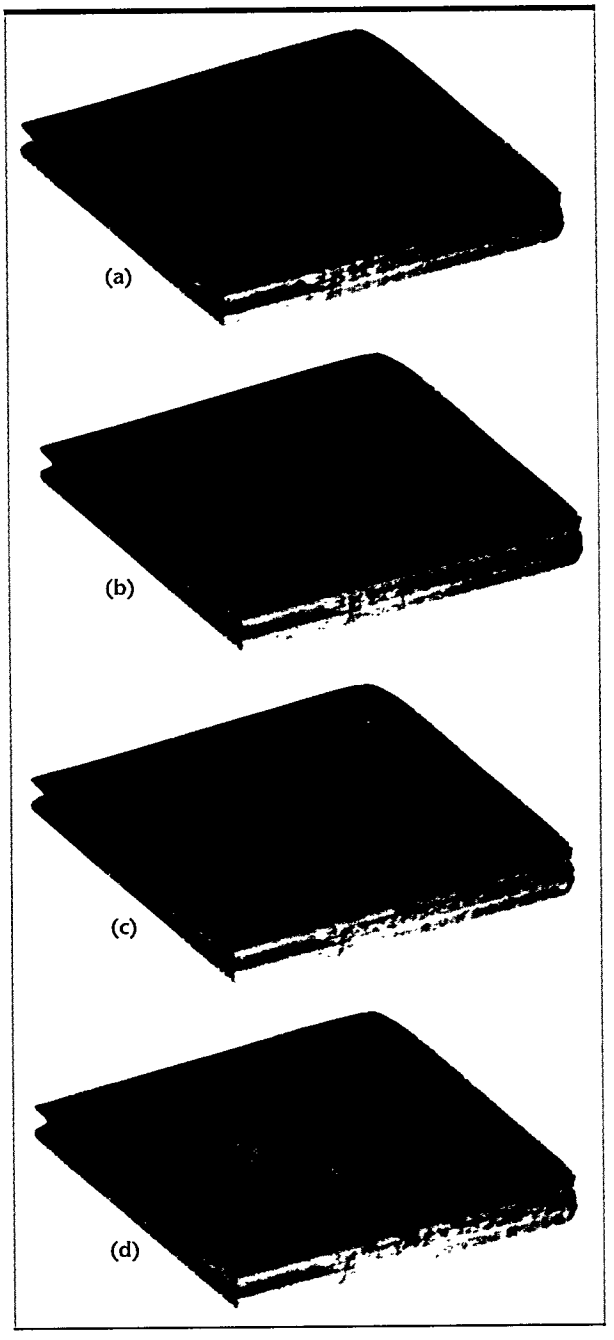


Figure 4. Rendering of volume (a) undeformed; (b) after first loading; (c) after second loading; and (d) after third loading. Voids are shown in shades of gray.

In performing the XTM scans, the x rays from a synchrotron source (Beamline 4-2, at the Stanford Synchrotron Radiation Laboratory) were passed through a monochrometer to select 25 keV photons. After passing through the specimen, the x rays strike a $CdWO_3$ scintillator. The light emitted from the scintillator passes through magnifying optics and is collected by a charge-coupled device (CCD) camera at a magnification sufficient to give the volume elements (voxels) in the reconstructed image the size 5.3 μm on a side. No resolvable voids were detected in the as-notched samples.

The beam was then loaded as shown in **Fig. 2** until the first sign of non-linearity in the load vs displacement plot was observed. In this case, the displacement was measured at the cross-head of the hydraulic testing machine. The mechanical test was interrupted, the specimen removed, and an XTM scan performed. The mechanical test was subsequently resumed to strain the beam further. This iterative sequence was continued for as long as was practical.

Figure 3 shows plots of the load vs cross-head displacement for three successive loadings of the notched bend specimens. **Figure 4** shows the rendering of the voids according to the XTM and post-processing graphics in the thin aluminum foils, as a result of plastic deformation.

The succession of void formation demonstrates that at the onset of plasticity, lenticular voids form at the metal/ceramic interface ahead of the notch tip. The voids grow in number and size as the deformation is forced by the application of load cycles. The small debond regions play an important role in the maximum attained interface strength, and dominate the failure process as crack growth is suppressed. As the voids begin to link up, as seen in the final loading step, final failure occurs.

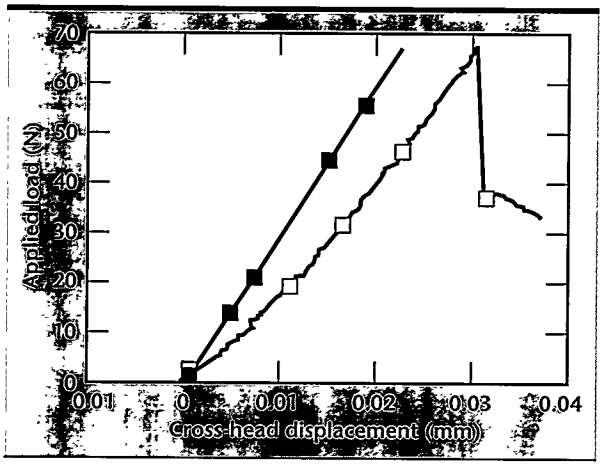


Figure 5. Comparison of measured and predicted cross-head displacements in the four-point bend bar interface fracture specimens.

Measurement of the Ductile Behavior of Constrained Aluminum Thin Foil

Consideration of the plots of load vs cross-head for the bend bar specimens reveals the difficulty of resolving the elastic/plastic constitutive behavior of such plots. In these plots, the compliance and plastic deformation of the 2-or 3-μm aluminum foils is added to the compliance of the 28 mm of alumina beam between the lower supports. **Figure 5** shows a plot of load vs cross-head displacement, as measured for a sandwich beam bend sample, and as predicted from beam theory for an alumina beam. The agreement between the predicted and the measured response helps validate the measurement technique, and shows that the compliance of the foil adds significantly to the overall compliance, even though it is only a small fraction (7×10^{-5}) of the total length.

In order to measure the elastic/plastic response of the as-bonded aluminum foil, a microextensometer system was developed to measure the opening of the notch. Conventional or commercial extensometers typically have too large a gauge length, and too much spring force to be applicable for measurement of the extension of the aluminum foil under the applied bending loads.

Figure 6 is a schematic of the microextensometer system that was developed. For the measurement system, two straight steel rods are aligned in a fixture and attached to the bottom side of the bend sample. A laser projects a beam of light through the rods, and illuminates a region of a CCD camera. The region of illumination corresponds to the distance between the rods. The laser and CCD camera are mounted on a precision movable stage, so that as the specimen is loaded and the central section

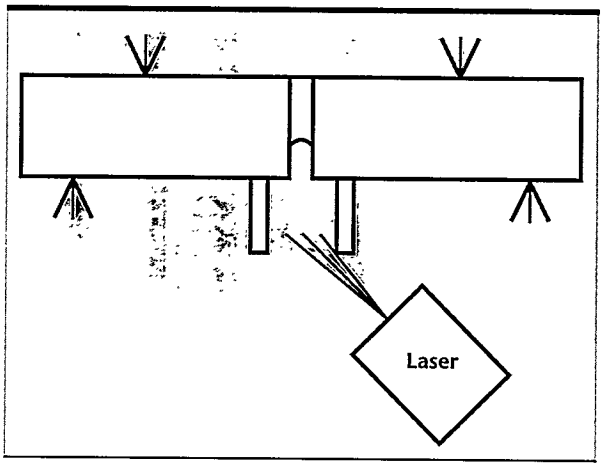


Figure 6. Schematic of microextensometer system.

moves downward, the same points on the rod are used to measure the notch opening, using the bottom of the specimen as the reference plane. The data taken by this system is logged into a computer for retrieval and reduction.

Figure 7 shows that the onset of plastic deformation occurs at much lower loads than would be predicted from a plot of load vs cross-head displacement. Here, judging from the nonlinearity of the deformation, plastic deformation of the foil has occurred at a load of 20 N. Large plastic deformation has occurred in the foil at the maximum load. The opening displacement of 30 μm is on the same order as the foil thickness (50 μm). For comparison, a plot of the corresponding notch opening, as predicted from a finite element analysis, is shown as the solid circles in **Fig. 7**. The constitutive material properties for the alumina and aluminum are given in **Table 1**.

The excellent agreement between the finite element predictions and the measured results at low loads, where the system response is nearly elastic, once again validates both the measurement technique and the analysis. There is significant divergence in the plastic response. This is explained by the fact that the aluminum plastic material behavior used in the finite element analysis is taken from sample data, where the gauge length of the deformation is on the order of tens of millimeters. In this size region, the plastic response is nearly independent of gauge length. The plastic response of thin foils is known to depend on foil thickness[3], and here the plastic response of a 50-μm thin film shows much larger strains for a given load than would be predicted using the large gauge length aluminum behavior.

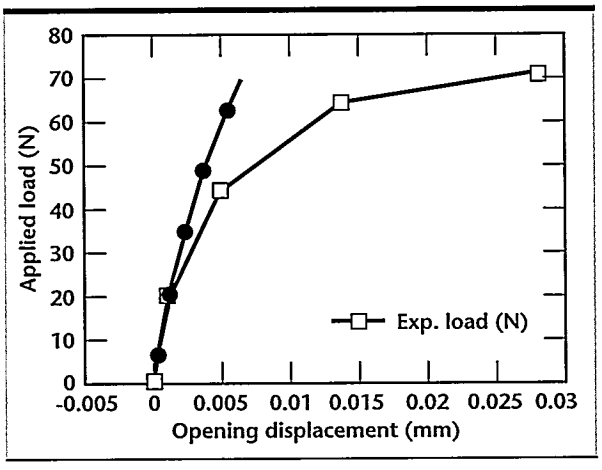


Figure 7. Comparison of finite element results vs experimental results.

A number of similar specimens of different foil thicknesses have been tested, and the plastic response of the larger group of specimens is yet to be quantified, both with respect to foil thickness and variability within a group of specimens with the same foil thickness.

However, in the preliminary data analysis, the maximum stress at failure has been correlated with foil thickness. For both the notched and unnotched specimens, the maximum stress at failure is calculated from the maximum load, P, using the equation:

$$\sigma_{max} = \left(\frac{3}{2}\right) \bullet P \bullet \frac{(l_b - l_t)}{b \bullet h^2}$$

where l_b and l_t are the lengths of the upper and lower spans, b is the thickness, and h the height of the rectangular specimen cross-section. **Figure 8** shows a plot of maximum stress at failure as compared to foil thickness for unnotched samples. For the 50-μm thick foil the average strength is 150 MPa, with a standard deviation of 57 MPa. For the 75-μm thick foil the corresponding numbers are 220 MPa and 38 MPa. It is seen that the average stress at failure increases with increased thickness, and that the data scatter is larger in the 50 μm foil thickness.

Table 1. Aluminum and alumina material properties.

Material name	Constitutive model	Modulus of elasticity (MPa)	Poisson's ratio	Values of effective stress (MPa)	Values of effective plastic strain
Aluminum	Isotropic elastic/plastic	68.94×10^3	0.33	3.5, 12.2, 16.5, 24.0, 34.4, 48.5, 68.5, 97.4	0.0, 0.002, 0.005, 0.015, 0.045, 0.125, 0.35, 1.0
Alumina	Elastic	345.0×10^3	0.21		

The stress at failure for the notched specimens is within the scatter band of the results for the unnotched samples in the 50-μm foil. **Figure 9** shows the plot for notched samples of both foil thicknesses. That removal of 30% of the area of the bond does not significantly decrease the overall

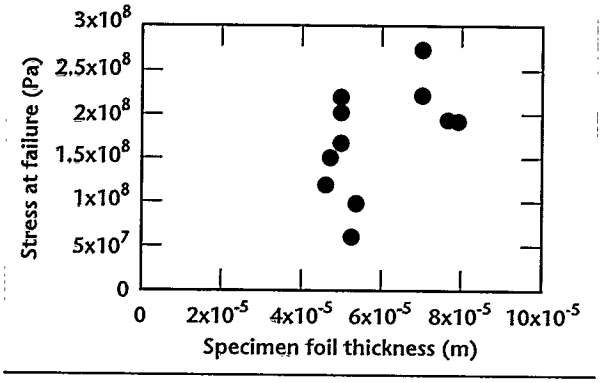


Figure 8. Interface four-point bend unnotched specimens.

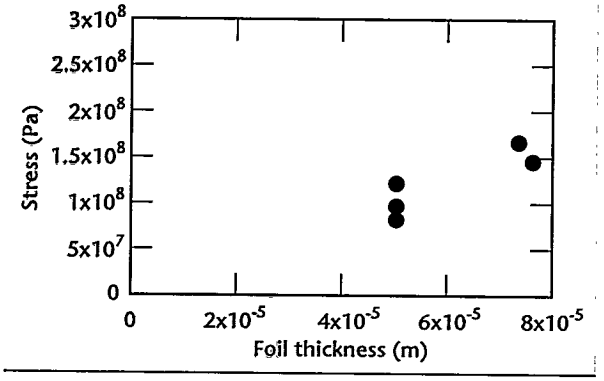


Figure 9. Interface four-point bend notched specimens.

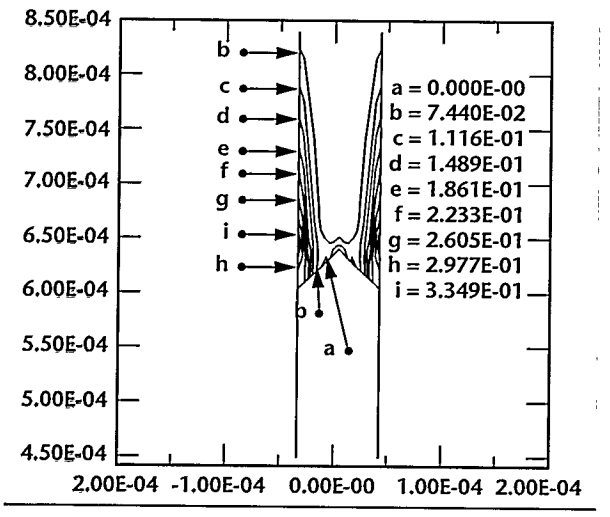


Figure 10. Contours of effective plastic strain, from finite element analysis of the notched region. A maximum is reached at the metal/ceramic interface, within about 25 μm of the notch tip.

interface strength with the thinner bonding foil has at least two possible explanations. One is that the ductile deformation and void growth of the foil are as significant as the metal/ceramic interface strength in determining overall sample strength. Another is that the metal foil is removed from the bottom side, in the region where the ceramic is most likely to be damaged, so that removing this bond area has a proportionately smaller decrease in overall bond strength. In the 75-μm thick foil bond, the reduced strength from the notch is proportional to the reduced area.

Finite Element Analysis of Ductile Behavior in Constrained Aluminum Thin Foil

The finite element analysis of the failure process in the sandwich bend bar experiments is used to more fully understand the interplay between the effective plastic strain and the hydrostatic stress. The effective plastic strain-to-failure in a ductile material is known to depend on the hydrostatic stress.[4] The precise relationship between the effective plastic strain-to-failure and the hydrostatic stress varies from material to material, and other factors may complicate the relationship as well.

Finite element analysis of the notched region shows that contours of effective plastic strain reach a maximum at the metal/ceramic interface, within about 25 μm of the notch tip. These contours are shown in **Fig. 10**, at a stress level that corresponds in the specimen to a load of 70 N. In contrast, the maximum hydrostatic stress reaches a maximum

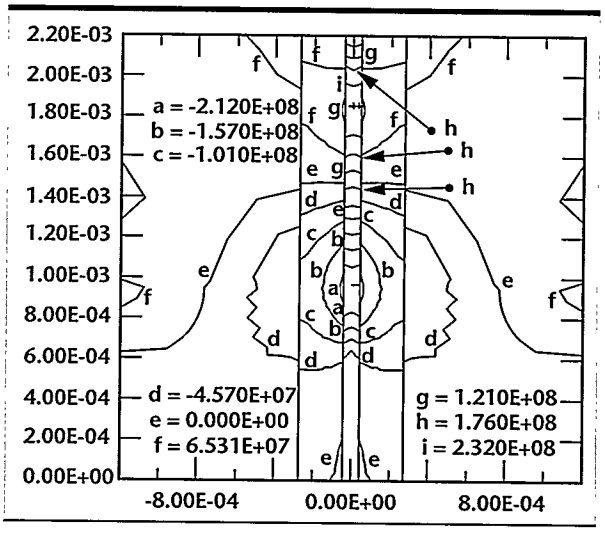


Figure 11. Contours showing maximum hydrostatic stress at about 350 μm in from notch tip.

about 350 μm in from the notch tip in a 50-μm foil, as shown in **Fig. 11**. From the XTM analysis of void growth, it is seen that the voids begin to initiate and coalesce at a distance from the notch tip which is more centered around the region of maximum hydrostatic stress.

Crack growth and void growth are competing failure mechanisms at notch tips in general. Here, the void growth in the ductile aluminum foil dominates.

Future Work

This failure process in the ductile aluminum foil is dominated by void growth in regions of large hydrostatic stress. The derivation or specification of the correct material parameter to describe this failure mode is being investigated. The energy release rate for void growth is a prime candidate in this regard.

Further development of this finite element model is planned, using a tie-break slideline to model the damage initiation and progression along the bondline. The effective plastic strain-to-failure for the slideline will be modified, based on the value of the hydrostatic stress in the element, as a first cut at a definitive failure criteria.

Reproducing the effect of notch depth and foil thickness on specimen strength will be the major focus of the finite element modeling.

In the experimental arena, the influence of the metal/ceramic interfaces within the MMC will be studied, using additional interface strength and toughness of specimen configurations. The bonds will be prepared in the ultrahigh vacuum diffusion bonding system, as well as by more conventional bonding methods. Specimens will also be prepared from commercially available MMC materials.

References

1. 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 Met*, (in press).

2. Campbell, G. H., R. A. Riddle, and W. E. King (1995), "Bonding of Aluminum Foils to Sapphire and the Effect of the Foil Thickness on Mechanical Behavior," *The American Ceramic Society, 97th meeting, Westerville, Ohio*, p. 177.

3. Nix, W. D. (1989), "Mechanical Properties of Thin Films," *Metallurgical Transactions*, Volume **20A**, November, 1989, pp. 2217–2245.

4. Mackenzie, A. C., J. W. Hancock, and D. K. Brown (1977), "On the Influence of State of Stress on Ductile Failure Initiation in High Strength Steels," *Engineering Fracture Mechanics*, Vol. **9**, pp. 167–188.

Dynamic Mechanical Behavior of Fiber Composites

Steve J. DeTeresa and Scott E. Groves
Manufacturing and Materials Engineering Division
Mechanical Engineering

We have determined the effects of deformation rate in the range of 10^{-4} to over 10^{3} s^{-1} on the compressive behavior of fiber composites. Rate is seen to have only a minor effect on the matrix-dominated shear modulus, and the trend is identical to that of the pure epoxy matrix. Fiber-direction compression strengths of laminates are only mildly effected by rate, showing increases of less than 10% with five orders of magnitude increase in rate. Matrix-dominated failure modes are more significantly affected by rate and the increases in strength are similar to those seen for the epoxy matrix. Through-thickness compression behavior is strongly dependent on fiber orientation and rate. Laminates exhibit through-thickness strengths in excess of 150 ksi, despite having no fiber oriented in the load-direction. Cross-ply laminates lose strength, while quasi-isotropic laminates exhibit some strength improvement at higher rates. The failure mechanism for through-thickness compression appears to be related to fiber tensile strength in multidirectional laminates.

Introduction

Advanced fiber composites are increasingly used in defense and commercial applications that subject the material to dynamic loads. Examples include munitions, such as high-pressure gun barrels, sabots for KE long-rod penetrators, artillery shells, and warhead penetrator cases. These materials are also used for personal security measures, such as bomb containment and body armor. Furthermore, composite structures in military aircraft, land vehicles and ships are subjected to shock loads by threats like anti-aircraft flak and mine blasts. In commercial vehicle markets, the effort to replace primary metal structures with lightweight fiber composites requires the ability of lighter structures to absorb impact energy during crash and to protect occupants from excessive deceleration. In most of these applications, the structures will experience strain rates within the range of 1 to 100 s^{-1}. The efficient use of the materials and the survivability of the structures requires a more complete knowledge of high-rate loading behavior than currently exists.

Previous work conducted on unidirectional carbon and glass fiber composites at intermediate and high (1 to 500 s^{-1}) rates have shown that fiber-direction tensile strength and stiffness are virtually unaffected by loading rate.[1,2] These observations suggest that the fiber is an elastic, rate-insensitive material.

In contrast, matrix-dominated tensile behavior has been shown to exhibit rate dependency in both stiffness and strength, but the trends are not clear. Because of the well-known rate-dependent properties of polymers, it is expected that the composite will also exhibit rate-dependence for the matrix-dominated loading directions. While this may be obvious for shear loads and normal loads applied transverse to the fiber, fiber-direction compression is also potentially coupled to matrix properties[3] and may exhibit more rate-dependence than the corresponding tensile loading condition.

Documented work in fiber-direction compression has been conducted up to Hopkinson bar rate (~1000 s^{-1}).[4–8] However, the failure modes in these tests may not be pure compression since nontraditional test methods were used, and quasi-static rate behavior measured using these techniques isn't always in agreement with data obtained from standard methods. Though the results of these studies all indicate that there are increases in compression strength with rate, the magnitudes of these increases are not consistent. Reported improvements in compression strength range from 9 to 42% at intermediate rates[4–7] and from 23 to over 170% at high rates[4,5,8]. These discrepancies are probably due to both material and test method contributions. The bulk of the high-rate compression work has been performed on laminates of the cross-ply and

quasi-isotropic families. Because these laminates contain a multiplicity of fiber directions, it is not certain that the rate effects are due to the fiber- or the matrix-dominated plies.

We have performed a systematic study of the in-plane compression behavior of several laminate subsets of a typical carbon fiber/epoxy system in order to determine the relative contribution to laminate behavior. In this study we have investigated both traditional and nontraditional specimen geometries and fixturing for in-plane compression tests. Many of the high-rate loading applications such as gun barrels, body armor and mine blast also involve high-rate compression through the thickness of the composite laminate. Little is known about this loading mode for any deformation rate. Consequently, we have also investigated the material response to out-of-plane compression as a function of rate, fiber orientation in the plane and material composition.

Progress

In-Plane Compression Study

Constitutive Behavior. Although moduli have previously been reported for intermediate and high strain-rate tests, there are normally stress wave reflections at initial sample impact with most methods. This "ringing" makes accurate moduli determination nearly impossible. To circumvent this problem, we conducted separate experiments to compare the rate-dependence of shear moduli exhibited by an epoxy and its composite with carbon fiber. Moduli were determined from oscillatory torsion tests of cylindrical specimens using a dynamic mechanical spectrometer. The time-temperature superposition principle was applied to data collected at several temperatures in the linear viscoelastic regime to obtain predictions of moduli over a wide frequency range.

The dynamic properties of the composite and epoxy exhibited the same dependence on temperature with a glass transition at 105 °C and a secondary beta transition at –87 °C. Values of complex shear moduli were collected over a frequency range of 0.01 to 100 rad/s at several temperatures in the vicinity of 20 °C. Master curves referenced to ambient temperature are shown in **Fig. 1** for the epoxy and unidirectional composite.

Composite moduli can be calculated from constituent properties, using the micromechanics solution for a rigid fiber and a viscoelastic matrix[9]

$$G_c = G^* \frac{(1+V_f)}{(1-V_f)} \qquad (1)$$

where G^* is the complex matrix shear modulus, and V_f is the volume fraction of fiber. Using the nominal fiber volume fraction of 62%, we find that the calculated composite shear moduli are indistinguishable from measured values. The corresponding shift factors for both the epoxy and composite are shown in **Fig. 2**, again indicating similar behavior over a wide range of deformation rates.

Clearly the *in-situ* viscoelastic properties of the resin in the composite are identical to those in the bulk state. More important, the rate-dependency of shear moduli over several orders of magnitude of deformation rate is identical, and only amounts to increases on the order of 1.3% per order of magnitude of strain rate. Larger increases would be

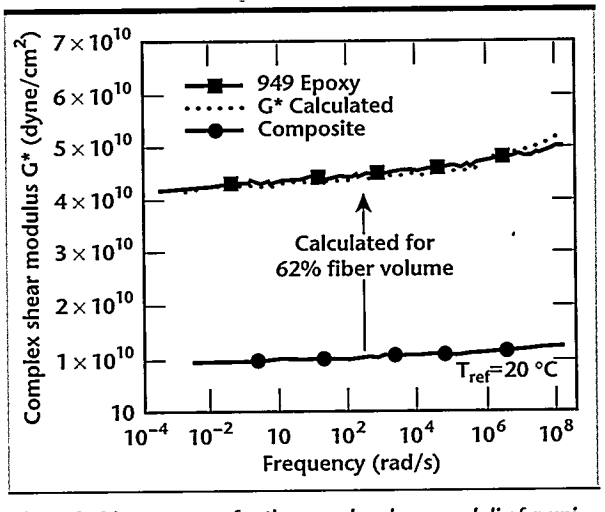


Figure 1. Master curves for the complex shear moduli of a unidirectional carbon fiber composite and the epoxy matrix material.

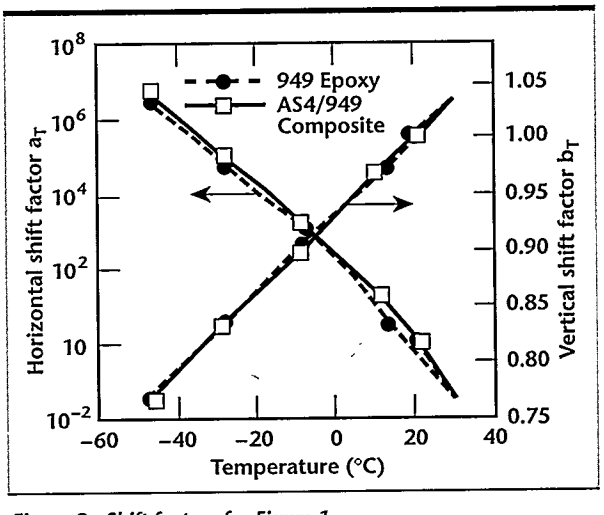


Figure 2. Shift factors for Figure 1.

expected in the vicinity of the transition temperatures; and for polymers having beta transitions close to ambient temperature, more significant increases in moduli with rate have been measured.[10] These findings suggest that the constitutive behavior of a composite may be adequately determined from the matrix properties. Thus, the battery of tests required to determine the rate-dependence of the composite moduli can be avoided. Simple rheological tests like those performed here for one material system can be used to check this correlation for others.

Previous work has indicated that there can be more significant changes in composite moduli with rate, and this might be attributed to material differences. However, many of those studies were performed using techniques which yielded distorted stress/strain curves due to stress wave effects, and therefore a large uncertainty in the calculated moduli. Perhaps the best supporting evidence for the mild dependence of moduli on rate found in our study is the moduli determined using acoustic techniques, where frequencies are on the order of 3×10^7 rad/s. This technique yields measured elastic constants that are not significantly greater than values determined from mechanical tests performed at quasi-static rates.[11]

Table 1. Designations and fiber orientations for in-plane compression tests.

Specimen designation	Fiber orientation
Unidirectional ("Uni")	0_{16}
Transverse	90_{16}
Shear	$[+45/-45]_{4S}$
Cross-ply	$[90/0]_{4S}$
Quasi-isotropic ("Quasi")	$[+45/90/-45/0]_{2S}$

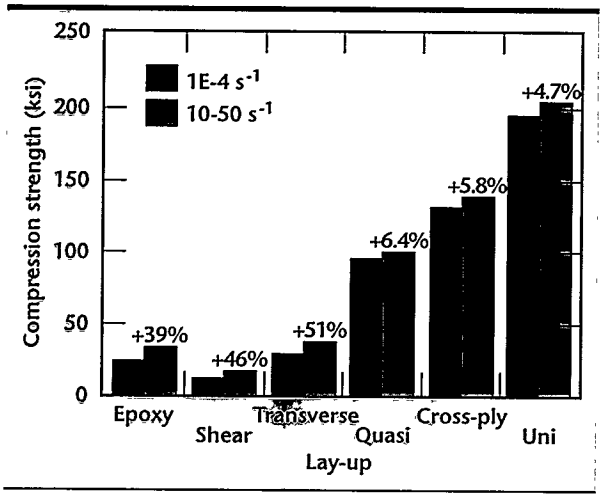


Figure 3. In-plane compression strengths at slow and intermediate rates for different lay-ups of a carbon fiber/epox composite.

Ultimate Properties. To avoid some of the problems and concerns normally associated with the compression testing of fiber composites, and to perform higher rate testing, we have made minor modifications to a traditional compression test method. Because of the sample geometry and fixturing, we are limited to strain rates on the order of 10 to 100 s^{-1}.

To generate sufficient energy to fail these specimens, we used an instrumented drop tower developed previously.[5] Consequently, the fixturing needed to be designed to stop the drop weight in a benign manner. The specimen has bonded fiberglass tabs to prevent end crushing and to allow some load transfer via shear forces. Clamping the tabs helps to delay premature adhesive failure. In our specimen, the gauge length is large enough to allow for bonded strain gauges which are needed to determine the effective strain rate. Precise alignment is maintained by fitting the fixtures in a cylindrical sub-press. Preliminary tests at quasi-static rates using this fixture and the related Boeing-modified ASTM D695 method yielded identical results, within experimental error (151 ksi for the new fixture, vs 143 ksi for the D695 fixture). We could not use the latter method, since it does not allow for strain gauging or for a simple way to stop the drop weight.

A Fiberite prepreg system 60U49, which consists of AS4 carbon fiber and a 250F cure toughened epoxy resin, designated 949, was used for all in-plane testing. The material was fabricated into 16-ply panels of various lay-ups, as shown in **Table 1**.

The convention used to describe these laminates consists of the fiber orientation angle, which is measured with respect to the load direction (0°), the subscript numbers, which are the numbers of pattern repeats, and the "S" subscript, which indicates symmetrical lamination about the center plane. These laminates represent some portion of the "quasi-isotropic" lay-up, which exhibits isotropic properties in the plane. This fiber architecture is the most significant in practice. The tests on the individual subsets of this lay-up provide a means of delineating the individual contributions of each to the strength/rate behavior. In addition, compression tests were conducted on the pure epoxy at the same rates, to determine if there is correlation between the rate behavior of the composite and the matrix phase, which is presumed to exhibit all the rate effects.

To determine the magnitude of any rate effects, tests were conducted at nominal strain rates of 5×10^{-4} s^{-1} using a screw-driven test machine, and 10 to 50 s^{-1} using the instrumented drop tower. Higher rates could not be achieved without overwhelming stress wave distortions. Also, because of this initial ringing, we were unable to accurately determine the initial moduli. Ultimate strengths were readily measured, and the results for tests at both rates are summarized in **Fig. 3**.

Only minimal increases are seen in the strengths of any laminate which contains fiber oriented in the load direction, i.e., the quasi-isotropic, cross-ply and unidirectional laminates. These strength improvements are less dramatic than those that have been reported for other test methods and materials. As expected, larger strength increases are found for the purely matrix-dominated laminates, namely, the transverse and shear tests. Both of these laminates fail in a shear-type mode, which is presumably governed by the yield strength of the matrix. The reasonably good agreement between the percentage increase in strengths of the pure epoxy and these two matrix-dominated laminates supports this hypothesis. These softer plies do not contribute much to the strength of the fiber-dominated lay-ups, and it appears that most of the increase in strength of the quasi-isotropic lay-up is due to the minimal increase in the strength of the unidirectional ply.

In order to achieve higher strain rates, new specimen geometries having reduced mass and size are required. For example, using the split Hopkinson bar for rates exceeding 1000 s^{-1}, the ideal length-to-diameter ratio for isotropic materials is claimed to be ~1:2.[12] It would be difficult to modify any in-plane fiber composite test method to this geometry and not be plagued by anomalous end failures. We have initiated a study of potential candidate specimens for higher rate compression and performed quasi-static tests to determine their utility. The various geometries are shown in **Fig. 4**. The goal has been to reduce specimen size and guarantee failure away from the ends. To produce these shapes, a relatively thick cross-ply panel was carefully fabricated to maintain quality similar to the thinner 16-ply panels already described.

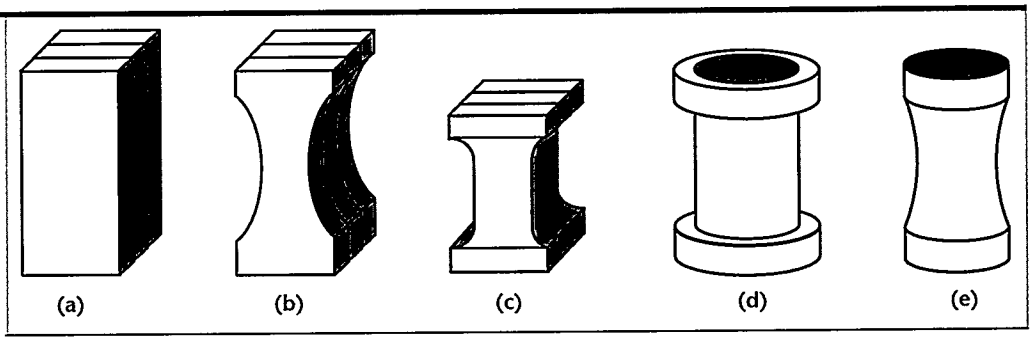


Figure 4. Geometries investigated for high-rate compression testing: (a) rectangle, (b) dog bone, (c) tapered cube, (d) cylinder with steel end rings, and (e) tapered cylinder.

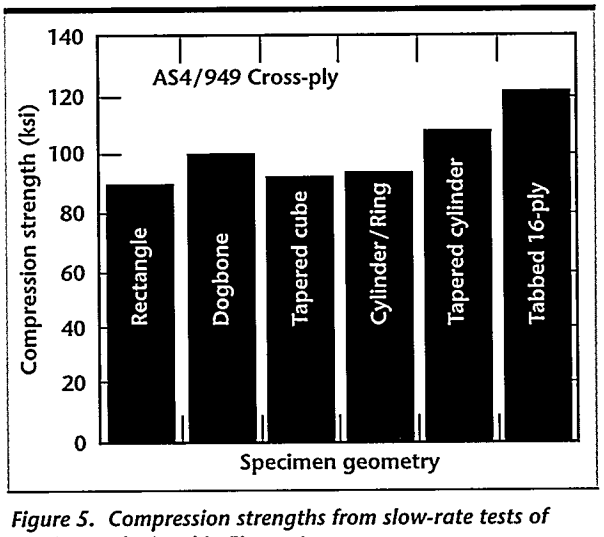


Figure 5. Compression strengths from slow-rate tests of specimens depicted in Figure 4.

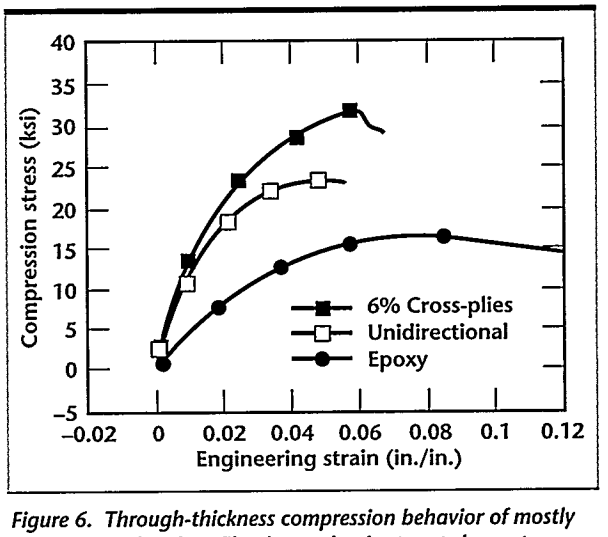


Figure 6. Through-thickness compression behavior of mostly unidirectional carbon fiber/epoxy laminates at slow rates.

Results of slow-rate compression tests for the different geometries are shown in **Fig. 5**. Included in this figure is the average strength obtained using the tabbed 16-ply specimens and the fixture described earlier. A "smeared" properties stress analysis using NIKE 3D was conducted for several geometries, and revealed problems with the double tapered cube. Although calculated stress concentrations were three times greater for this shape, the measured strengths are not significantly lower than for the other specimens.

The strengths of the tapered cylinder come closest to the thin panel test, but this specimen failed away from the point of minimum area, due to delamination of surface plies. Compression at a rate of 50 s^{-1} using a specially modified hydraulic machine[13] did show a 12% increase in strength. Again, failure occurred away from the minimum gauge area. While strength values for this specimen are encouraging, the observed failure mode is not pure compression. Future work will determine if some variant of this specimen geometry provides true compression failure so that higher strain rates can be obtained.

Out-of-Plane Compression Study

Some studies have shown that there can be improvements in through-thickness strength over a unidirectional lay-up by orienting fiber in a cross-ply configuration[14,15], but the information is incomplete. We have conducted a study to examine how through-thickness compression behavior is determined by lay-up and rate of loading.

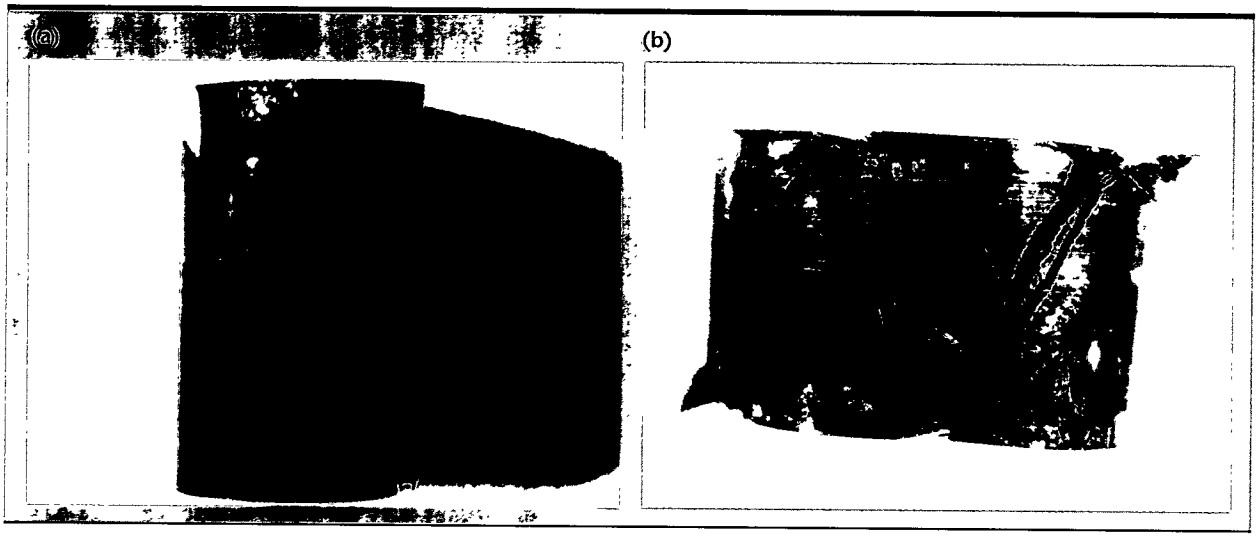


Figure 7. Shear slip lines in unidirectional laminates failed under through-thickness compression.

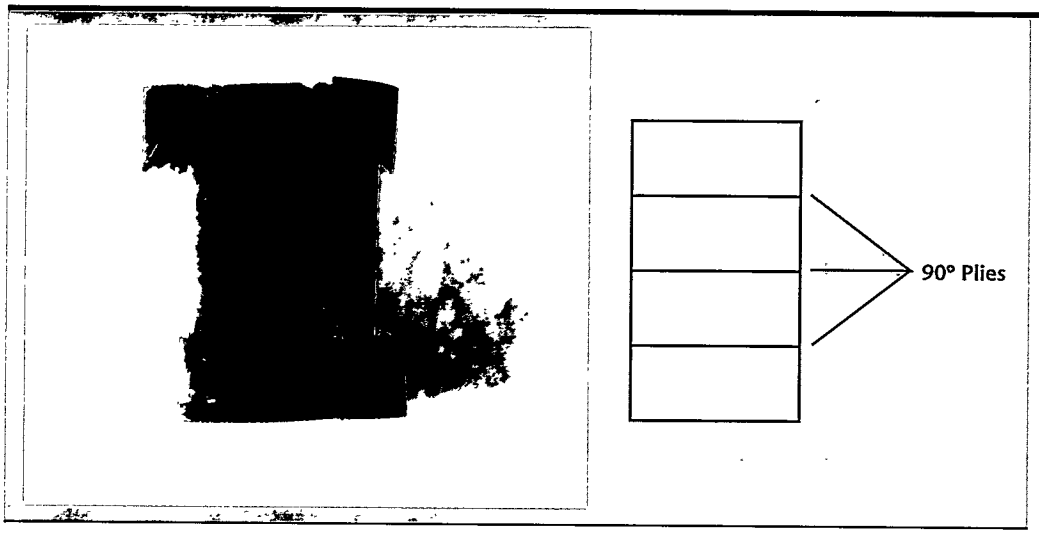


Figure 8. Effect of adding a few cross-plies to a unidirectional laminate on the through-thickness compression failure mode.

Relatively thick (3/8 in.) panels of carbon fiber/epoxy and glass fiber/epoxy composites were core-drilled to right circular cylinder specimens. Ends were ground flat to within 0.0001 in. and compression tests were conducted using precision subpresses. In order to compare results at quasi-static and high rates, both 1:1 and 2:1 aspect ratio (length:diameter) specimens were tested. Static tests usually require 2:1 or greater aspect ratios, while dynamic tests are usually conducted at ratios of 1:1 or less. The latter becomes a problem if there is significant friction at the load surface because the stress state is no longer uniaxial. To minimize these effects, all tests were conducted using lubricated loading platens.

At quasi-static rates, the unidirectional material fails in shear, with a slip plane oriented at about 30° away from the load axis. The stress/strain behavior is similar to that of the pure epoxy, as shown in **Fig. 6**. The smaller aspect ratio unidirectional specimens are slightly stronger, but also fail in shear along similar slip lines, as shown in **Fig. 7**. The addition of only a few layers of cross-ply material to the unidirectional laminate restricts the slip line formation, as shown in **Fig. 8**, and increases strength, as shown in **Fig. 6**. The strengths of both the pure epoxy and unidirectional composite increase with deformation rate, as shown in **Fig. 9**. The trends are similar, again showing the direct influence of the epoxy properties on the matrix-dominated behavior of the composite.

The dramatic effect of alternating the angles between adjacent plies of a laminate by at least 45° is seen with the cross-ply and quasi-isotropic lay-ups. Both undergo nearly elastic compression at quasi-static rates up to ultimate stress and strain levels of ~140 ksi and 10%, respectively. For the quasi-isotropic lay-up, some strength increase is seen with rate. However, the cross-ply lay-up begins to show a significant drop-off in strength, beginning in the intermediate

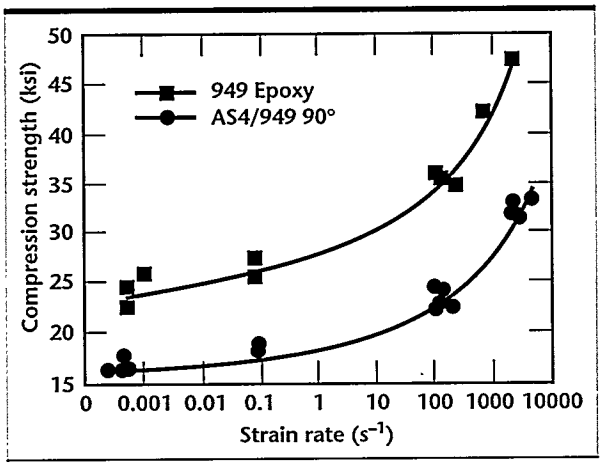


Figure 9. Effect of rate on compression strengths of epoxy and unidirectional carbon fiber laminate, loaded perpendicular to the fiber axis.

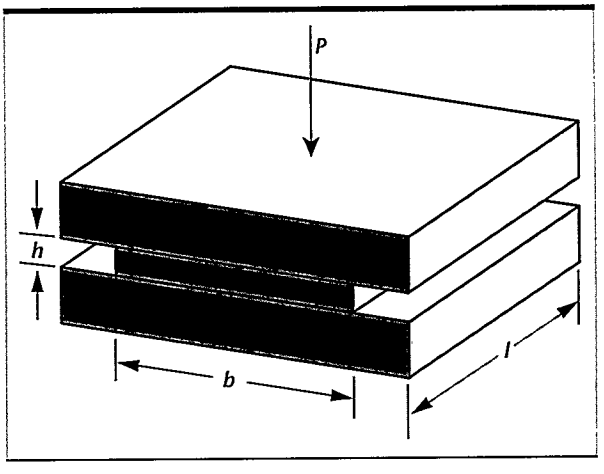


Figure 11. Definition of parameters used in limit load analysis of through-thickness strength.

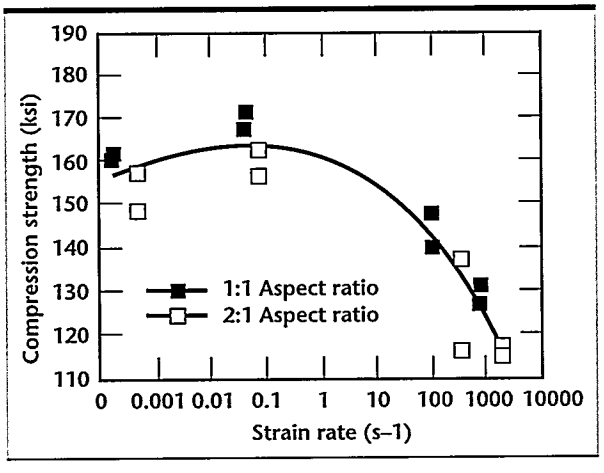


Figure 10. Dependence of through-thickness strength on rate, for a cross-ply laminate.

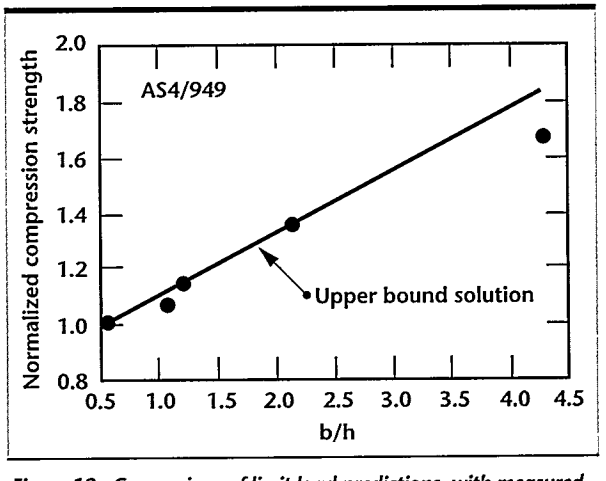


Figure 12. Comparison of limit load predictions, with measured through-thickness strengths of mostly unidirectional laminates.

rate range, as shown in **Fig. 10**. This strength loss was observed consistently for intermediate and high rate tests and for both aspect ratios, which indicates that the poorer performance at higher rates is a material property, as opposed to a test anomaly. This result is pertinent to composite ballistic armor and blast shields which are usually cross-plied materials. The ability of the quasi-isotropic lay-up to retain its high strength at high rates would suggest that it would be a better laminate for these applications.

The effect of aspect ratio and ply orientation on through-thickness strength appears similar to the limit load behavior predicted for plastic materials. The effect can be modeled using an approximate plane strain analysis derived for the forging of a bar, as depicted in **Fig. 11**.[16] The analysis yields the upper bound prediction

$$\sigma = \sigma^o\left(1+\frac{b}{4h}\right) \qquad (2)$$

where σ^o is the uniaxial compressive yield stress and the dimensions b and h are defined in **Fig. 11**. For the composite through-thickness compression specimens, the large disparity between fiber-direction and transverse direction moduli and strength results in a nearly plane strain state within a ply. For comparison with **Eq. 2**, the value b is the specimen diameter and h is the thickness of the largest group of like-oriented plies. Using the results from the different aspect ratio unidirectional materials and the laminates having only a few (6%) cross-plies, results are compared to the limit load solution, in **Fig. 12**. Good agreement is obtained for the small diameter/length ratios, but measured strengths fall below predictions when the plies become relatively thin. The disparity is much worse for the cross-ply and quasi laminates, where predicted strengths are several times greater than measured values. These discrepancies are most likely due to the deviation from the plane strain condition in the case of the laminates.

The question then remains regarding the governing failure mode for the high-strength quasi-isotropic and cross-ply laminates. Failure of these materials is catastrophic, since the large elastic energy is released abruptly. The failure always involves significant fiber fracture, and led us to determine if fiber tensile strength played a role. Tests were conducted on quasi-isotropic laminates of S2-Glass, AS4 carbon and T1000G carbon fibers, which have composite tensile strengths of 250, 325 and 480 ksi, respectively. All three laminates exhibited the same linear behavior to relatively large strains. A plot of through-thickness strength vs fiber-direction strength, shown in **Fig. 13**, reveals a strong and linear relationship between the two parameters. It appears then that the laminates fail in through-thickness compression in a mode which depends on fiber strength.

Future Work

We will complete the study of the through-thickness compression behavior by assessing the material response as a function of lay-up in $\pm\theta$ laminates with angles varying from 0 to 45°. The results obtained in the current work will be used to optimize composite materials and fiber architecture for ballistic armor and blast shield applications, as well as for improving the design of mass-loaded flywheels and high-pressure gun barrels. We will begin a study to determine the effect of damage on residual properties of composite laminates with the ultimate goal of developing a 3-D continuum damage material model for our structural codes.

Acknowledgments

We greatly appreciate the assistance given by W. Andrade and I. Fine in devising methods to machine the novel compression specimens; by I. Chiu in preparing the epoxy resin specimens; by P. Harwood in fabricating the tabbed composite specimens; by R. Lum in taking the photographs; and by R. Lum, R. Sanchez, P. Harwood, D. Freeman and M. LeBlanc in conducting tests and reducing data. We also acknowledge the fruitful discussions with E. Zywicz, who pointed out the similarity between the through-thickness test results and the limit load analysis.

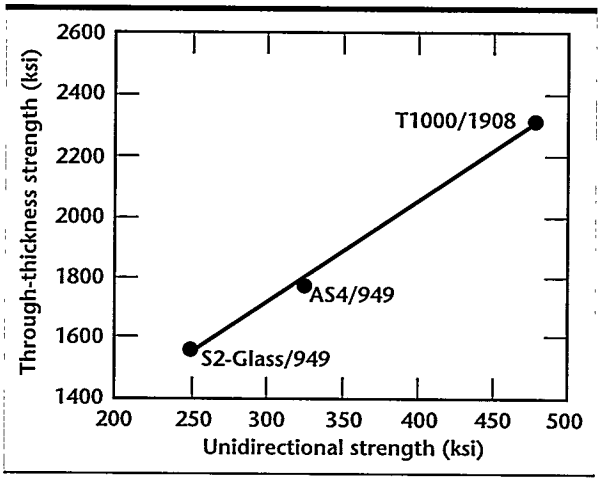


Figure 13. Relationship between fiber-direction tensile strength and through-thickness compression strength of quasi-isotropic laminates, based on different fiber reinforcement.

References

1. Daniel, I. M., W. G. Hamilton, and R. H. LaBedz (1982), "Strain-Rate Characterization of Unidirectional Graphite/Epoxy Composite," in *Composite Materials: Testing and Design (Sixth Conference)*, *ASTM STP 787*, I. M. Daniel, Ed. (ASTM), pp. 393–413.

2. Daniel, I. M., and T. Liber (1978), "Testing of Fibre Composites at High Strain Rates," *2nd Int. Conf. on Comp. Matls., ICCM/2, Toronto, Canada* April, 1978.

3. Christensen, R. M., and S. J. DeTeresa (1966), "The Kink Band Mechanism for the Compressive Failure of Fiber Composite Materials," in press.

4. Lankford, J. (1993), "Micromechanisms of Compressive Failure in a Glass Fibre-Reinforced Amorphous Thermoplastic," *J. Mater. Sci.* **28**, pp. 3027–3034.

5. Groves, S. E., S. E. Sanchez, R. J. Lyon, and A. E. Brown (1993), "High Strain Rate Effects for Composite Materials," in *Composite Materials: Testing and Design (Eleventh Volume), ASTM STP 1206*, E. T. Camponeschi, Ed., (ASTM), pp. 162–176.

6. Montiel, D. M., and C. J. Williams (1992), "A Method for Evaluating the High Strain Rate Compressive Properties of Composite Materials," *Composite Materials: Testing and Design (Tenth Volume), ASTM STP 1120*, G. C. Grimes, Ed. (ASTM), pp. 54–65.

7. Tzeng, J. T., and A. S. Abrahamiam (1995), "Dynamic Compressive Properties of Composites at Interior Ballistic Rates of Loading—Experimental Method," *Comp. Eng.* **5**(5), pp. 501–508.

8. Choe, G. H., W. W. Finch, Jr., and J. R. Vinson (1988), "Compression Testing of Composite Materials at High Strain Rates," *Proceedings, 4th Japan-U.S. Conf. Comp. Mater., Washington, D.C.*, June 27–29, 1988, pp. 82–91.

9. Christensen, R. M. (1979), *Mechanics of Composites Materials*, John Wiley & Sons, New York, N.Y.

10. Koppelman, J. (1958), *Rheol. Acta* **1**, p. 20.

11. Zimmer, J. E., and J. R. Cost (1969), "Determination of the Elastic Constants of a Unidirectional Fiber Composite Using Ultrasonic Velocity Measurements," *J. Acoust. Soc. Amer.*, **3**(2), pp. 795–803.

12. Davies, E. D., and S. C. Hunter (1963), *J. Mech. Phys. Solids* **11**, p. 155.

13. LeBlanc, M. M., and D. H. Lassila (1995), *A Hybrid Technique for Compression Testing at Intermediate Strain Rates*, Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-JC-118740).

14. Abdallah, M. G., T. O. Williams, and C. S. Muller (1990), *Experimental Mechanics of Thick-Laminates: Flat Laminates Mechanical Properties Characterization*, Hercules Technical Report, Hercules, Calif., (DDR:153253).

15. Satoh, Y., H. Ikezawa, and I. Hara (1994), "Compressive Behavior of Thick GFRP Laminates," *26th Inter. SAMPE Tech. Conf., Atlanta, Ga.*, pp. 107–111.

16. McClintock, F. A., and A. S. Argon (1966), *Mechanical Behavior of Materials*, Addison-Wesley, Reading, Mass., pp. 380–382.

Participating Media Radiant Transport at High Temperature: NIF Flashlamps and Investment Casting Molds

Mark A. Havstad
Laser Science Engineering Division
Mechanical Engineering

Charles W. Dingus
New Technologies Engineering Division
Mechanical Engineering

We have designed two optical systems around a high vacuum furnace capable of maintaining optically transmissive glass and ceramic samples at temperatures in excess of 1100 °C. These systems have allowed us to measure the visible and near infrared transmissivity of the cerium-doped quartz in Lawrence Livermore National Laboratory (LLNL)'s National Ignition Facility (NIF) flashlamp envelopes, and the transmission, scattering and absorption of a ceramic shell, which is the mold material for investment casting. Our studies of the NIF flashlamp glass have quantified the temperature dependence of its transmission in the laser pump band, 0.4 to 1 μm, over its operating temperature range, ambient to 800 °C. We find a decrease in transmission at 0.4 μm, of ~30% between ambient and 1000 °C, and little change at other wavelengths. Our studies of the ceramic shell have led to estimates of its optical scattering and absorption. These measurements are key to obtaining closure between calculation and measurement of the cooling rate of investment castings.

Introduction

Historically, optical elements and materials have rarely operated much above ambient temperature. Consequently, optical transmissivity, optical constants and other optical properties are seldom of interest at temperatures even as high as 100 °C. The sophisticated and precise instruments evolved by the optics industry are rarely capable of treating samples at elevated temperatures. However, development work at LLNL in two seemingly very different technologies has led to the need for transmission (and, in one case, scattering) measurements with samples at high temperature.

The flashlamp envelopes proposed for LLNL's NIF are a key component, because they transmit both the pump band energy (0.4 to 1 μm) for the laser amplifiers, and also the thermal (beyond 1 μm) and UV (below 0.4 μm) energy which heats the laser slab and forces a wait of many hours before the slab is sufficiently cool and uniform in temperature to fire again. Adequate beam quality and no risk of damaging the slab are the concerns. A high temperature optical transmission system was required to quantify flashlamp performance, because the glass envelopes heat as the lamps fire. During the 350 μs firing period, the lamp's envelope heats to approximately 800 °C. The room temperature transmissivity of the flashlamp envelope (quartz, doped with ~1% cerium) determines the pumping efficiency at the onset of firing, but by the end of firing it is the 800 °C transmissivity that is relevant.

Here, we present the temperature dependence of the transmission of the glass in the pump band. However, the optics we have built around the axis of the furnace are also well suited to transmission measurements out of the pump band.

Investment casting molds contrast sharply with the smooth, non-scattering, and transparent (in the visible region) NIF optics noted above. These molds are composed of slurries of glassy particulates (mostly silicates and aluminates) that are strongly scattering at room temperature, and in normal light appear quite opaque. However, when the mold

surrounds a volume of liquid metal and is removed from a furnace and set aside to cool, the profile of the metal beneath the shell is clearly visible, as are thermocouples occasionally embedded in the middle of the thickness of a mold. The cooling rate at the onset of casting cooling (which can also occur inside a vacuum furnace, in the special case of Bridgeman growth of uniaxial or single crystal castings for turbine blades), is key to determining the metallurgical properties of a casting and to obtaining closure between modeling of the casting process and observations of it. The radiant transport through and from the casting mold is the boundary condition to the transient heat transfer problem, which is casting cooling. Efforts to advance investment casting into new markets and technologies, reduce its waste stream, and improve the performance of parts made by this process are depending on the application of finite element modeling for development insight and improved techniques.

Our work this year has focused on assembling a unique experimental capability from existing equipment and expertise in optical properties and high temperature materials development. Once the techniques and tools were in place we measured 1) the magnitude of the change in transmission with temperature in the pump band of the cerium-doped quartz, and 2) the absorption and scattering of a single investment casting mold material at wavelengths less than 1 μm. We have left for future work the measurement of transmission outside of the pump band (for cerium-doped quartz) and absorption and scattering coefficients at longer wavelengths for the casting mold.

Progress

The two optical systems designed around the high temperature furnace are shown in **Figs. 1** and **2**. The systems share two light sources to cover the broad spectral band of interest, 0.2 to 10 μm. A carbon filament lamp provides sufficient emission for studies from the middle of the visible to 10 μm. This source is positioned off the furnace axis (also the optical axis for the hot samples under study), and a spherical mirror in a kinematic mount is used to produce a nearly collimated beam down the furnace axis. A 200-W mercury-xenon arc lamp allows investigations in the 0.2 to 2 μm range. This source and its UV-grade fused silica lens are mounted on the furnace axis, and when it is in use, the spherical mirror used with the carbon lamp is removed from the optical path.

The ends of the furnace are enclosed by BaF_2 and CaF_2 windows (both with ~0.2 to 10 μm bandpass) mounted in conflat flanges. A carbon-carbon composite heating element (4 in. long, .025 in. thick and 2.5 in. in diameter) surrounds the samples held in graphite capture rings supported by alumina rods. Five layers of refractory metal shielding isolate the heating element from a water cooling jacket which fits snugly in the vacuum vessel. Shielding is also used at the ends of the hot central zone of the vacuum vessel. These shields are bored with 0.5-in. holes on their centers, for the optic axis, and the shields at the optical exit end are slitted in the horizontal direction so that scattering as a function of angle can be obtained for the measurements with the investment casting molds.

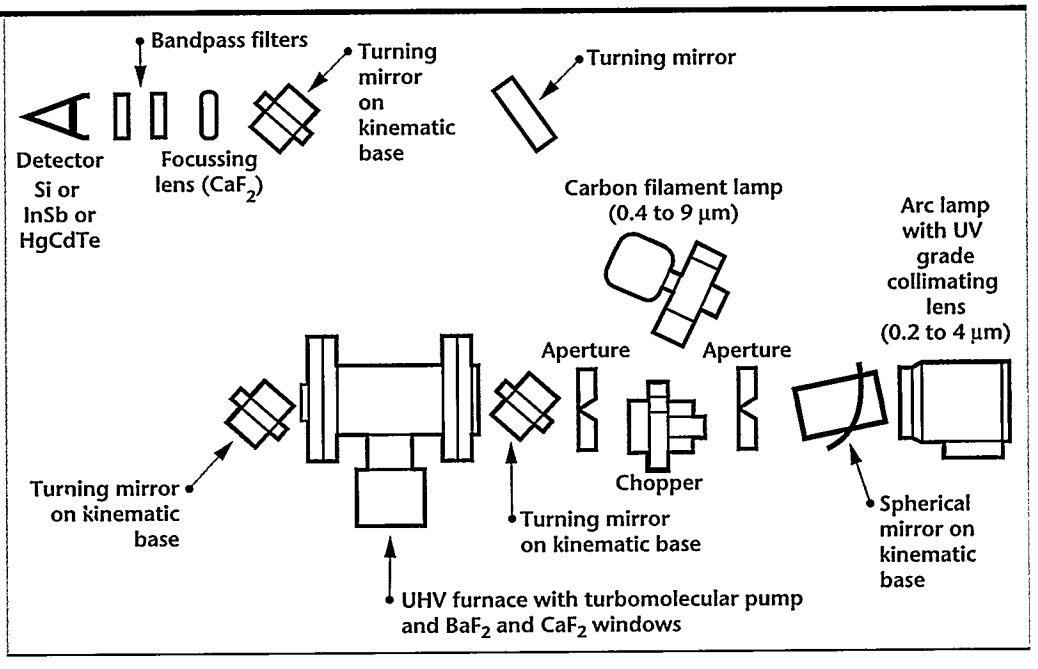


Figure 1. Optical system for measurement of transmission of NIF flashlamp glass.

Cerium-doped Quartz

For the transmission measurements through flashlamp glass, there are two optical paths, a reference path and a sample path. The reference path is averted from the furnace by a turning mirror in a kinematic mount (**Fig. 1**). A second turning mirror folds the reference path into the detector. The sample path is folded just after exiting the furnace by the same turning mirror used in front of the furnace by the reference path. A second turning mirror (also in a kinematic mount) in the sample path folds the sample path into the detector. A CaF_2 lens, common to both beam paths, focuses either beam at the detector. The bandpass filters for spectral selection are placed between the detector and the CaF_2 lens. Separate optical enclosures surround both the input optics (items to the right of the furnace) and the output optics (to the left) to minimize stray light entering the optical paths.

When investigating the transmission of cerium-doped quartz, the beam alternates between the sample and reference paths by manual movement of the turning mirrors in the kinematic mounts. For each pair of sample and reference readings, a different bandpass filter is in place in front of the detector. The response at the detector for the two beam paths can be given by the following two relations:

$$V_{DET.REF.} = B_s\rho_{T1}\rho_{T2}\tau_L\tau_f R_D \qquad (1)$$

$$V_{DET.SAM.} = B_{s'}\tau_{w1}\tau_{SAM}\tau_{w2}\rho_{T1'}\rho_{T3'}\tau_{L'}\tau_{f'}R_{D'} \qquad (2)$$

Where ρ is mirror reflectivity, τ is transmissivity, R_D is detector responsivity, and B_s is lamp output through the chopper and apertures. Subscripts L, f, T1, T2 and T3 refer to the lens, the filters and the turning mirrors, respectively. Subscripts w1 and w2 refer to the entrance and exit windows on the furnace. The use of a prime symbol with several of the subscripts indicates a second order deviation from the corresponding term without the prime. For example, the reflectivity of the turning mirror that moves between the two ends of the vacuum furnace (designated T1) is, to first order, the same for either position. It is not identically the same for the two positions, because the beam size, location and angle of incidence are not absolutely identical for the two positions, but they are close enough so that differences in reflectivity are small compared to the magnitude of the loss in transmission due to heating the sample in the vessel. Similarly, there are other second order effects that apply at other positions in the optical system. These effects combine to produce small systematic errors in the results shown below. Drift in the lamp output, non-repeatability in the positioning of the kinematic mirror mounts and, to a lesser extent, drift in the detector responsivity also combine to generate systematic and random errors, but these are small compared to the drop in transmission of the sample due to temperature rise.

Measurements of the spectral transmission for each optical path were made for sample temperatures of 40 °C, 400 °C, 600 °C, 800 °C and 1000 °C. The results at any given temperature and wavelength can be represented as:

$$\tau_{SAM.} \doteq \frac{V_{DET.SAM.}}{V_{DET.REF.}} \frac{1}{\tau_{CaF_2}\tau_{BaF_2}} \doteq \frac{B_{s'}\tau_{w1}\tau_{SAM}\tau_{w2}\rho_{T1'}\rho_{T3}\tau_{L'}\tau_{f'}R_{D'}}{B_s\rho_{T1}\rho_{T2}\tau_L\tau_f R_D} \frac{1}{\tau_{w1}\tau_{w2}} \qquad (3)$$

Use of this relation requires separate measurements of the transmissivity of the CaF_2 and BaF_2 windows on the vacuum vessel ends. Alternatively,

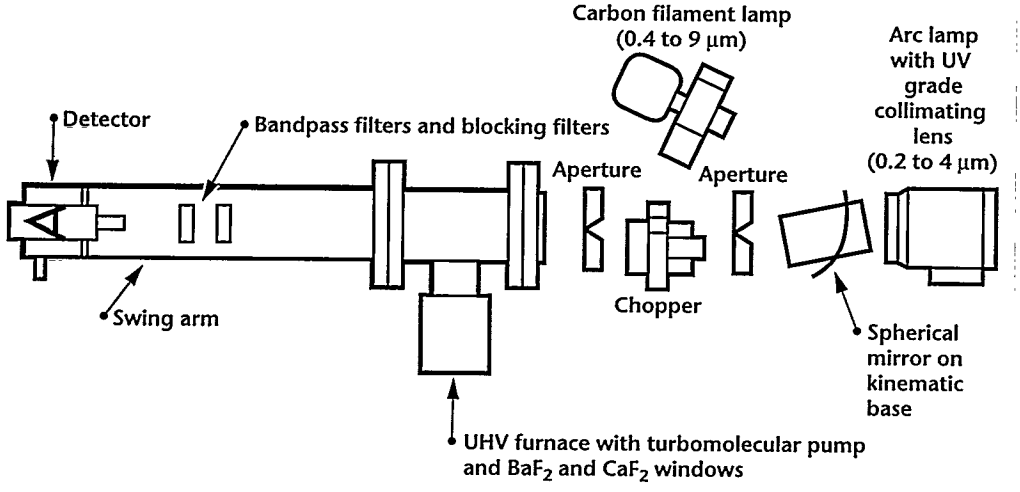


Figure 2. Optical system for measurement of transmission and scattering in investment casting molds.

by taking the ratio of **Eq. 3** above, applied at elevated temperature to **Eq. 3** applied at 40 °C , we avoid the window measurements, and at the same time reduce the so-called "second order" errors described above. This approach uses the measurement system to yield a ratio of transmittances, τ(hot)/τ(ambient), rather than an absolute transmittance. Since the transmission of cerium-doped quartz at ambient is well known from precise transmission measurements made in commercial instruments, it is sufficient here to derive a ratio of transmittances:

$$\frac{\tau(T>amb.)}{\tau(T=amb.)} \doteq \frac{\left[\frac{B_S\tau_{W1}\tau_{SAM}\tau_{W2}\rho_{T1}\rho_{T3}\tau_{L'}\tau_{f'}R_{D'}}{B_S\rho_{T1}\rho_{T2}\tau_L\tau_f R_D}\right]_{T=T>amb.}}{\left[\frac{B_S\tau_{W1}\tau_{SAM}\tau_{W2}\rho_{T1}\rho_{T3}\tau_{L'}\tau_{f'}R_{D'}}{B_S\rho_{T1}\rho_{T2}\tau_L\tau_f R_D}\right]_{T=amb.}}$$

$$\doteq \frac{\left[\frac{V_{DET.SAM.}}{V_{DET.REF.}}\right]T>ambient}{\left[\frac{V_{DET.SAM.}}{V_{DET.REF.}}\right]T=ambient} \qquad (4)$$

By taking the approach of **Eq. 4** above, one is not forced to assume that the primed and unprimed quantities above are equal. They separately cancel with their corresponding value at the different temperature in **Eq. 4**. There are fewer and smaller errors this way. For example, in **Eq. 4**, one assumes that the CaF_2 and BaF_2 window transmissivities are independent of temperature. Although this is not strictly true, for a sample temperature of 1000 °C, the exit window only reaches about 120 °C and the entrance window only 50 °C. Other small errors are incurred at all the optical elements, but in general, all are yet smaller than in **Eq. 3**. If the absolute value for transmission of cerium-doped quartz at high temperature is desired, one merely multiplies the ratio determined from **Eq. 4** by a value obtained by a precise commercial instrument at ambient temperature.

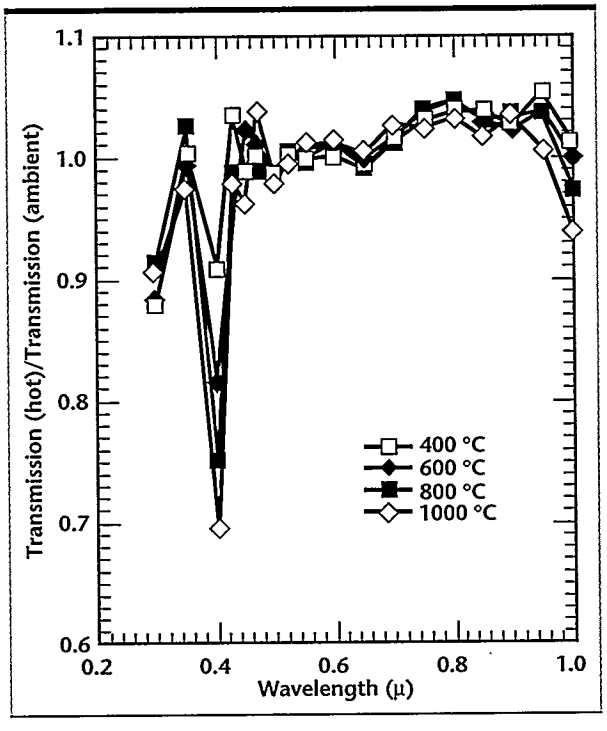


Figure 3. Transmission of cerium-doped quartz in the laser pump band.

Figure 3 shows the ratio of hot transmission to ambient transmission for cerium-doped quartz in the NIF pump band for temperatures of 400 to 1000 °C. The strong temperature dependent absorption at 0.4 μm was expected from unpublished measurements made on a different type of cerium-doped quartz. We suspect that as the quartz host is heated, the cerium atoms populate a different distribution of energy states, one that interacts more strongly with 0.4 μm light. This may imply that the envelope interacts less strongly with UV light at elevated temperature. If so, the flashlamp envelopes may have a compound decrease in performance at high temperature, poorer transmission in the pump band, and weaker blocking in the UV region. There is a similar, but weaker, temperature dependent absorption at 1 μm. This may not have significant influence on pump band performance, but absorption effects between 1 and 5 μm (where the ambient temperature transmissivity of cerium-doped quartz cuts off) can influence both flashlamp heating of the laser slab and flashlamp envelope heating, which is the probable cause of flashlamp failure.

On either side of 0.4 μm, there are changes in transmission which are not monotonic with temperature. These may be due to random errors in the data, or to shifts in the population of the energy levels of the cerium dopant atoms. Further work with reduced experimental uncertainty and data over the full spectral range of interest to flashlamp performance and heating (0.2 to 9 μm) is desired.

Investment Casting Molds

The design of the optical system applied to scattering by investment casting molds, shown in **Fig. 2**, was influenced by the experimental works of Glicksman *et al.*,[1] and McCarthy,[2] but is unique for its ability to make spectral measurements and to maintain samples at high temperature. The output optics (after the furnace) are mounted on a swing arm with a pivot point that coincides with the back face of the sample. The detector collects the light transmitted by the sample into an arc in a horizontal

plane behind the sample back face. The limits to the swing arm motion are 50° on either side of normal to the sample back face. The detector is mounted on a two-axis translation stage so that it can be moved in both directions perpendicular to the beam axis. This motion allows characterization of the input beam size, divergence and power variation with position, when no sample is present in the optic path. Once the beam profile and divergence are mapped, these translators are used to center the detector on the beam axis. As before, the bandpass filters are mounted in front of the detector but here no lens is required to focus the transmitted intensity into the detector.

The purpose of the detector on the swing arm is to measure the angular dependence of the light transmitted by the sample. This transmission is determined by how the incident beam is scattered and absorbed once it enters the mold material. The scattering and absorption coefficients are key to determining the modes and magnitudes of high temperature heat transfer through casting molds. These coefficients are best understood by their use in the radiative transfer equation:[3]

$$\frac{dI_\lambda(s)}{ds} = -(\alpha_\lambda + \sigma_{s\lambda})I_\lambda(s) + \alpha_\lambda I_{b\lambda}(s) + \frac{\sigma_{s\lambda}}{4\pi}\int_{\omega'=4\pi} I_\lambda(s,\omega')\Phi_\lambda(\omega,\omega')d\omega' \quad . \tag{5}$$

Spectral quantities are signified by λ; b connotes a Planckian spectral distribution. The phase function, $\Phi_\lambda(\omega, \omega')$, describes the scattering of wavelength λ from direction ω' into direction ω for the scattering coefficient $\sigma_{s\lambda}$.

The term on the left hand side is the rate of change in intensity of a monochromatic beam as it propagates through a participating media. The first term on the right hand side is the decrease in intensity of the beam due to scattering ($\sigma_{s\lambda}$) and absorption (α_λ) during propagation. The second term on the right hand side gives the increase in intensity due to emission. At equilibrium, absorption and emission are equal, and α_λ can be used for the monochromatic emission. The last term on the right hand side is the increase in intensity with propagation due to scattering into the direction of propagation from other directions. This term is an integral over all possible directions for scattering into the given propagation direction. The phase function (a bidirectional quantity) is normalized (with 4π) so that when integrated over all incoming directions (ω') it gives unity. The emission term is not important here because the incident beam is chopped. Only transmitted light is chopped and then amplified by the detection electronics by means of a lock-in amplifier.

The ideal apparatus provides data that would permit computation of α_λ, $\sigma_{s\lambda}$ and the phase function. In practice, determining the phase function of transmitting and scattering materials from the radiative transfer equation is presently quite difficult.[4,5] Fortunately, the physical structure of casting molds makes it probable that they scatter isotropically (equally in all directions). Further, parameter studies have shown a weak dependence of heat transfer on the phase function for moderate and stronger scatterers.[6] However, even with the assumption of isotropic scattering, it is still not trivial to obtain the scattering and absorption coefficients from the transfer equation.

The apparatus measures scattered intensity leaving the sample as a function of angle relative to the normal to the back surface. Clearly this intensity declines for thicker samples because the intensity entering the sample has a longer path over which to be scattered away or absorbed. For weakly scattering samples, the scattered intensity vs angle will be strongly peaked around the incident direction. As the albedo ($\sigma_{s\lambda}/(\sigma_{s\lambda} + \alpha_\lambda)$) increases, the angular dependence of measured intensity becomes more uniform. For very strongly scattering materials, no variation in scattered intensity is expected. The directional character of the incident beam is entirely lost before reaching the exit face of the sample.

The measured and calculated variations in sample signal with detection angle are shown in **Fig. 4.** The rate of decrease in signal magnitude with angle indicates that the material is moderately scattering, and properties can be deduced from the data. The computed variation was obtained by Monte Carlo calculation of the radiant transport through the sample (using TRIM3D), and the scattering and absorption coefficients were varied iteratively to obtain good agreement. Specifics of the computational effort are given in an article by Maltby *et al.*, in the Computational Mechanics section of this report. To confirm the validity of our combined measurement and computational approach to obtaining the scattering and absorption coefficients, we must make additional measurements and calculations on samples of differing thickness, use the same computational approach and obtain similar values of $\sigma_{s\lambda}$ and α_λ. In short, we must independently arrive at the same properties from additional experiments. Further, efforts are underway to deduce $\sigma_{s\lambda}$ and α_λ from closed form calculations with the radiative transfer equation. This work is also described in the Computational Mechanics section cited above.

Future Work

In FY-96 we will obtain the transmissive and emissive properties of cerium-doped quartz over the full spectral range relevant to flashlamp envelope performance and failure modes (0.2 to 10 μm). With these properties we can model and investigate how and why envelopes explode. Detailed spectral radiant transport modeling (using the radiation transport code TRIM3D) will be coupled to finite element thermal and stress codes to compute the thermal stress transient on the inner diameter of the quartz envelopes. The data shown in **Fig. 3** for transmission of the envelope as a function of temperature indicates that there is a positive feedback loop between envelope absorption and temperature rise: absorption by the material causes a temperature rise, which causes more absorption. We suspect that flashlamps fail by this coupling mechanism and that finite element computations can reveal more detailed aspects than are presently known. Success here could reduce the rate of flashlamp failures or allow increased operating output.

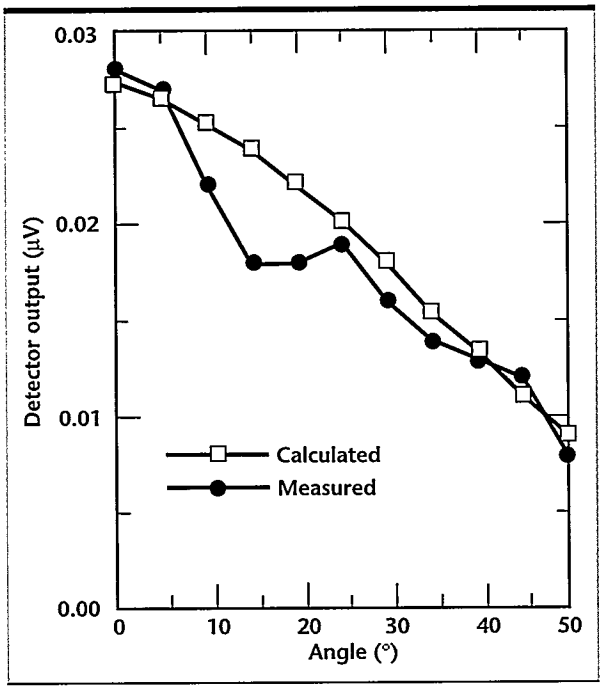


Figure 4. Measured and calculated variation in scattered and transmitted light through investment casting mold.

Additional scattering experiments with differing thickness mold samples will be performed in FY-96, to confirm our approach to determining the scattering and absorption coefficients. These tests are only required in a single wavelength band and at a single sample temperature. Once independent measurements concur we will have established the method, and we can obtain mold properties over the full spectral and temperature ranges of concern to the investment casting process. The properties so defined will then be used in finite element models of casting cooling. Closure between computed casting cooling rates and experiment (presently not achieved) is the goal, so that computational tools can be used to reduce waste, decrease dimensional uncertainty and improve the properties of investment castings.

References

1. Glicksman, L., M. Scheutz, and M. Sinofsky (1987), *IJHMT* **30**, p. 187.

2. McCarthy, T. (1989), *Measurements of infrared radiative transfer for porous ceramic materials*, M.S. thesis, University of Illinois, Urbana-Champaign, Ill., p. 235.

3. Siegel, R., and J. Howell (1981), *Thermal Radiation Heat Transfer*, McGraw-Hill, New York, N.Y., p. 862.

4. Hendricks, T., and J. Howell (1994), *HTD-vol. 276*, ASME, p. 105.

5. Skocypec, R., D. Walters, and R. Buckius (1987), *JHT* **109**, p. 151.

6. Maltby, J. D. (1993), *HTD-vol. 276*, ASME, P. 161.

Synthesis and Processing of Nanocrystalline Aluminum and Hydroxyapatite

T. G. Nieh
Materials Division
Chemistry and Materials Department

Donald R. Lesuer
Manufacturing and Materials Engineering Division
Mechanical Engineering

We have performed dynamic compaction of nanocrystalline aluminum (n-Al). Although Al powder particles were severely deformed during dynamic compaction, the presence of an amorphous surface oxide prevented a metallurgical bond between particles. Theoretical simulations of the dynamic compaction of n-Al were also carried out using an Eulerian hydrocode. The results were in good agreement with experimental observations.

In a second task, a technique was developed to synthesize hydroxyapatite (HA, Ca_{10} $(PO4)_6$ $(OH)_2$) with nanocrystalline structure. The method involved the initial preparation of an aqueous solution using inorganic water-soluble compounds. A precursor powder mixture was subsequently formed via atomization, by spray-drying the aqueous solution. Thermo-conversion of the spray-dried precursor powder resulted in the formation of a nanocrystalline structure. Structural and morphological characterizations of the synthesized material were performed using x-ray diffraction and electron microscopy.

Introduction

Nanocrystals, because of their extremely fine sizes, are thermally unstable. Grain growth occurs readily during high-temperature consolidation. Therefore, retaining these fine grain sizes during the consolidation of nanocrystals also becomes one of the major challenges for material scientists. Dynamic compaction is a viable consolidation method to retain the initial grain size of a material.[1] Pressures above the yield strength are readily achieved to deform and densify particles. In addition, consolidation is achieved by dynamic high pressures in very brief times (~ns). The fast time scale offers the opportunity to consolidate ultrafine-sized powders so quickly that they do not have time to increase their grain size. The fast time scale also produces heterogeneous surface heating and high quench rates on particle boundaries, which offers the possibility of interfacial bonding between particles, while maintaining relatively low temperatures in the interiors of the particles. Thermal quench rates are estimated to be about 10^9 °C s^{-1}.[2] At such a high quenching rate, grain growth is extremely limited. In the first part of this project, therefore, our goal was to study the dynamic compaction behavior of n-Al, and compare the results with computer simulations of grain deformation.

In the second part of this project, we investigated the synthesis and processing of HA nanocrystals. HA is a major mineral component in animal and human bodies. It has been widely used not only as a biomedical implant material, but also as biological chromatography supports in protein purification and DNA isolation. For example, a typical wet cortical bone is composed of 22 wt% organic matrix, 69 wt% mineral, and 9 wt% water. The major subphase of the mineral consists of submicroscopic crystals of an apatite of calcium and phosphate, resembling HA crystal structure. The apatite crystals are usually formed as slender needles, 20 to 40 nm in length, by 1.5 to 7 nm in thickness.[3] The mineral phase is not a discrete aggregation of the calcium phosphate mineral crystals *per se*; rather, it is made of a continuous cellular structure which gives a good mechanical strength. The apatite family of minerals, $A_{10}(BO_4)_6X_2$, crystallizes into the hexagonal rhombic prism (a = 0.9432 to 0.9418 nm and c = 0.6881 to 0.6884 nm). The ideal Ca/P ratio of HA is 10/6, and the calculated density is 3.219 g/cm^3.

Calcium phosphate-based bioceramics have been in use in medicine and dentistry for over 20 years, because of their excellent biocompatibility with human tissues. Applications include chromatography supports in the purification of protein and DNA; dental implants; percutaneous devices; and use in periodontal treatment, alveolar ridge augmentation, orthopedics, maxillofacial surgery, otolaryngology, and spinal surgery.[4]

Many synthesis methods for HA have been published, but they are all essentially chemical precipitation methods.[5–7] Controlled powder morphology with enhanced nanocrystalline structure was not the primary concern. Nanocrystalline HA is expected to provide better coatings for hip implants, bulk structures with controlled morphology, and chromatographic separation.

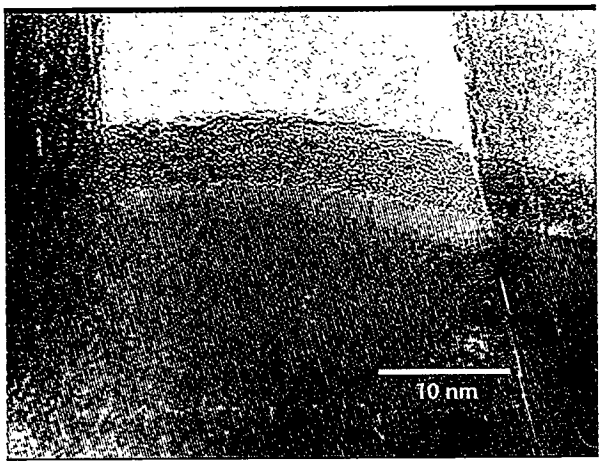


Figure 1. Transmission electron micrograph of the microstructure of a dynamically compacted Al sample (impact pressure of 4 GPa).

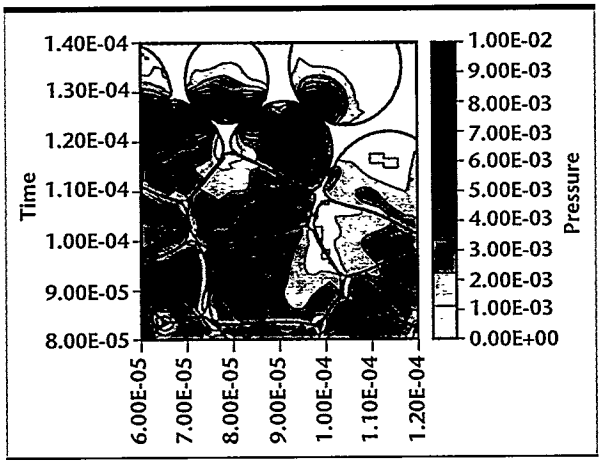


Figure 2. Contours of the pressure within the particles at the shock front at 2.1 ns. Pressures peak at approximately 1 GPa at the particle impact points, with the nominal pressure behind the shock being approximately 0.5 GPa.

Progress

Experimental Results and Discussion

Aluminum. We have already reported the experimental observations of shock compaction of n-Al at 2 to 3 GPa (20 to 30 kbars). We found that, despite substantial deformation during dynamic compaction, the surface oxide (**Fig. 1**) remained intact, and prevented a metallurgical bond between the nanosized particles. As a result, compacted samples showed a near-full density, but contained small voids (~5 nm), mainly at grain trip junctions.

In FY-95, a 2-D, planar calculation of the dynamic compaction of n-Al powders was carried out using an Eulerian hydrocode. The initial distribution of the particles with 54% of solid density was calculated, using a pseudo-gravity method. A simple Monte Carlo method generated the particle-size distribution, assuming a uniform distribution from 150 to 250 nm. A velocity boundary condition along the bottom surface imposes a particle velocity of 400 m/s, which is the estimated particle velocity of the Al powder at 2 GPa. A transmitting boundary condition is used on the upper surface to ensure that only a single shock passed through the particles. The left and right boundaries have symmetry conditions imposed. The initial spatial domain is 1500 nm on a side, with a uniform mesh of 100 zones in each direction. A Gruneisen equation of state with the Steinberg-Guinan[8] plasticity model approximates the material response. Each particle is assumed to be homogeneous and isotropic; no attempt was made to model the alumina on the surface of the particles.

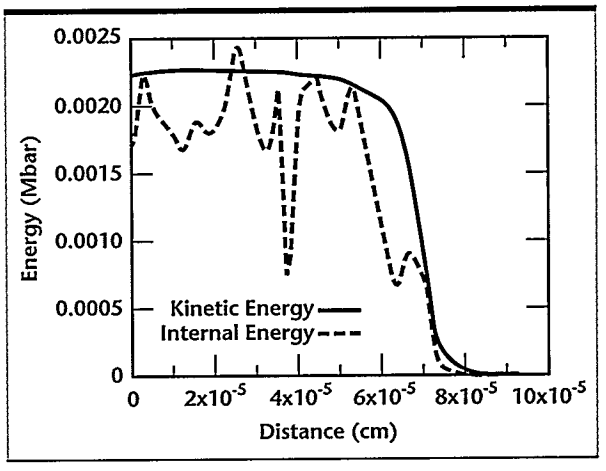


Figure 3. Profile of the internal energy and kinetic energy in the y direction, averaged in the x direction, at 2.1 ns. Because the internal and kinetic energies are approximately equal, the shock has nearly achieved its steady-state condition. The distance is measured relative to the piston at the bottom of the calculation.

The reverberation of the shock wave, which is physically important, was approximated by adjusting the impedance in the transmitting boundary condition. During the compression of the powder, the nominal pressure behind the shock was approximately 0.5 GPa, and after the densification, it increased to 6.5 GPa, which is slightly higher than in the experiment (2 to 3.4 GPa). Contours of the pressure near the shock front at 2.1 ns are shown in **Fig. 2**. As shown in **Fig. 3**, the kinetic energy and the internal energy are approximately equal behind the shock at 2.1 ns. This indicates that the shock has nearly achieved its steady-state conditions after traveling only a few particle diameters, and the numerical solution is representative of the powder in the interior of the capsule. The powder is fully dense at approximately 2.5 ns, and by 3.0 ns, the pressure in the powder has equilibrated to 6.5 GPa.

Similar compacted density values were obtained in both the simulations and the experiments (~2.7 gm/cm^3). This occurred despite the fact that the simulations had larger particle sizes (approximately 200 nm) than the experiments (approximately 50 to 73 nm). These results suggest that the compacted density may not have a strong dependence on initial particle size over the range of sizes used in the present study. Although a fully dense sample is predicted by the calculation, the experimental results indicate the presence of nano-sized pores. The discrepancy between having 5-nm voids and full compaction is quite small. Three factors account for this difference. First, the pores are on the order of 5 nm, while the mesh resolution is 15 nm. Second, the stiffening effect of the alumina on the particle surface was not modeled. Third, the energy per unit volume to collapse a spherical void is approximately 13% greater than that for a cylindrical void, for a perfectly plastic material.

Temperature profiles through the particles along the centerline of the calculation are shown in **Fig. 4** for 2.1 and 2.7 ns. The vertical lines represent gaps between the particles. As expected, the temperature increase is the greatest at the particle boundaries. The peak temperatures on the particles range from 600 to 800 K (the initial temperature is 300 K), and are in agreement with the experimental observation that the peak temperature stayed below melting (i.e., 933 K). Note that there is a substantial temperature jump between the initial shock and the final pressure state. Contours of the temperature in the particles show a highly heterogeneous distribution (**Fig. 5**). While the radially symmetric analytical models may give the correct average value of the temperature on the surface of a particle, the variation of the temperature on the surface may be quite large.

In summary, the theoretical calculations using an Eulerian hydrocode indicate that the powder was fully dense at approximately 2.5 ns, and by 3.0 ns, the pressure in the powder equilibrated to 6.5 GPa. The discrepancy between having 5-nm voids (experiment) and full compaction (theory) is quite small. The peak temperature on the particles was also calculated to range from 600 to 800 K, in agreement with the experimental result that the peak temperature stayed below melting.

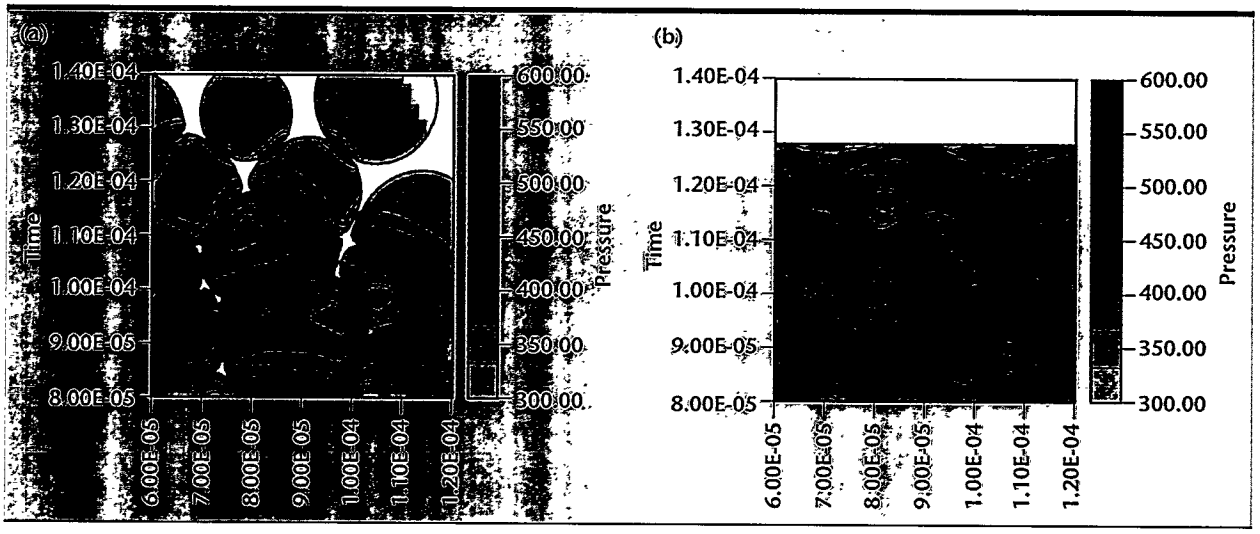


Figure 4. The temperature (K) profile along the center line of the calculation at (a) 2.1 and (b) 2.7 ns. Full densification of the powder occurs at 2.5 ns. Note that a significant increase in the temperature occurs after the full densification, indicating that modeling the reverberations is important for this problem.

Hydroxyapatite. In this task, spherical HA powders with nanocrystalline structure are synthesized by spraying an aqueous solution via atomization. A solution was initially made of a water-soluble compound containing constituent elements in HA, i.e., calcium nitrate and ammonium hydrophosphate. The ratio of calcium to phosphorus was 10:6. The pH was adjusted to 10, using an ammonium hydroxide aqueous solution. The resulting precipitates were washed eight times with distilled water using a centrifuge, until the ammonium nitrate was completely removed.

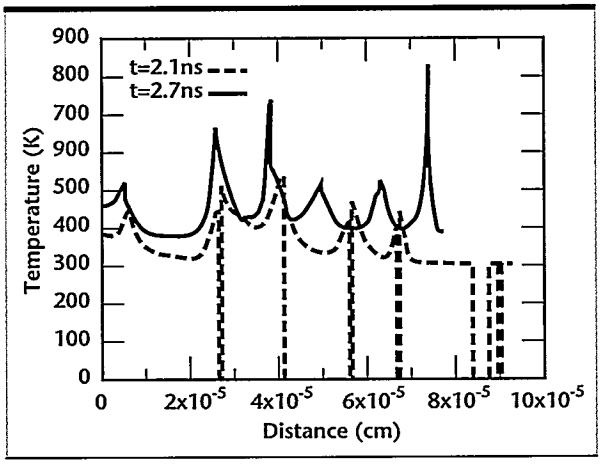


Figure 5. The temperature (K) contours in the particles at (a) 2.1 ns and (b) 2.7 ns.

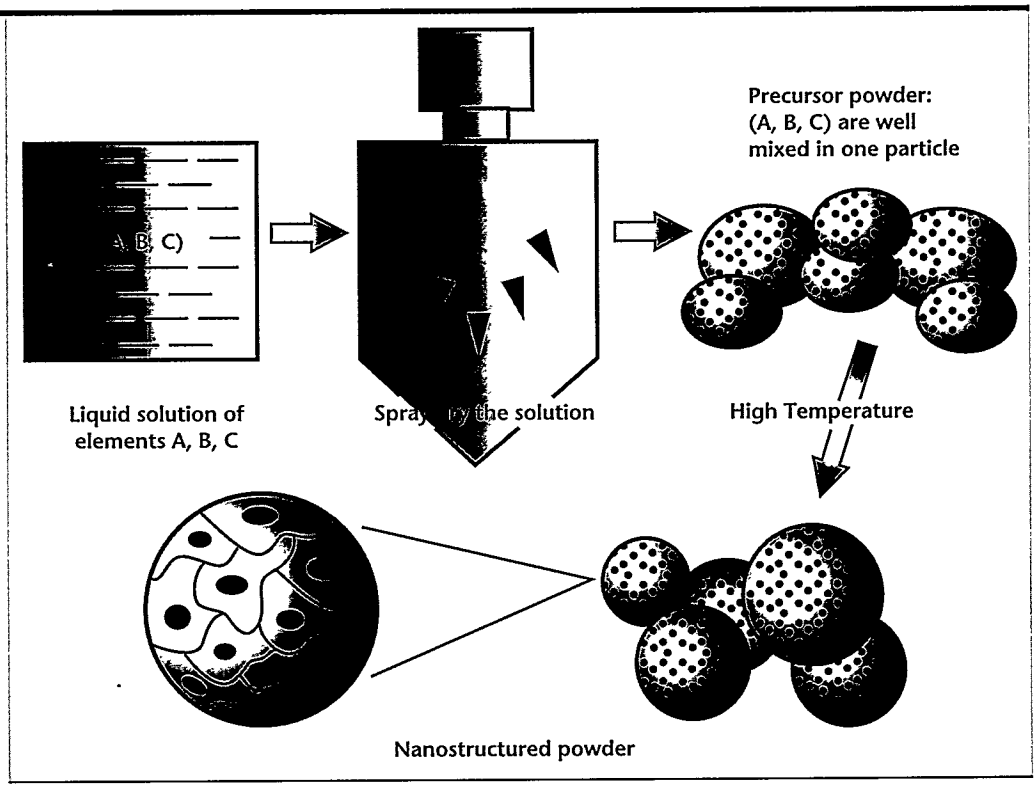
An overview of the synthesis procedure is given schematically in **Fig. 6**. It consists primarily of three steps: 1) solution preparation, 2) precursor synthesis, and 3) thermal treatment.

In the first step, a solution is made of several water soluble compounds containing constituent elements in HA, i.e., calcium, phosphate, and hydroxy ions (Ca^{2+}, $(PO_4)^{-3}$ and OH-). Ca/P is in the ratio of 10/6. The solution is mixed thoroughly in a flask with a magnetic stirrer. Diluted acid and basic solutions are added to the mixture (without changing the concentrations of the constituent elements) to adjust its pH-value to prevent from precipitation.

In the second step, spray-drying of the solution mixture is carried out with a spray dryer. The solution is extracted from a feed tank and passed through the spray nozzle. Compressed air with a selected pressure was used during spraying (atomization). The atomized liquid is rapidly dried by a coaxial flow of air (flow rate = 1.0 m^3/min), preheated to a temperature above 200 °C. The dried powder (i.e., precursor powder) with a submicron size is cyclone-separated from the flowing air stream.

In the third step, the precursor powder mixture is heat-treated into nanocrystalline HA. Heat-treating was performed by placing the precursor powder mixture in an alumina boat in a furnace preheated to a given temperature.

Figure 6. Schematic illustration of the synthesis procedure used in the present study.



Shown in **Fig. 7** is a transmission electron micrograph with the corresponding diffraction pattern from powders synthesized at 700 °C for 1.5 h. The powders are of the order of nanometer in size, and most of them are rod-shaped. HA is naturally crystallized in rectangular (or rod) shape, but coarsens after sintering. The corresponding diffraction pattern indicates the crystallinity of the powders with certain preferred orientations. Shown in **Fig. 8** is a rod-shaped particle (50 nm in lateral dimension). The corresponding diffraction pattern suggests that the particle is single crystal (hexagonal) with a $\langle 2\bar{1}\bar{1}0 \rangle$ zone axis. Preliminary results indicate that the average size of the as-synthesized particles is 23 ± 7 nm. X-ray diffraction from the sprayed powders also confirmed that these particles are crystalline.

To characterize the bond structure, we performed Fourier transform infrared analysis (FTIR) of the powder synthesized at 700 °C for 1.5 h (**Fig. 9**). Most of the peaks in **Fig. 9** are attributed to two kinds of vibration modes: three $(PO_4)^{-3}$ stretching modes, occurring at 962, 1045, and 1090 cm^{-1}; and two $(PO_4)^{-3}$ bending modes, at 569 and 603 cm^{-1}. The stretching modes of OH^- occur at around 3572 and 3641 cm^{-1}, where the powder exhibits relatively sharp absorption peaks. The flexural mode of OH^- at

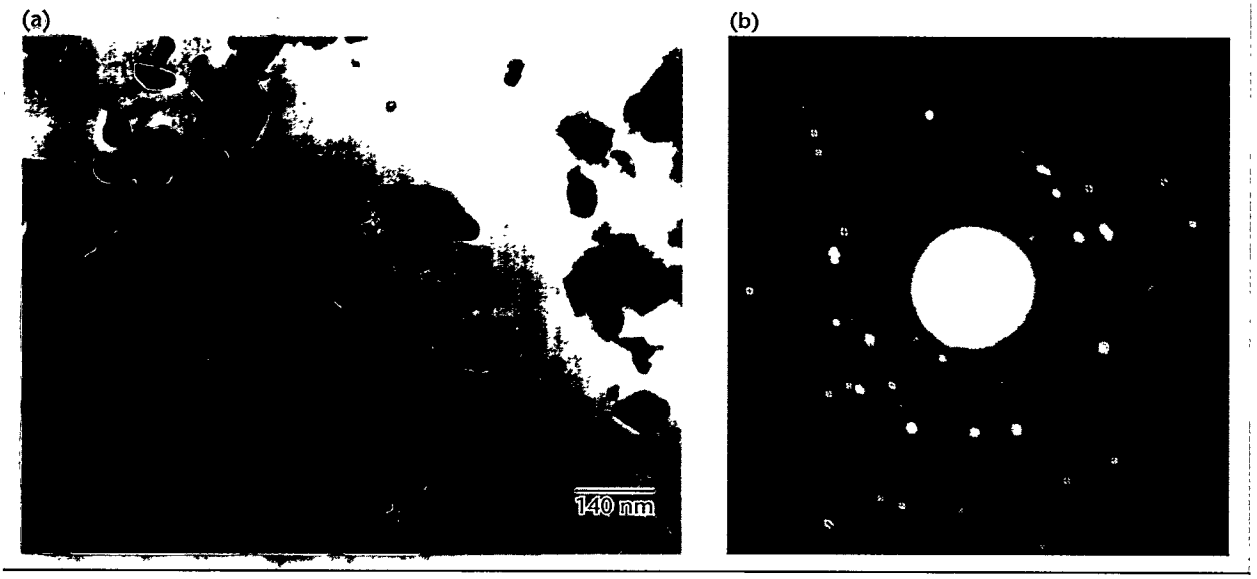


Figure 7. Transmission electron micrograph (a), with the corresponding diffraction pattern (b), from powders synthesized at 700 °C for 1.5 h.

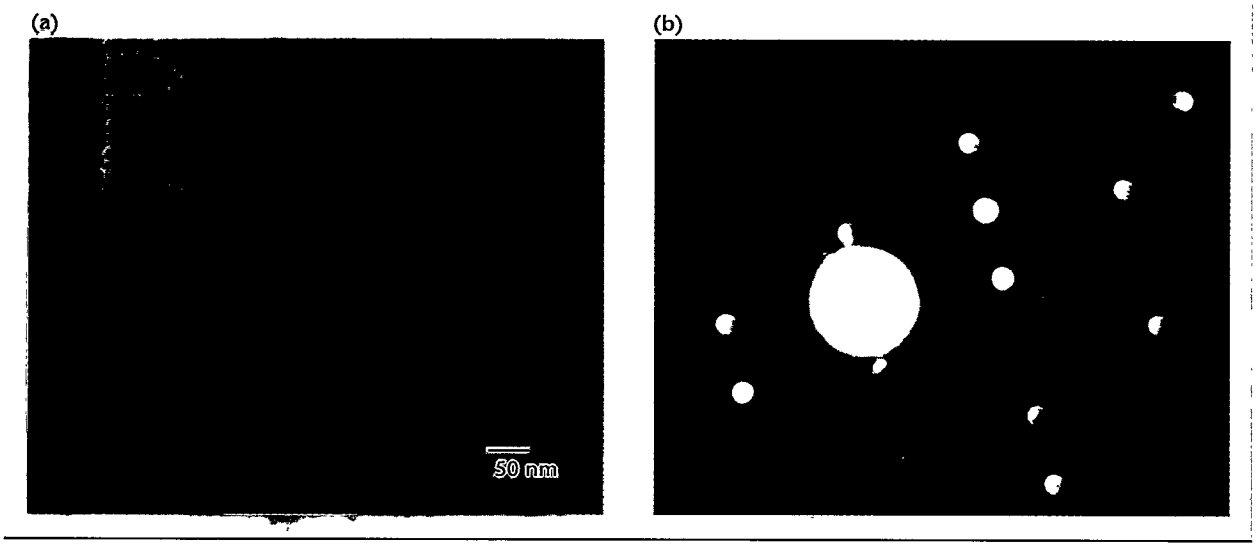


Figure 8. Transmission electron micrograph (a), with the corresponding diffraction pattern (b), for a rod-shaped single-crystal hydroxyapatite particle (50 nm in lateral dimension).

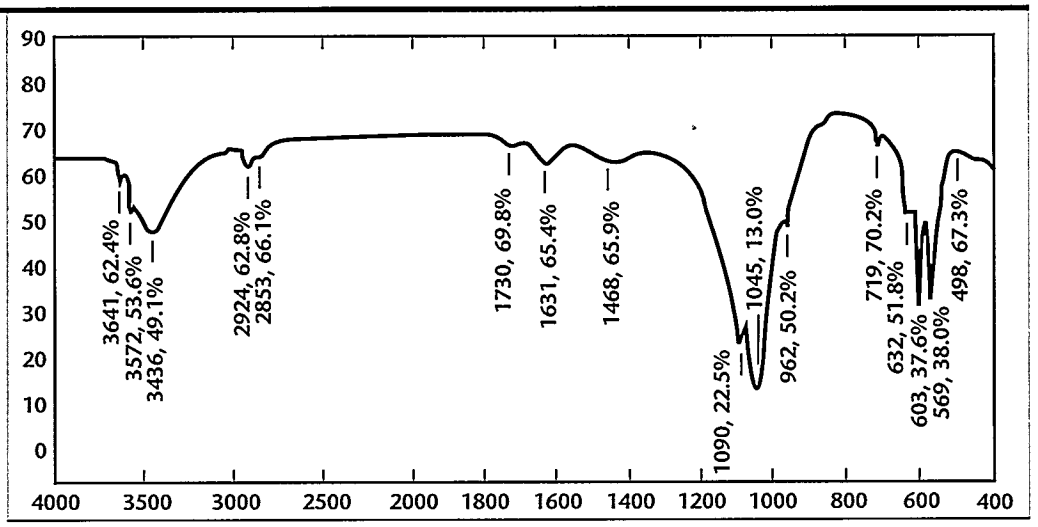


Figure 9. Fourier transform infrared analysis of the powder synthesized at 700 °C for 1.5 h.

632 cm^{-1} is also observed. In addition, small amounts of $(CO_3)^{-3}$ are recorded at 1400 to 1460 cm^{-1}, but their intensities are relatively low.

Future Work

Our plans for FY-96 include 1) to fully develop and characterize HA nanocrystals, and to further understand their thermal stability; 2) to investigate the sintering/consolidation kinetics of HA nanocrystals, in order to control bulk structures with various densities and morphologies; 3) to use various coating techniques to produce graded coatings for Ti alloy implants; and 4) To collaborate with medical clinics/institutes (e.g., Norian, Zimmer, Bio-Rad, Stanford University, and MIT) to identify applications for specific HA structures developed under this project.

Acknowledgments

The authors are indebted to Dr. P. Luo for her innovative work. We also would like to thank M. Wall, and A. Schwartz for their TEM work.

References

1. Hoke, D. A., M. A. Meyers, L. W. Meyer, and G. T. Gray III (1992), "Reaction Synthesis Dynamic Compaction of Titanium Diboride," *Metall. Trans. A*, **23**, pp. 77–86.

2. Gourdin, W. H. (1984), "Energy deposition and microstructural modification in dynamically consolidated metal powders," *J. Appl. Phys.* **55**(1), pp. 172–181.

3. Hohling, H. J., B. A. Ashton, and H. D. Koster (1974), "Quantitative electron microscopic investigations of mineral nucleation in collagen," *Cell Tissue Res.* **29**, pp. 11–26.

4. Hench, L. L. (1991), "Bioceramics—From Concept To Clinic," *J. Am. Ceram. Soc.* **74**, pp. 1487–1510.

5. Tanahashi, M., K. Kamiya, T. Suzuki, and H. Nasu (1992), "Fibrous Hydroxyapatite Grown in the Gel System—Effects of PH of the Solution On the Growth Rate and Morphology," *J. Mater. Sci.—Medicine* **3**, pp. 48–53.

6. Yubao, L., J. d. Wijn, C. P. A. T. Klein, S. v. d. Meer, and K. d. Groot (1994), "Preparation and Characterization of Nanograde Osteopatite-like Rod Crystals," *J. Mater. Sci.—Medicine* **5**, pp. 252–255.

7. Li, P., C. Ohtsuki, T. Kokubo, K. Nakanishi, N. Soga, T. Nakamura, and T. Yamamuro (1993), "Process of Formation of Bone-like Apatite Layer on Silica Gel," *J. Mater. Sci.—Medicine* **4**, pp. 127–131.

8. Steinberg, D. J., and M. W. Guinan (1978), *High-strain-rate Constitutive Model for Metals*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-80465).

Fatigue Behavior of Al Metal Matrix Composite at Elevated Temperatures

Chol K. Syn, Donald R. Lesuer, and David S. Hiromoto
Manufacturing and Materials Engineering Division
Mechanical Engineering

The fatigue properties of a 6090/SiC/25p-T6 metal matrix composite have been evaluated at room and elevated temperatures. In some cases, initiation sites such as cracked SiC particles, could be identified. However, in most cases, fatigue cracks initiated within the matrix, or at random sites less than 2 μm in size. Fatigue crack growth occurred through plastic deformation of the matrix without particle fracture or particle/matrix decohesion. At 300 °C and 450 °C, test samples exhibited large ductility during high-cycle fatigue; deformation strain increased with increasing cycles-to-failure.

Introduction

Discontinuously-reinforced metal matrix composites (MMCs) are attractive for many structural applications, because the materials exhibit unusual combinations of structural, physical, and thermal properties. These properties include low density, low thermal expansion,[1] high modulus and strength, and good creep[2] and wear resistance.[3] The good wear properties of MMCs have led to the commercial production of Al-Al_2O_3 composites as automotive pistons and engine blocks.[4] Applications in many defense and weapons systems have been initiated and promoted because of the materials lightweight, thermal stability and the ability to tailor the thermal expansion coefficient (CTE) of the MMC to that of other materials. These characteristics will lead to applications as structural materials, as well as materials for microelectronic and optical components. In fact, the commercial-scale development of aluminum-based particulate and whisker reinforced MMCs was funded by the Air Force Title III Program. In many of these applications, components are subjected to large, fluctuating thermal-mechanical stresses. Therefore, a clear understanding of the fatigue behavior under such thermo-mechanical conditions is important for more extensive applications.

Room-temperature fatigue properties of the Al-matrix MMC have been investigated extensively.[5–13] However, only limited high-temperature data are available.[9,14–16] Most of these data were obtained by using a small amount of experimental materials produced by individual researchers; hence, the quality of the material was poor, and data have shown large scatter. We have chosen a 6090 Al composite, reinforced with 25 vol.% SiC particulate, a material produced on a commercial scale. We used it as a model material in characterizing fatigue behavior at elevated temperatures under isothermal as well as thermo-mechanical fatigue conditions. We measured fatigue properties at various temperatures. These properties will be discussed and further compared with those measured from other Al-matrix composites.

Progress

Experimental Procedure

Material. The composite used in the present study, 25 vol.% SiC particulate-reinforced 6090 Al (hereinafter, 6090/SiC/25p, after the Aluminum Association nomenclature) was made using powder metallurgy techniques by DWA Composite Specialties (Chatsworth, California). The composition of 6090 base alloy was very close to that of Al 6061, and properties of the 6061 alloy will be used for the properties of the matrix of the 6090 MMC. The material was a U.S. Air Force Title III material,[17] received in the form of a hot rolled 1-cm-thick plate that had a recrystallized microstructure. As a result, microstructure as well as mechanical properties of the material are expected to be consistent. The SiC particles were typically 2 to 5 μm in diameter; however, particles of 10 μm and up in their largest dimension could be detected.

Sample Preparation. Both smooth-bar axial and flat-tapered cantilever fatigue tests were conducted at room and elevated temperatures. The uniaxial samples were fatigue-tested isothermally; the cantilever samples were tested isothermally and non-isothermally. The loading axis was oriented parallel to the rolling direction of the plate for the uniaxial samples, and perpendicular to the plate for the cantilever samples. The uniaxial samples had a uniform gauge section, and the flat cantilever samples had a tapered gauge section with a uniform thickness, as shown in **Fig. 1.** The uniaxial samples were machined on a CNC-equipped lathe with carbide cutting tools from the as-received plate, and lightly ground to remove machining damage. The cantilever samples were electro-discharge machined (EDM) from the plate and lightly ground likewise. The as-ground samples were solution treated at 530 °C for 70 min, quenched in water and then aged at 160 °C for 16 h to achieve a T6 condition for the matrix. The T6 treated samples were polished until no scratches transverse to the long axis of the samples were evident at 30× magnification. Typically, the root mean square surface roughness (Ra) of 0.02 µm was obtained on the samples.

Test Procedures. The uniaxial fatigue tests were conducted in air with commercially-available servo hydraulic equipment, hydraulic grips, and an electric resistance-heating furnace. All uniaxial fatigue testing was done in fully reverse loading (R = –1) under load control (sinusoidal waveform) with an operating frequency of 20 Hz. Some of the uniaxial samples were tensile-tested in air at constant true strain rate, and strain was measured with an extensometer. Cantilever samples had two thermocouples laser-welded on one side. The thermocouples were welded in each of the gauge shoulder areas. One was used to monitor temperature; the other was used for control. Samples were heated or heat-cycled using two infrared flash lamps shining on the side opposite to the thermocouple welded side.

The thermo-mechanical fatigue tests using the cantilever samples were conducted in air with equipment that was designed and built at Lawrence Livermore National Laboratory (LLNL). The apparatus holds the wide end of a cantilever sample stationary, and exerts a cyclic bending force on the narrow end of the sample, generating cyclic tensile stress on the flash-heated side. All cantilever samples were tested in cyclic tension-tension loading (R = 0.1) under sinusoidal load control at a frequency of 20 Hz for both isothermal and thermal cycled tests. For the non-isothermal fatigue tests, typically thermal cycles had a period of 5 min, with 1 min hold at the lower bound

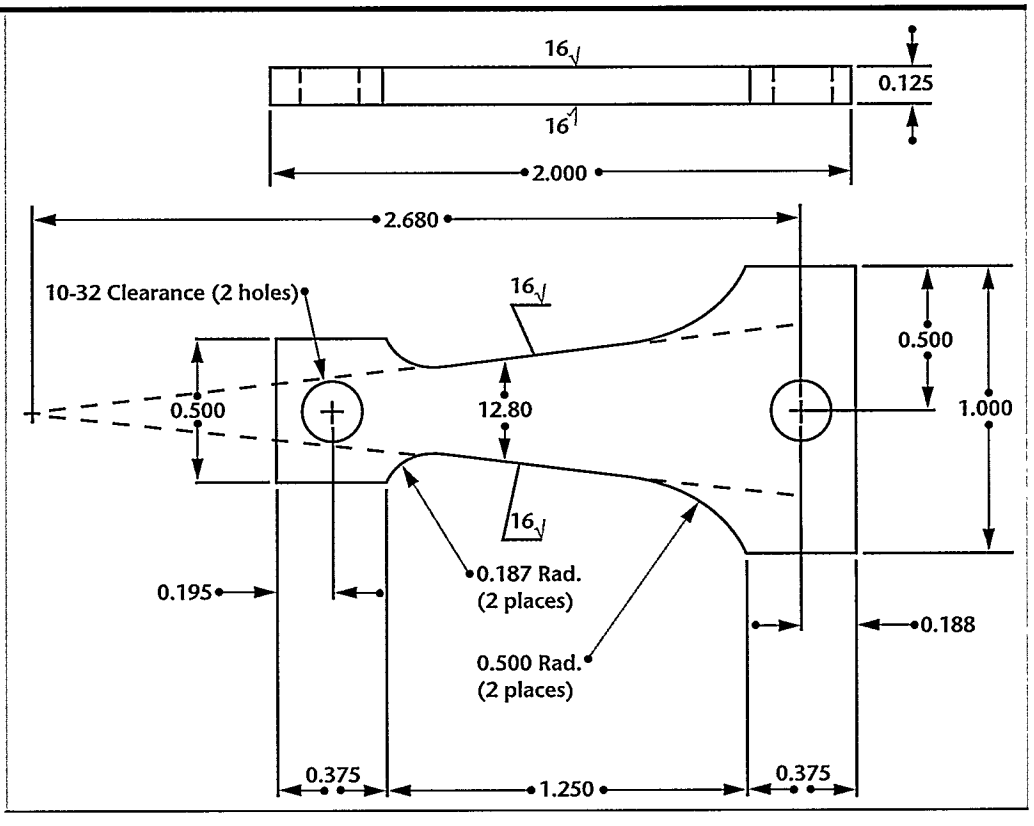


Figure 1. Tapered cantilever sample for bending fatigue tests. Dimensions are in inches.

temperature, 1.5 min heat-up and cool-down between the lower and upper bound temperatures, and 1 min hold at the upper bound temperature.

Fractography. All fatigued and broken samples were put in a scanning electron microscope, and examined to locate the initiation site and to identify the source of the fatigue fracture.

Results and Discussion

Uniaxial Fatigue Tests. The results of fatigue tests performed at room temperature, 200 °C, 250 °C, and 300 °C are shown in **Fig. 2**, where the maximum stress level, S, used in each test, was plotted against the number of fatigue cycles, N. The S-N data for room temperature include the results (open circles) by Harrigan[17] on the same material, and show good agreement with the data from this study. The curves for each test temperature were obtained using a power-law function. We see that the fatigue life (number of cycles to failure) under a given stress or fatigue strength of the 6090/SiC/25p is reduced as the temperature is increased. As the test temperature increases, the slope of the curves gradually decreases, indicating that the curves, especially the three curves for the elevated temperatures, may converge at some point beyond the endurance limit (typically at 10^7 cycles). To see how the S-N behavior of this MMC at room temperature compares with that of the matrix alloy, the room temperature data were replotted together with those of Al 6061-T6 alloy,[18–20] as shown in **Fig. 3**.

It is clear that the fatigue strength is much higher for the MMC than for the 6061 alloy. The data in **Fig. 3** were replotted in **Fig. 4**, after the maximum stress was divided by Young's modulus, to calculate an equivalent elastic strain. **Figure 4** clearly shows that the two curves (one for the MMC, and one for the base alloy) overlap almost completely, indicating that the fatigue behavior of the MMC material is basically controlled by the elastic strain applied to the material. The fatigue strength in the MMC is enhanced essentially by the stiffness increase from the reinforcing SiC.

The room-temperature fatigue data were analyzed further using the measured tensile stress-strain curves of the present MMC, as mentioned earlier. A constitutive relationship of the following type was used to describe the measured stress-strain curves of the MMC.

$$\varepsilon = \sigma/E + \alpha \cdot \varepsilon_0 \cdot (\sigma/\sigma_0)^n \qquad (1)$$

where ε and σ are strain and stress, E is Young's modulus, n is the strain hardening coefficient, σ_0 and ε_0 are stress and strain at the proportional limit on the stress-strain curve, respectively, and α is a constant. For the present MMC at room temperature, the values E = 114 GPa, α = 1.08, n = 13.9, ε_0 = 0.0001, and σ_0 = 345 MPa were used.

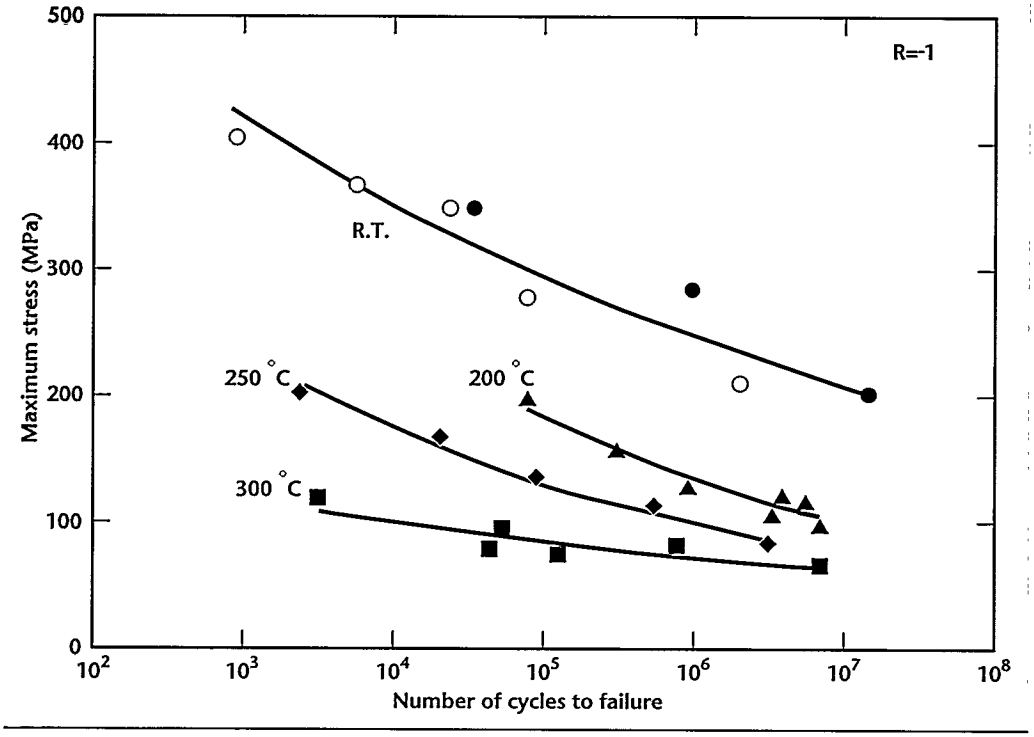


Figure 2. Maximum stress vs number of cycles to failure (S-N curve) for 6090/SiC/25p-T6 Al MMC at room and elevated temperatures.

The constitutive equation **(Eq. 1)** enabled calculation of elastic and plastic strain components of the fatigue sample at a given stress. The equivalent strain for the MMC at room temperature shown in **Fig. 4** was thus divided into the elastic and plastic strain components, as shown in **Fig. 5. Figure 5** clearly shows that the MMC fatigue is controlled by the elastic component in the high cycle region (beyond 10^4 cycles). A similar analysis is being attempted for elevated-temperature fatigue. The data shown in **Fig. 2** were replotted in **Fig. 6**, in a similar manner to **Fig. 4**. The three elevated

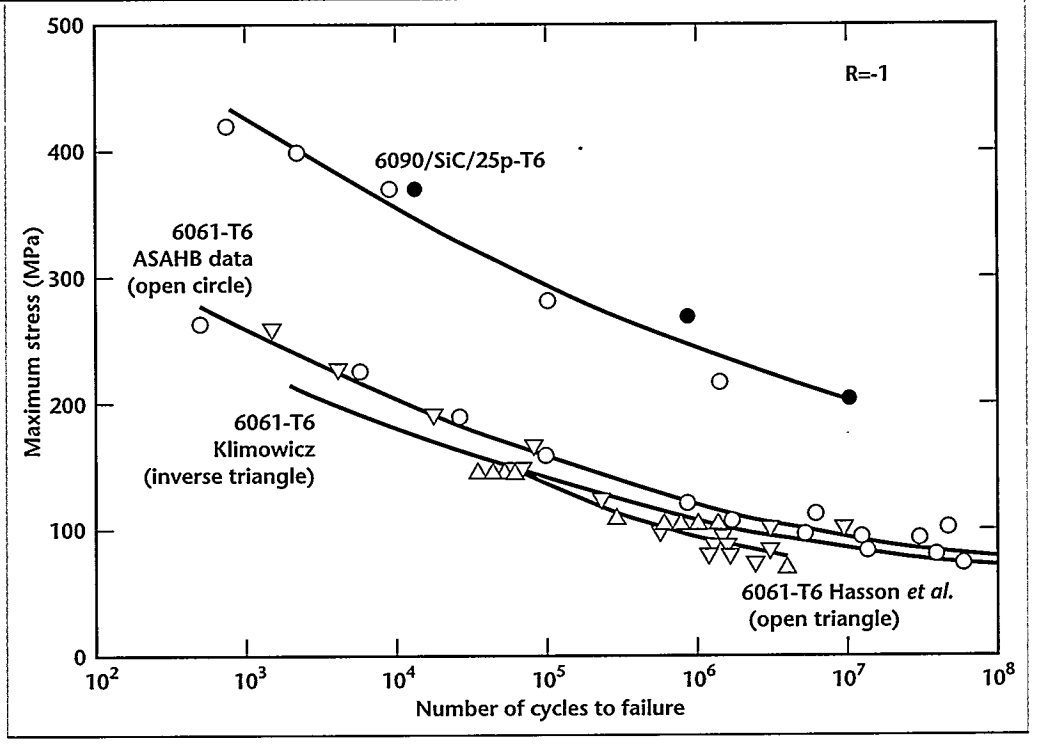


Figure 3. S-N curves for 6090/SiC/25p-T6 and 6061-T6 at room temperature. Data for 6061-T6 were from other sources, as indicated.

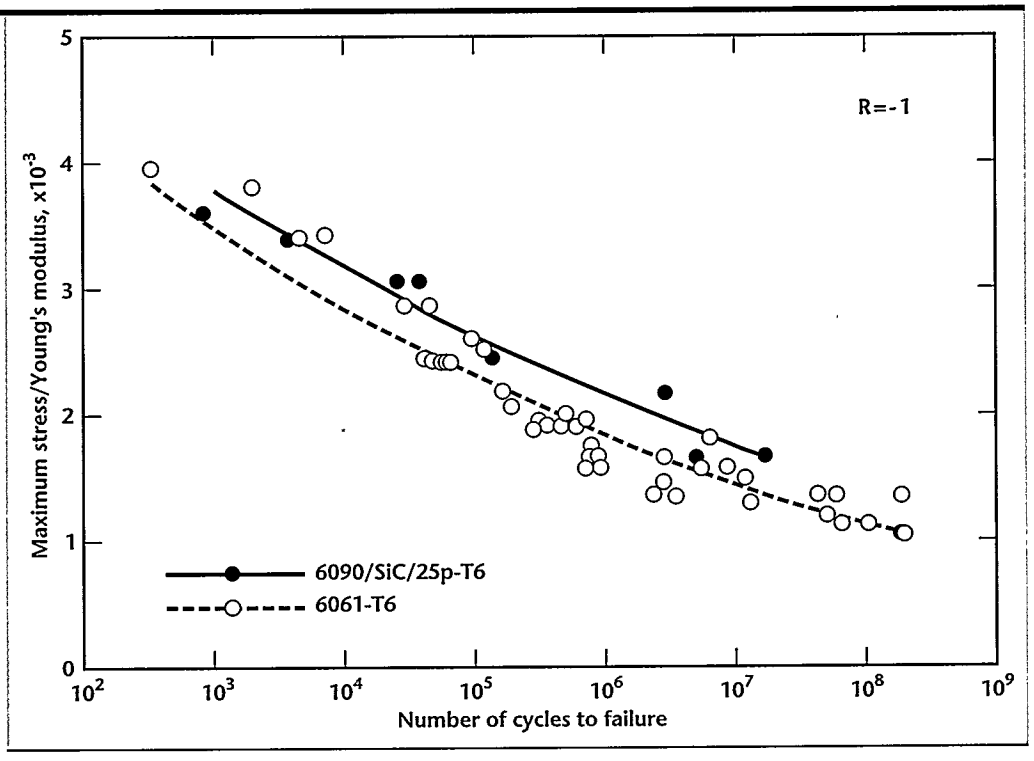


Figure 4. Maximum stress/Young's modulus vs number of cycles to failure (S/E-N curves) for 6090/SiC/25p-T6 and 6061-T6 at room temperature calculated from data shown in Figure 3.

temperature curves in **Fig. 6** show a converging trend. When they were extrapolated to the right, they were found to converge at about 0.00025 strain. This convergence indicates that, at very small stresses, the elevated-temperature fatigue may be controlled by a single critical failure strain, at least, for the 200 to 300 °C temperature range.

Cantilever Bending Fatigue Tests. Only a limited number of bending fatigue tests including both isothermal- and thermal-cycled tests were performed at R = 0.1. The results are summarized in **Fig. 7**. The isothermal bending S-N curve for 300 °C shows higher stress levels and a slightly steeper slope than the one for 450 °C, as observed in the uniaxial tensile

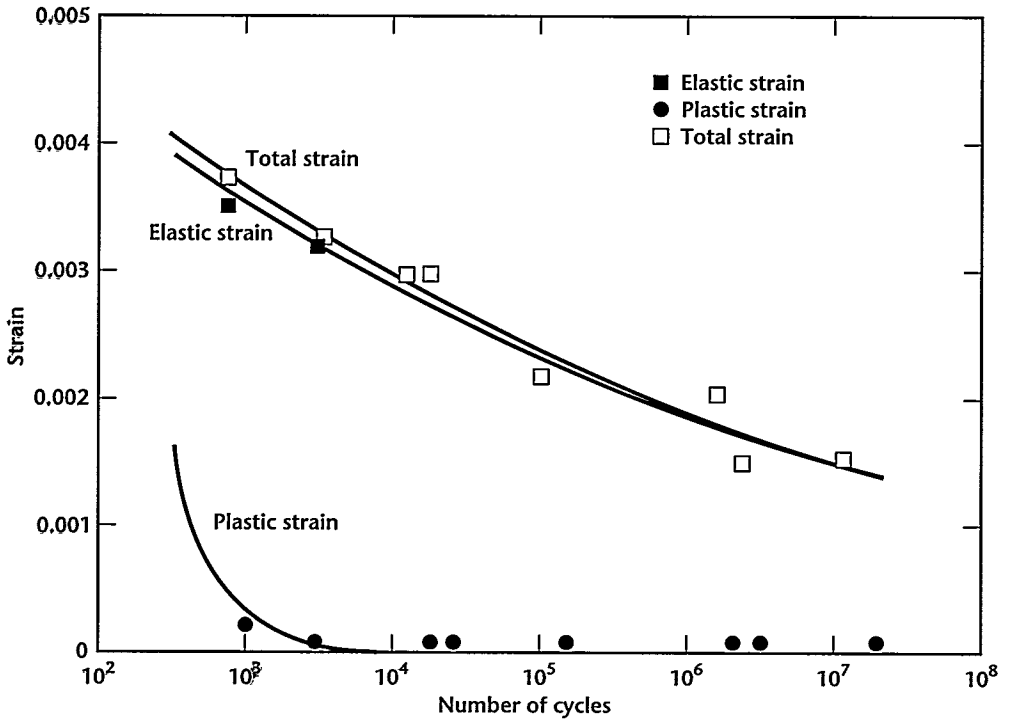


Figure 5. Strain vs number of cycles to failure curve for 6090/SiC/25p-T6 Al MMC at room temperature. Strains were calculated from data shown in Figure 4 using Equation 1.

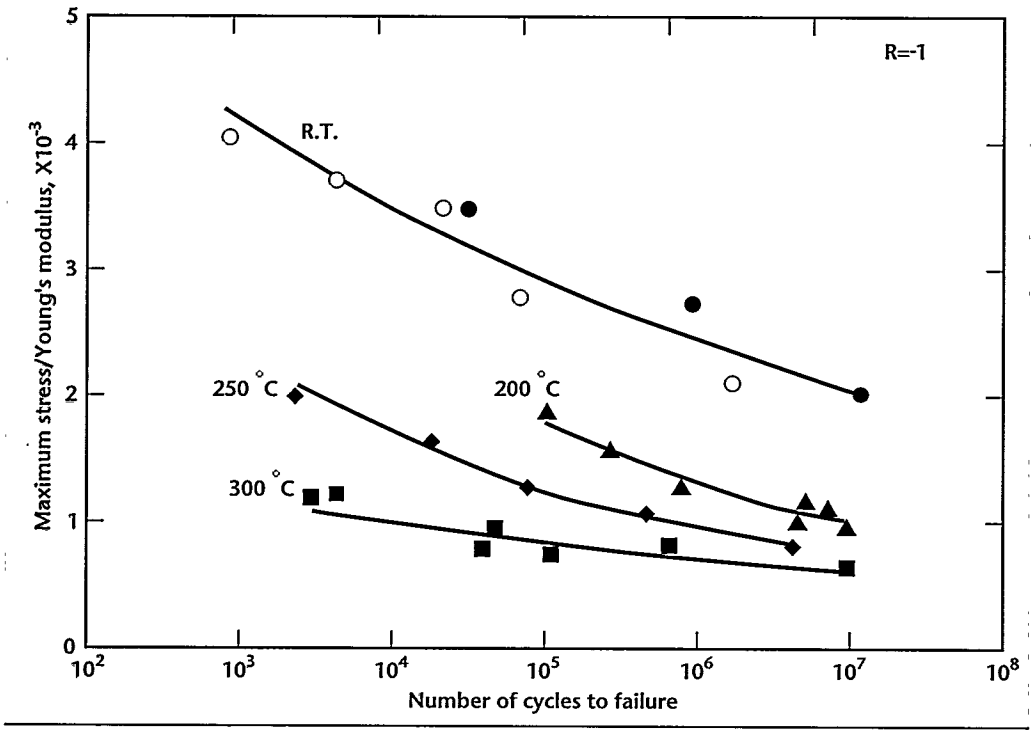


Figure 6. S/E-N curves for 6090/SiC/25p-T6 at room temperature and elevated temperatures, calculated from data shown in Figure 2.

fatigue tests. However, the direct comparison between the results of cantilever bending and uniaxial tensile fatigue tests are not possible, because different R ratios (R = 0.1 and –1) were used.

The three data points representing the thermal-cycled fatigue tests (one for 100 to 450 °C, and two for 25 to 375 °C thermal cycling) show interesting trends. The three tests were done under the same temperature excursion of 350 °C, while their lower and upper bound temperatures and their average temperatures (275 °C and 200 °C) are quite different. The data points for these two thermal-cycling tests are roughly the same and overlap with those of the 300 °C isothermal S-N curve. Thus, if the average

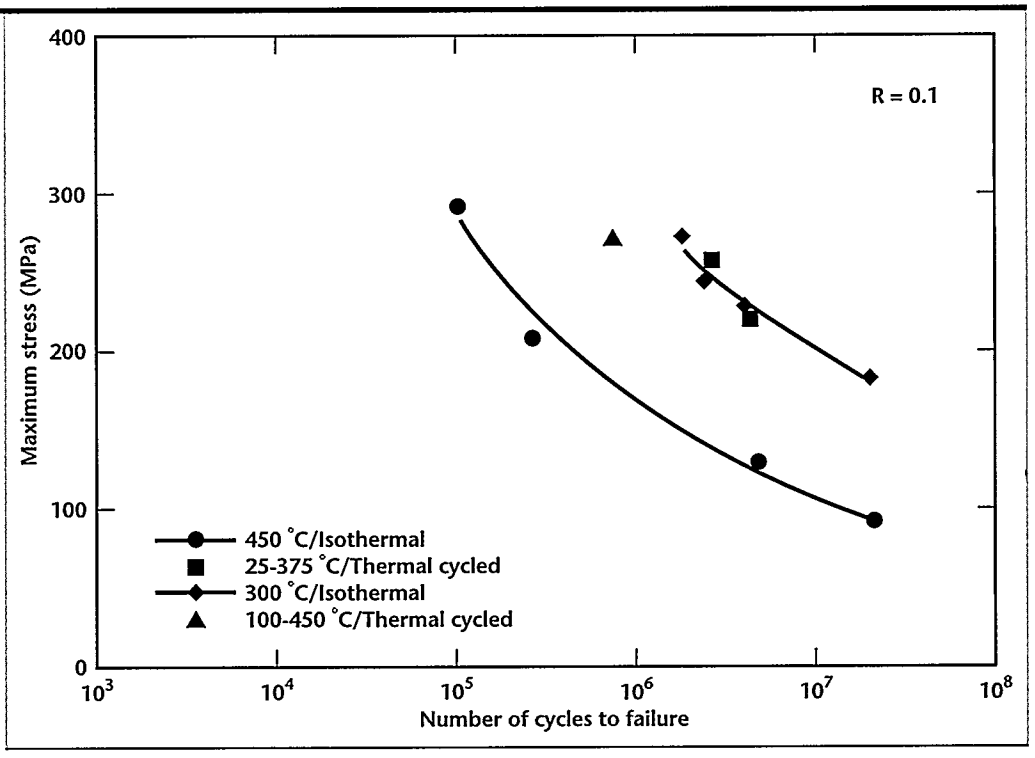


Figure 7. S-N curves and data for cantilever bending fatigue test samples. Three data points indicated by arrows represent one thermal-cycled test at 100 to 450 °C and two tests at 25 to 375 °C.

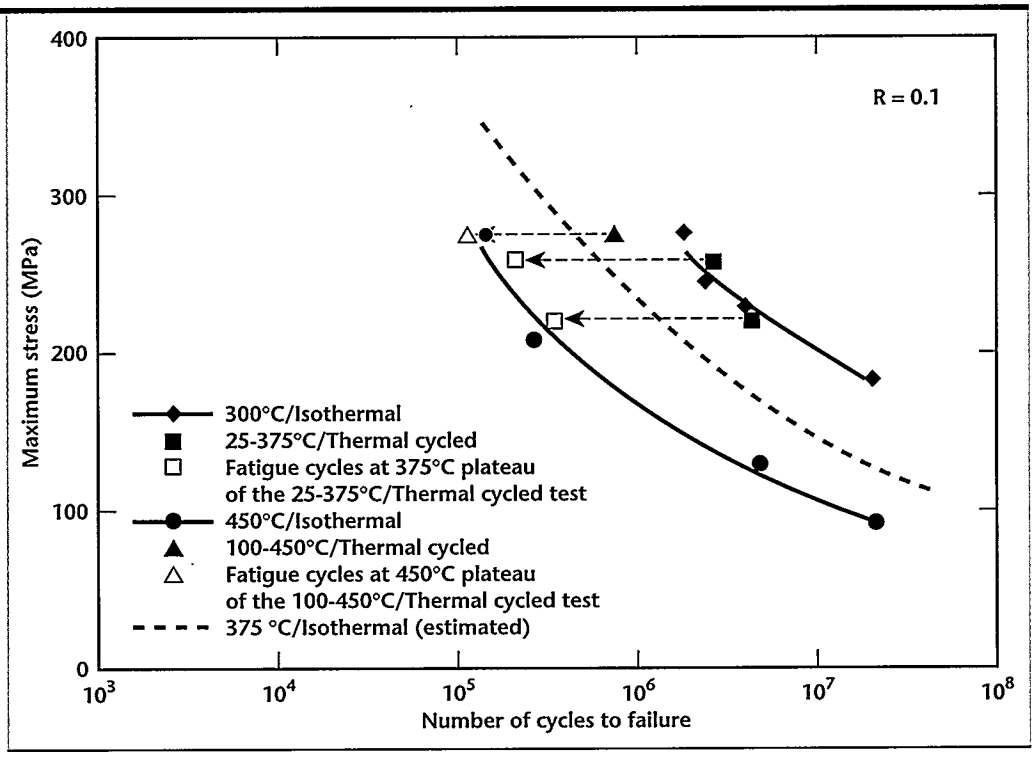


Figure 8. S-N curves, as in Figure 7. Dotted arrows represent the numbers of cycles at the upper bound temperature portion (upper plateau) of the thermal cycling of the three thermal cycled tests shown in Figure 7.

temperature of the cycling is compared with that of isothermal test temperature, it is clear that thermal cycling accelerated the development of fatigue damage in the material.

Analysis of the data point for the 100 to 450 °C thermal-cycled test shows that the number of fatigue cycles during the 450 °C plateau portion of the thermal cycling falls on the 450 °C isothermal fatigue S-N curve, as indicated by an arrow in **Fig. 8**.

On the other hand, the data points for the 25 to 375 °C thermal-cycled tests show a rather different behavior. The number of fatigue cycles during the 375 °C plateau portion of the thermal cycling were much lower than the number of cycles for the corresponding stress on the 375 °C/isothermal fatigue S-N curve. The 375 °C/isothermal S-N curve was estimated by taking the average of the 300 °C/isothermal and 450 °C/isothermal curves. This difference between the 100 to 450 °C and 25 to 375 °C thermal-cycled tests may be explained by considering the fatigue damage generated during the lower temperature period of the thermal cycling, and its recovery, or healing, during the upper bound temperature period of the cycling. For the 100 to 450 °C-cycled test, most of the fatigue damage might have been completely healed, and the damage occurring only during the 450 °C period accumulated and led to failure. On the other hand, in the 25 to 375 °C-cycled tests, the damage accumulated below the 375 °C plateau might not have been completely healed and eliminated, because the slower recovery kinetics at 375 °C might not have been as effective as that at 450 °C.

Inspection of all thermal-cycled and isothermal fatigue test samples at 300 °C and 450 °C showed that a large measurable amount of bending occurred in each of the samples, and the amount increased with increasing number of cycles to failure. From measurement of the bending strain in each sample, the amount of creep at elevated temperatures can be estimated and related to the fatigue characteristics of the material.

Origin of Fatigue Fracture. SEM fractography of the fatigue-broken samples showed that the fracture initiation point could not be clearly identified in the majority of the samples tested at room temperature to 300 °C, indicating that fracture was initiated in the matrix, or from SiC particles less than about 2 μm in size. The lack of defect-originated fracture also reflects the defect-free, high-quality of the MMC material used in this study. Fatigue crack growth must have occurred through plastic deformation of the matrix, without particle fracture or particle/matrix decohesion. In a few cases, extra large SiC particles were either fractured or debonded from the matrix. In a few other cases, fracture was shown to initiate at a site where the consolidation of the starting powders was incomplete, creating a pocket of microporosity. Cantilever bending fatigue samples tested at 450 °C showed simultaneous fracture initiation at two to four different sites, and at each site, MgO and other contamination were detected.

Acknowledgments

The authors would like to thank M. Thiry, J. Groth and W. Andrade for machining test samples; E. Ziebarth for laser-welding thermocouples; and J. Ferreira and E. Sedillo for SEM fractography.

References

1. Vaidya, R. U., and K. K. Chawla (1994), "Thermal Expansion of Metal-Matrix Composites," *Comp. Sci. Technol.* **50**, pp. 13–22.

2. Nieh, T. G. (1984), "Creep Rupture of a SiC Reinforced Aluminum Composite," *Metall. Trans. A* **15A**, pp. 139–146.

3. Wang, A., and H. J. Rack (1992), "A Statistical Model for Sliding Wear of Metals in Metal/Composite Systems," *Acta Metall. Mater.* **40**(9), pp. 2301–2305.

4a. Takenaka, T., and M. Fukazawa (1988), "Fabrication of SiC Whisker Reinforced Al Alloys and Their Applications," in *Future Transportation Technology Conference and Exposition* (SAE Technical Paper **881189**, SAE International, Warrendale, Pa.).

4b. Dinwoodie, J. (1987), "Automotive Applications for MMCs Based on Short Stable Alumina Fibers," in *International Congress and Exposition* (SAE Technical Paper **870437**, SAE International, Warrendale, Pa.).

5. Ackermann, L., J. Charbonnier, G. Desplanches, and H. Koslowski (1985), "Properties of Reinforced Aluminum Foundry Alloys," in *5th Inter. Conf. on Composite Materials*, W. C. Harrigan, Jr., J. Strife, and A. K. Dhingra, Eds. (The Metallurgical Society, Warrendale, Pa.), pp. 687–698.

6. Harris, S. J. (1988), "Short Fiber and Particulate Reinforced Metal Matrix Composities," NATO AGARD-CP-444, U.S. Department of Commerce, pp. 25.1–25.13.

7. Llorca, N., A. Bloyce, and T. N. Yue (1991), "Fatigue Behavior of Short Alumina Fiber Reinforced AZ91 Magnesium Alloy Metal Matrix Composite," *Mater. Sci. Eng.* **A135**, pp. 247–252.

8. Bonnen, J. J., J. E. Allison, and J. W. Jones (1991), "Fatigue Behavior of a 2XXX Series Aluminum Alloy Reinforced with 15 Vol PCT SiCp," *Metall. Trans. A* **22A**, pp. 1007–1017.

9. Hunt, Jr., W. H., C. R. Cook, and R. R. Sawtell (1991), "Cost-Effective High Performance P/M Aluminum Matrix Composites for Automotive Applications," in *International Congress and Exposition*, (SAE Tech. Paper **910834**, SAE, Warrendale, Pa.).

10. Saito, N., Y. Nishida, M. Nakanish, and Y. Sakai (1991), "Fatigue of Aluminum-Borate Whisker-Reinforced Aluminum Alloy," *Science and Engineering of Light Metals*, K. Hirano, H. Oikawa, and K. Ikeda, Eds. (Japan Inst. of Light Metals, Tokyo) pp. 507–512.

11. Levin, M., and B. Karlsson (1993), "Fatigue Behavior of a Saffil-reinforced Aluminum Alloy (AA 6061)," *Composites* **24**, pp. 288–295.

12. Vyletel, G. M., J. E. Allison, and D. C. VanAken (1993), "The Influence of Matrix Microstructure and TiC Reinforcement on Cyclic Response and Fatigue Behavior of 2219 Al," *Metall. Trans. A* **24A**, pp. 2545–2557.

13. Maruyama, N., M. Sumita, and K. Nakazawa (1993), "Fatigue and Fretting Fatigue Behavior of Al 2024-T6 Composite Materials Reinforced with 20 Vol% SiC Particles," *J. JIM* **57**(11), pp. 1268–1274.

14. Spengler, W. G., and W. B. Young (1986), "Techniques to Upgrade Heavy Duty Aluminum Pistons," in *International Congress and Exposition*, (SAE Technical Paper **860162**, SAE International, Warrendale, Pa.).

15. Donomoto, T., N. Miura, K. Funatani, and N. Miyake (1983), "Ceramic Fiber Reinforced Piston for High Performance Diesel Engines," in *International Congress and Exposition*, (SAE Technical Paper **830252**, SAE International, Warrendale, Pa.).

16. Soave, R., J. M. M. Leites, G. Ferran, C. W. Kuhn, and G. A. Gegel (1994), "Development of a Fiber Reinforced Aluminum Piston for Heavy-Duty Deisel Engines," in *International Congress and Exposition*, (SAE Tech. Paper **940584**, SAE International, Warrendale, Pa.).

17. Harrigan, Jr., W. C. (1991), "Scaling-Up Particulate-Reinforced Aluminum Composites for Commercial Production," *JOM* **43**(8), pp. 32–35.

18. Metals & Ceramic Information Center (1992), *Aerospace Structural Alloys Handbook*, J. H. Holt, H. Mindlin, and C. Y. Ho, Eds. (CINDAS/Purdue University, West Lafayette, Ind.).

19. Klimowicz, T. F. (1989), data obtained from Dural Aluminum Composites Corp., San Diego, Calif.

20. Hasson, D. F., D. R. Crowe, J. S. Ahearn, and D. C. Cooke (1984), "Failure Mechanisms in High Performance Materials," *Proceedings from 29th Mechanical Failures Prevention Group Meeting*, J. G. Early, T. R. Shives, and J. H. Smith, Eds., (NBS), pp. 147–156.

Pearlite in Ultrahigh Carbon Steels: Heat Treatments and Mechanical Properties

Eric M. Taleff
University of Texas
Austin, Texas

Chol K. Syn and Donald R. Lesuer
Manufacturing and Materials Engineering Division
Mechanical Engineering

Oleg D. Sherby
Stanford University
Stanford, California

Two ultrahigh carbon steel (UHCS) alloys containing 1.5 and 1.8 wt.% carbon, respectively, were studied. These materials were processed into fully spheroidized microstructures, and were then given heat treatments to form pearlite. The mechanical properties of the heat-treated materials were evaluated by tension tests at room temperature. Use of the hypereutectoid austenite-cementite-to-pearlite transformation enabled us to achieve pearlitic microstructures with various interlamellar spacings. The yield strengths of the pearlitic steels are found to correlate with a predictive relation based on interlamellar spacing and pearlite colony size. Decreasing the pearlite interlamellar spacing increases the yield strength and the ultimate strength, and decreases the tensile ductility. We have shown that solid solution alloying strongly influences the strength of pearlitic steels.

Introduction

Ultrahigh carbon steels (UHCS), which are hypereutectoid steels, have been a topic of much interest for their unique mechanical properties. Particularly of recent interest are the high strengths that UHCS materials exhibit. A majority of the studies concerning room-temperature strength in UHCS alloys have concentrated on fully spheroidized microstructures.[1,2] Another microstructural condition yielding interesting mechanical properties at room temperature is pearlite, which can be formed by heat treatment from a spheroidized microstructure. Because UHCS materials lie in the hypereutectoid region of the Fe-C phase diagram, the pearlite transformation is different from that in hypereutectoid steels. In austenitizing unalloyed UHCS, the amount of carbon dissolved in austenite can be varied from the eutectoid composition of 0.77 wt.% at the A_1 temperature, to a maximum of 2.1 wt.% at the A_{cm} temperature at the E point (the point of maximum solubility of carbon in austenite). Upon air cooling of UHCS below the A_1 temperature, austenite with dissolved carbon is converted into pearlite. As the amount of carbon dissolved in austenite is increased, fewer cementite particles will be retained from the original spheroidized microstructure, and their sizes will be smaller. An increase in the amount of carbon dissolved in austenite prior to transformation will result in a decrease in the interlamellar spacing of the transformed pearlite. In austenite at or below the A_1 transformation temperature, an increase in dissolved carbon increases the driving force for pearlite formation, resulting in a decrease of the interlamellar spacing for nucleated pearlite. This process also allows for the creation of various ratios of spheroidized carbide to lamellar pearlitic carbide, from an initially spheroidized microstructure.

Progress

The two alloys studied contain 1.5 wt.% C (UHCS-1.5C) and 1.8 wt.% C (UHCS-1.8C). Each alloy also contains 1.6 wt.% Al, 1.5 wt.% Cr, and 0.5 wt.% Mn.

The addition of Mn minimizes the deleterious effects of sulfur and phosphorus. The addition of Cr helps to prevent graphitization and stabilizes the carbides, making coarsening more difficult. The addition of Al inhibits the formation of a hypereutectoid carbide network and stabilizes the ferrite phase. The addition of 1.6 wt.% Al raises the Al transformation temperature from 727 °C to approximately 780 °C. **Figure 1** shows an approximation of the Fe-C phase diagram for alloys containing 1.6 wt.% Al, 1.5 wt.% Cr and 0.5 wt.% Mn[3]. The two UHCS alloys of this study are denoted in **Fig. 1** by vertical lines at their respective carbon contents.

Experimental Procedure

Each of the two UHCS materials was cast into an ingot and then hot-forged into bars of 51 mm × 51 mm cross section. Sections of 51-mm length were cut from each bar and processed by hot-and-warm working (HWW) into plates for making tensile samples. The HWW processing involves rolling the material from above the A_{cm} temperature until it reaches the A_1 temperature. This processing procedure substantially refines the microstructure. The UHCS-1.8C material was subsequently given a divorced-eutectoid transformation (DET) in order to fully spheroidize the microstructure. DET involves heating just above the A_1 temperature and air cooling.[4] The UHCS-1.5C material was given a divorced-eutectoid transformation with associated deformation (DETWAD) treatment in order to create a spheroidized microstructure. DETWAD involves heating slightly above the A_1 temperature, as in a DET, and rolling while cooling.[4] The results for both materials were fine, spheroidized carbide microstructures.

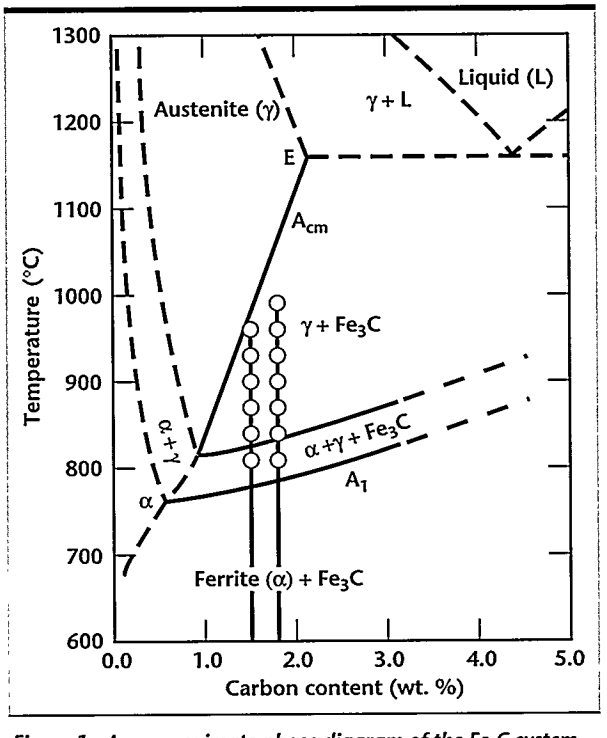


Figure 1. An approximate phase diagram of the Fe-C system for alloys containing 1.6 wt.% Al, 1.5 wt.% Cr, and 0.5 wt.% Mn. Open circles represent austenitizing temperatures selected in this study.

The processed plates of each material were machined into tensile coupons with gauge lengths of 25.4 mm. Machined samples were sealed in stainless-steel bags for heat treatment. Powders of titanium were inserted into each bag, in order to absorb oxygen during heating and prevent oxidation of the samples. Samples were austenitized for 20 to 30 min at temperatures between the A_1 and A_{cm} temperatures. Each sample was air-cooled after austenitizing, in order to form pearlite. Samples of UHCS-1.5C were austenitized at temperatures of 810, 840, 870, 900, 930, and 960 °C. Samples of UHCS-1.8C were austenitized at temperatures of 810, 840, 870, 900, 930, 960, and 990 °C. The austenitizing temperatures for each material are indicated by open circles in the approximated phase diagram of **Fig. 1**. The samples were tested in tension to failure, and strain was measured by an extensometer. After failure, microscopy samples were taken from the undeformed grip region of specimens from each heat treatment. Microscopy samples were etched with nital, given a conductive coating, and analyzed in a scanning electron microscope (SEM).

Results

The microstructures resulting from each heat treatment in the UHCS-1.5C material are shown in **Fig. 2**. **Figures 2(a)** through **2(f)** show the effects of progressively higher austenitizing temperatures, from 810 °C for **Fig. 2(a)** to 960 °C for **Fig. 2(f)**. These microstructures show an increase in the amount of pearlitic carbide, and a decrease in the amount of spheroidized carbide with increasing austenitizing temperature. After austenitizing at 810 °C and air cooling, the microstructure is almost completely spheroidized, because of the DET process [**Fig. 2(a)**], and after austenitizing at 960 °C and air-cooling, the microstructure is nearly all pearlite [**Fig. 2(f)**]. As the quantity of pearlitic carbide increases in comparison with spheroidized carbide, the pearlite interlamellar spacing decreases significantly. Because some spheroidized carbides remain undissolved during all austenitizing treatments below the A_{cm} line, grain growth is inhibited, and a fine pearlite colony size is obtained after transformation.

The microstructures resulting from each heat treatment of UHCS-1.8C are shown in **Fig. 3**. **Figure 3(a)** shows the DET microstructure. **Figures 3(b)** through **3(h)** show the transformed microstructures from austenitizing treatments at 810 °C [**Fig. 3(b)**] through 990 °C [**Fig. 3(h)**]. The same trends observed in UHCS-1.5C are observed for the UHCS-1.8C material. The UHCS-1.8C material is fully spheroidized after the 810 °C heat treatment [**Fig. 3(b)**], because of the DET process. After austenitizing at 990 °C the microstructure is fully pearlitic, with extremely fine interlamellar spacings [**Fig. 3(h)**]. The pearlite interlamellar spacings of each heat treatment for the UHCS-1.5C and UHCS-1.8C materials were measured from the finest observable interlamellar spacing under the SEM, and are given in **Table 1**. The pearlite colony sizes are also listed in **Table 1**. For these two UHCS alloys, the pearlite colony size is considered effectively the same as the prior austenite grain size. The relationship between pearlite interlamellar spacing, λ, and austenitizing temperature is shown in **Fig. 4**. Of note is the sharp initial drop in interlamellar spacing with increasing austenitizing temperature, until fine spacings of less than 0.1 μm are reached.

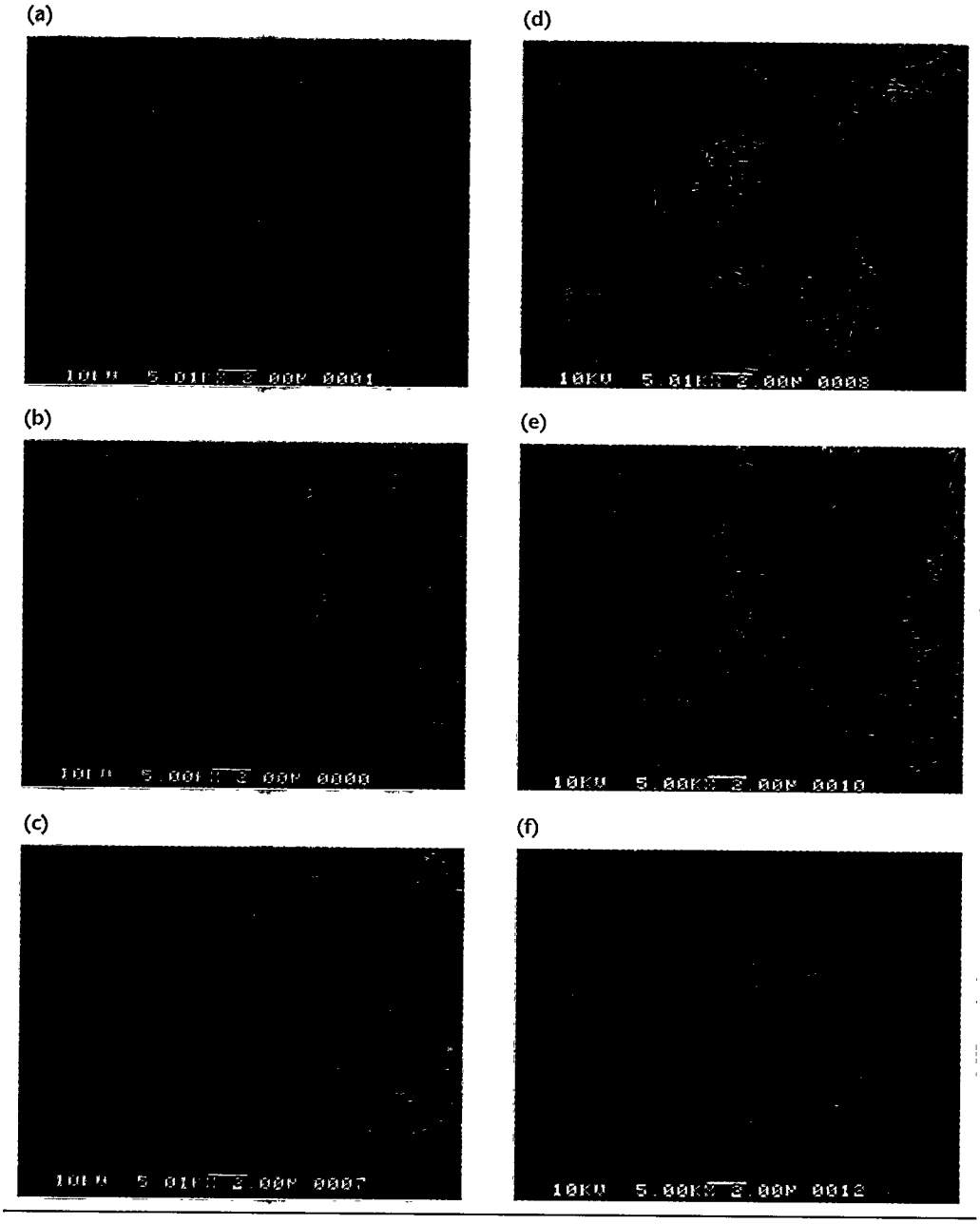


Figure 2. SEM photomicrographs of the UHCS-1.5C samples transformed to pearlite after austenitizing at (a) 810 °C, (b) 840 °C, (c) 870 °C, (d) 900 °C, (e) 930 °C, and (f) 960 °C.

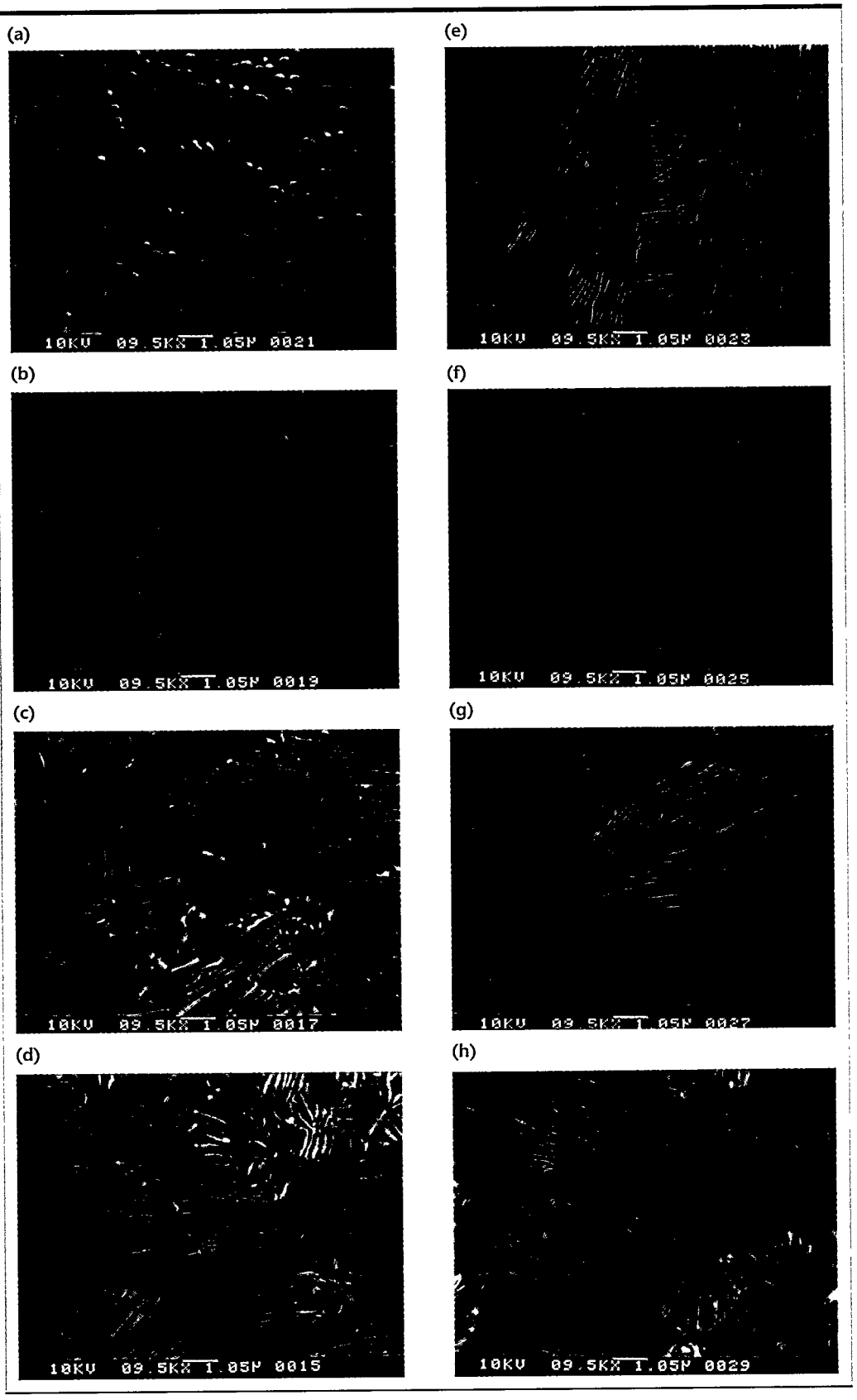


Figure 3. SEM photomicrographs of the UHCS-1.8C samples transformed from (a) the DET structure to pearlite after austenitizing at (b) 810 °C, (c) 840 °C, (d) 870 °C, (e) 900 °C, (f) 930 °C, (g) 960 °C, and (h) 990 °C.

Data from tensile tests of the heat-treated UHCS-1.5C samples are given in **Fig. 5** as a plot of engineering stress vs engineering strain. Of note is the maximum ductility of 15% obtained for the sample austenitized at 840 °C. Also of note are the very high yield and ultimate tensile strengths observed, with yield strengths of over 1000 MPa for the samples austenitized at 930 °C and 960 °C. Data from tensile tests of the heat-treated UHCS-1.8C samples are given in **Fig. 6** as a plot of engineering stress vs engineering strain. These data show a maximum ductility of between 13% and 14% for the samples austenitized at 810 °C and 840 °C. Remarkably high yield strengths of 1200 MPa are observed for the samples austenitized at 960 °C and 990 °C. Very high ultimate strengths are observed for both the UHCS-1.5C and UHCS-1.8C materials, and are limited by fracture to a maximum value of 1600 MPa, for the case of the sample austenitized at 990 °C.

Discussion

Yield strength was found to correlate well with the inverse square root of interlamellar spacing, λ. **Figure 7** shows this good correlation on a plot of yield strength vs inverse square root of interlamellar spacing for both the UHCS-1.5C and UHCS-1.8C materials. The observed relationship has been used in other studies describing pearlite strength, and follows a Hall-Petch type relation, given as:

Table 1. The finest measured pearlite interlamellar spacing (λ), pearlite colony size (L), measured yield strength $(\sigma_o)_{ss}$ inverse square root of interlamellar spacing ($\lambda^{-1/2}$), and D_s^* for each UHCS material and soaking temperature (T). The standard deviations calculated from microstructural measurements are given.

Material	T (°C)	λ (μm)	L (μm)	σ_y (MPa)	$\lambda^{-1/2}$ ($\mu m^{-1/2}$)	D_s^* (μm)
UHCS-1.5C	840	0.18 ±0.02	2.2 ±0.2	864	2.36	0.25
UHCS-1.5C	870	0.18 ±0.04	2.6 ±0.2	889	2.36	0.25
UHCS-1.5C	900	0.13 ±0.01	2.7 ±0.4	950	2.77	0.18
UHCS-1.5C	930	0.11 ±0.01	2.8 ±0.4	1038	3.02	0.16
UHCS-1.5C	960	0.09 ±0.01	3.6 ±0.6	1038	3.33	0.13
UHCS-1.8C	840	0.25 ±0.05	2.2 ±0.4	800	2.00	0.35
UHCS-1.8C	870	0.19 ±0.04	2.2 ±0.2	934	2.29	0.27
UHCS-1.8C	900	0.11 ±0.01	2.6 ±0.6	957	3.02	0.16
UHCS-1.8C	930	0.095 ±0.006	3.6 ±0.9	1099	3.24	0.13
UHCS-1.8C	960	0.069 ±0.003	4.8 ±1.8	1204	3.81	0.10
UHCS-1.8C	990	0.065 ±0.003	5.4 ±3.9	1204	3.92	0.09

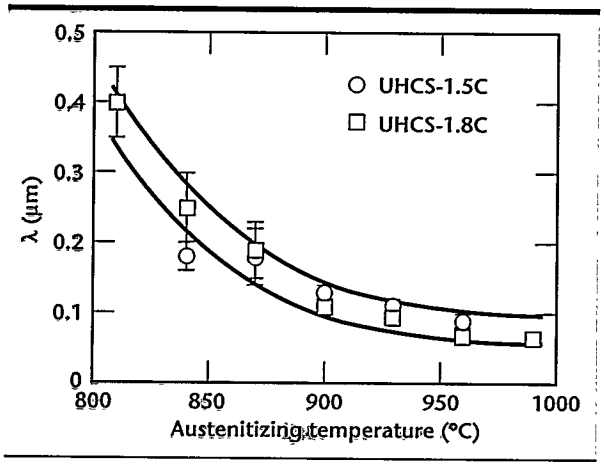


Figure 4. Pearlite interlamellar spacing vs austenitizing temperature, for both UHCS-1.5C and UHCS-1.8C. The value of λ given at 810 °C is the interparticle spacing.

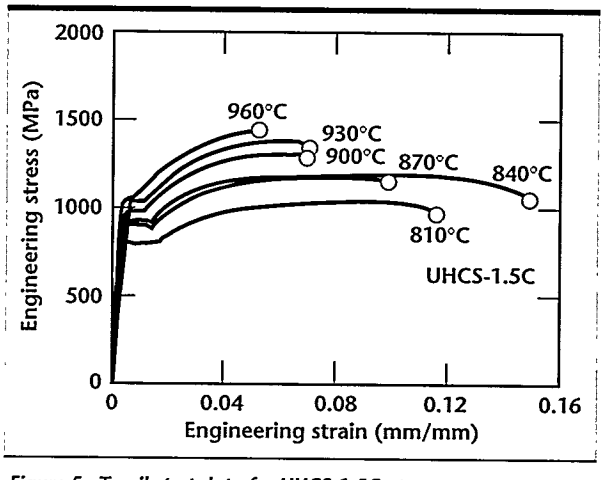


Figure 5. Tensile test data for UHCS-1.5C at room temperature.

$$\sigma_y = \sigma_0 + k_y\lambda^{-1/2} \quad (1)$$

where σ_y is the yield stress, σ_0 is the friction stress, k_y is a material constant and λ is the interlamellar spacing. Data from Hyzak and Bernstein[5] and from Gladman *et al.*[6] on eutectoid steels also show strengths which vary linearly with the inverse square root of interlamellar spacing, as shown in **Fig. 7**. This data show lower strengths than the UHCS materials of the present study. These differences in strength are strongly related to differences in solid solution content of the steels and will be discussed in detail later. An additional variable, not considered in the analysis of data in **Fig. 7**, is the pearlite colony size, which will also be discussed in detail later.

A comparison was made of the yield strengths of the pearlitic UHCS materials investigated here with earlier studies on the spheroidized UHCS materials. In an earlier study on spheroidized steels, the Hall-Petch relation was used to predict yield strength.[2] The prediction for yield strength was based on ferrite grain size, L, and interparticle spacing of carbide particles, D_s^*, as barriers for dislocation motion. The predictive phenomenological relation derived from these dependencies is given as:

$$\sigma_y = 310(D_s^*)^{-1/2} + 460L^{-1/2} \quad (2)$$

where lengths are in units of μm and stress is in units of MPa. In **Eq. 2** the first term, $310(D_s^*)^{-1/2}$, is considered related to the friction stress in the Hall-Petch relation, and the second term, $460L^{-1/2}$, is related to the grain size term. In order to relate the microstructural parameters L and D_s^*, to pearlitic microstructures, corresponding parameters were considered. The grain size, L, for a spheroidized microstructure is most logically correlated with the pearlite colony size in a pearlitic microstructure. The interparticle spacing, D_s^*, in a spheroidized microstructure should correspond to the average dislocation barrier spacing involving carbides in a pearlitic microstructure. As carbide lamella are the most obvious barriers to dislocations, the pearlite interlamellar spacing, λ, should relate to D_s^* The average glide distance for a dislocation is not, however, equal to λ, as measured by the finest observable interlamellar spacing. This is because calculation of D_s^* from the measured value of λ in a pearlitic microstructure requires knowledge of the slip directions in the ferrite matrix between carbide lamella.

The slip directions in body-centered cubic (BCC) metals are normally along the $\langle 111 \rangle$ directions.[7] If the habit plane of carbide platelets in ferrite is known, then the length along the slip direction between carbide platelets can be calculated from λ to give D_s^*. Three dominant orientation relationships have been observed between ferrite and carbide lamella in pearlitic steels of eutectoid composition.[8] These three orientation relationships between habit planes are[8–14]:

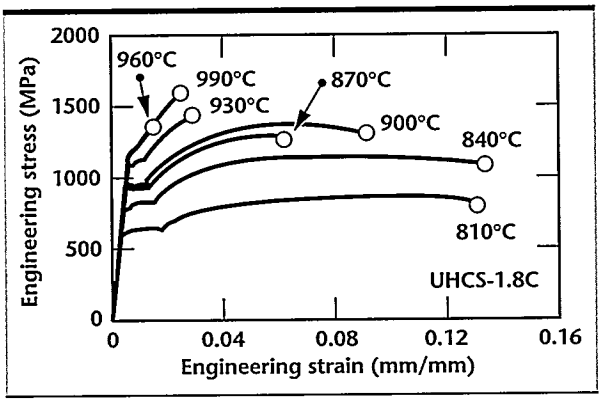


Figure 6. Tensile test data for UHCS-1.8C at room temperature.

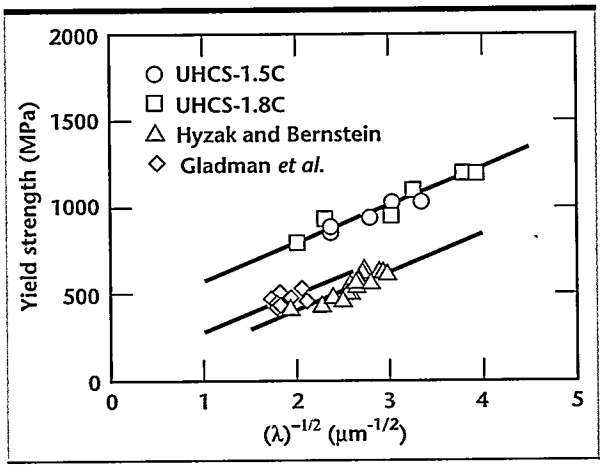


Figure 7. Yield strength vs the inverse square root of pearlite interlamellar spacing for both UHCS-1.5C and UHCS-1.8C.

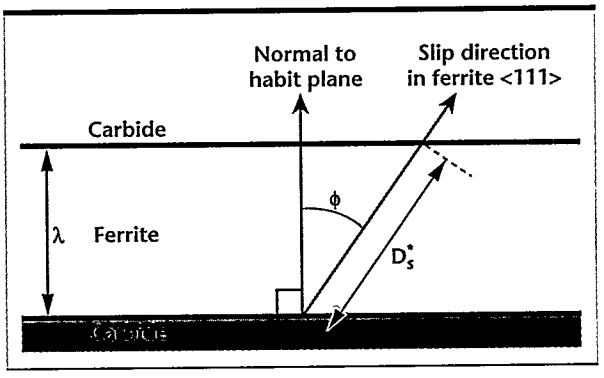


Figure 8. Schematic showing calculations for finding D_s^* from λ using the angle ϕ.

$(001)_c // (11\bar{2})_f$ (Bagaryatsky),

$(103)_c // (01\bar{1})_f$ (Isaichev), and

$(101)_c // (\bar{2}15)_f$ (Pitsch - Petch),

where a subscript of "c" denotes carbide orientation, and a subscript of "f" denotes ferrite orientation. The angle between the normal of each habit plane and each of the $\langle 111 \rangle$ slip directions in ferrite, ϕ, was calculated as shown schematically in **Fig. 8.** For each orientation relation, the angles of all slip directions which intersect the habit plane were calculated. The average of angle ϕ was then found by giving equal weight to each slip direction. The average angle from the normal to the habit plane, ϕ, can be used to find the average slip distance between carbide platelets from the relation:

$$(D_s^*) = \lambda/\cos\phi \ .$$

The Bagaryatsky orientation relation gives an average value of ϕ = 47.7°, while the Isaichev and Pitsch-Petch orientation relationships give values of 56.6° and 35.3°, respectively. It should be noted that the measured values of these orientation relations are generally within only a few degrees of those quoted above.[8] For example, Zhou and Shiflet found the measured values of the Bagaryatsky orientation relation to be always 3.8° from the ideal orientation given above, a deviation which was attributed to misfit strain[8]. For this reason, ϕ should be taken as accurate to within no more than approximately 5°. Assuming that each of the three orientation relations is of equal importance gives an average value of ϕ = 45°. Using this value yields the following relation for calculating D_s^*, from the measured pearlite interlamellar spacing, λ:

$$(D_s^*) = (2)^{1/2}\lambda \ . \qquad (3)$$

Using this relationship for D_s^*, and values for L equal to the measured pearlite colony sizes, yield strengths were predicted, using **Eq. 2.** The predicted (**Eq. 2**) and observed yield strengths from microstructures in which pearlite interlamellar spacings and colony sizes were measured are shown in **Fig. 9**, which is a plot of measured yield strength vs predicted yield strength. Data and calculations used in **Fig. 9** for the pearlitic UHCS materials are given in **Table 1**. The data show a remarkably good fit to **Eq. 2**, using the same preconstants deduced from data for spheroidized hypereutectoid steels. Also shown in **Fig. 9** are data for a spheroidized steel of the same composition as

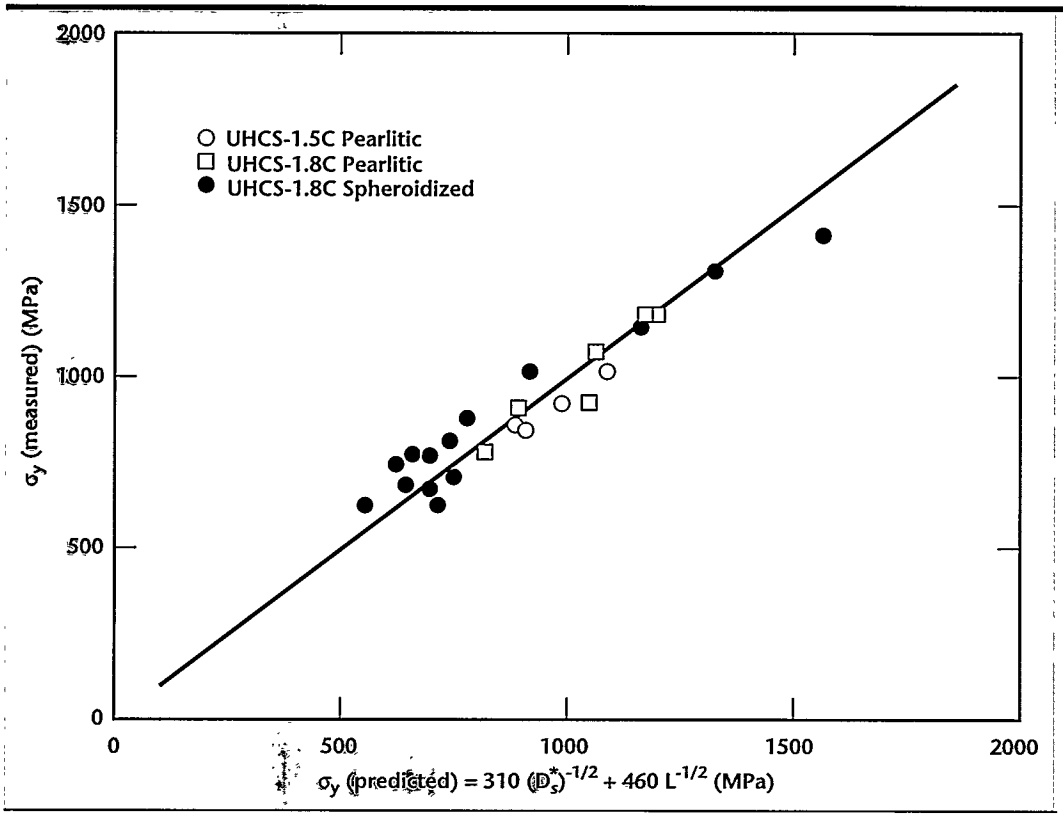


Figure 9. Measured yield strengths vs predicted yield strengths from Eq. 2 for pearlitic UHCS materials and spheroidized UHCS steels. Data for spheroidized UHCS are from Reference 2.

the UHCS-1.8C in the present investigation.[2] Remarkably, both spheroidized and pearlitic microstructures fit the predictions of **Eq. 2**.

The predictive aspect of **Eq. 2** does not explain the large strength differential between the eutectoid composition steels of other studies and the present hypereutectoid composition steels in the pearlitic condition, (**Fig. 7**). These differences are explained by taking into account the contributions of the term from **Eq. 2** containing L, and of solid solution alloying on the strength of pearlitic and spheroidized carbon steels. Thus, an additional term for **Eq. 2** is proposed, $(\sigma_o)_{ss}$, the contribution from solid-solution strengthening, which will be a constant for any given alloy composition. **Eq. 2** now takes the form:

$$\sigma_y = (\sigma_o)_{ss} + A(D_s^*)^{-1/2} + BL^{-1/2} \qquad (4)$$

where A and B are constants to be determined from available data. While $(\sigma_o)_{ss}$ is a constant for only a particular alloy or group of alloys, the constants A and B should hold for any eutectoid or hypereutectoid steel, in either spheroidized or pearlitic condition.

In order to determine the constant B of **Eq. 4**, it is necessary to plot yield stresses at a constant value of D_s^* against $L^{-1/2}$, and determine the slope of the data. Data with a constant value of D_s^* is available (Syn *et al.*, **Fig. 8**) and shows that $B = 460$, which is also valid for **Eq. 2**. In order to determine the constant A of **Eq. 4**, it is necessary to plot $(\sigma_y - 460L^{-1/2})$ vs $(D_s^*)^{-1/2}$, and measure the slope of the data. Such a plot is given in **Fig. 10**, where A is the slope measured from data on the yield strengths of several compositional groups with pearlitic or spheroidized microstructures. The single value of slope selected was determined principally from the data for spheroidized UHCS-1.8C material giving A = 145. The y-intercepts of the data in **Fig. 10** yield values of $(\sigma_o)_{ss}$ for each group of alloys, yielding the following predictive equation:

$$\sigma_y = (\sigma_o)_{ss} + 145(D_s^*)^{-1/2} + 460L^{-1/2} \ .$$

The sources of the data shown in **Fig. 10** are listed in **Table 2**, together with the values of $(\sigma_o)_{ss}$ found from **Fig. 10** for each of the four alloy groups that were created, based on solid-solution alloy composition.

In plotting data in **Fig. 10**, the values of L and D_s^* for the spheroidized materials were taken from those tabulated by Syn *et al.* (**Table 2**),[2,5,6,13-17] with the exception of data from Davidson and Ansell, where a value of $L = 16$ μm was taken from the micrograph in **Fig. 2**.[16] The values of L for the pearlitic steels were taken from the values given for pearlite colony size (Hyzak and Bernstein, **Table 2**[5]; Gladman *et al.* **Table 2**[6]). The values for interlamellar

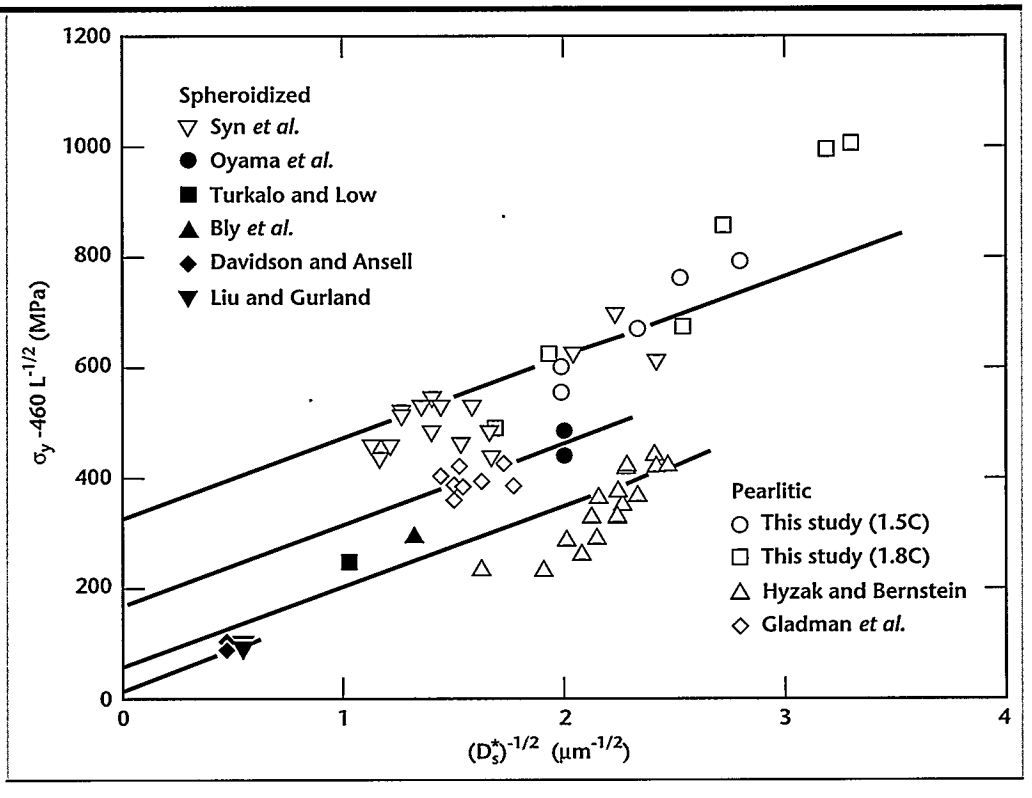


Figure 10. Plot of $\sigma_y - 460L^{-1/2}$ vs $(D_s^*)^{-1/2}$ for the determination of $(\sigma_o)_{ss}$ for each of the four alloy categories.

spacing given by Hyzak and Bernstein (**Table 2**)[5] were used in calculating D_S^* by means of **Eq. 3**. Gladman *et al.*, having used the method of Birkbeck and Wells[18] for evaluating interlamellar spacings, reported interlamellar spacing as S_0, which is not equivalent to the finest measurable interlamellar spacing reported as λ in the present investigation. Birkbeck and Wells reported the ratio of S_0 to the finest measurable interlamellar spacing to be approximately 1.3 for their data. This value of $S_0/\lambda = 1.3$ was used to convert the interlamellar spacings reported by Gladman *et al.* for calculation of D_S^* using **Eq. 3**.

Table 2 gives the composition of each alloy shown in **Fig. 10**. The alloys in **Table 2** are categorized into four groups, and one value of $(\sigma_0)_{ss}$ was calculated from **Fig. 10** for each group. These four groups are:

1) highly alloyed UHCS materials (this study; Syn *et al.*), where $(\sigma_0)_{ss} = 330$ MPa;
2) moderately-alloyed UHCS and eutectoid steels (Oyama *et al.*; Gladman *et al.*), where $(\sigma_0)_{ss} = 170$ MPa;
3) low-alloy eutectoid steels (Hyzak and Bernstein; Turkalo and Low; Bly *et al.*), where $(\sigma_0)_{ss} = 60$ MPa; and
4) unalloyed eutectoid and hypereutectoid steels (Davidson and Ansell; Liu and Gurland), where $(\sigma_0)_{ss} = 20$ MPa;.

The highly-alloyed UHCS materials give a high value for $(\sigma_0)_{ss}$ of 330 MPa, while the unalloyed eutectoid and hypereutectoid steels give a low value for $(\sigma_0)_{ss}$ of 20 MPa, which comes from solid-solution strengthening by interstitial carbon. Using the $(\sigma_0)_{ss}$ values given in **Table 2**, a predicted flow stress may be calculated from **Eq. 5** for each material. In **Fig. 11**, the predicted yield strength from **Eq. 5** is compared with the measured yield strength for each material of **Table 2**. **Figure 11** shows a very good correlation

Table 2. Composition in weight percent of eutectoid and hypereutectoid steels and the calculated values for $(\sigma_0)_{ss}$. Both pearlitic (P) and spheroidized (S) microstructures are given. The method of determination is shown in Figure 10.

Investigator	$(\sigma_0)_{ss}$ (MPa)	Structure	Composition, wt.% C	Al	Cr	Mn	Si	Ni	Cu	N
This investigation	330	P	1.8	1.6	1.5	0.5	...	...	...	...
This investigation	330	P	1.5	1.6	1.5	0.5	...	...	...	...
Syn *et al.*[2]	330	S	1.8	1.6	1.5	0.5	...	...	...	...
Oyama *et al.*[13]	170	S	1.5	...	1.5	0.5	0.5	...	...	...
Gladman *et al.*[6]	170	P	0.81	...	...	0.92	0.40	...	...	0.010
	170	P	0.81	...	...	0.95	0.91	...	...	0.008
	170	P	0.81	...	...	1.54	0.43	...	...	0.007
	170	P	0.81	...	...	1.56	0.92	...	...	0.010
	170	P	0.81	...	...	0.88	0.34	...	...	0.018
	170	P	0.89	...	...	0.97	0.99	...	...	0.019
	170	P	0.78	...	...	1.48	0.25	...	...	0.017
	170	P	0.81	...	...	1.51	0.95	...	...	0.017
Hyzak and Bernstein[5]	60	P	0.81	...	...	0.87	0.17	...	...	...
Turkalo and Low[14]	60	S	0.75	...	...	1.0	0.17	...	...	...
Bly *et al.*[15]	60	S	1.1	...	...	0.75	0.24	0.43	0.56	...
Davidson and Ansell[16]	20	S	0.83	...	...	...	...	...	...	...
	20	S	1.1	...	...	...	...	...	...	...
Liu and Gurland[17]	20	S	1.23	...	...	...	...	...	...	...

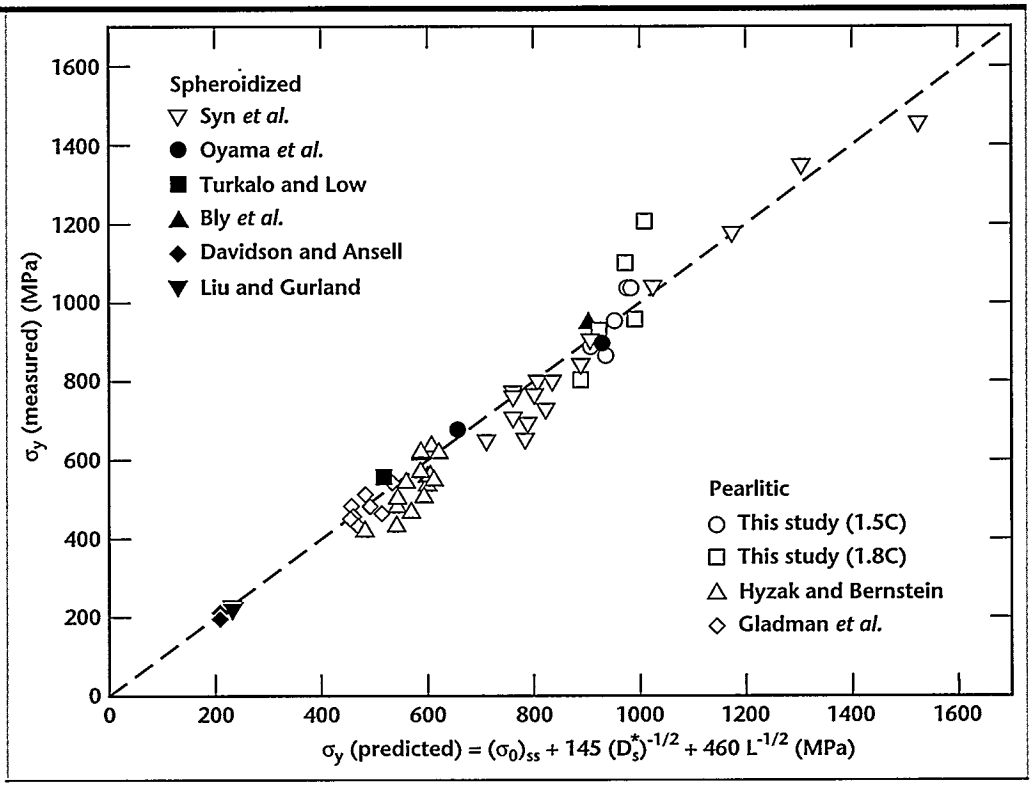


Figure 11. Measured yield strengths vs predicted yield strengths from Eq. 5 for pearlitic and spheroidized hypereutectoid and eutectoid steels.

of the predicted values (**Eq. 5**) with the measured values, over a wide range of yield strengths. The results indicated that spheroidized and pearlitic structure steels are influenced in a nearly identical manner by carbide spacing, grain size and solid solution alloying. This is surprising, since it is intuitively expected that a lamellar structure would lead to a stronger condition than a particulate structure. This expected difference is not evident from the correlations given, although a few data points for the pearlitic UHCS materials do fall above the spheroidized UHCS material in strength at fine interlamellar spacings (**Fig. 10**).

Conclusions

Using the transformations in the hypereutectoid region of the Fe-C phase diagram allows for the formation of pearlitic microstructures with various interlamellar spacings and pearlite colony sizes. By varying the ratio of undissolved spheroidized carbide to pearlitic carbide, the interlamellar spacings in pearlite can be controlled to yield a range of strengths and ductilities.

A phenomenological equation, found to predict the yield strength for spheroidized UHCS materials based on microstructural parameters, can also the predict yield strength for pearlitic UHCS materials. The microstructural parameters used to predict yield strength in the pearlitic microstructure steels are the pearlite colony size and the interlamellar carbide spacing. An extension of this phenomenological equation was made to account for differences in solid-solution strengthening between various alloys, adding an additional parameter, $(\sigma_0)_{ss}$, which depends on alloy content. This extended analysis accurately predicts yield strength for several pearlitic and spheroidized eutectoid and hypereutectoid steels of various alloy content.

Acknowledgments

One author, E. M. T., would like to express his gratitude to the Lawrence Livermore National Laboratory for assisting in the advancement of his Ph.D. thesis program at Stanford University.

References

1. Lesuer, D. R., C. K. Syn, A. Goldberg, J. Wadsworth, and O. D. Sherby (1993), "The Case for Ultrahigh-Carbon Steels as Structural Materials," *JOM*, vol. **45**, pp. 40–46.

2. Syn, C. K., D. R. Lesuer, and O. D. Sherby (1994), "Influence of Microstructure on the Tensile Properties of Spheroidized Ultrahigh-Carbon Steels," *Metall. Trans.*, vol. **25A**, pp. 1481–1493.

3. Sherby, O. D., and T. Oyama (1993), *Microstructures and Mechanical Properties of an Ultrahigh Carbon Steel Processed by the Divorced Eutectoid Transformation*, Stanford University, Stanford, Calif., unpublished data.

4. Sherby, O. D., and T. Oyama (1985), *UHCS Alloy and Processing Thereof*, U.S. Patent No. 4,533,390, August 6, 1985, pp. 1481–1493.

5. Hyzak, J. M., and I. M. Bernstein (1992), "The Role of Microstructure on the Strength and Toughness of Fully Pearlitic Steels," *Metall. Trans.*, vol. **7A**, pp. 1217–1224.

6. Gladman, T., I. D. McIvor, and F. B. Pickering (1972), "Some Aspects of the Structure-Property Relationships in High-Carbon Ferrite-Pearlite Steels," *J. Iron Steel Inst.*, vol. **210**, pp. 916–930.

7. Hull, D., and D. J. Bacon (1984), *Introduction to Dislocations, 3rd ed.*, Pergamon Press, New York, N.Y., p. 121.

8. Zhou, D. S., and G. J. Shiflet (1992), "Ferrite: Cementite Crystallography in Pearlite," *Metall. Trans.*, vol. **23A**, pp. 1259–1269.

9. Petch, N. J. (1953), "The Orientation Relationship Between Cementite and a-Iron," *Acta Cryst.*, vol. **6**, p. 96.

10. Pitsch, W. (1962), "Der Orientierungszusammenhang zwischen Zementit und Ferrit im Perlit," *Acta Metall.*, vol. **10**, pp. 79–80.

11. Bowden, H. G., and P. M. Kelly (1967), "Deformation Twinning in Shock-Loaded Pearlite," *Acta Metall.*, vol. **15**, pp. 105–111.

12. Boswell, P. G., and G. A. Chadwick (1977), "The Crystallography of Herringbone Pearlite," *Scripta Metall.*, vol. **11**, pp. 1001–1002.

13. Oyama, T., U. Ritter, B. Walser, and O. D. Sherby (1995), "Microstructures and Mechanical Properties of an Ultrahigh Carbon Steel Processed by the Divorced Eutectoid Transformation," *Trans. Iron Steel Inst. Jpn.*, in press.

14. Turkalo, A. M., and J. C. Low, Jr. (1958), "The Effect of Carbide Dispersion on the Strength of Tempered Martensite," *Trans. TMS-AIME*, vol. **212**, pp. 750–758.

15. Bly, D., O. D. Sherby, C. Young, and Y. Mischenkov (1973), "Influence of Thermal Mechanical Treatments on the Mechanical Properties of a Finely Spheroidized Eutectoid-Composition Steel," *Mater. Sci. Eng.*, vol. **12**, pp. 41–46.

16. Davidson, T. E., and G. S. Ansell (1968), "The Structure Sensitivity of the Effects of Pressure upon the Ductility of Fe-C Materials," *Trans. ASM*, vol. **61**, pp. 242–254.

17. Liu, C. T., and J. Gurland (1968), The Strengthening Mechanisms in Spheroidized Carbon Steels," *Trans. TMS-AIME*, vol. **242**, pp. 1535–1542.

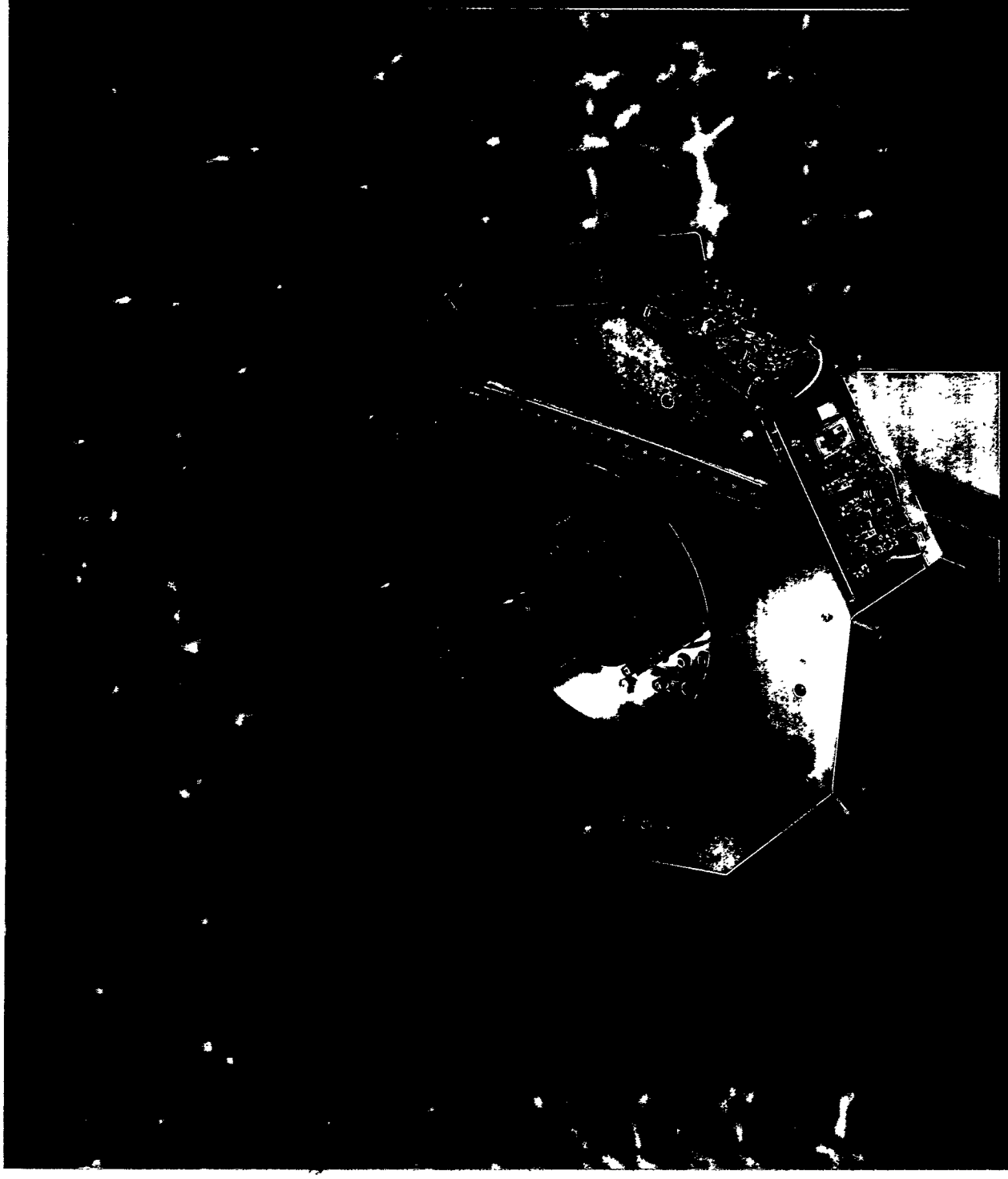
18. Birkbeck G., and T. C. Wells (1968), "Pearlite Morphology in Three Low-Carbon Steels," *Trans. AIME*, vol. **242**, pp. 2217–2220.

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 intelligent 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-93, we concentrated our research efforts on the three project areas described in these reports: 1) Development of Advanced Modulator Technology for Heavy Ion Recirculators, 2) Modeling and Testing of Nanostructure Multilayer Capacitors, and 3) Microradar Development.

Power Conversion Technologies



6

6. Power Conversion Technologies

Overview
Mark A. Newton, Thrust Area Leader

Development of Advanced Modulator Technology for Heavy Ion Recirculators
Mark A. Newton, Hugh C. Kirbie, and Roy L. Hanks *6-1*

Modeling and Testing of Nanostructure Multilayer Capacitors
Gary W. Johnson and Wang C. Ng *6-7*

Microradar Development
Stephen G. Azevedo, Thomas E. McEwan, and John P. Warhus *6-17*

Development of Advanced Modulator Technology for Heavy Ion Recirculators

Mark A. Newton, Hugh C. Kirbie, and Roy L. Hanks
Laser Engineering Division
Electronics Engineering

We are designing and building a small-scale recirculator to demonstrate feasibility of the recirculator concept. This report provides background on the recirculator concept and describes the development of a new solid-state modulator technology that is the foundation of the recirculator and other advanced accelerator concepts.

Introduction

Recirculating induction accelerators are being investigated as potential low-cost drivers for inertial fusion energy. A recirculator is a circular induction accelerator where beams of heavy ions are accelerated and bent in a closed path, as illustrated in **Fig. 1.**

To use the recirculator as an inertial fusion driver, the ions are then extracted from this closed path and focused on an inertial fusion target. The recirculator appears to be an economically attractive driver option when compared to other induction accelerator technologies, because, unlike linear machines, the acceleration sequence reuses the most expensive components of the induction accelerator. However, the recirculator also presents new technical challenges.

The heavy ion fusion (HIF) project at Lawrence Livermore National Laboratory (LLNL) is designing and building a small-scale recirculator to demonstrate feasibility of the recirculator concept.[1] This small-scale recirculator is a 4.5-m diameter ring that will accelerate singly ionized potassium ions to 320 keV, after 15 laps around the ring. This ring will be used to investigate the physics and demonstrate the technology associated with accelerating and transporting a space-charge-dominated beam in a circular induction accelerator.

Description of the Small Recirculator

The small-scale recirculator (**Fig. 2**) consists of four main elements: 1) an injector for generating the ion beam; 2) a matching section to transition the circular beam into the magnetic quadrupole focusing of the ring; 3) an insertion/extraction section for transporting the beam into and out of the ring; and 4) a circular accelerator for accelerating the beam.

The beam is generated in a high voltage diode with a potassium xeolite source. Singly-ionized potassium ions are extracted from the diode with an 80 kV pulse generated by a standard thyratron line-type pulser. An ion beam current of approximately 2 mA is accelerated into an electrostatic matching section, where it is conditioned for insertion into the recirculator ring. The beam exits the matching section into the insertion region, where it is bent with a series of electric kickers to insert the ions into the ring.

Once the beam is in the ring, the ions are accelerated, focused and bent as they pass through the 40 half-lattice modules which make up the ring.

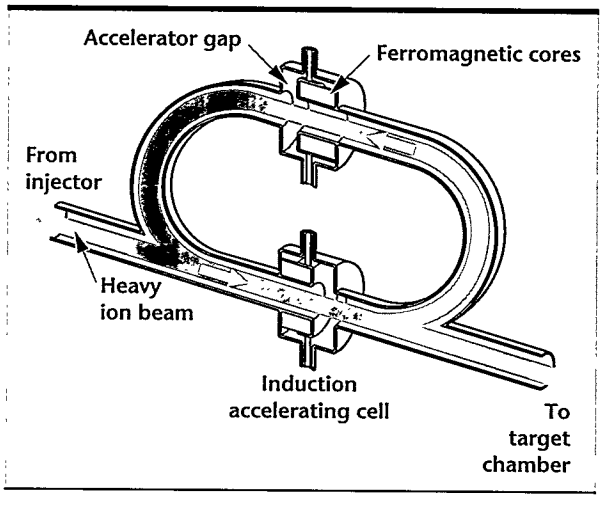


Figure 1. Illustration of recirculator.

The half-lattice modules are the basic building blocks of the recirculator. Each half-lattice module has three primary components: an induction accelerating cell for accelerating the beam, an electric dipole for bending the beam, and a magnetic quadrupole for focusing the beam as it traverses the ring. In addition, each half-lattice period provides access to the beam for diagnostics and ports for vacuum. A sectional view of the half-lattice module is shown in **Fig. 3.**

In the ring, the ions are accelerated in 470-V increments as they pass through 34 induction accelerating cells. There is one induction cell for every half-lattice period, with the exception of the six half-lattice periods in the insertion/extraction region, where there is no acceleration.

Electric dipoles provide a time-varying electric field to bend the beam around the ring. The field on the dipoles must be precisely coordinated with the beam energy during an acceleration sequence to maintain a constant bend radius. The beam will gain nominally 16 keV per lap, and after 15 laps will have an energy of approximately 320 keV. Electric kickers in the insertion/extraction stage will redirect the ion pulse out of the ring.

Progress

Status and Description of Modulator Development

Electrical performance objectives. One of the most difficult and critical challenges faced in building a recirculator is driving the induction accelerator cells that couple pulsed electrical energy into the ion beam. A sequence of pulses must be generated by each modulator to accelerate the beam on each lap around the recirculator. The time between each pulse in a sequence is determined by the time required for the beam to complete one lap around the ring. In our 4.5-m diameter ring, this time varies from approximately 25 μs to 10 μs, which corresponds to pulse repetition rates from 40 to 100 kHz. One of the objectives of the recirculator experiment is to demonstrate beam compression. As a result, not only will the velocity of the beam be increasing through each lap, but the duration of the beam will be decreasing. The modulator must be agile enough to generate a wide spectrum of pulses and have the ability to change their shape and duration from pulse to pulse. The modulator that drives the induction cells must perform three

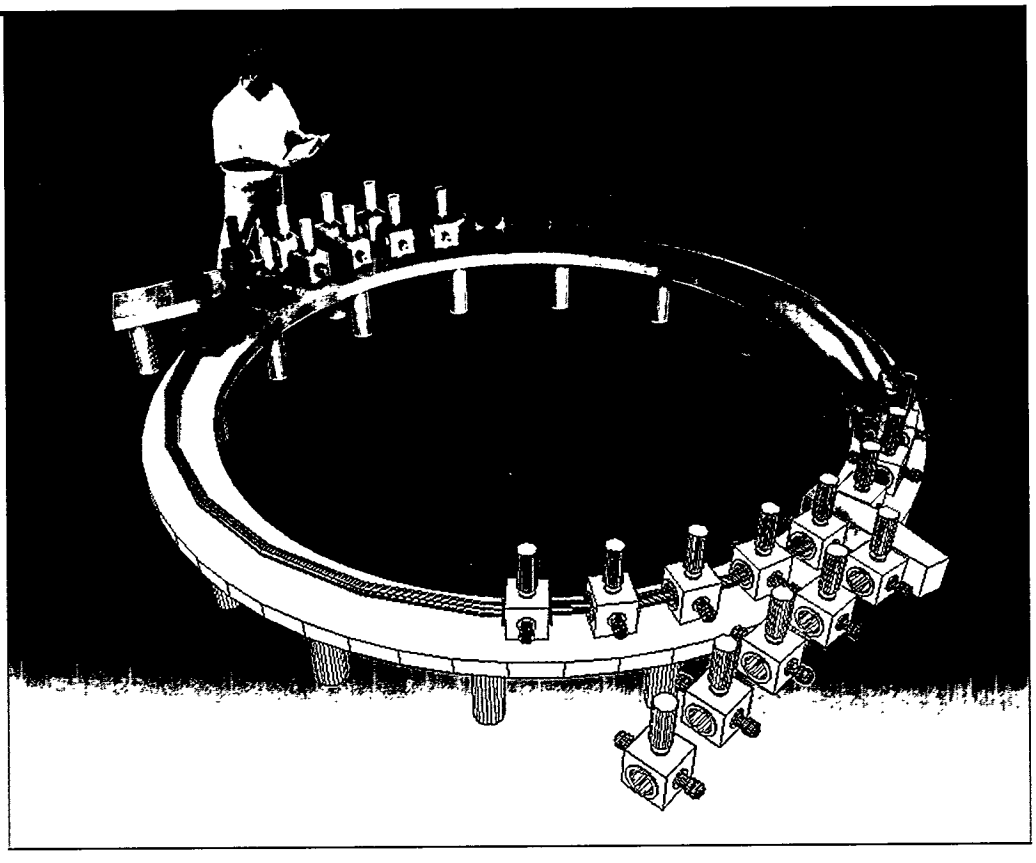


Figure 2. Conceptual illustration of small recirculator.

primary functions, each of which is critical to the operation of the recirculator. These primary functions are acceleration, longitudinal confinement and cell reset.

For the small recirculator, acceleration of the beam requires a sequence of nominally 470-V pulses that are 4 to 2 μs in duration. The acceleration pulses are nominally square pulses with slight "tilts" in the voltage to compress the beam longitudinally. These tilts also vary from pulse to pulse within a single acceleration sequence. **Figure 4** illustrates the range of pulse shapes and pulse widths required for acceleration on the small recirculator.

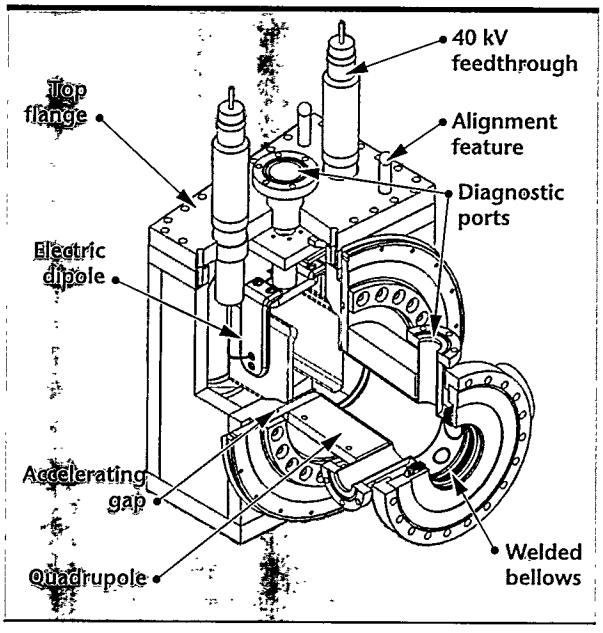


Figure 3. Sectional view of the half-lattice module, the basic building block of the recirculator.

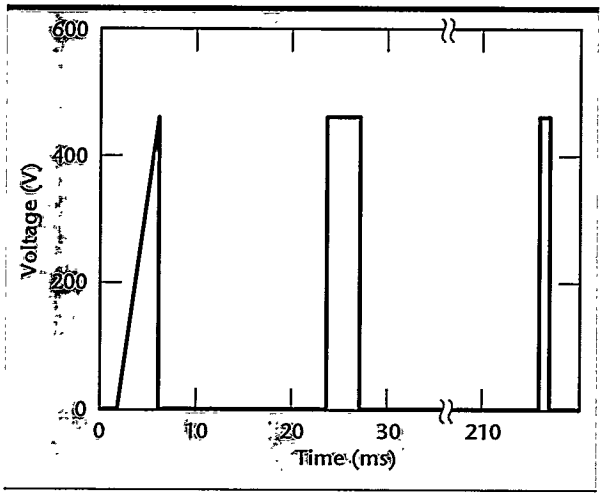


Figure 4. Illustration of the range of pulse shapes and pulse widths required for acceleration on the small recirculator.

Longitudinal confinement of the ion beam is necessary to compensate for the space-charge forces, which are acting to expand the beam in the axial direction. Short "ear" pulses must be generated that accelerate the tail of the beam, and decelerate the head of the beam, just enough to balance the space-charge forces. As with the acceleration pulses, the shape and amplitude of the ear pulses must also vary within a single acceleration sequence. **Figure 5** shows an example of the bipolar ear pulses that will be required for the small recirculator.

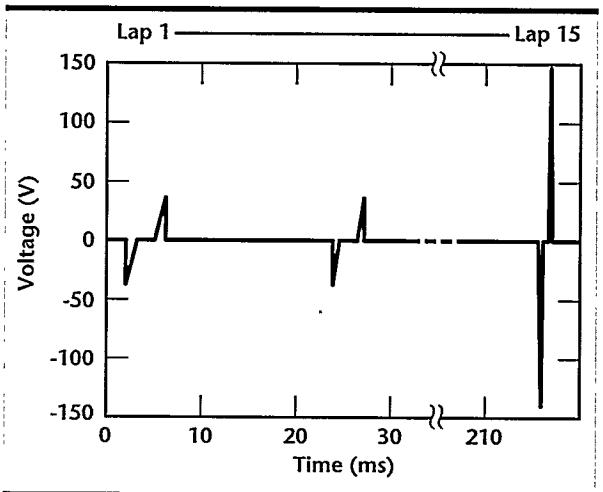


Figure 5. Example of bipolar pulses required for the small recirculator.

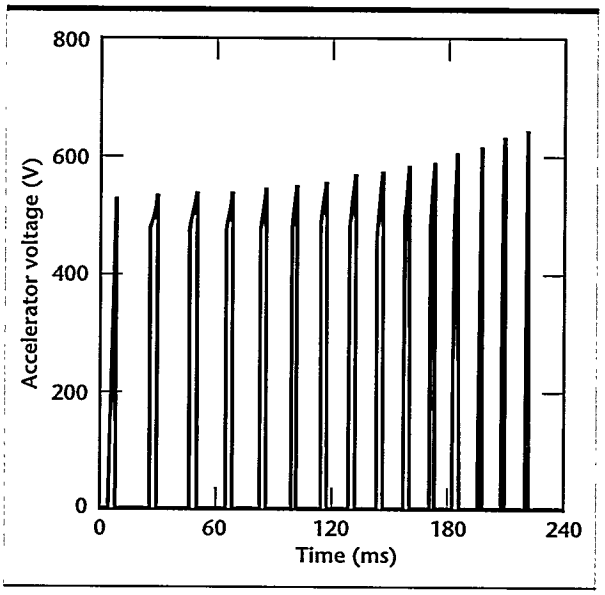


Figure 6. Waveform equivalent to the summation of acceleration and ear pulses.

Reset of the magnetic core material in each cell is required to return the magnetic material to its initial state in preparation for the next pulse. If the magnetic core material is not adequately reset after each pulse, the flux density in the core will gradually build up until the core saturates, shorting out the modulator. Aside from a potentially disastrous impact on the modulator, a saturated core will inhibit further acceleration of the ion beam.

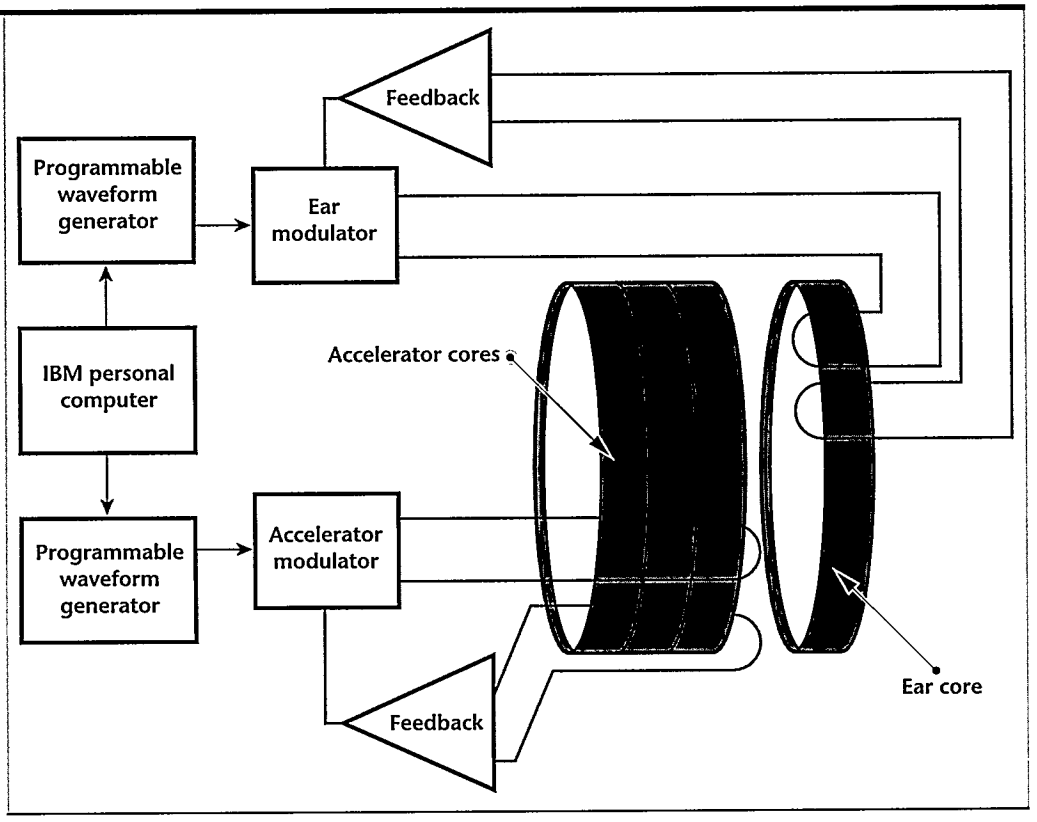
Mechanical objectives. Packaging of the modulator into a half-lattice module is a very challenging task. It is physically constrained to a toroidal volume, with approximately 25 cm inner diameter, by approximately 60 cm outer diameter, by 10 cm axial dimension. This volume houses the induction cores, energy storage, and pulse shaping circuits required for agile control of the accelerating waveform. The inner diameter of the modulator is constrained by the need to allow access to the diagnostic ports and magnetic quadrupole without disturbing the quadrupole alignment. Radial growth of the induction modulator cell is limited by interference with external hardware. Along the beam axis, the induction modulator must fit between the dipole housing and the diagnostics ports of the half latice period. The axial length of the modulator and cell is constrained to approximately 10 cm.

Circuit description. The modulator requirements for the small recirculator are challenging, but the basic requirements of repetition rate and reset have been demonstrated in previous development efforts.[2,3] However, those development efforts focused on using the MOSFETS as on/off switches. This mode of operation is inadequate for the small recirculator because of the radical changes in pulse shape that are required from lap 1 to lap 15, as seen in **Fig. 4**. Our present development efforts are focused on developing a high power MOSFET amplifier to generate the required waveforms.

There are a number of modulator/cell configurations that could be used to generate the appropriate waveforms for acceleration of the ion beam. The most flexible approach is to generate the acceleration and the bipolar ear pulses separately. However, this would require a bipolar amplifier for the ears. Ultimately, the beam needs to see the waveform shown in **Fig. 6**, which is the summation of the acceleration and ear pulses.

It is important to note that the resultant waveform is not bipolar. The approach being pursued on the small recirculator is to inductively sum the outputs of a 500-V main acceleration amplifier and a 200-V ear pulse amplifier, so that the required waveform appears across a single accelerating gap.

Figure 7. Block diagram of modulator.



MOSFET arrays are being used as amplifier elements to tailor the waveshape and regulate the voltage that appears across the gap. The amplifier configuration allows the generation of arbitrary waveshapes. Programmable waveform generators drive the parallel MOSFET arrays to generate the desired acceleration waveforms for the beam. A block diagram of the modulator is shown in **Fig. 7.**

Modeling. The circuit analysis program, Micro-Cap IV is being used to model the electronic portion of the design, including the cell and amplification circuits. **Figure 8** is a typical simulation, showing an input waveform compared with the output voltage at the accelerator cell, including reset.

Benchtop testing. Initial circuit operation has been validated with a benchtop test of a small-scale version of the modulator. The benchtop test is driving a linear, air-core inductor instead of the metglas induction cell that is the ultimate load.

Progress on prototype. A drawing of the cell modulator package is shown in **Fig. 9. Figure 10** is a photograph of the recently completed hardware.

Unlike more conventional system designs for induction accelerators, the modulator is integral to the cell, to eliminate the need for connecting cables to carry the pulses to the induction cell. This approach was taken to eliminate the possibility of unwanted reflections occurring because of impedance mismatches between the cable and the time-varying load. The long pulse durations make it impractical to rely on time isolation of the load and the modulator to insure that reflections don't disturb the desired pulse fidelity.

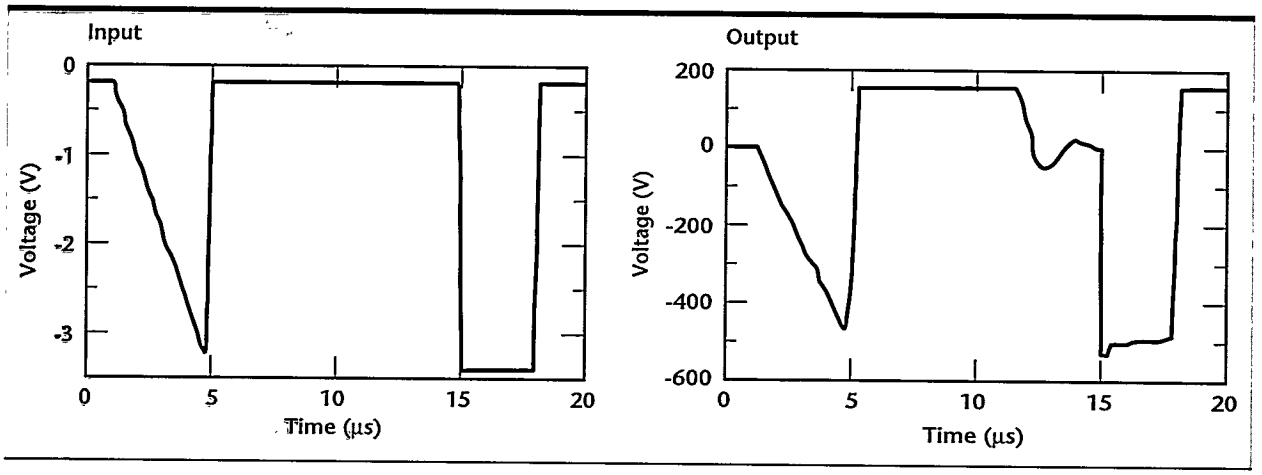


Figure 8. Typical simulation by circuit analysis program, Micro-Cap IV.

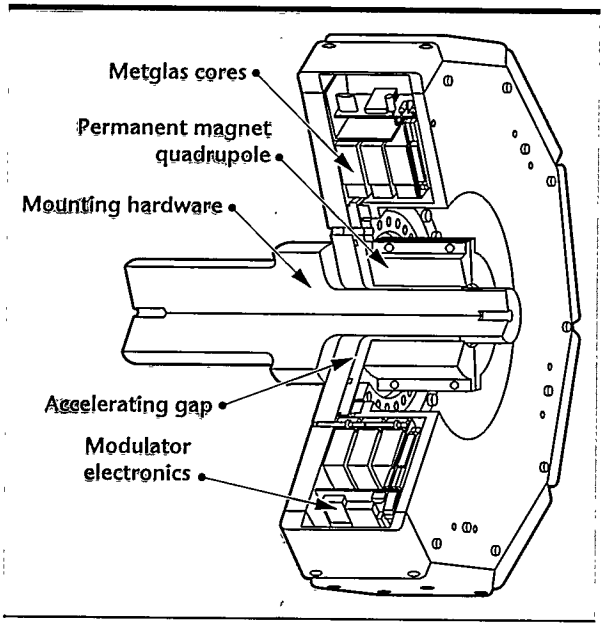


Figure 9. Drawing of cell modulator package.

Figure 10. Photograph of recently completed hardware.

Future Work

This year's development efforts have resulted in the design and fabrication of an integrated solid state modulator for advanced accelerators. We have performed circuit modeling and benchtop tests to validate the design prior to fabrication and testing of a fully integrated modulator. Testing and evaluation of the fully-integrated solid state modulator technology will be conducted in FY-96.

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 Toward a Prototype Recirculating Induction Accelerator for Heavy-Ion Fusion," *Proceedings of the 1995 Particle Accelerator Conference, Dallas,Tex.*, (UCRL-JC-119538).

2. Kirbie, H., W. Cravey, S. Hawkins, M. Newton, and C. Ollis (1993), "A FET-Switched Induction Accelerator Cell," *Proceedings of the 9th IEEE Pulsed Power Conference, Albuquerque, N. Mex.*, p. 415.

3. Kirbie, H., S. Hawkins, B. Hickman, B. Lee, M. Newton, C. Ollis, C. Brooksby, D. Gilbert, D. Longinotti, and E. Eubank (1995), "Development of Solid State Induction Modulators for High PRF Accelerators," *Proceedings of the 10th IEEE Pulsed Power Conference, Albuquerque, N. Mex.*

Modeling and Testing of Nanostructure Multilayer Capacitors

Gary W. Johnson and Wang C. Ng
Defense Sciences Engineering Division
Chemistry and Materials Science

This project has assisted in the development of nanostructure multilayer capacitors (NMC's) at Lawrence Livermore National Laboratory (LLNL). Two topics were included in this effort: dielectric and electromagnetic field modeling, and electrical performance testing. In the modeling area, an electrostatic field model was used to quantify the geometric field intensification and the effects of microcontamination. An interface work function modification technique which may lead to increased performance has also been modeled. We also studied the dynamic effects of rapid discharge and measured the basic performance parameters of capacitance and voltage breakdown. These models and measurements are crucial to the success of the NMC program.

Introduction

Our controlled nanostructure deposition techniques are capable of synthesizing extraordinarily smooth, sub-micron-thick layers of dielectric and conductor materials, resulting in high-voltage, high-energy-density capacitors. We expect that, with this technology, high-voltage capacitors with an order of magnitude improvement in energy density are achievable. NMC's are inherently solid state, exhibiting superior mechanical and thermal properties. Their compact parallel-plate form and robust conductor cross-section guarantees low effective series inductance and resistance, which are important parameters for power electronic and pulse power applications. We investigated dielectrics consisting of mixtures of metal oxides (e.g., Al_2O_3), as well as several conductor compositions.

High-energy-density capacitors are strategic to many LLNL programs, including diode lasers, pulsed plasma processing, accelerators, advanced high-frequency switching power supplies, and high-intensity electromagnetic field generation. An order of magnitude improvement in capacitor volume and mass will drastically shrink the size and weight of pulse power and power electronic equipment. Also, the form factor of NMC's is well-suited to state-of-the-art integrated packaging. For instance, snubber capacitors on high-power switching devices could be integrated onto the semiconductor's substrate. These same factors make our capacitors attractive to the electronics industry in general, indicating that this is a dual-use technology with broad marketability.

Our current efforts are dedicated to the fabrication of dielectric test films and the demonstration of simple multilayer capacitors. Synergistic work under a CRADA with Pratt and Whitney also contributed to progress on the capacitor project. Two additional tasks were undertaken: dielectric and electromagnetic field modeling, and electrical performance testing. Using a static field model, geometric field intensification and the effects of microcontamination have been investigated. An interface modification technique which may lead to increased performance has been modeled. We studied the dynamic effects of rapid discharge and measured the basic performance parameters of capacitance and voltage breakdown. These models and measurements are crucial to the success of the NMC program.

Reduced funding resulted in limitations in the number and variety of test devices that could be fabricated, and there was insufficient time to make devices suitable for more advanced testing, such as partial discharge studies. Continued funding is expected in FY-96, which should allow us to complete many of these important efforts.

Below is the list of deliverables as specified in our original proposal. We completed the tasks on the list, except for those items marked with an asterisk.

1) Capacitors fabricated for testing:
   Single-layer test films, metal-oxide dielectrics (e.g., Al_2O_3);
   * Single-layer test films, binary metal-oxide dielectrics (e.g., $CaTiO_3$);
   Multilayer capacitors, metal-oxide dielectrics.
2) Electrical measurements performed on test films:
   Breakdown voltage (DC);
   Capacitance and ESR (using HP4194A network analyzer);
   Long-term voltage withstand;
   * Partial discharge (maybe; samples are very small).
3) Electrical measurements performed on multilayer capacitors:
   * Breakdown voltage over temperature;
   Capacitance over temperature (using HP4194A network analyzer);
   ESR over temperature and frequency (using HP4194A network analyzer);
   ESL (using HP4194A network analyzer);
   * Long-term voltage withstand over temperature;
   * Partial discharge;
   Rapid (single-shot) discharge performance.
4) Dielectric modeling and validation:
   Predict and verify breakdown and degradation mechanisms;
   Suggest improvements in dielectric composition or application, verify results;
   * Analyze partial discharge data, attempt to correlate with breakdown mechanisms;
   * Determine if partial discharge can be used as an R&D and/or production QA tool.
5) EM field modeling and validation:
   Develop idealized model of EM fields in NMC structures;
   Suggest optimal geometry, predict performance, and verify results
   – increase breakdown voltage
   – minimize ESR and ESL.

NMC Device Fabrication

The basic NMC structure is a parallel plate capacitor consisting of alternating thin layers of metal and dielectric. Each layer is masked and sputter-deposited. The mask shuttles back and forth for each new layer, successively building a layered structure where alternate metal layers overlap and connect at the edges. Dielectric layers are totally sealed inside metal.

Metal thickness should be small to minimize its contribution to the total mass. A practical minimum thickness is about 0.1 μm, and an upper limit might be 5 μm.

Choice of metal is flexible with this process. Copper is generally not acceptable because it diffuses into many dielectrics and often does not form an exceptionally smooth interface. Aluminum suffers from extreme roughness unless special deposition procedures are implemented. Low-copper alloys, such as Monel, are acceptable, as are nickel alloys, such as stainless steel. We can also use very thin layers of precious metals to modify the work function of the conductor surface, while using a cheaper material as the bulk conductor. This subject was investigated from a theoretical standpoint and will be discussed later.

As with the metals, dielectric material choice and thickness are also flexible. Results from our initial test films in 1994 indicated that simple metal oxides, such as zirconia and alumina, provide good energy density and are easy to deposit. Thickness can be varied as desired to meet voltage standoff requirements. Our FY-95 deliverable for DOE calls for nominal 500 working volts DC.

Capacitor Fabrication

Arrays of test capacitors were deposited on high-quality silicon wafers and on float glass. Both of these substrates provide very smooth surfaces. Silicon wafers were thermally oxidized to 1.0 μm of SiO_2, to prevent conduction through the substrate, which would cause interactions among capacitive elements. Conductors were copper, stainless steel, or proprietary alloys. Stainless steel was used preferentially because of its very smooth surface when sputtered. However, its resistivity is very high, hundreds of times greater than copper, due to its microcrystalline structure. This results in very high ESR in a capacitor, which, of course, is undesirable. Other composites will be investigated in future work.

Normal conductor thickness was 0.2 μm. Dielectrics were alumina, which has a bulk dielectric constant (k) between 8 and 9, and proprietary mixtures, with k ranging from 10 to 18. Dielectric layer thicknesses were varied from 0.5 to 3.5 μm. Multiple small capacitors are deposited to permit us to destructively test a "set" of capacitors on one wafer, and to avoid total failure of the sample in case a defect is encountered. **Figure 1** shows the layout of a typical sample. Large capacitors were also fabricated with a 5 cm square mask. Capacitors with one and two dielectric layers were fabricated.

Extensive modification of the deposition chamber was required before these devices could be fabricated. The chamber was rotated to a horizontal position to facilitate substrate handling. Automated masking equipment with process gas control and

three axes of motion control was designed, installed, programmed, and tested to suit the conditions of these experiments. Large (10 cm × 25 cm) sputter sources were installed to assure uniform film thickness profiles over the entire substrate area. Several runs were made to verify deposition rate, uniformity, and dielectric stoichiometry. Procedures for substrate preparation and handling were developed and gradually improved during the experimental series. All of these steps are time consuming, but necessary for the production of high-quality capacitor structures.

Electrical Testing Procedures

High-Frequency. Capacitance, ESR, and ESL were measured with a Hewlett-Packard HP4184A analyzer using low-inductance probes installed in a fixture with three-axis precision positioners. This apparatus has a residual inductance, when nulled, of less than 0.3 nH and a resistance of less than 0.5 mΩ. Frequency response extends from 100 Hz to 15 MHz. A DC bias of up to ±40 V can be applied to observe C-V characteristics.

Routine calibration and probe zero/offset adjustments were performed prior to each measurement session. The analyzer displays impedance magnitude and phase vs frequency, and computes equivalent circuit parameters using a selected model. The model for these capacitors is a series RLC network, where R is the equivalent series resistance (ESR), and L is the equivalent series inductance (ESL). Because the test dots are very small, the ESR and ESL measurements were often dominated by parasitic effects.

Breakdown Voltage. Breakdown voltage was determined by applying a variable DC high voltage as shown in **Fig. 2.** The specimen was placed in a plastic dish filled with 3M Fluorinert FC-72 to prevent corona-induced damage or breakdown. A series resistor limits the applied power. Polarity was switched during some tests, but no significant difference in results was observed. Voltage was increased manually in 100 V steps, or by smooth, continuous adjustment. The latter technique results in a higher apparent breakdown voltage. The operator observes the current during the test, and notes the presence of current spikes that indicate self-healing breakdowns, and large step increases in current, indicating device failure. A record of current vs voltage is kept during the test. Just prior to breakdown, the measured current is required to calculate the actual capacitor voltage according to the relation:

$$V_c = V_s - IR \qquad (1)$$

where V_c is the capacitor voltage, V_s is the source voltage, I is the measured current, and R is the series resistance, 100 kΩ.

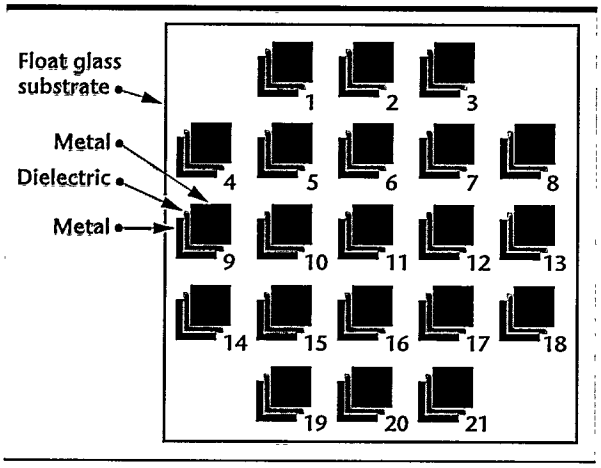


Figure 1. Capacitor deposition pattern.

Progress

Electrical Testing Results

Following is a general discussion of the results and the associated models.

Capacitance and Dielectric Constant. Capacitance was measured for each functional dot. Values ranged from 0.5 to 5 nF for all samples, the variation due to changes in dielectric thickness and composition. The value for thickness used in our calculations is that of one of the center dots, which are in the most uniform deposition region. From capacitance, we calculate the effective dielectric constant, or relative permitivity, for the multilayer composite material according to the relation:

$$k = \frac{Cd}{\varepsilon_0 a} \qquad (2)$$

where C is capacitance, d is dielectric thickness, a is the dot area, and ε_0 is the permitivity of free space, 8.854×10^{-12} F/m. The published value of k for bulk Al_2O_3 is 8.8, and this data falls within that range.

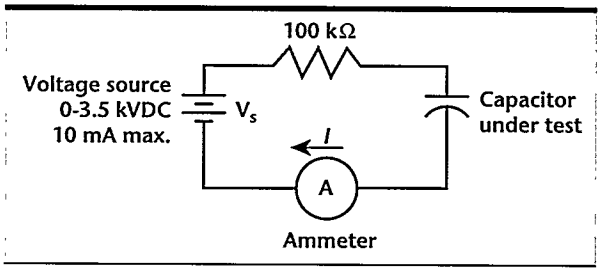


Figure 2. DC breakdown voltage test circuit. The resistor limits current to a safe value.

There is little information available on oxides of tungsten; WO_3 is reported as having an (anisotropic) dielectric constant of 6,000 to 90,000, which is very appealing.[2] We are still analyzing deposited samples and suspect that $WO_{2.5}$ is the nominal stoichiometry present for the samples with $k = 48$. Dielectrics built from our proprietary mixtures provided the best results in this experimental series.

From test films with multiple dots, capacitance can be used as a way to infer thickness. Thickness of the center dot was measured and used to correlate d vs C. A contour plot (**Fig. 3**) was interpolated from the measurements obtained.

Capacitance over temperature was measured by placing the test wafer on a small electric heater. Results for one sample are shown in **Fig. 4**. This metal oxide dielectric exhibits a linear and positive temperature coefficient of 467 ppm/°C, which is satisfactory performance for many applications.

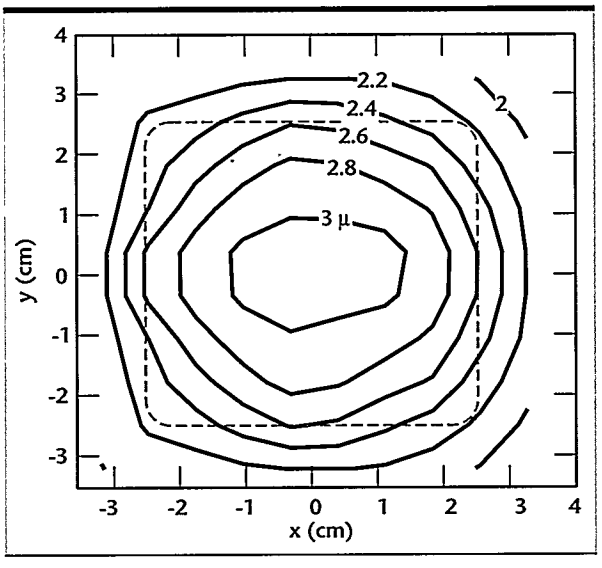


Figure 3. Dielectric thickness contour plot for representative sample.

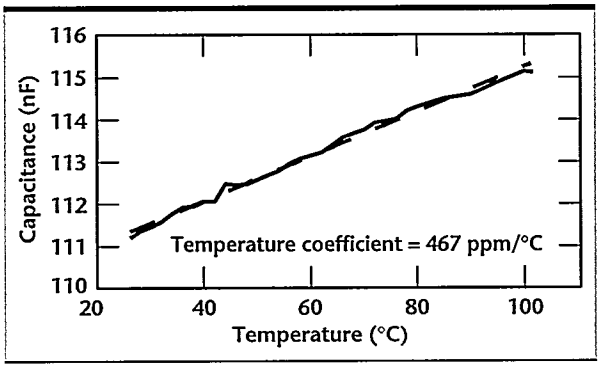


Figure 4. Temperature test for representative sample.

Breakdown Voltage. A wide variation in breakdown voltage was observed among test films and among individual dots. We attribute much of this variation to the presence of particulate contamination. This clearly indicates a need for careful substrate preparation, cleanroom practices, and careful design of the deposition system to minimize contamination.

While ramping the DC voltage during breakdown testing, all samples showed momentary current spikes, sometimes accompanied by an audible click and a small flash of blue light, indicative of an arc. These events were generally followed by a reduction in steady-state leakage current. This is typical self-healing behavior, where a defect in the dielectric causes localized electric field intensification, followed by breakdown and vaporization of material in the area of the defect. After the defect is cleared, the potential can be increased until another weak point fails in a similar manner. Eventually, a catastrophic failure occurs, permanently shorting the film.

The maximum breakdown voltages observed were 1.6 MV/cm for a single dot with an alumina dielectric, and 0.64 MV/cm for a 23 cm^2, two-layer capacitor with proprietary dielectric. Most of the dots exhibited much lower breakdown voltage, on the order of 0.5 MV/cm. Virtually all breakdowns occurred within the dielectric rather than at the conductor edges, which indicates the presence of microcontamination or substrate roughness. Scaling the dielectric area by a factor of 200 to the largest device fabricated did not degrade the average value. This is very good news because it indicates that breakdown is limited by a process parameter (cleanliness) rather than by an intrinsic material parameter. Future work will therefore be directed toward process improvements.

Energy Storage. Knowing the breakdown voltage, capacitance, and dielectric volume, we calculate peak energy density from

$$W_d = \frac{Energy}{Volume} = \frac{0.5\,CV_b^2}{ad} \quad . \tag{3}$$

Maximum energy density, considering only the dielectric material, was 0.27 J/cm^3. A practical multilayer capacitor would have up to 10% of its volume occupied by conductor material, reducing theses figures to 0.24 J/cm^3. Additional reductions in working energy density are required for safety factor derating. Process improvements will rapidly improve voltage breakdown figures; 10 J/cm^3 was achieved with small test films in 1992 and 1993.

Electromagnetic Field Modeling

Effective use was made of LLNL electromagnetic modeling capability. Static field models were applied to simple parallel plate capacitors with tapered conductors, to multilayer capacitors, and to microcontamination studies where small particles are placed in the dielectric. Additional studies covered the electrostatic effects of interfacial layers, and the electrodynamic forces produced during transient discharge.

Modeling Tool. The modeling tool used in this effort is the Maxwell 2-D Field Simulator (ANSOFT), a finite-element, frequency-domain code running on a UNIX machine and currently residing in a Sun/SPARC work station. This X Window-based field simulator is capable of computing electromagnetic fields (DC and AC) for structures with uniform cross-section or full rotational symmetry. It can handle objects of vastly different sizes, up to 1000:1, and thus, is ideal for this modeling effort. This tool also comes with a drawing package for the development of geometric models, and a post-processor for the analysis of the modeling results.

Static Field Geometries. Figure 5 shows the structures of a single-layer capacitor. To conserve computational resources and time, only the left half of the structure was modeled and an even-symmetry plane was placed at the right-hand boundary of the model. A 'balloon' boundary condition, which simulates an infinitely large space beyond the boundary, is applied to the other three boundaries of the model. The wedge-shaped edge of the upper electrode approximates the structure of the shape of a real capacitor's conducting plates. The field intensification factor has been computed for structures with different edge slope.

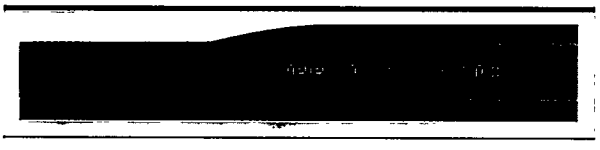


Figure 5. Single-layer model.

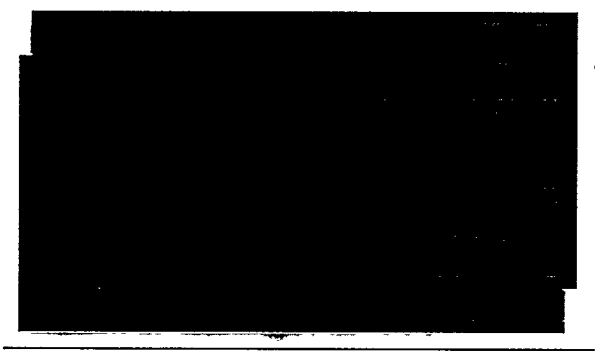


Figure 6. Multilayer model.

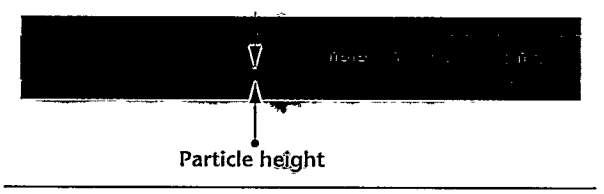


Figure 7. Microcontamination model, showing triangular particle.

The 'balloon' boundary condition was applied to all boundaries of the model of the multilayer capacitor depicted in **Fig. 6**. Electrically, the only difference in this structure is that all electrodes are embedded in the dielectric material, whereas in the single-layer structure, the electrode is sandwiched between free-space and the dielectric material.

The effect due to microcontamination was also studied in this effort. **Figure 7** shows the model of a generic capacitor with a small piece of triangular metallic material attached to the bottom electrode. The E-field was computed for contaminants of different shapes and sizes.

Static Field Results. Figure 8 is representative of the results displayed by the Maxwell 2-D model for a single-layer capacitor. The plot showing relative electric field magnitude vs distance is overlaid on a cross-section of the capacitor device. As expected, field intensity grows rapidly at the sharp discontinuity, and dielectric breakdown is expected at such a location. The magnitude of enhancement is important, because it allows us to approximate the true dielectric strength of the material under test.

Table 1 summarizes the modeling results for the single-layer structure. As shown in the table, the intensification factor is about 5 at the edge of the

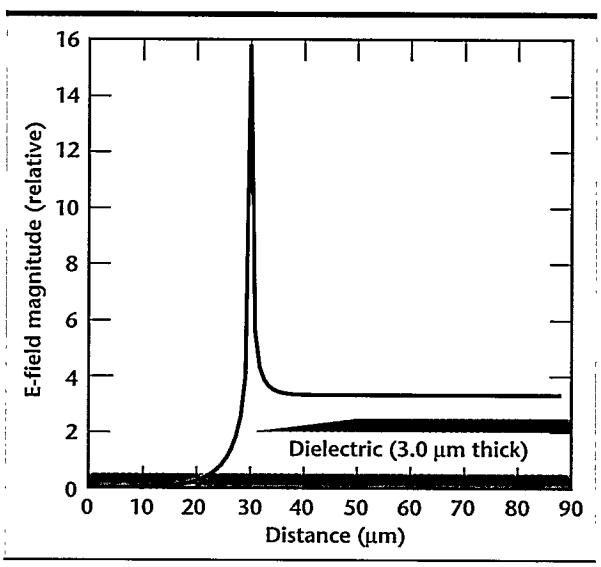


Figure 8. Electric field intensity along a single-layer capacitor. Note intensification at the sharp tip.

electrode, and is almost independent of the slope of the transition region. The change from a smooth transition (slope = 11.3°) to an abrupt transition (slope = 90°) is only 6%.

The intensification factor of the multilayer structure is about 3.5, approximately 30% less than that of the single-layer structure. This reduction is reasonable, since the electrode in the single-layer structure is sandwiched between dielectric material with high permitivity and free space, meaning that electric flux lines at the edge of the electrodes are more likely to be confined inside the dielectric. On the other hand, electrodes in the multilayer structure are embedded in the dielectric and flux lines are thus more spread out, resulting in lower field intensity.

Table 1. Edge effect.

Slope	Intensification
1 μm / 15 μm (3.81°)	4.8
1 μm / 8 μm (7.13°)	4.8
1 μm / 0 μm (90.0°)	4.5

The effects of field intensification due to microcontaminants are summarized in **Tables 2** and **3**. It is obvious that the sharpness, not the size, of the contaminants is the dominant factor of the field intensification. **Table 2** shows that varying the sharpness can increase the field by as much as 500%, while **Table 3** shows that increasing the size of the contaminants 25 times only increases the field by less than 40%. The results on contaminants with semi-circular cross-sections show even less size effect (< 10%).

Dynamic Effects Due to Rapid Discharge. It has been observed that during transient discharge, a force is exerted on the electrodes of the capacitor, and physical damage to the structure of the capacitor may result if the discharging process is very rapid. In general, there are two counteracting forces, an attractive electric force and a repulsive magnetic force. The magnetic force, when dominant, is responsible for the familiar destruction of pulse power equipment. In the case of capacitors, the outer foils may be explosively thrown off.

According to Ampere's law, a magnetic field will be generated around a closed path if current flows through the loop. Although in this case no charge

Table 2. Microcontamination effect (sharpness of triangles).

Height (nm)	Width (nm)	Angle (°)	Intensification
500	200	22.6	14.8
500	1000	90	7.9
500	2000	126.8	6.2
500	3000	143.2	4.2
500	4000	152	3.2
500	5000	157.4	2.7

Table 3. Microcontamination effect (size).

Height (nm)	Width (nm)	Angle (°)	Intensification	Ratio
Triangles				
500	1000	90	7.9	
100	200	90	5.76	1.36
500	5000	157.4	2.7	
100	1000	157.4	2.02	1.34
Semicircles				
500 (radius)	N/A	N/A	2.36	
100 (radius)	N/A	N/A	2.19	1.08

carrier crosses the dielectric between electrodes, displacement current (I_d) exists inside the dielectric during the discharging process. Because of continuity, I_d equals the discharge current. With this information we can calculate the magnetic field at the edge of a circular-plate capacitor with radius r as:

$$H = \frac{I_d}{2\pi r} \quad . \tag{4}$$

To find the magnetic force exerted on the electrodes, we need to know the current distribution inside the electrodes,

$$\vec{F}_m = \vec{I} \times \vec{B} \quad . \tag{5}$$

Calculating the current distribution has been the most difficult part in the process of estimating the force, because the general-purpose EM codes do not directly compute the transient current at the surface. Some LLNL- developed dynamic codes could provide an answer, but these codes would require special set-up and modification.

For a reasonable approximation, we can estimate the charge distribution inside the electrodes (plates) by ignoring the edge effect. If we assume that discharge current is removed from the center of the plate, the initial magnetic force (repulsive) is:

$$F_m = \frac{\mu_0 \, Q_0^2}{8 \, \pi \, \tau_{RC}^2} \tag{6}$$

where τ_{RC} is the RC time constant of the circuit and Q_0 is the total charge. On the other hand, the electric force (attractive) is :

$$F_e = \frac{Q_0^2}{Cd} \tag{7}$$

where C is the capacitance and d is the separation between the plates. Using typical numbers such as d = 3 μm, the radius of a circular capacitor, r = 3 cm, and a voltage of 1200 V, we calculate C = 0.127 μF and Q_0 = 0.155 μC. The total force is thus equal to

$$F = F_m + F_e = \frac{1.2 \times 10^{-15}}{\tau_{RC}^2} - 6.3 \times 10^4 \quad . \tag{8}$$

The total force equals zero when τ_{RC} = 0.14 ns. The initial forces calculated here are the maximum forces, because they will decay as charges are being removed from the system. As a result, the plates will begin to experience a repulsive force when the RC time constant is smaller than 0.14 ns. This force is inversely proportional to the square of the time constant.

We have developed a more accurate model, including the edge effect. In this model the charge distribution at the edge is approximated with a step function (**Fig. 9**). From the modeling results, we estimated the transition region to be about 3 μm. Thus, the width of the step was set to 1.5 μm. Using this model, a parallel, but much more complicated formula is derived. The details of the derivation are documented in **Table 4**. The result of this model does not deviate significantly from the result of the model without edge effect, which is expected, because the charge at the edge is only a small portion of the total charge.

In conclusion, the magnetic force is negligible, compared to the electric force, unless the RC time constant of the discharging circuit is close to, or less than 0.14 ns. If this short τ_{RC} is exceeded, the magnetic force could be devastating, since it is inversely proportional to the square of τ_{RC}. For instance, at 0.014 ns the repulsive (magnetic) force will be about 100 times the attractive (electric) force, about 6 MN. In practical capacitors, the ESR, ESL, and circuit parasitics will probably exceed the required discharge resistance. For the example above, 0.14 ns and 0.127 μF implies R = 1.2 mΩ, which is about equal to the expected ESR. Any inductance will disallow the short discharge time.

No attempt was made to model a multilayer structure, due to time constraints. Such structures will provide much larger capacitance while maintaining low ESR and ESL. Thus, it is probable that a destructive discharge could occur just as it does in the larger, commercially-available capacitors. A project incorporating a time-variant 3-D field model may be subsequently proposed.

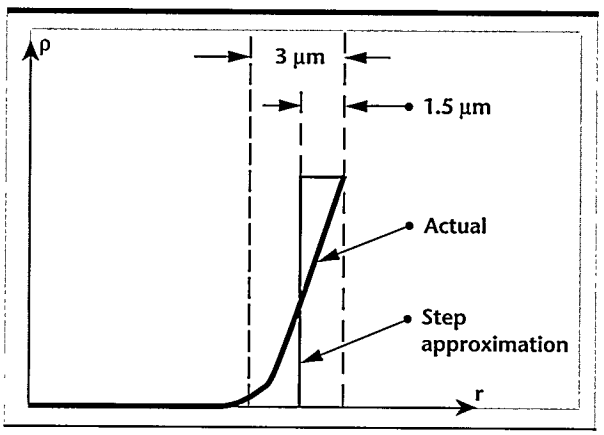


Figure 9. Approximation of the charge distribution at the edge.

Effects of Interface Electron Injection. Adding a layer of conductive material with high work function at the metal-dielectric interface can significantly change the breakdown threshold. The actual problem is extremely difficult to solve, since it involves tunneling of the work function barrier and avalanche multiplication in the dielectrics. For a 1-D approximation, O'Dwyer[1] has formulated this solution: the breakdown field strength is proportional to the 3/4 power of the work function of the metal (e.g., doubling the work function will increase the breakdown field by a factor of 1.68). Yee (LLNL) has suggested that a thin layer, on the order of 100 Å (30 to 40 atoms), can probably be modeled reasonably well using the bulk properties of the material. Barbee (LLNL) indicates that a radius of about 5 lattice constants, which is about 50 Å, would be the minimum. We would be conservative and deposit 100 Å, to avoid small defects.

Using the 3/4 power effect for work function, we consulted a table of metal data and found a few numbers that look encouraging. Assuming that aluminum is a likely choice of conductor metal for conductors, we can see if there is a significant advantage to adding an interfacial layer of another metal to improve breakdown voltage. Note that there might be other reasons to use a particular metal as an interfacial layer, besides work function; e.g., diffusion, electromigration, or oxygen transport management are all important factors. Cost must also be considered. Though there is only a small amount of material required, the extra processing step may not be cost-effective. **Table 5** lists the work functions for several candidate metals, the expected decrease in E-field intensity (compared to aluminum), and the resulting increase in energy density, which is proportional to the square of the voltage.

In the best case, where a thin layer of platinum is deposited, energy density may increase by a factor of 1.4. This 40% improvement may well be worth the investment in equipment and materials. Unfortunately, time constraints did not permit us to verify this finding in the laboratory. This effect has been noted in a recent patent disclosure. We are not aware of its use in other dielectric applications.

O'Dwyer also indicates that traps and defects that are close to the metal/dielectric interface tend to intensify the avalanche multiplication, much more

Table 4. Mathcad solution for dynamic effects due to rapid discharge.

$$\mu 0 := 4 \cdot \pi \cdot 10^{-7} \cdot \frac{\text{henry}}{\text{m}}$$

$$a := 3 \cdot 10^{-2} \cdot \text{m}$$

$$\tau := 1 \cdot 10^{-9} \cdot \text{sec}$$

$$Q0 := 156 \cdot 10^{-6} \cdot \text{coul}$$

$$V := 1200 \cdot \text{volt}$$

$$p0 := \frac{Q0}{\pi \cdot a^2}$$

$$d := 3 \cdot 10^{-6}$$

$$Fm = \frac{\int_0^{2\cdot\pi} \int_{0\cdot m}^{a} \left(\frac{p0 \cdot \pi \cdot r^2}{2 \cdot \pi \cdot r} \right) \cdot \frac{p0 \cdot \pi \cdot \left(a^2 - r^2\right)}{2 \cdot \pi \cdot r} \cdot r\,dr\,d\theta \cdot \mu 0}{\tau^2}$$

$$Q0 := 156 \cdot 10^{-4} \cdot \text{coul}$$

$$Fm = \frac{1}{8} \cdot p0^2 \cdot a^4 \cdot \pi \cdot \frac{\mu 0}{\tau^2} = 1.217 \cdot 10^3 \cdot \text{kg} \cdot \text{m} \cdot \text{sec}^{-2}$$

$$Fe = V \cdot \frac{Q0}{d} = 6.24 \cdot 10^4 \cdot \text{kg} \cdot \text{m}^2 \cdot \text{sec}^{-2}$$

$$t_eq := \sqrt{\frac{1.205 \cdot 10^3 \cdot \text{newton} \cdot 10^{-18} \cdot \text{sec}^2}{6.24 \cdot 10^4 \cdot \text{newton}}}$$

$$t_eq = 1.39 \cdot 10^{-10} \cdot \text{sec}$$

Table 5. Effects of surface work function on capacitor energy storage.

Metal	W.F. (eV)	E-field *	Energy *
Al	4.2	1	1
Au	4.8	1.11	1.22
C (diamond)	4.8	1.11	1.22
Co	5.0	1.14	1.30
Ni	4.9	1.12	1.26
Pd	5.0	1.14	1.30
Pt	5.3	1.19	1.42
Re	5.0	1.14	1.30
Zr	3.8	0.93	0.86

* aluminum = 1.0

than traps in the bulk of the dielectric do. Indeed, our biggest contamination problems occur when moving a mask or changing sources, so contaminants are most likely to end up at the interface, the worst possible location. Again, contamination control proves to be a key factor in improvement of breakdown voltage.

Future Work

Most of the work remaining in the NMC program is related to materials science and to process improvements. Particular items of interest are high-k dielectrics, high-conductivity metals with excellent compatibility with dielectrics, improved masking techniques, and microcontamination control.

In the area of electromagnetic modeling, a proposal has been submitted in association with the development of advanced 3-D modeling codes. Useful models include a 3-D time-variant follow-on to the simple transient discharge analysis performed here, and a 3-D analysis of electric field intensification.

Additional electrical measurements, as originally proposed, will also be performed when suitable capacitor samples are available.

Acknowledgments

The authors would like to thank T. Barbee, for providing his expertise in materials science and multilayer synthesis; and B. Olsen, who built and operated the deposition system.

References

1. O'Dwyer, J. J. (1973), *The Theory of Electrical Conduction and Breakdown in Solid Dielectrics*, Oxford Univ. Press, New York, N.Y.

2. Young, K. F., and H. P. R. Frederikse (1973), "Compilation of the Static Dielectric Constant of Inorganic Solids," *J. Physical and Chemical Reference Data* 2(2), pp. 313–410.

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 develop compact low-cost impulse radar systems for various commercial and government applications. Integration of these hardware systems with other expertise at Lawrence Livermore National Laboratory (LLNL) in modeling, rapid prototyping, signal processing, and antenna design have led to numerous projects in support of LLNL programs, as well as generating new outside funding. Radar arrays of transmit/receive elements in support of imaging applications were key to a number of our system concepts.

Introduction

The LLNL-developed Micropower Impulse Radar (MIR) technology has quickly gone from laboratory concept to embedded circuitry in numerous government and commercial systems in the last few years.[1] 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 catalyst for a new generation of compact radar systems. Some of the past commercial successes have been widely publicized,[2,3] though products are only now (fall, 1995) starting to become available for market. More than 30 patents have been filed and more than 10 licenses have been signed on various aspects of the MIR technology. In addition, higher performance systems are under development for specific LLNL programs and government reimbursables.

In spite of the above successes, there are still many new directions that we plan to explore, to continue our leadership role in MIR. As an entirely new technology area, rather than a single circuit, MIR has the potential to address a rich set of new applications for which there is not yet program, government, or commercial support. The work to-date has inspired numerous novel concepts in the last year. Many of these have a 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.

The objective of this project was to expand and advance our current capabilities in MIR technology along several directions, described below.

Progress

We briefly present here some of the key technical developments over the last year in MIR. Many diverse systems were built, each having its own circuit board, antennas, packaging, and (sometimes) computer interface. Nearly 200 complete radar systems were assembled, tested, and fielded during FY-95 alone. The systems can be roughly categorized by four sensor types: 1) motion sensors; 2) distance sensors; 3) imaging sensors; and 4) communication devices. Tables in each section below give general background on the sensors. (For more information on these systems, and on the overall MIR Project, see our world-wide web page at http://www-lasers.llnl.gov/lasers/idp/mir/mir.html.)

Motion Sensors

The simplest MIR circuit that has been designed and fabricated is the Motion Sensor (see the block diagram and photograph in **Fig. 1**). It generates an

ultra-wideband impulse (see specifications in **Table 1**) that can penetrate normal walls and detect small motions of people, at ranges up to 6 m or more. It has a sharply-bounded detection radius, defined by a narrow echo-acceptance range, or range gate, that is adjustable. Only

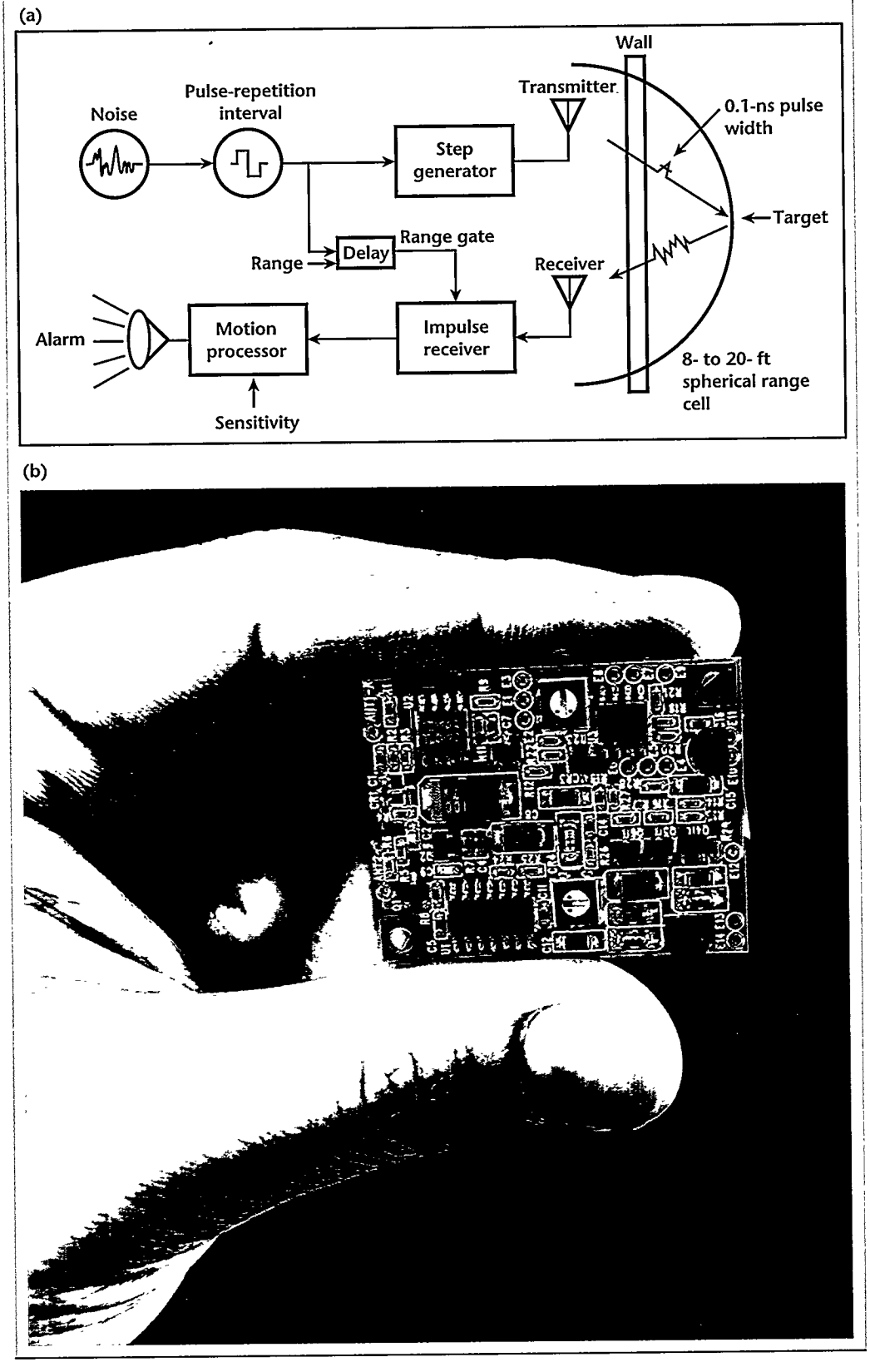


Figure 1. Block diagram (a) and photograph (b) of basic MIR Motion Sensor.

motion-modulated signals or changes from a baseline measurement are detected, thereby eliminating false triggers from stationary room "clutter." Alternate pulse types can be transmitted that can generate a filled volume of sensitivity. The Motion Sensor does not respond to objects outside of its range gate, and it does not false-trigger on small nearby objects, such as insects.

Some of properties of the MIR Motion Sensor are common to all other MIR devices described in this report. The frequency band allows penetration through many common materials such as concrete, most wall materials, ground, mud, and snow, making it environmentally robust. It normally operates on a standard 9-V battery, while some extremely low power versions have been built that can operate continuously for several years on ordinary alkaline batteries. Pulse emission times are randomized, and many pulses are averaged to eliminate interference with other sources or MIRs, and also to make them virtually undetectable. Component costs are very low, so that in mass quantities entire radar units should cost less than ten dollars. The MIR Motion Sensor is an excellent security and safety device that has numerous applications in remote sensing.

The entire circuit, minus some drive transistors, could be customized to single-chip integration on an application-specific integrated circuit (ASIC). We have been discussing this possibility with several commercial and military chip foundries, and we expect the Motion Sensor to be the first such IC. The sampler circuit is essentially identical to the original transient digitizer circuit now in use by LLNL's NOVA laser, and field trials have shown it to be the chosen technology for NIF diagnostics. Continued efforts to support NIF will point to new innovations for MIR, and vice versa. Other direct applications for the basic MIR Motion Sensor have been found in NAI, W-Division, Special Forces and other parts of the Nuclear Weapons Complex, as well as in the commercial sector.

Table 1. MIR Motion Sensor Specifications.

Parameter	Value
Antenna pattern (H-plane)	360° with a dipole antenna, 160° with a cavity-backed monopole, and narrower with horn/reflector/lens
Center frequency	1.95 or 6.5 GHz +/-10%
Emission bandwidth	500 MHz @ 1.95 GHz center
Average emission power	~ 1 μW (measured)
Duty cycle	<1%
PRF (average)	2 MHz +/- 20%
PRF coding	Gaussian noise, or low coherence swept FM, or pseudo-noise
Receiver noise floor	<1 mV rms
Receiver gate width	250 ps for 1.95 GHz system
Range delay	RC analog, pot/DAC controllable
Range delay jitter	<1 ps rms
Range delay stability	RC component-limited over temperature (drift in range delay due to RC expands/shrinks the detection shell)
Detection range	Adjustable from 5 cm to >8 m
Motion passband	0.3 to 10 Hz, Doppler-like signature
Analog output	~0.1 to 2 V peak on motion sensing, hand motion at 2 m gives ~300 mV peak
Receiver gain	70 dB
Power	5 V @ 8 mA, normal power mode, 2.5 V @ 20 μA, long battery life version
Size	1.5-in. square SMT PCB with 1.5-in. long wire dipole elements
Semiconductors	74AC00 CMOS (1ea), bipolar or CMOS op amps (2 ea, quads), Si-bipolar RF transistor @ >4 GHz f_t (2 ea), Schottky diodes, C_j (0) < 1 pF (2 ea)

Table 2 lists several separate MIR devices, and their potential uses, that are all based on the standard motion sensor circuitry. All have the same basic specifications for the pulser and sampler circuits, and all are measuring nearby movement, but each has its own operating characteristics. We have detailed information on each of these systems, but we will briefly mention only one of them here—the MIR Organ Sensor **(Fig. 2)**. By setting the range to be very short and altering the antenna, we have modified the Motion Sensor to detect subtle motion of internal body organs. The emissions from such a device are medically safe (at least three orders of magnitude lower than international safety standards), and yet it is very sensitive to muscle motion, such as the heart, lungs, or vocal cords. It penetrates through clothing and is being considered for battlefield medical units by Walter Reed Army Institute of Research. As a non-acoustic

Table 2. MIR Motion-Sensing Hardware. Note that all of these devices have a single fixed range gate.

Name	Description	Potential uses
Motion Sensor	Measures motion at a fixed range-gate	Security, automotive warning devices safety, robotics and manufacturing, proximity fuses
Presence Sensor	Same, but with background subtract; detects change from background	Same as above, but for unmoving objects; e.g., automatic lighting, faucet shut-off
Field Disturbance Sensor	Low-freq FM chirped homodyne radar with range-gating	Same as above, but at conventional radar frequencies for longer range, larger objects
Stud Finder/Metal Detector	Like short-range Presence Sensor; detects changes from background	Locate (not image), studs in walls, rebar in concrete, metal; any changes in baseline
Trip-line	Separate transmitter and receiver; detects small changes in field between them	Perimeter or border surveillance, automatic switches, industrial automation
Smart Wire	Same, but the pulse is launched on a wire or other conductor	Power window shut-off, touch switches, movement near metals
Fluid-level Switch	Measures presence of nearby fluid	Placed on tank wall, will provide auto shut off, overflow protection; IV bubble detection
Organ Sensor	Short range, senses movement inside the body (non-localized)	Heart monitor, muscle physiology, breathing motion, speech measurement

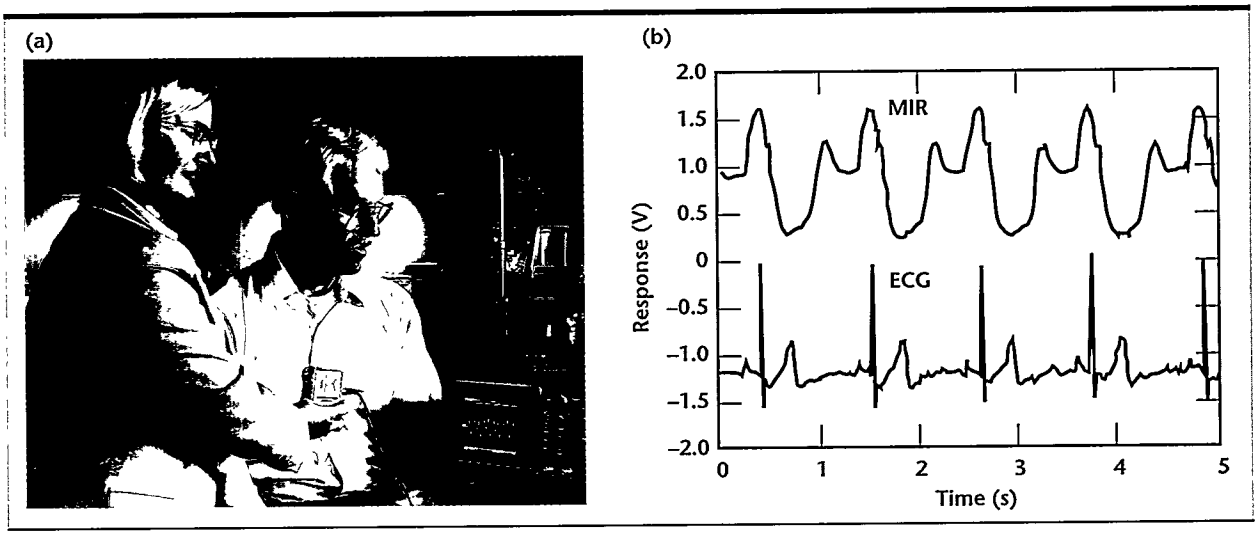


Figure 2. Photograph of the MIR Organ Sensor (a) and the return signals from the heart (b), compared to a standard electrocardiograph (ECG).

microphone, it can potentially detect and amplify motion in the acoustic range, even in very noisy environments, such as a helicopter or ambulance.

In FY-95, we have modified the basic Organ Sensor to have 16 possible range gates and "quadrature" settings (see the Heart Monitor in **Table 3**). The quadrature measurements come about when a damped sinusoidal radiated waveform is used. Pairs of motion sensitive regions can be offset by 1/4 wavelength, so vector processing can produce improved signals. The other range gates can be set for different depths, depending on the size of the person or the distance from the organ. Gate and quadrature settings are currently preset in the device by a programmed logic device (PLD) on the board.

Distance Sensors

Expanding on this idea of using multiple range gates, we can now build MIR devices with the ability to measure distances in various ways. By sweeping through different ranges, still averaging many pulses at each range, one can get a slowed-down or equivalent-time reading of all return pulses within the swept distance. Further signal processing can then extract features about the "scene" surrounding the radar. Several MIR devices have been built with this concept, and are summarized in **Table 3**. The MIR Heart Monitor mentioned above is one such device, and we will briefly mention others.

Table 3. MIR Distance-Sensing Hardware. All of these devices have multiple adjustable range gates.

Name	Description	Potential uses
Heart Monitor	16 programmable range gates, high-gain antennas	Heart monitor, breathing sensor, system for detecting range of incoming targets
Dipstick	Measures reflections of pulse caused by dielectric disturbance along wire	Fluid level sensing for cars, reservoirs, vats, silos; remote environmental monitor
Time-of-flight Positioner	Measures time/distance between separate transmit and receive antennas	Accurate positioner, virtual reality, video conferencing, localizing rescue personnel
Time-of-arrival Locator	Wireless triangulation system from one transmitter and multiple receivers	Alternate relative positioning applications; "local GPS" for centimeter accuracies
Rangefinder	Ultra-wideband radar sweep measurement at equivalent time	Imaging, altimeters, wireless level sensors, navigation, traffic control
Laser Rangefinder	Same, but uses a small laser diode and photodiode to measure range	Line of sight distance measurement; "laser tape-measure," very narrow beam ranging

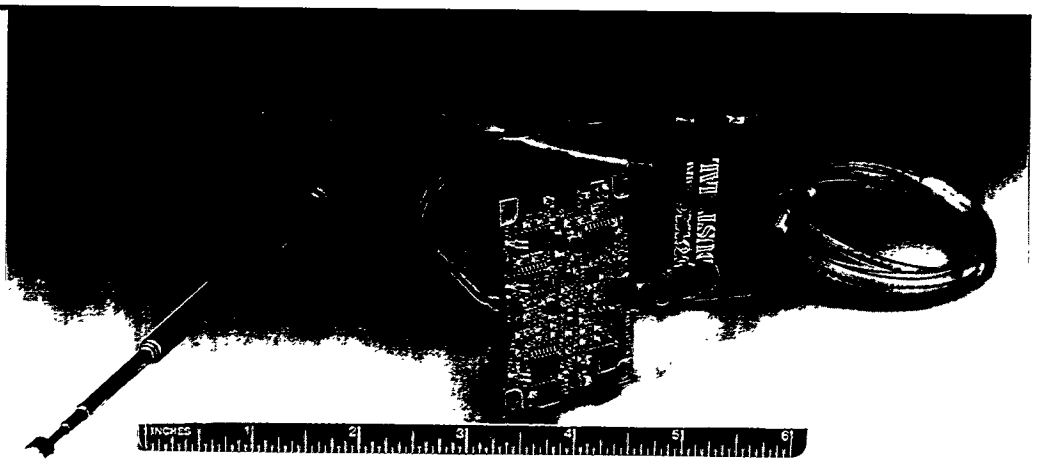


Figure 3. Photograph of electronic MIR dipstick.

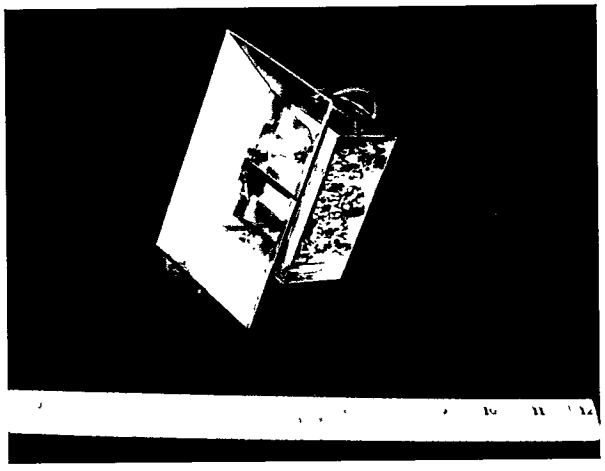


Figure 4. Photograph of MIR Rangefinder.

Rather than launching the micropower pulse in air (or in the body), we can launch it along a wire or transmission line. This is the principle behind the MIR Dipstick (**Fig. 3**). Interference by a dielectric interface contacting the wire, such as an air-fluid boundary, causes a reflected pulse that can be measured in the same way as with the previous devices. Here again, multiple range gates are needed to precisely determine the time-of-flight of the return pulse. The MIR then becomes a low-cost time-domain reflectometer (TDR). The transmission line can be a simple non-moving metal strip, or dipstick-like wire, with a length of several centimeters to many meters. It requires a volume of fluid to reflect from, and as such, it is impervious to minor wetting, condensation,

Table 4. MIR Rangefinder Specifications

Parameter	Value
Antennas	3-cm aperture each, transmit and receive, 160° beamwidth, a horn/ reflector/lens can be added
Center frequency	3 GHz, typical
Emission bandwidth	3 GHz, typical
Average emission power	~1 μW
Duty cycle	<0.001
PRF (average)	2 MHz +/- 20%
PRF coding	Gaussian noise, or low coherence swept FM, or pseudo-noise
Receiver noise floor	3 μV rms (without preamp)
Receiver gate width	100 ps
Receiver STC	~60 dB
Range sweep	RC analog, pot/DAC controllable
Range timing jitter	<1 ps rms
Range sweep stability	RC component limited over temperature
Sweep range	10 cm to 3 m, 30 m possible
Sweep rate	40 Hz, typical; can be controlled by external source or computer
Sensitivity	20 dB S/N on 4 × 6-in. metal plate at 10 ft
Signal to clutter	30 dB from 10 cm to < 3 m, without digital subtract
Analog output	+/-2 V_{max} equivalent-time replica spanning 25 ms
Digital output or input	CMOS sync, 40 Hz square wave; CMOS range, scale = 1 ft/ 2 ms pulse width
Receiver gain	100 dB (STC = max); 5 V @ <20 mA, normal power mode, long battery life version possible
Size	3 × 6-cm SMT PCB
Cost	$10 in PCB components/unit quantity
Semiconductors	74HCxx, 74ACxx, CMOS op amps, Si-bipolar RF transistors @ > 4 GHz f_t (2 ea), Schottky diodes, C_j (0) < 0.5 pF (2 ea), General purpose NPN (1 ea)

corrosion or grime on the sensor element. An improved version of the dipstick (Version 2.0) is currently being fabricated with a more accurate crystal time-base, and therefore, finer distance resolution is expected (0.1% resolution, or 1 cm sensitivity at ranges up to 10 m, as opposed to the previous 1% resolution). This same time-base will be incorporated into the other devices using swept range-gating.

Time-of-flight and time-of-arrival systems are low-cost short-range (less than 30 m) radar systems for locating an object in 3-D space. They measure the propagation time of a short RF burst traveling between a transmitter and one or many receiver units. The time-of-flight system currently requires a timing cable between sensors, and produces absolute distance measures, while the time-of-arrival system is wireless but produces relative distances that must be triangulated to find position. They can be thought of as "local GPS" devices for tracking movements of targets, robots, head position for virtual reality systems, active vehicle suspension, and others. Fire and rescue personnel may wear the transmitter while support personnel monitor their position. A wireless time-of-flight system is currently under development that will replace both of the above devices.

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. A photograph of the Rangefinder is shown in **Fig. 4**. The device has separate cavity-backed monopole antennas for transmit and receive, and operates as a compact, low-cost ultra-wideband radar with a swept range gate (see the operating specifications in **Table 4**). It generates an equivalent-time A-scan (echo amplitude vs range, similar to a WW-II radar) with a typical range sweep of 10 cm to 3 m, and an incremental range resolution, as limited by noise, of 0.3 mm. Rangefinder II, a new version, is currently under design with antenna specialists from the Computational Electromagnetics and Electronics Thrust Area. It will have improved features important to many projects, such as lower noise, more accurate time-base, and integrated wideband antennas.

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. Some versions transmit RF packets rather than a short impulse, and can provide swept-range quadrature Doppler information. A major application for this microradar is imaging (see next section). It operates in spectral regions that readily penetrate walls, control panels, and to an acceptable extent, concrete and human tissue.

The basic antennas have a very broad beamwidth and corresponding 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. Some examples of different types of antennas associated with the MIR Rangefinder are shown in **Fig. 5**.

The Laser Rangefinder is a recent invention that uses the same circuitry, but instead of emitting a microwave signal, it emits light. It is similar in

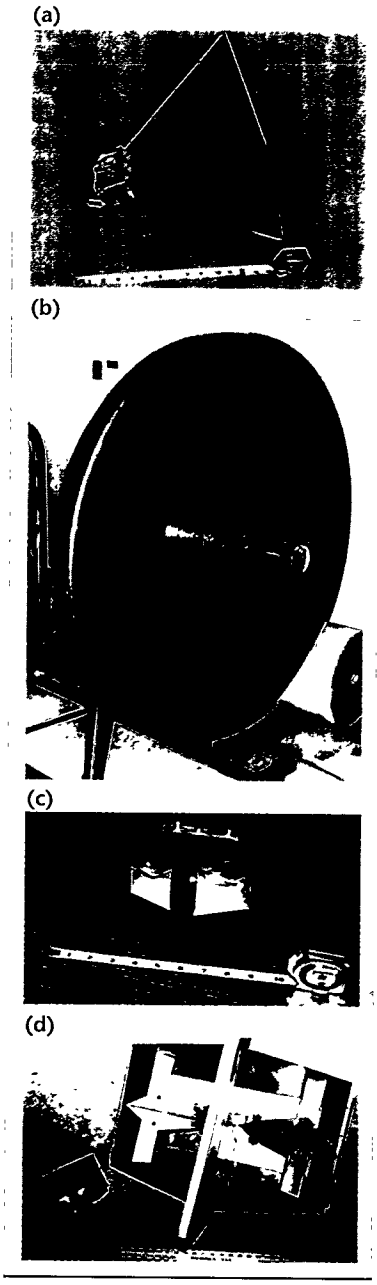


Figure 5. Antenna systems developed for use on MIR Rangefinder systems: a) altimeter for unmanned air vehicles (Pacific Northwest Laboratory); b) dish antennas for ocean wave imaging (Navy ASW); c) wall surveyor radar (Navy Special Forces); and d) underground imager for locating utility pipe.

operation to the radar rangefinder, except that a laser diode replaces the transmit antenna, and an avalanche photodiode replaces the receive antenna. The direct pulse-echo flight time of a collimated visible or infrared laser beam is used to measure the range to the spot of light projected by the laser. Thus, non-contact range measurements can be made from a single location, and the precise measurement location is pinpointed by the laser spot. Its advantages over the radar rangefinder are: 1) it has a very narrow beam to facilitate operation in tight locations; and 2) the projected spot of light tells the user where the measurement is being taken. The cost of eye-safe visible beam laser diodes is under $10, in volume. The cost of avalanche photodiodes may currently range up to $100, due to low volume usage.

The Laser Rangefinder has countless industrial safety, measurement and object profiling applications. A truly low-cost laser rangefinder will revolutionize the measurements taken; i.e., it could replace the yardstick and tape measures that are in everyday use.

Imaging Systems

The basic MIR Rangefinder described above provides the perfect vehicle for ultra-wideband imaging applications. These radars are coupled to a portable computer to form complete radar systems with either synthetic aperture or real-array imaging. Radar return signals are digitized and stored in a portable lap-top computer. Reconstruction of cross-sectional images from B-scan or waterfall type data is performed by diffraction tomography software on the lap-top.[4] 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.

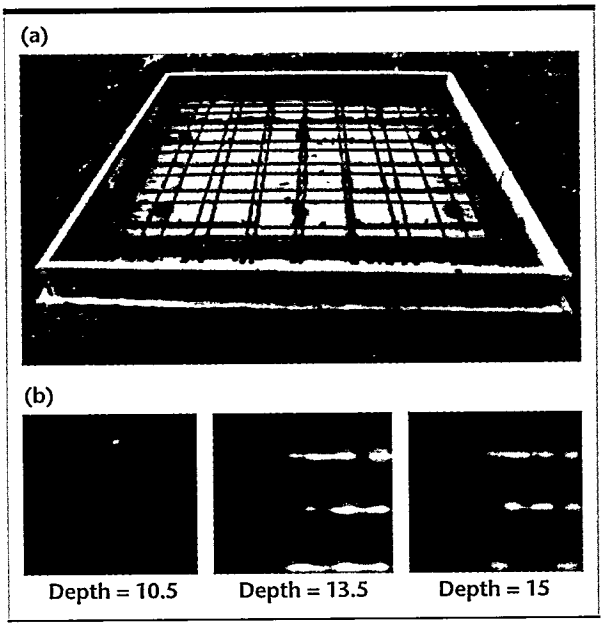


Figure 6. Imaging of concrete delaminations: (a) a concrete slab, before pouring, with simulated delaminations, and (b) MIR radar images of final cured slab covered with asphalt.

We have developed two arrays, one containing 32, and the other eight, Rangefinders for test and evaluation, and others are being designed. The imaging system detects differences in the dielectric constant of the materials under inspection. Larger differences cause larger radar return signals. Resulting images, then, have high intensities at interfaces between materials of differing electrical properties. Effective resolution of the current system is about 2 to 7 cm, depending on the medium and depth. The current array uses monostatic imaging, but future versions will be capable of multistatic operation.

There are a number of applications for these imaging systems, and we will briefly describe some of them here:

1) **Road-bed and bridge-deck inspection.** Delamination of concrete structures due to corrosion of rebar layers is a major problem in our national transportation infrastructure. We have demonstrated the capability to image the delaminations below asphalt and concrete layers, as an alternative to coring or other destructive methods. **Figure 6** shows a concrete slab (before being poured) with simulated delaminations, and the resulting radar-image through asphalt. We have a contract with the Federal Highway Administration to construct a truck-mounted imaging radar array (called the HERMES project) that can inspect the roadway for damage at near-highway speeds. The initial prototype system is expected to be deployed in the summer of 1996.
2) **Land mine and buried ordnance detection.** With 100 million mines in the world, and more being placed daily, detection and removal of them in an effective, reliable way is a major problem both for military and humanitarian objectives. Non-metallic mines pose a particular problem, because standard metal detectors are incapable of finding them. The MIR is almost the perfect radar for both hand-held and vehicle-mounted situations, because of its light weight and low power consumption.[5,6] Pictures of various horizontal subsurface slices through plastic anti-tank mines are shown in **Fig. 7**.

Image reconstruction is necessary to ascertain the shape and size of the buried object to distinguish it from ground clutter. We are

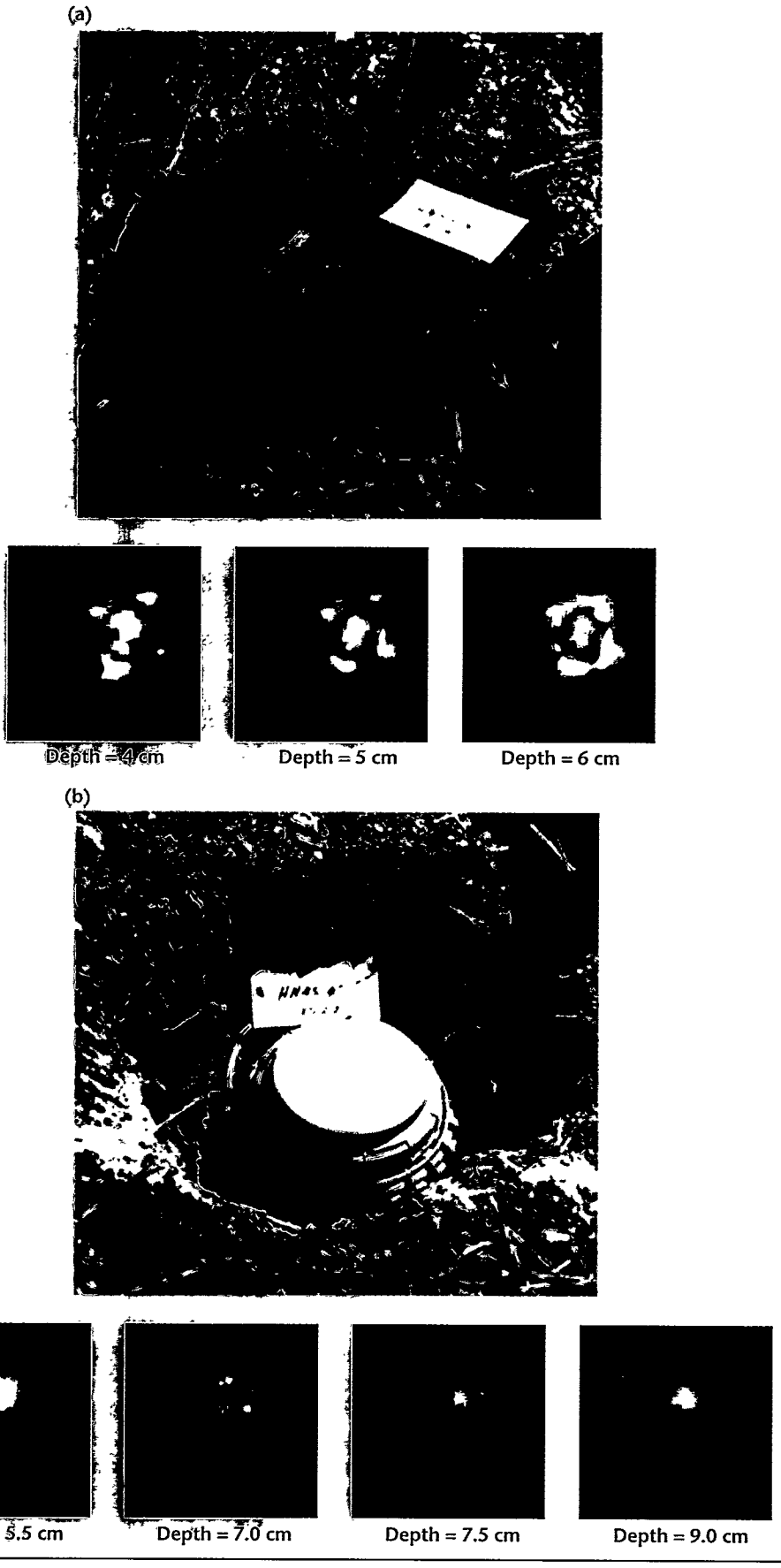


Figure 7. Non-metallic mine detection at the Nevada Test Site minefield, with MIR imaging systems. Several horizontal slices, at different depths in the ground, show images of two types [(a) and (b)] of plastic anti-vehicular mines.

currently negotiating with various military and non-governmental agencies for funding to continue refining this technique, and eventually for deployment. Buried unexploded ordnance presents a similar hazard to civilian rehabitation of an area after war, and MIR imaging could potentially address this problem also, though the specifications are slightly different.

3) **Detection of underground utility lines.** Utility companies, and others who must trench underground, are constantly concerned about damaging buried pipes or cables while digging; and often they would like to find emplaced lines for repair. Plastic piping or fiber optic lines are especially difficult to locate if detailed maps are not previously drawn (and frequently they are not). The MIR imaging system has been used to find a buried plastic water line below 25 cm of reinforced concrete and ground (**Fig. 8**), and studies are on-going to extend its capabilities. Several of our licensees are particularly interested in this problem, and one is funding us to develop a system for deeper and more sensitive imaging. The antenna in **Fig. 5d** will be used to attain greater penetration, and will have application in other areas, such as sink-hole detection, buried bunker location, and hazardous waste burial.
4) **Ocean imaging.** In support of a large Navy contract for anti-submarine warfare in the Imaging and Detection Program, we are developing small MIR-driven radar imaging systems to study ocean waves at low grazing angles. Specialized MIR transmit and receive modules with directional high-gain antennas (**Fig. 5b**) have been used at the University of California, Santa Barbara, wave simulation facility to generate and validate models of ocean-wave behavior.
5) **Wall thickness and composition.** Special Forces and rescue personnel are interested in the thickness and composition of walls in order to breech them without large collateral damage. We are building a hand-held imaging system (**Fig. 5c**) that can be scanned over walls, and rapidly produces images for estimates of the wall contents.
6) **Nondestructive evaluation (NDE).** Like the road-bed and wall imaging systems above, there are many general NDE applications, where microwaves could produce information about materials or structures that is complementary to current techniques, such as x rays or ultrasound. The MIR is particularly well-suited to detecting flaws in 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. 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.).
7) **Through-wall detection of people.** Military, law enforcement, and search and rescue teams are often confronted with the need to "see" through walls without being detected. Radar systems have been proposed in the past, but clutter from other sources generally degrades any attempt to image the scene. However, an array of MIR sensors with both range-gating and motion sensing capabilities may be able to detect and locate small movements behind walls in normal-sized rooms. Simulations and experiments are being performed to determine the dimension and element-spacing of such arrays needed to achieve accurate detection. We are actively pursuing the development of such arrays for various government agencies.

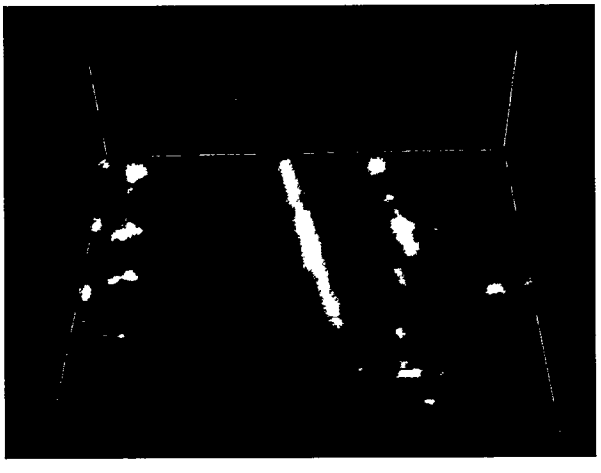


Figure 8. Underground utility imaging in ground and concrete. A plastic water line is clearly visible at 25 cm below the surface and through rebar.

It will be necessary to integrate the electronics with the antenna array in future versions of the imaging radar systems. This is needed to keep the 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 imaging arrays may also be required for specific applications.

Communication Devices

Several MIR devices were developed in FY-95 that can be classified as short range communication devices. While our patent position is being determined for some of these, we will describe only two that will open up many new applications, the Micropower Receiver, and the MIR Comm Link.

The Micropower Receiver is similar in operation to those used in garage door remote controls, and other low-cost, short-range RF remote control applications. The difference between the MIR device and the conventional receiver is typically lower power consumption, by a factor of 100 or more. The micropower receiver draws only 6 μA of supply current, so it can run continuously for a quarter of a century from a single penlight battery (neglecting shelf life).

Low current operation is vital to battery-operated remote control functions, where a receive capability is needed. Typically, a remote control transmitter will have long battery life because the transmitter is normally off until a button is pushed. A key fob "clicker" transmitter typically draws only a few milliamperes for a few tenths of a second for each push, so battery life can easily exceed one year. However, the receiver does not know when to expect a transmission, and needs to be on continuously. In order to get a year of battery life from a small lithium cell, current draw needs to be well under 100 μA.

The Micropower Receiver makes possible important new applications, such as 1) keyless entry systems; 2) stand-alone, instantly installed MIR proximity sensors with a 2-way handshake to a central controller; 3) inventory and attractive-asset tracking ID cards; 4) medical implant data links; and 5) avalanche rescue transponders. A common scenario will be where an electronic device in a remote location remains dormant for long periods of time, except for the Micropower Receiver which remains active, until it is awakened by a burst from the transmitter and proceeds to function. For all these applications, the Micropower Receiver can be powered by a long-life lithium battery. One supplier currently specifies 0.5%/year self-discharge rate. The transmitter normally used with the Micropower Receiver is comprised of a single low-cost RF transistor that is driven directly from a CMOS encoder. During its pulsed-on state, it draws < 5 mA at 3 to 15 V. Encoder/decoder pairs, and complete key fob transmitters, are commercially available with over 100 million address/data combinations.

The MIR Comm Link is actually a novel use of the basic MIR transmitted pulse chain as a spread-spectrum data channel. By synchronizing multiple MIR units and encoding information in the pulse stream, low-baud-rate data streams can be generated to communicate detections, current status, faults, or other pertinent information. In this way, the same system that measures motion can also be used to communicate a detection to neighboring remote systems. For the case of many MIR sensors in remote areas, a message-passing daisy chain system of communication could be established, where information is passed from one unit to the next for long distances. Such a system would still be low power, and would be virtually undetectable by standard electronic RF sensors. This is the concept behind the GUARDIAN system that has been proposed for border surveillance and law enforcement.

Future Work

New ideas for MIR developments show every sign of continuing in the next five years. Nearly every interaction with an LLNL program scientist, government

Table 5. MIR Future Development Areas

Parameter	Current Value	Extensions	Applications
Frequency	1 to 4 GHz	up to 96 GHz	Hi-res imaging, directional motion sensing
Frequency	1 to 4 GHz	down to 200 kHz	Deep imaging (e.g., buried waste/bunkers)
Range (round-trip)	50 m	400 m	Deep penetration, long-range detection
Sensitivity (SNR)	20 dB	40 dB	Generally more robust, high SNR, smaller objects improved imaging
Sweep rate	40 Hz	20,000 Hz	Vehicle imaging (e.g., for Federal Highway Administration), projectile tracking

agency, industry representative, or community group inspires new application areas for MIR. 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, in order to reduce it to practice. As proprietors of this technology, we anticipate performing much of the "proof-of-principle" development work, while attracting private industry participation for mass production. Some development areas are shown in **Table 5** below.

One direction in which we plan to make progress next year is toward higher frequencies. We would like to go from the current maximum of about 4 GHz, to 24-, 48-, or even 96-GHz systems. 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. We plan to build on the Rangefinder 2.0 innovations and develop a high-frequency version (Rangefinder II-HF) for the next step in functionality.

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 current MIR operating frequency band. Reasonably-sized millimeter-wave antennas, with improved gain and directivity could 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; airport ground traffic surveillance/tracking; radar ocean imaging; high-resolution personnel imaging; radar aids for the blind; and NDE of composite materials.

Acknowledgments

The authors would like to acknowledge the contributions of the scientists, engineers and technicians involved with the Microradar Project. These include R. Cavitt, G. Dallum, C. Fetterman, D. Gavel, L. Haddad, H. Jones, J. Mast, D. Mullenhoff, S. Nelson, T. Rosenbury, R. Stever, T. Story, J. Trebes, P. Welsh, and M. Wieting. Administrative help came from M. Bell, C. Bothwell, M. Lynch, F. Reyna, R. Sachau, and S. Turner-Perry. We would also like to thank M. Hall of New Jersey Institute of Technology.

We also acknowledge additional support allowing us to complete this project, from the Electronics Engineering Department, particularly S. Trost.

References

1. 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).

2. Popular Science (March 1995), *Radar on a Chip: 101 Uses in your Life.*

3. New Scientist (August 1995), *The Pocket Radar Revolution.*

4. 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.

5. 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.

6. Azevedo, S. G., D. T. Gavel, J. E. Mast, 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.

Harry E. Martz, Thrust Area Leader

The Nondestructive Evaluation (NDE) thrust area at Lawrence Livermore National Laboratory (LLNL) supports the development of inspection science and technology. Our goal is to provide cutting-edge technologies that show promise for inspection tools that will be needed in the future.

The NDE thrust area supports a multidisciplinary team consisting of mechanical and electrical engineers, physicists, computer scientists, chemists, technicians, and radiographers. These team members include personnel from various 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 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 part-time technology that strongly impacts issues of certification, life prediction, and life extension.

The primary contributions of the NDE thrust area in FY-95 are described in the five reports that follow: 1) Infrared Computed Tomography For Characterizing Structural Defects; 2) Techniques to Improve the Sensitivity of Laser-based Ultrasonics; 3) Progress in 3-D Quantitative DR/CT; 4) Computed Tomography Software and Camera Upgrades; and 5) Applications of Micropower Impulse Radar to Nondestructive Evaluation.

Nondestructive Evaluation

7

7. Nondestructive Evaluation

Overview
Harry E. Martz, Thrust Area Leader

Infrared Computed Tomography For Characterizing Structural Defects
Nancy K. Del Grande, Philip F. Durbin, and Dwight E. Perkins *7-1*

Techniques to Improve the Sensitivity of Laser-based Ultrasonics
Diane Chinn, Graham Thomas, James Candy, and James Spicer *7-13*

Progress in 3-D Quantitative DR/CT
Hal R. Brand, Daniel J. Schneberk, Harry E. Martz, Pierre-Louis Bossart, and Stephen G. Azevedo *7-19*

Computed Tomography Software and Camera Upgrade
Hal R. Brand, Daniel J. Schneberk, and Michael R. Gorvad *7-31*

Applications of Micropower Impulse Radar to Nondestructive Evaluation
Jeffrey E. Mast and Stephen G. Azevedo *7-39*

Infrared Computed Tomography For Characterizing Structural Defects

Nancy K. Del Grande, Philip F. Durbin, and Dwight E. Perkins
Manufacturing and Materials Engineering Division
Mechanical Engineering

We have met all of our feasibility-study goals to characterize defects in metal and composite structures. We use a uniform pulsed heat source to stimulate structural defects, a dual-band infrared camera to detect them, and smart defect-recognition algorithms to characterize their thermal signatures. Our patented infrared computed tomography (IRCT) method maps precise temperatures, which specify defect "x" and "y" coordinates, and time-varying thermal inertia, which specifies defect "z" (depth) coordinates, for voids, inclusions, metal-loss, delaminations and unbonds. We quantify thickness-loss thresholds above 5 to 8%, for 1- to 10-mm-thick aluminum and 6- to 8-mm-thick steel. Hardware improvements increased the precision of the IRCT method by ten-fold. Software enhancements reduced defect-characterization times by one-hundred-fold. Our feasibility studies prompted interest in the potential of IRCT applications to support several programs at Lawrence Livermore Laboratory (LLNL).

Introduction

Potential for Thermal NDE

We are investigating potential applications of thermal nondestructive evaluation (NDE) in support of LLNL programs: 1) Energy Manufacturing and Transportation Technology (EMATT) has expressed interest in thermal NDE to inspect pavement integrity and characterize heat transfer in glass manufacturing processes; 2) the DOE Fission Energy Systems Safety Program (FESSP) is concerned with the potential of thermal NDE to inspect the steel-container integrity of 50,000 exposed steel cylinders, which store 700,000 tons of UF_6 waste, at DOE facilities in Kentucky, Ohio and Tennessee; 3) the DOE International Thermonuclear Experimental Reactor (ITER) Program at LLNL needs to know the capability of thermal NDE to inspect the integrity of plasma-facing components; 4) the Advanced Development and Production Technologies (ADAPT) Program at LLNL is considering thermal NDE to characterize the uranium spin-forming process; and 5) the National Ignition Facility (NIF) Program at LLNL has requested thermal NDE support to test the uniformity of tank-coating depositions, the efficacy of table-top polishing, the thermal-stress mechanisms which damage KDP crystals, the bond-integrity of coated aluminum coupons, and the mechanisms of fracture-initiation in optical components.

Limitations of NDE

Typical NDE methods for corrosion detection are slow, visual, operator-dependent, subjective, qualitative and imprecise. Thermal NDE is an emerging technology with an impressive potential as a fast, safe, effective, wide-area tool to quantify corrosion damage and characterize structural defects in three spatial dimensions. Such a tool does not presently exist.

Single-band thermal NDE is used as a tool for corrosion detection and imaging in two spatial dimensions.[1–5] Single-band infrared (SBIR) images typically depict shallow corrosion defects, and corrosion-like defects which have similar thermal signatures. SBIR thermal NDE methods cannot distinguish temperature variations from spatially-varying surface emissivity noise. This often results in false corrosion-defect calls. False detection of corrosion defects in military aircraft results in costly, unnecessary, and destructive exploratory maintenance.[6]

Dual-band thermal NDE was pioneered at LLNL as a quantitative tool for corrosion detection and imaging in 3-D.[7–17] We use a uniform pulsed heat

source to stimulate structural defects, a dual-band infrared (DBIR) camera to detect them, and smart defect-recognition algorithms to characterize their thermal signatures. DBIR imaging provides a ten-fold improvement in temperature-difference accuracy and precision compared to SBIR imaging.[7–17]

We used DBIR image ratios to decouple temperature patterns from spatially-varying surface-emissivity noise. We depicted weak heat flow anomalies from 1- to 3-mm-deep corrosion metal losses, in 1- to 10-mm-thick aluminum panels; from exterior and interior wall-thinning in 6- to 8-mm-thick, 10-cm-diameter steel pipes, and 1.2-m-diameter, steel-cylinder walls; from 1- to 3-mm-deep defects in 1- to 8-mm-thick composite panels; and from 5- to 10-cm-deep delaminations in 25-cm-thick, asphalt-concrete slabs.

Infrared computed tomography (IRCT) uses temperature maps to specify defect "x" and "y" coordinates, and time-varying thermal inertia maps to specify defect "z" (depth) coordinates, for voids, inclusions, metal-loss, delaminations and unbonds. ICRT uses expert defect-recognition algorithms to distinguish the thermal signatures of a rough surface, bare metal, uneven paint, excess sealants, rivets, tape, insulation, and skin ripples, from the thermal signatures of corrosion metal-loss defects.

Physical Rationale for Precise, Emissivity-Corrected Temperature Maps

The patented DBIR method makes use of a property of Planck's radiation law, that near 288 K (15 °C), the DBIR (radiant emittance) signal intensities vary as the emissivity times absolute temperature (T), to the power (50/λ) where λ is the wavelength in microns.[18] This expression is based on a power-series expansion of Planck's radiation law, which provides a convenient approximation at wavelength intervals centered near 5 μm and 10 μm. The signals near 5 μm vary as the tenth power of T, whereas the signals near 10 μm vary as the fifth power of T. By taking the appropriate DBIR image ratios, we produce images of high-contrast temperature (T^5) maps and surface emissivity-noise (E) maps, where:

$$\left(T/T_{av}\right)^5 = \left(S/S_{av}\right)/\left(L/L_{av}\right) \qquad (1)$$

$$E = \left(L/L_{av}\right)^2/\left(S/S_{av}\right) \quad . \qquad (2)$$

In **Eqs 1** and **2**, S is the short-wavelength intensity signal; S_{av} is the average value of the pixels in S; L is the long wavelength intensity signal; and L_{av} is the average value of the pixels in L. We use DBIR image ratios, from DBIR cameras that scan targets at infrared wavelengths of 3 to 5 μm and 8 to 12 μm, to enhance surface temperature contrast and remove the mask of surface emissivity noise. Typically, the DBIR method provides from five to ten times improved defect signal-to-noise ratios, and better interpretability, compared to SBIR methods. Improved temperature contrast, clutter removal, and quantification of material losses are features of the DBIR method that clarify interpretation of defect sites by classifying defect types and by eliminating false-defect calls.

Physical Rationale for Time-Varying Thermal Inertia Maps

We developed thermal inertia maps to characterize defective materials that have bulk thermal property differences with respect to their surroundings. Thermal inertia maps have been used previously for other applications.[19,20] We solve the heat transfer equation for a thick target with an instantaneous surface heat flux[21]:

$$T(z,t) = \frac{q}{\sqrt{4\pi k\rho c t}} \exp\left(-\frac{z^2}{4\alpha t}\right) \qquad (3)$$

where T is temperature, z 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 the surface heat flux. For a semi-infinite solid approximation, the surface temperature is proportional to the inverse square root of time. Thermal inertia is the square root of the product of conductivity, density and heat capacity. It provides a measure of the resistance to bulk material temperature changes which differ for defective and normal structures. In practice, we map the thermal inertia, $(k\rho c)^{1/2}$, based on the (inverse) slope of the surface temperature vs the inverse square root of time. Time-varying thermal inertia maps provide defect depth information. Early-time thermal inertia maps depict shallow defects. Late-time thermal inertia maps depict deep defects. Early-time and late-time thermal inertia maps characterize defective materials which have bulk thermal properties differing from their surroundings.

DBIR Camera System Operating Modes

The DBIR camera operates in both the image-scan and line-scan modes. In the image-scan mode, we typically map flash-heated structures as large as 25 cm high, and 36 cm wide. We store up to 80 images at 0 to 6.4 s after the heat flash. These images resolve defects as small as .5 mm on a side. In the line-scan mode, we typically map a time sequence of horizontal profiles taken across a flash-heated structure, called a timegram. The timegram has time increasing downward, from 0 to 80 ms, along the vertical axis; a profile of the structure positioned along the horizontal axis; and the cooling-rate signature, depicted by a color-coded temperature scale at the left. Pulsed-thermal metal structures have cooling rates at early times, i.e., less than 10 ms after the flash, that are 15 or 20 times faster than they are at late times.

Progress

During FY-95, we successfully met all of our feasibility-study goals to characterize defects in metal and composite structures. Using a uniform flashlamp source and a DBIR camera, we characterized the thermal signatures of structural defects. Our patented IRCT method produced precise temperature maps to specify the defect "x" and "y" coordinates and time-varying thermal inertia maps to specify the defect "z" (depth) coordinates.

Using LLNL/VIEW algorithms on a Silicon Graphics (SGI) Workstation, we developed precise temperature, time-varying thermal inertia, and timegram cooling-rate maps. The unique combination of these maps characterized structural defects by their location, size, depth, and type. We exploited the spatial (0.5 mm), thermal (0.1 °C), spectral (3 to 5 μm and 8 to 12 μm), and temporal (0.5 ms per line) resolution of the Bales Scientific Inc. (BSI) TIP-200 prototype DBIR camera, and 12-bit digital image processing system that inputs DBIR images to the SGI Workstation. We successfully characterized the various defect types, e.g., voids, inclusions, thickness losses, laminar displacements (delaminations), and bonding weaknesses (unbonds) for metal and composite structures, as shown in **Figs. 1–10**.

Hardware and Software Improvements

We developed hardware to improve the capability of IRCT for new applications. The improved hardware provided a uniform flashlamp source with a 4.2-ms heat flash. This improvement removed the hot spots from our thermal images, which reduced heating variations from 30% to 3%. The ten-fold increase in precision lowered the threshold to 5% for quantifying corrosion metal losses and reduced measurement uncertainties to 3%.

Our software enhancements greatly reduced the time needed to produce a thermal inertia map. The original method for making the thermal inertia calculations involved a pixel-by-pixel calculation of a least squares fit to the time sequence data. This method of calculating the fit was slow, and led to calculation times in excess of 4 h, the actual time depending on the size of the area chosen for the fit. We recognized that VIEW has the ability to perform these same calculations on a complete image at once. The memory requirements are considerably more for the complete image method than for the pixel-by-pixel method, but our workstations have more than adequate memory. This change in approach yielded a factor of 100 or more increase in speed. Thermal inertia mapping is now done in minutes instead of hours. This increases the efficacy with which we achieved our goal, to map corrosion defects and eliminate false corrosion-defect calls.

Improved IRCT System

The improved IRCT system offers a unique combination of thermal, temporal and spatial resolution to classify defect types, quantify corrosion thickness losses and improve interpretability of structural defects for metallic and composite structures. The enhanced, high-contrast temperature map of **Eq. 1**, resolves 0.2 °C emissivity-corrected temperature variations from surface preparation, fabrication defects and structural defects. Using the emissivity-noise map of **Eq. 2**, we identify sites with surface reflected IR variations, emissivity noise, or clutter. If similar patterns are present on the co-registered temperature and surface emissivity-noise maps, they are removed, lest they surface be misinterpreted as subsurface structural defects.

Visual inspections of bulges to find "pillowing" from expansion of corrosion by-products in large aluminum panels is highly subjective, since surface bulges vary only slightly for different types of material defects (e.g., sealant globs, skin ripples). To gain improved precision, definition and interpretability over visual inspections, we measured the dynamic temperature, timegram and thermal inertia responses to characterize defects and to quantify the corrosion damage.

We quantified thickness-loss thresholds above 5%, for 1- to 10-mm-thick aluminum, and 8%, for 6- to 8-mm-thick steel structures. We made hardware improvements and software enhancements to develop new application capabilities. Our results prompted a small business CRADA with BSI, a Technical Assistance Agreement with Anchor Glass Containers, and a Work-For-Others Agreement to develop DBIR technology for bridge-deck inspections for the Federal Highway Administration. By exploiting the unique features of IRCT, we have characterized defect sites, depths, volumes, and types, for four metal and two composite structures, shown in **Figs. 1–8**. Additional studies of glass and asphalt-concrete structures are not described here.

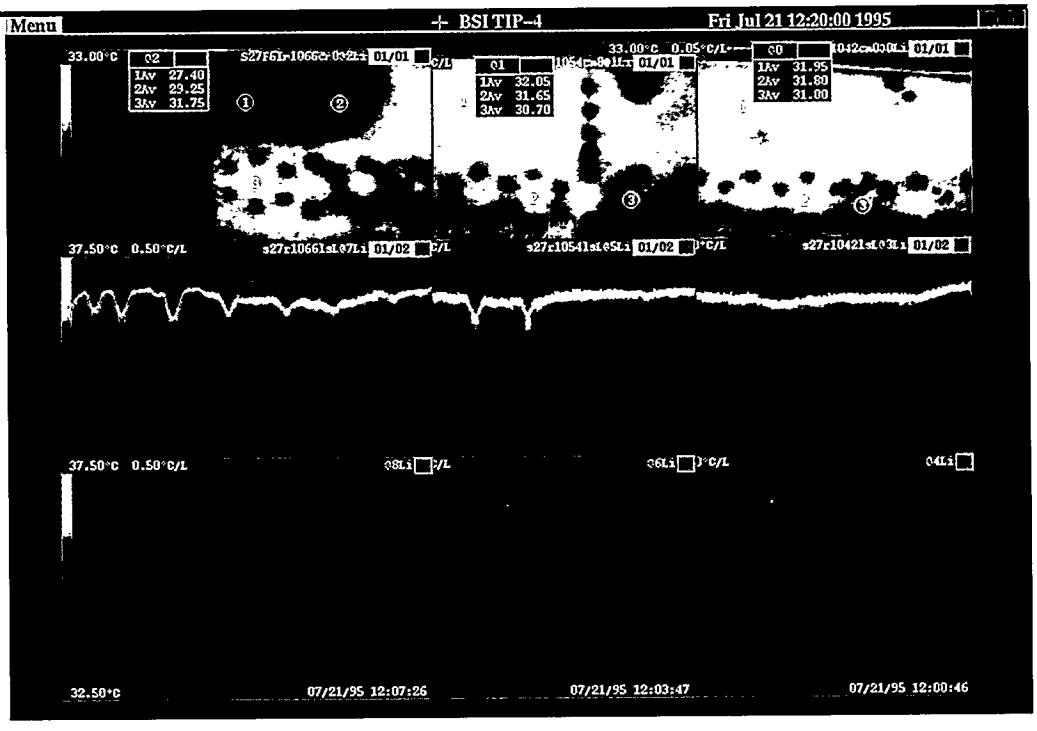


Figure 1. Aircraft skin temperature map (above) and timegram (below), showing 10% corrosion-loss defects within the third-from-left hexagonal rivet pattern, and corresponding timegram directly below.

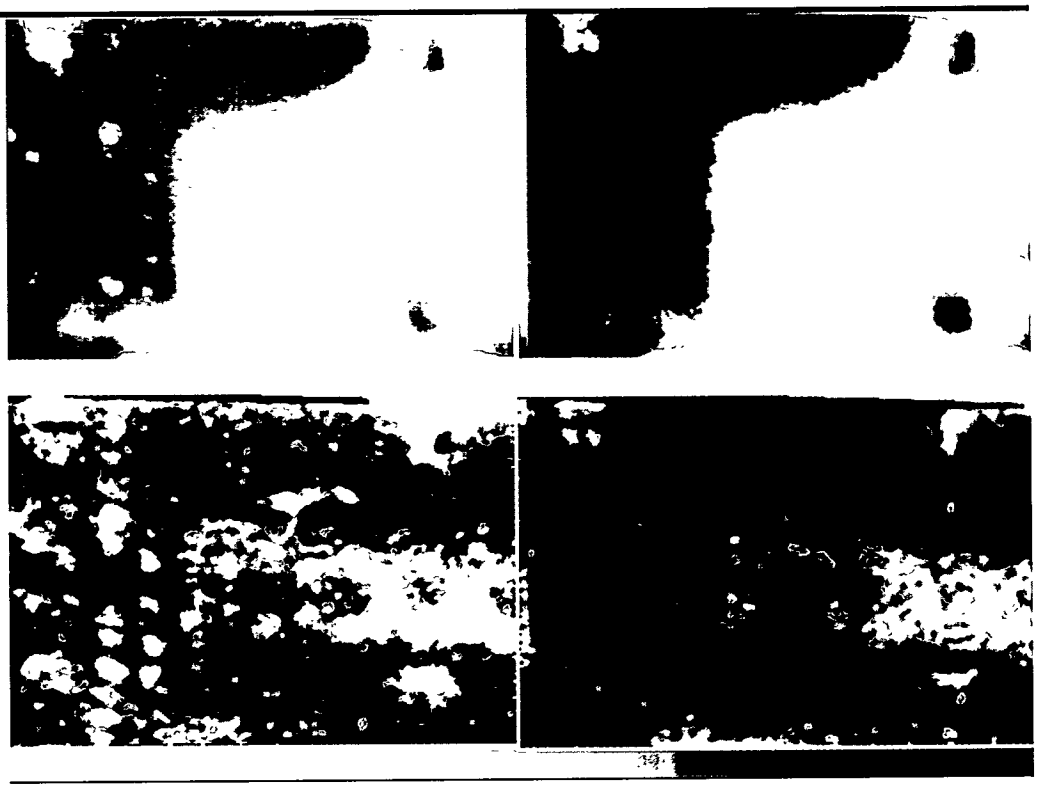


Figure 2. Top left to right: aircraft skin temperature maps at 0.12 s and at 0.43 s after the flash. Bottom left to right: early-time thermal inertia map of outer skin and late-time thermal inertia map of deep stiffener. Note 10% corrosion-thinned skins are lighter than normal near the bottom right corner of the early-time temperature and thermal inertia maps.

Detection of Defects in Metal and Composite Structures

The Boeing Commercial Airplane Group made available a Boeing 727 airframe, which they will disassemble. This gives us a unique opportunity to verify how accurately thermal NDE measures low thresholds of corrosion damage. Our aluminum aircraft panel proof-of-principle experiment measured a 1-m-wide, 0.25-m-high, section of the flash-heated aircraft, with 5% to 10% corrosion thinning, as shown in **Fig. 1**. We took three single-skin temperature maps, shown in the top row, and corresponding timegram cooling-rate maps, shown in the middle and bottom rows, directly below. The timegrams mapped time sequences of linescan cooling rates taken across the center of the hexagonal pattern of rivets, shown on the temperature maps at 0 to 80 ms (middle row), and at 80 to 160 ms (bottom row).

The temperature scale at the left of the images was 5 °C warmer at top (white) than at bottom (black). The third from the left hexagonal pattern of rivets enclosed an area with 10 ± 2.5 percent corrosion thinning, at 0.8 ± 0.2 °C above-ambient temperatures, based on our measurements. Pillowing from the volume expansion of corrosion by-products was seen where we measured the 10% corrosion thinning. Six reference standards with known thicknesses provided the basis for our metal-loss thickness calibrations. Doubler and tripler panels were riveted to the outer 0.063-in.-thick skin, at the left side of **Fig. 1**. The thick multi-layered panels had below-ambient temperatures which were off-scale on temperature maps and timegrams. Thinner skins reached higher temperatures, and took longer to cool to ambient temperatures, as noted on the timegrams.

Figure 2 shows temperature and thermal inertia maps for the left-sided images in **Fig. 1**. Temperature maps were recorded at 0.12 s, top left image, and at 0.43 s, top right image, after the flash. Temperature maps taken at 0.43 s had greater thermal contrast than temperatures taken at 0.12 s after the flash. The Boeing 727 temperature maps show skin thicknesses of 0.107 in. as dark grey, 0.070 in. as medium grey, and 0.063 in. as light grey. The thinner skins were chemically milled. Outer skins backed by doublers, triplers and stiffeners, which had a total thickness of about 0.50 in., appeared darkest at 0.43 s after the flash. The skin thicknesses were measured using an ultrasonic thickness gauge. We determined that a typical 1 °C temperature rise at 0.43 s after the flash corresponded to a 12.5% metal loss for skins ranging in thickness from 0.030 to 0.125 in.

The thermal inertia maps at the bottom of **Fig. 2** show the near surface structures, at bottom left, as having bulk thermal properties different from those of the deep stiffener, at bottom right. Thermal inertia is the square root of the target conductivity, density, and heat capacity. It is a measure of the resistance to temperature change for bulk materials within the target. At early times, from 0.1 to 0.4 s, the thermal inertia map at the left depicts thermal property differences for the outer surface skin. At later times, from 0.4 to 0.7 s, the thermal inertia map at the right shows corrosion metal losses within a deeper structure, such as the 0.267-in.-thick stiffener which was used to strengthen the outer 0.063-in.-thick skin. Thermal inertia maps clarify interpretation of defect types, provide defect depth information, and eliminate false-defect calls.

Thermal NDE Inspection of UF_6 Steel Container Integrity

We conducted a steel container integrity proof-of-principle experiment. Our goal was to determine if thermal NDE detected metal thinning which reduced a steel cylinder wall thickness from 0.335 in. (8.5 mm) to less than 0.25 in. (6.4 mm), the design-safety margin for steel cylinders which DOE facilities use to store UF_6 waste products. We measured a section of the 4-ft (1.2-m)-diameter, steel wall, after machining three flat-bottom holes each on the outside convex, and the inside concave surfaces. **Figure 3** shows a schematic, and **Fig. 4** shows photographs of a section of the steel cylinder that we inspected.

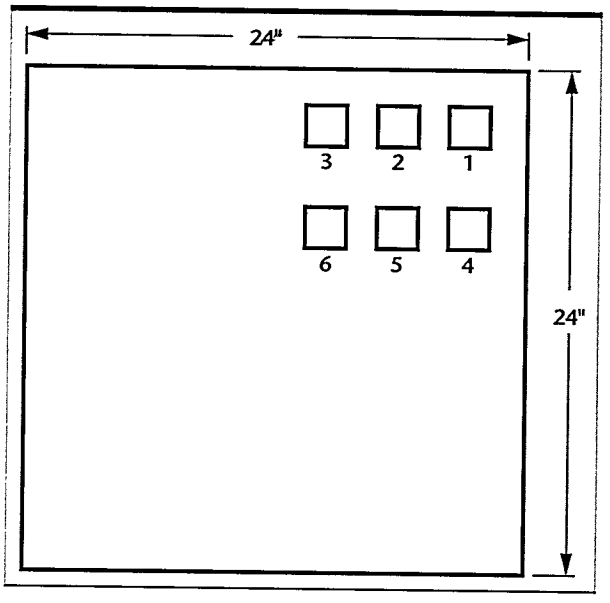


Figure 3. Schematic of UF_6 waste container inspected by thermal NDE. Square hole sites are 2 in. × 2 in. machined areas, separated by 2 in. horizontally and 3 in. vertically.

Figures 5 and **6** show temperature, thermal inertia and timegram cooling-rate maps. These maps depict metal losses for the steel wall which was flash-heated on the outside, convex surface. The maps depict hole sites with residual wall thicknesses, left to right, of 0.085 in. (2.2 mm), 0.15 in. (3.7 mm), and 0.21 in. (5.3 mm). The images show inside holes at the top, and outside holes at the bottom. In **Fig. 5**, we used IR images recorded at 3 to 5 μm. The left-sided images show early-time, temperature maps above, and thermal inertia maps below. The right-sided images show late-time, temperature maps above, and thermal inertia maps below. At early times, less than 0.5 s after the flash, we measured the effect of shallow paint sealant which was present, except where it had been removed when we machined the holes, prior to adding a coat of water-soluble black paint, before conducting the test. In **Fig. 6**, the left side images were recorded at 3 to 5 μm and the right side images at 8 to 12 μm. **Figure 6** shows timegrams at 0 to 80 ms, top to bottom, for inside holes at the top row, and outside holes at the center row. Temperature maps at the bottom row depict the inside surface holes above and the outside surface holes below.

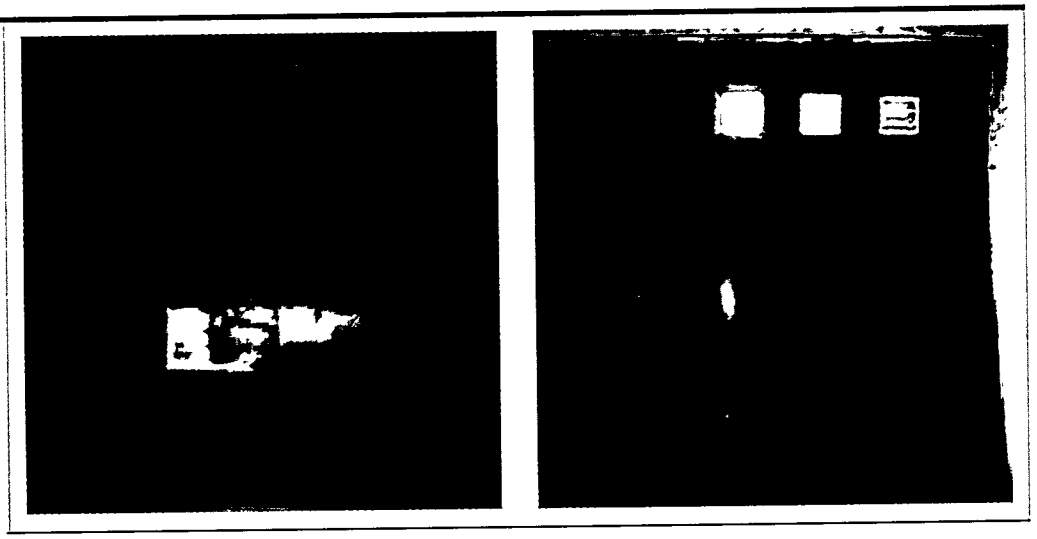


Figure 4. Photographs of a section of the UF_6 waste container inspected by thermal NDE, showing the flash-heated outside convex surface at the left, and the inside concave surface at the right. The outside surface was coated with a sealant, except at the machined hole sites, which were not visible after we applied a top coat of water-soluble black paint.

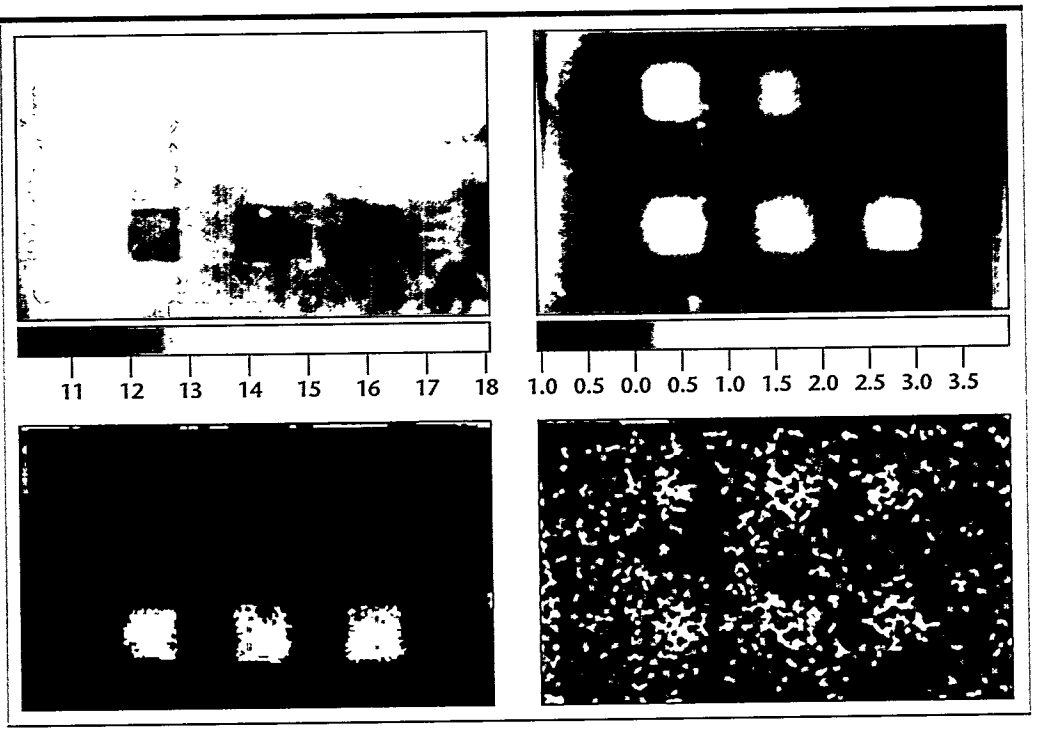


Figure 5. Left images: early-time, temperature map (above), and thermal inertia map (below), of a steel container wall with shallow outside, metal-loss holes. Right images: late-time, temperature map (above), and thermal inertia map (below), of a steel container wall with deep inside metal-loss holes at the top, and shallow outside metal-loss holes at the bottom.

Thermal NDE for Petrochemical Product Pipeline Inspections

Our goal for the pipeline integrity proof-of-principle experiment was to determine if thermal NDE could measure low levels of corrosion thinning from petrochemical product pipelines. We measured a flash-heated, 10-cm-diameter, 6.65-mm-thick, corroded steel pipe wall, with 0.5 mm (7.6%) interior metal loss, provided by Russel Mack of Dow Chemical International. This implied that pulsed-thermal NDE provided a technique that addressed the OSHA 1910 requirements for nondestructive testing (NDT) of exposed petrochemical product pipelines. Typical NDT methods used to inspect pipeline corrosion are visual. Visual methods are slow, costly, and do not detect incipient corrosion damage. Our results depicted the 0.5 mm-thick, 7.6%, incipient corrosion metal-loss damage in a corroded pipe section, marked by a rectangle, and surrounded by uncorroded 6.35-mm-thick steel pipe with surface scale build-up, which had a total thickness of 6.65 mm (**Fig. 7**). Also, for purposes of comparison, we show pulsed thermal NDE temperature maps for a 0.25-in. (6.4 mm)-thick reference standard with 10%, 20% and 40% metal-loss holes in **Fig. 8**.

Thermal NDE Characterization of Defect Types and Depths

We conducted a proof-of-principle experiment to determine if thermal NDE uniquely characterized the different defect types and depths, for voids, inclusions and unbonds, in a flash-heated, boron-epoxy composite panel which was adhesively attached to an aluminum aircraft skin. We successfully characterized the different defect types, e.g., voids, inclusions and unbonds. Our thermal NDE maps are shown in **Fig. 9** for a 6-ply, 0.0335-in. (0.85-mm)-thick composite panel.

In **Fig. 9**, we show temperature maps above, and thermal inertia maps below, for a 6-in. (15-cm)-wide, 9-in. (23-cm)-high, boron-epoxy, composite-panel, adhesively attached to a Boeing 737 aircraft skin. The temperature maps at top left and right specify defect "x" and "y" spatial coordinates, and the time-varying thermal inertia maps at bottom left-to-right specify defect "z" depth coordinates. The two temperature maps show the defect sites for nine inclusions, and four voids marked by thirteen spots, of the sixteen spots where the average temperature values are tabulated.

Three shallow inclusions, near the panel bottom at spots marked 1 to 3 on the upper right temperature map, are 1-ply deep, based on their low (white), 0.8 s

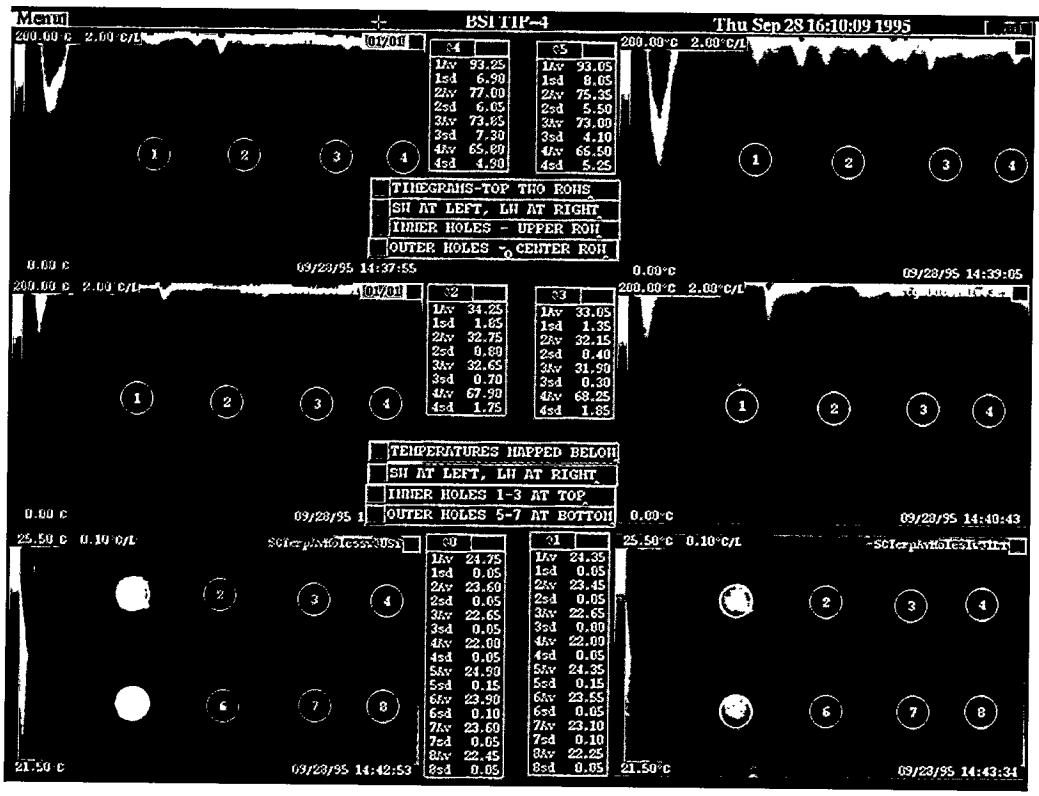


Figure 6. Left side, 3 to 5 μm and right side, 8 to 12 μm of timegrams at 0 to 80 ms, top to bottom, for inside holes, top row and outside holes, center row. Temperatures at the bottom row depict inside-surface holes above, and outside-surface holes below.

early-time thermal inertia values. Three deeper inclusions, near the panel center, at spots marked 5 to 7 on the upper left temperature map, are 3-ply deep. The three deepest inclusions, near the panel top at spots marked 1 to 3 on the upper left temperature map, are 5-ply deep, based on their high (dark grey), 0.8 s early-time thermal inertia values. Four voids, positioned along the panel perimeter at spots marked 5 to 8 on the upper right temperature map, have the highest (black) 0.8 s early-time thermal inertia value on the

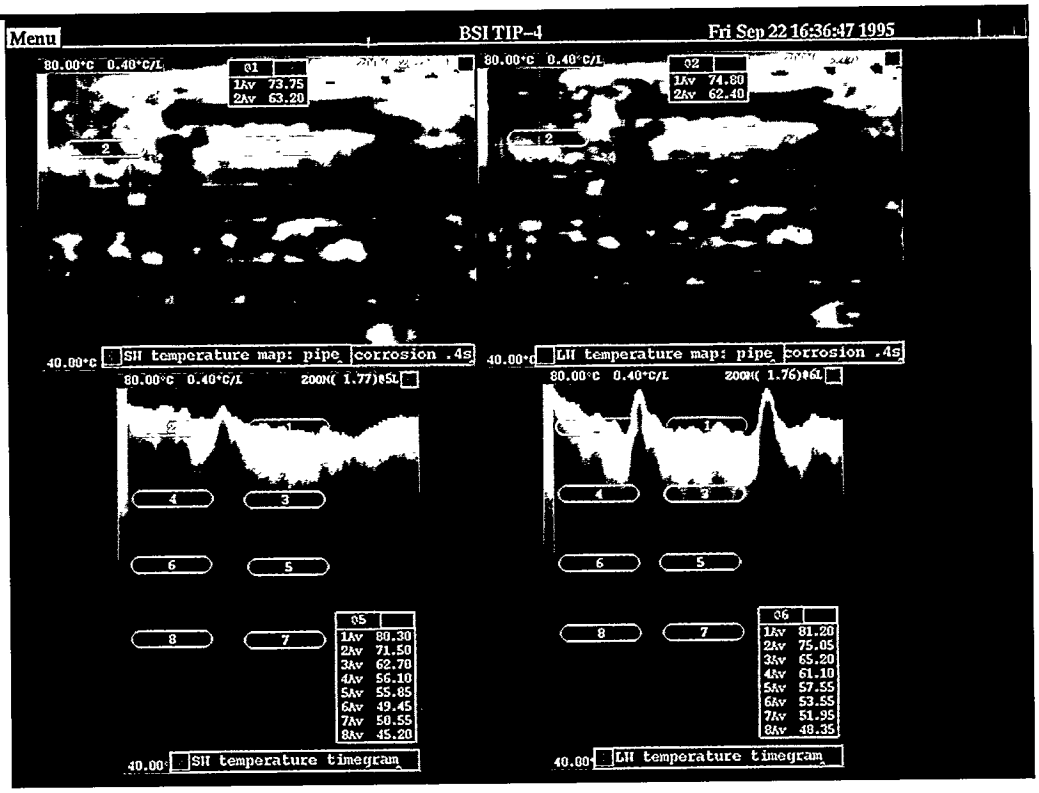


Figure 7. Top row temperature maps and bottom row timegram cooling-rate maps, showing 1/4-in. steel pipe with 7.6% interior metal-loss corrosion near the top of the rectangular marker, at 3 to 5 µm, left, and 8 to 12 µm, right.

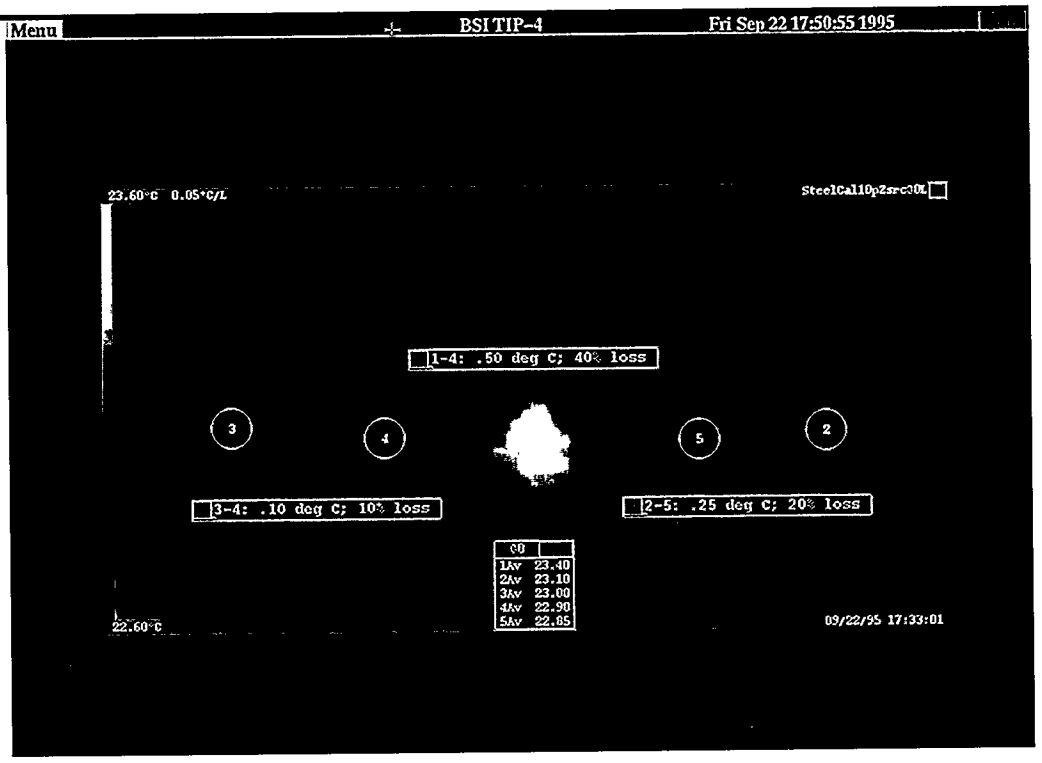


Figure 8. Late-time temperature map of a flat, 1/4-in. steel plate with metal-loss hole thicknesses of 10%, 40% and 20%, as seen from left to right. The steel plate with metal-loss holes is used as a reference standard.

lower left thermal inertia map. The deepest defects were at the interface of the composite panel with the aircraft skin, where there was a cross-like pattern of unbonds which have low (white), 2.5 s late-time thermal inertia values, where the aircraft skin rivets are not adhesively bonded to the composite panel.

Thermal NDE Characterization of Delaminations

We conducted a proof-of-principle experiment to determine if thermal NDE characterized delaminations in a flash-heated, thermoplastic, composite panel. In **Fig. 10**, we show temperature patterns recorded at 0.45 s after the heat flash, which depicted major temperature rises at delamination sites, in a multi-layered, composite thermoplastic panel. The composite panel has a diameter of about 5 in. (13 cm). It is bonded to a 7-in. (18-cm)-wide composite substrate. Note the laminar gap patterns show the sites of 0.10-in. (2.5-mm)-deep subsurface delaminations, in the thermoplastic panel, which has a total thickness of 0.315 in. (8 mm), including the substrate. Spot temperature averages and standard deviations are shown in the table at the right.

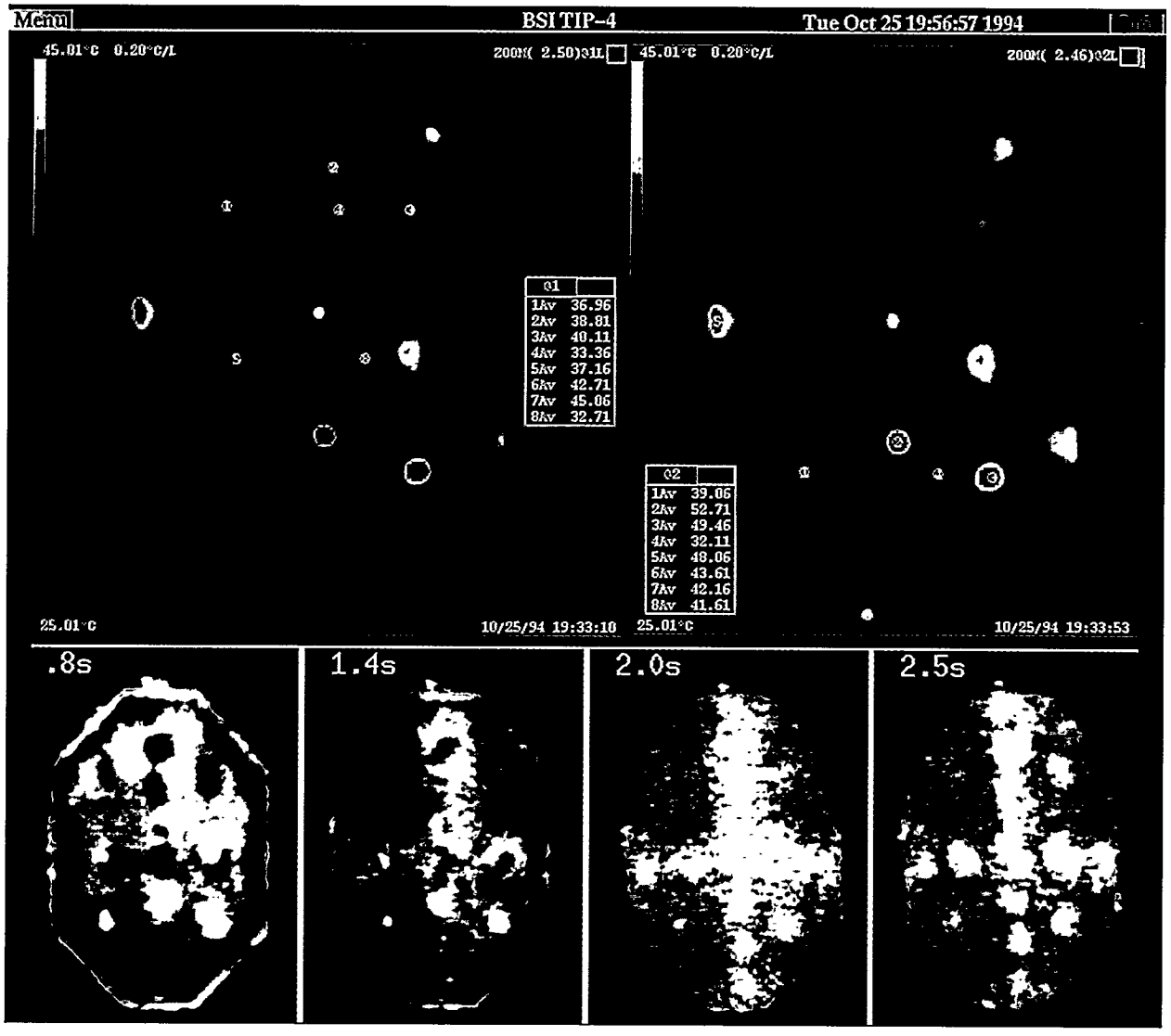


Figure 9. A 6-ply boron-epoxy composite-panel, adhesively attached to an aluminum aircraft skin. Temperature maps at top left and right have numbered sites, with tabulated average temperature values that specify defect "x" and "y" coordinates. Time-resolved thermal inertia maps at bottom, left to right, have low (white) and high (black) thermal inertia values that specify defect "z" depth coordinates. Note the three shallow 1-ply inclusions near the bottom, on the upper right temperature map; the three deeper 3-ply inclusions near the center, on the upper left temperature map; the three still deeper, 5-ply inclusions near the top, on the upper left temperature map; the four voids along the perimeter, on the upper right temperature map; and the dozen unbonds, shaped in a cross, on the 2.5 s late-time thermal inertia map at lower right, where aircraft-skin rivets were not adhesively bonded to the composite panel.

Summary

The proof-of-principle experiments conducted during FY-95, prompted much interest in the potential of IRCT applications to support LLNL programs, e.g., EMATT, FESSP, ITER, ADAPT, and NIF.

Also, the potential of applications which have not yet been exploited include 1) the Department of Defense Air Logistics Command's expressed need, e.g., at Warner Robbins Air Force Base, and Kelley Air Force Base, to quantify corrosion damage in military aircraft; 2) Dow Chemical International's expressed interest in pulsed-thermal NDE to inspect exposed petrochemical pipelines; 3) a Canadian Government-affiliated transportation company's request to license LLNL's thermal NDE methods for characterizing defects in civil structures, e.g., roadways, airstrips and bridges; and 4) a glass-container company's request for technical assistance to characterize their 1250 °C glass gob heat transfer processes in free-fall.

Future Work

We are planning additional proof-of-principle experiments to test the potential of new IRCT applications in support of LLNL programs:

1) NIF, for coating-tank, polishing-table, coating bond-integrity, crystal-damage and optics-fracture studies;
2) ADAPT, for monitoring the efficacy of uranium spin-forming processes;
3) ITER, for plasma-facing components integrity studies;
4) FESSP, for locating unreliable UF_6 waste cylinders, and
5) EMATT, for inspecting pavement integrity and monitoring glass-manufacturing processes.

Bearing this in mind, our FY-96 objectives for IRCT will be to reduce to practice DBIR pulsed-thermal NDE as a 3-D thermal imaging tool to characterize deep structural defects and heat transfer processes in support of LLNL programs. Our specific goals will be to conduct proof-of-principle

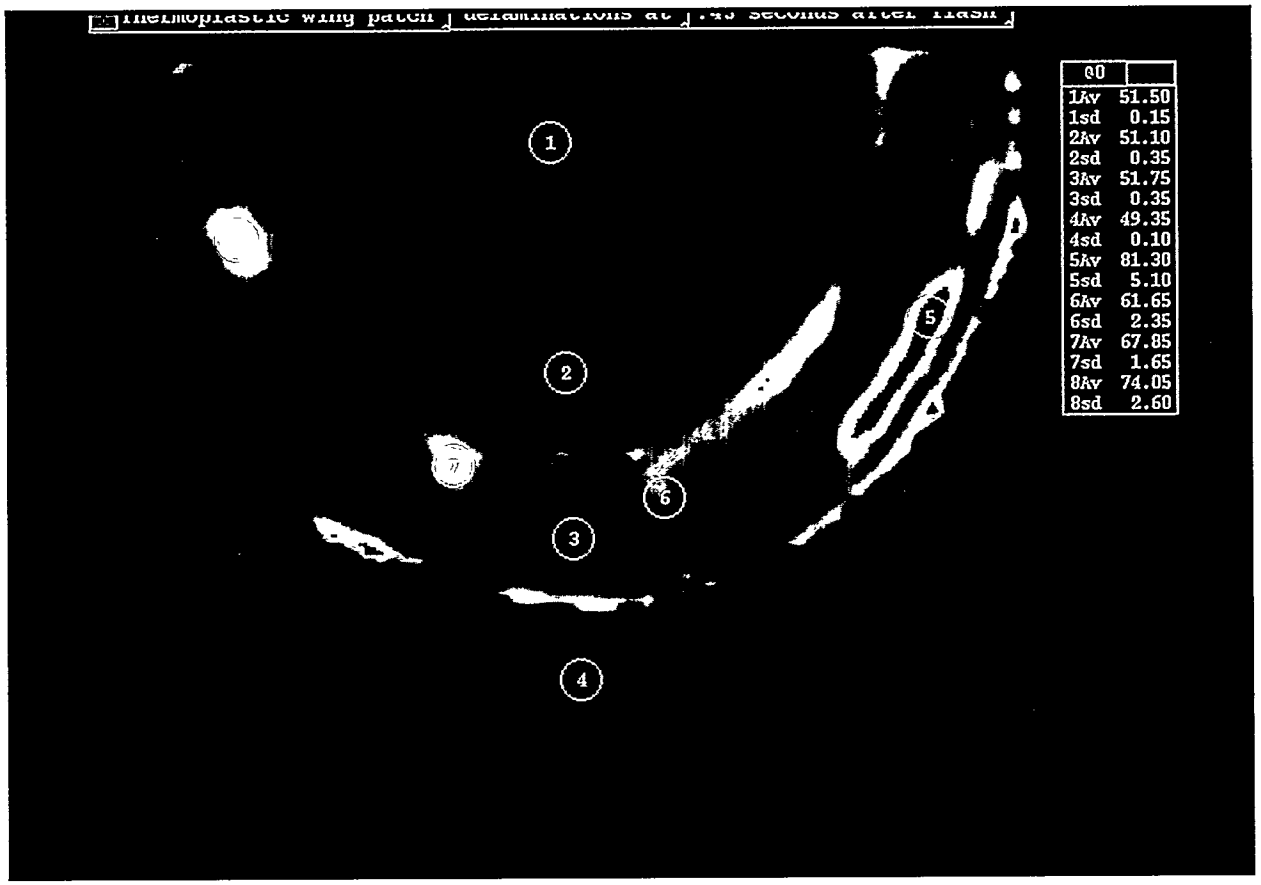


Figure 10. Thermoplastic composite patch with a diameter of 13 cm on a 18-cm-wide composite substrate. Note the laminar gap (delamination) patterns at 0.45 s after the flash, which show the sites of 2.5-mm-deep subsurface delaminations in the three-layered thermoplastic patch with a total thickness of 8 mm, including the substrate. Spot temperature averages, and standard deviations, are shown in the table at the right. Spot temperature 5, which is off-scale, appears dark, whereas other spot temperatures appear lighter as their temperatures increase.

experiments which evaluate the capabilities and limitations of IRCT for LLNL program applications. We will characterize defects in multi-layered materials, quantify damage within complex structures, and test the predictability of theory to characterize 3-D dynamic heat transfer processes.

Also, the results gleaned from proof-of-principle experiments conducted in FY-95 have prompted a small business CRADA with BSI. LLNL will rewrite the algorithms which are presently used on a SGI Workstation, in a language which BSI can adapt for their commercial infrared imaging system. Working jointly with BSI, we will produce a single evaluated-defect map. We will eliminate false corrosion-defect calls from the evaluated defects depicted by the combined temperature, thermal inertia and timegram cooling-rate maps. Our evaluated-defect map will be part of the LLNL/BSI image-processing system to quantify corrosion damage in metal structures, e.g., to inspect steel UF_6 waste storage containers, and exposed petrochemical pipelines.

Acknowledgments

We acknowledge support from the FAA Aging Aircraft Non-Destructive Inspection R&D Program, and the Federal Highway Administration. We are grateful for the loan of the BSI uniform flashlamp heating source; the Rockwell reference standards, developed by S. James; the F-15 corroded wing box, which we used as a reference standard, and the thermoplastic patch with delaminations, which we characterized from Owen Manning at Northrop; the composite and metal-loss panels developed by J. Thompson at Boeing, which we used as reference standards; and ultrasonic thickness-gauge measurements of the Boeing 727 taken by the staff of the Boeing Commercial Airplane Group. We are grateful for technical discussions with H. Martz, the NDE thrust area leader at LLNL, and NIF scientists H. Powell, J. Campbell, D. Aikens, A. Burnham, C. Stolz, M. Yan, and P. Hurst.

References

1. Spicer, J. W. M., W. D. Kerns, L. C. Aamodt, R. Oslander, and J. C. Murphy (1993), "Time-Resolved Infrared Radiometry (TRIR) using a Focalplane Array for Characterizations of Hidden Corrosion," *Proceedings of SPIE Conference 1933: Thermosense XV*, L. R. Allen, Ed., Orlando Fla., p. 148.

2. Favro, L. D., P. K. Kuo, R. L. Thomas, T. Ahmed, and Y. X. Wang (1993), "Thermal wave imaging of corrosion and disbonds in aircraft structures," *Proceedings of SPIE Conference 2001: Nondestructive Inspection of Aging Aircraft*, M. T. Valley, N. K. Del Grande, and A. S. Kobayashi, Eds., San Diego, Calif.

3. Selman, J. J., and J. T. Miller (1993), "Evaluation of a prototype thermal wave imaging system for nondestructive evaluation of composite and aluminum aerospace structures," *Proceedings of SPIE Conference 1933: Thermosense XV*, L. R. Allen, Ed., Orlando, Fla., p. 178.

4. Syed, H. I., and K. E. Cramer (1993), "Corrosion Detection in Aircraft Skin," *Proceedings of SPIE Conference 1933: Thermosense XV*, L. R. Allen, Ed., Orlando, Fla., p. 160.

5. Vavilov, V. P., and X. Maldague (1992), "Dynamic thermal tomography: a new promise in the IR thermography of solids," *Proceedings of SPIE Conference 1682: Thermosense X1V*, J. K. Eklund, Ed., Orlando, Fla., p. 194.

6. Alcott, J. (1994), "An Investigation of Nondestructive Inspection Equipment: Detecting Hidden Corrosion on USAF Aircraft," *Materials Evaluation*, American Society for Nondestructive Testing, January.

7. Del Grande, N. K., K. W. Dolan, P. F. Durbin, and D. E. Perkins (1995), *Emissivity-Corrected Infrared Method for Imaging Anomalous Structural Heat Flows*, United States Patent Number 5,444,241, August.

8. Del Grande, N. K., and P. F. Durbin (1995), "Mapping hidden aircraft defects with dual-band infrared computed tomography," *Proceedings of SPIE Conference 2455: Nondestructive Evaluation of Aging Aircraft, Airports, Aerospace Hardware, and Materials*, T. Cordell, Ed., Oakland, Calif., pp. 82–93.

9. Del Grande, N. K., and P. F. Durbin (1995), "Using emissivity-corrected thermal maps to locate deep structural defects in concrete bridge decks," *Proceedings of SPIE Conference 2456: Nondestructive Evaluation of Aging Bridges and Highways*, S. Chase, Ed., Oakland, Calif., pp. 250–261.

10. Del Grande, N. K., P. F. Durbin, and M. R. Gorvad (1994), "Applications Of Infrared Computed Tomography," in *Nondestructive Evaluation Thrust Area Report FY94, Engineering Research Development and Technology*, Lawrence Livermore National Laboratory, Livermore, Calif., H. E. Martz, Ed. (UCRL-ID-119059).

11. Del Grande, N. K., and P. F. Durbin (1994), "Multi-use applications of dual-band infrared (DBIR) thermal imaging for detecting structural defects," *Proceedings of SPIE Conference 2217: Aerial Surveillance Sensing Including Obscured and Underground Object Detection*, I. Cindrich, N. K. Del Grande, S. Gowrinathan, P. Johnson, and J. F. Shanley, Eds., Orlando, Fla., pp. 107–116.

transducers, such as glove boxes and furnaces. the detection system is the preferable option.

12. Del Grande, N. K., and P. F. Durbin (1994), "Dual-Band 17. Del Grande, N. K., K. W. Dolan, P. F. Durbin,

Nondestructive Evaluation

Progress

This year we explored several ways to improve the sensitivity of laser ultrasonic detectors. We made a minor hardware modification to our Michelson interferometer to increase its sensitivity. We explored the application of phase conjugation techniques for minimizing the effects of optical distortion. Finally, our main effort this year focused on developing model-based signal processing algorithms to improve the signal-to-noise (SNR) ratio in laser-detected ultrasonic data.

Michelson Interferometer Modification

The Michelson interferometer uses phase differences between reference and modulated beams to detect surface displacements resulting from laser ultrasonic excitation. A photo of our Michelson interferometer in its original configuration is shown in **Fig. 1**. The original skewed-stabilized configuration of our Michelson interferometer, sketched in **Fig. 2a**, has a 10 mW He-Ne laser source which usually delivers less than 1 mW of energy for detection. The interferometer takes the reference beam from detector B and the modulated beam from detector A to determine displacements on the sample surface. The reference mirror is mounted on a piezoelectric tube which is connected to a feedback loop that stabilizes the interferometer and discerns actual displacements from ambient vibrations. Skewing the beam helps prevent feedback to the laser cavity from the reference beam going to detector B. The sensitivity of this configuration depends on proper alignment.

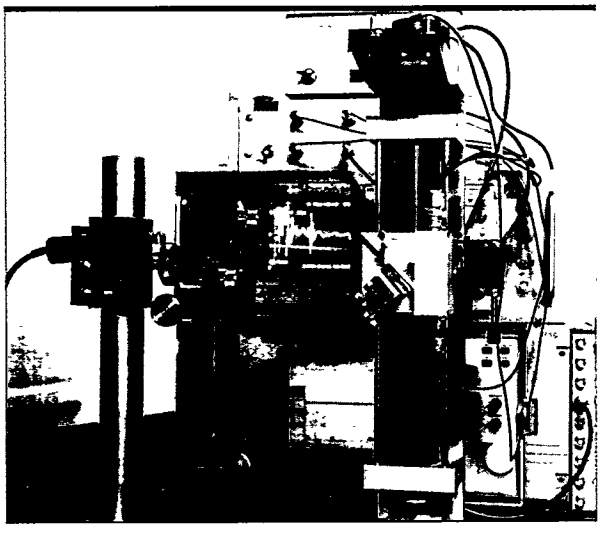


Figure 1. Michelson laser interferometer in the skewed-stabilized configuration before modification.

This year we modified our skewed-stabilized interferometer to the parallel-stabilized configuration shown in **Fig. 2b**. Parallel beams are easier to align than skewed beams. Easier alignment results in higher SNR for the interferometer. Further, the in-line design allows the interferometer beam to be focused with a standard lens, facilitating detection and increasing sensitivity on fairly rough reflecting surfaces. An additional beamsplitter cube separates the reference beam from the laser, avoiding feedback between the two. The quarter-wave plate polarizes the beam before it reaches the reference detector, also helping to increase the sensitivity of the interferometer. Easier alignment and focusing offered by the parallel-beam configuration has helped to increase the sensitivity of our interferometer.

Phase Conjugation

Phase conjugate cavities contain a medium which returns the conjugate of the incoming light beam. The conjugate wave propagates along the same

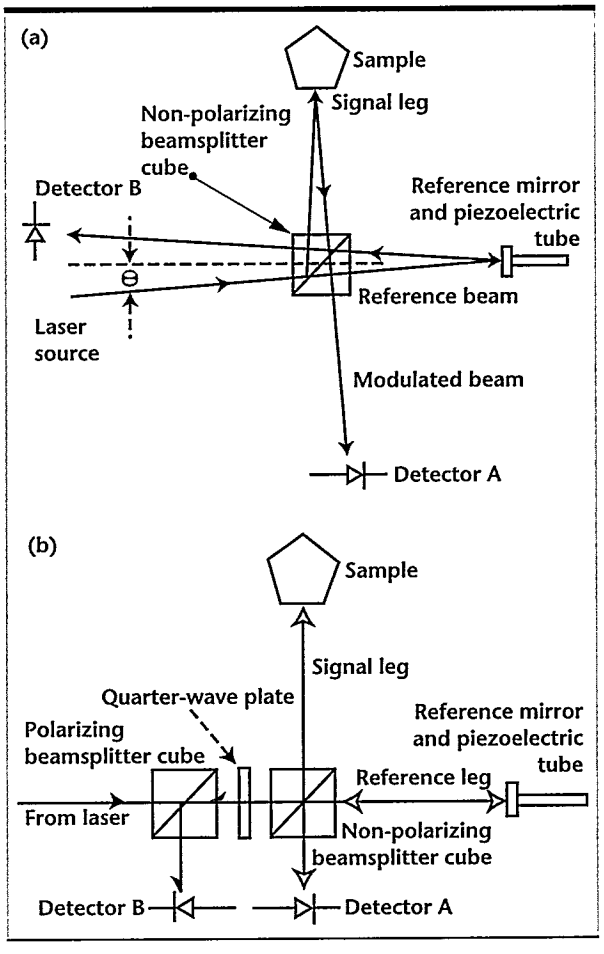


Figure 2. Schematic representation of (a) the skewed-stabilized Michelson interferometer; and (b) the parallel-stabilized Michelson interferometer after modification.

path as the incident beam, but with phase and propagation direction reversed. Phase conjugate cells, also called time-reversal mirrors, have gained popularity in the field of laser optics because cavities containing certain crystals can correct distortions in a wavefront that may be introduced by passing through inhomogeneous media. Corrections to the wavefront can remove degrading effects resulting from such system features as poor optics and laser aberrations.

Application of phase conjugate cavities to laser ultrasonics is a new area of research. By offering a correction for systematic distortions introduced into the laser beam, phase conjugation can increase the sensitivity of laser interferometers tremendously. Phase conjugate cavities have been shown to remove large amounts of distortion in high energy laser interferometers.[2,3] Low energy lasers used in laser ultrasonics require different conjugating materials. Double phase conjugate cavities containing barium titanate crystals have shown promising increases in sensitivity for detecting ultrasonic displacements on rough surfaces.[4] At this time, the technique of applying phase conjugation to laser ultrasonic wave detection is in its infancy, and would be prohibitively expensive to implement on our system. We recognize its potential in increasing detection sensitivity and will keep abreast of developments in the field.

Model-based Signal Processing

A model-based approach is being applied to enhance the laser ultrasonic signals acquired from our developmental system. Generally, this approach uses *a priori* information about the phenomenology (laser-induced ultrasonic wave propagation), measurements (interferometry) and noise (such as that from a detector or thermal noise) to develop a model-based processor (MBP) capable of increasing the SNR, and therefore the overall sensitivity (minimum detectable displacement) of the system. For the laser ultrasonic problem, we have developed the model-based processor depicted in **Fig. 3**, which consists of a preprocessor (averaging filter) coupled to an optimal (Wiener) model-based processor.[5] We generated a model response (displacement) using a thermoelastic laser ultrasonic propagation model[6] for the test specimen and used the response as the "desired" or model reference signal in the MBP to yield the enhanced displacement signal estimate.

The results of a typical enhancement are shown in **Fig. 4**. In **Fig. 4a**, we see the epicentral displacement measurement of a 6 mm thick aluminum alloy plate at SNR = 30.8 dB. The SNR is calculated using the standard definition of the ratio of input signal power (model) to measurement noise (measurement). **Figure 4b** shows the epicentral displacement

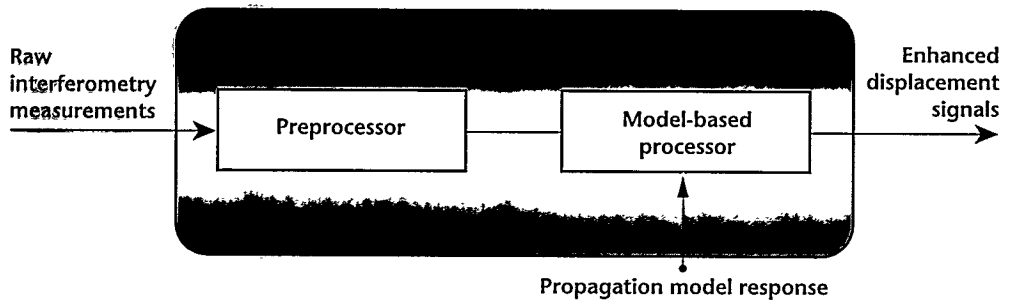


Figure 3. Model-based processor structure for laser ultrasonic displacement signal estimation.

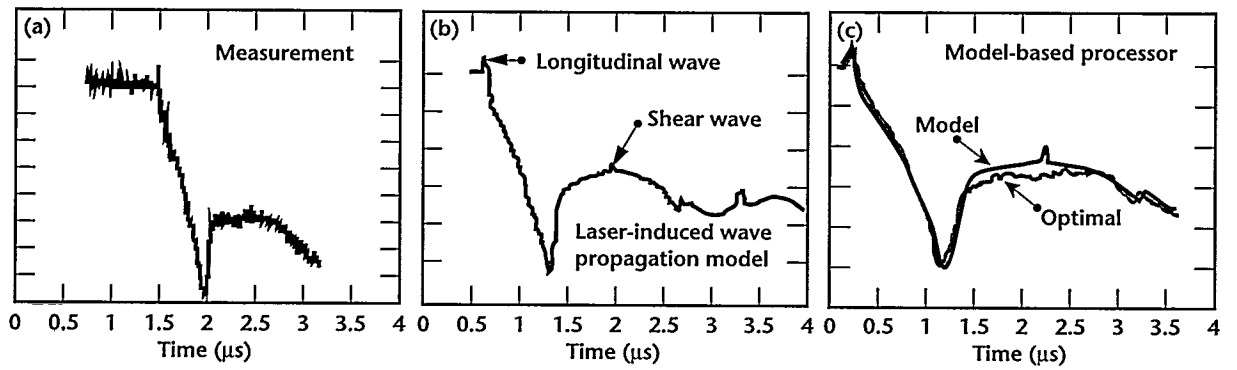


Figure 4. Model-based processor performance for laser ultrasonic measurements of epicentral displacement on a 6 mm aluminum plate: (a) measured interferometric response (SNR= 30.8 dB); (b) predicted propagation model response; (c) enhanced MBP displacement signal estimate (SNR= 52.1 dB).

response predicted by the thermoelastic propagation model, which is also used as the desired signal in the MBP design. Finally, in **Fig. 4c** we see the enhanced displacement produced by the processor, visually indicating the enhancement which results in SNR = 52.1 dB. In this example, MBP increases SNR by 21.3 dB.

Various experiments performed on the aluminum plate at various generation and detection laser settings produced the range of input SNR shown in column 2 of **Table 1**. The output SNR in column 3 indicates the SNR for the enhanced MBP signal. The SNR gain in column 4 clearly indicates the improved sensitivity using the model-based approach.

As part of this effort, we developed an interactive Laser Ultrasonic Toolbox to perform all of the analysis and required signal processing with the MATLAB[7] computational environment. The Toolbox performs the following basic operations: I/O, Processing, Analysis, and Applications, where each of these main operations is controlled by a Supervisor along with a set of pull-down menus developed using the MATLAB graphical user interface (GUI) tools. Under each of these main operations we have the capabilities shown in **Table 2**.

Summary

This year we directed our efforts at increasing the sensitivity of our detection systems. Hardware modifications to our Michelson interferometer have increased our ability to align the system, resulting in increased sensitivity. We are keeping abreast of developments in the application of phase conjugate cavities to laser ultrasonic interferometers. We recognize the potential of this technique in laser ultrasonic detection. However, the state of the technology cannot readily be implemented into our system. Our model-based processing algorithms have increased signal-to-noise ratios of laser ultrasonic measurements by approximately 20–30 dB.

Future Work

Our goal is to position LLNL to benefit from internal and external developments in the technology of laser ultrasonics. We have established a laser ultrasonic facility and plan to continually improve it by implementing new technology. We plan to collaborate with universities and other laboratories to incorporate the latest technology, as well as produce our own contributions to the technology. We are particularly sensitive to directing our resources at satisfying our customer requirements. Combining our strengths in NDE system engineering with the most recent developments, we can implement laser ultrasonics in situations where traditional ultrasonic techniques can not be used. Laser-based ultrasonics is particularly amenable to process control, high temperature environments, clean optical inspections and inaccessible monitoring of complex parts. In addition, we will develop laser scanning capabilities, which are required for imaging.

Process control increases quality for most manufacturing operations. Ultrasonics is a powerful sensor for monitoring processes such as composite curing, metal welding, and material thickness in a forming process. Process control often requires a remote, noncontacting way to perform the ultrasonic evaluation. Laser based ultrasonics satisfies most of these requirements. The LLNL weapons program is interested in monitoring the uranium spin-forming operation to measure properties of the formed material. Comparing the measured properties with computer predictions will validate the process models. Process monitoring and control are critical elements of ADaPT. We will continue to develop laser ultrasonics to meet ADaPT NDE needs.

Laser ultrasonics can scan large areas quickly with mirrors. This type of noncontact scanner is ideal for inspecting components of complex shape, because laser beam alignment to the surface is not critical. The International Thermonuclear Experimental Reactor (ITER) program has a laser ultrasonic

Table 1. Model-based processor performance for laser ultrasonic measurements. Each of the 11 runs measures epicentral displacement on an aluminum plate with different laser generation or detection settings. The input SNR varies according to the degree of system optimization. The output SNR resulting from model-based processing of the input signals shows SNR gains ranging from 20–30 dB.

Run No.	Input SNR (dB)	Output SNR (dB)	SNR Gain (dB)
1	28.1	50.3	22.1
2	28.5	50.0	21.6
3	26.9	48.9	21.9
4	27.8	49.7	21.9
5	25.3	49.8	24.6
6	30.8	52.1	21.3
7	23.1	50.4	27.3
8	25.4	52.0	26.6
9	21.1	47.6	26.5
10	22.9	49.4	26.5
11	23.8	57.6	27.8

Table 2. Capabilities of the interactive Laser Ultrasonic Toolbox.

I/O	Processing	Analysis	Applications
Find	Normalize	Spectrum	Interpolation
Read	Window	Blackman-Tukey	Wiener
Write	Cut	Periodogram	Adaptive
	Filter	Max Likelihood	
	IIR	Max Entropy	
	FIR	Correlation	
	Median	Spectrogram	
	Averager	Cepstrum	
	Resample		
	Align		

application inspecting the beryllium plasma facing layer. Laser ultrasonics is ideally suited to provide the thickness and bond quality information required without a contaminating couplant.

A laser interferometer can passively sense acoustic emissions without requiring contact with the component. A potential application is the continuous monitoring of laser optics for the National Ignition Facility to detect minor damage to the optics before a catastrophic failure occurs.

Fiber optic techniques can guide laser energy to generate and detect ultrasonics.[8] Fiber optics allow us to gain access to locations that traditional ultrasonic testing could not inspect. For example, fiber optic laser ultrasonics could detect cracks and measure thickness in weapons as part of surveillance activities.

Next year, we will combine our expertise in ultrasonic imaging with a laser scanning system to produce defect images. Many applications require an image of defects. We have extensive experience in designing and fabricating scanners for defect imaging using couplant ultrasonics. We plan to combine laser generation and detection of ultrasonics with our scanning and imaging technologies to build a facility for demonstrating the ability of laser ultrasonics to evaluate components. Also, we will assemble a fiber optics based laser ultrasonic system.

LLNL has the expertise and capabilities to be a leader in the field of laser ultrasonics. We have the ultrasonic, laser, and signal processing experience needed. We also have the design, engineering, and manufacturing expertise to produce state-of-the-art laser ultrasonic systems.

References

1. Monchalin, J.-P. (1993), "Progress Towards the Application of Laser-Ultrasonics in Industry," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. **12**, D. O. Thompson and D. E. Chimenti, Eds., Plenum Press, New York, N.Y., p. 495.

2. Munch, J., R. Wuerker, and M. LeFebvre (1989), "Michelson interferometer with a holographic memory for measuring fidelity of aberrated phase conjugated beams," *Applied Optics* **28**(9).

3. Basov, N., I. Zuberev, A. Moronov, S. Mikhailov, and A. Okulov (1980), "Laser interferometer with wavefront-reversing mirrors," *Sov. Phys. JETP* **52**(5).

4. DeLaye, P., A. Blouin, D. Drolet, and J. Monchalin (1996), "Heterodyne detection of ultrasound from rough surfaces using a Double Phase Conjugate Mirror," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. **15**, D. O. Thompson and D. E. Chimenti, Eds., (in press).

5. Candy, J. (1988), *Signal Processing: The Modern Approach*, McGraw-Hill, New York, N.Y.

6. Spicer, J. (1991), *Laser Ultrasonics in Finite Structures: Comprehensive Modelling with Supporting Experiment*, Ph.D. thesis, Johns Hopkins University, Baltimore, Md.

7. MATLAB (1995), "The MathWorks," Natick, Mass.

8. Huang, J., S. Krishnaswamy, and J. D. Achenbach (1994), "A Fiber-Optic Heterodyne Dual-Probe Interferometer for Laser-Ultrasonic Crack Detection," *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. **13**, D O. Thompson and D. E. Chimenti, Eds., Plenum Press, New York, N.Y., p. 485.

Progress in 3-D Quantitative DR/CT

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

Harry E. Martz and Pierre-Louis Bossart
Manufacturing and Materials Engineering Division
Mechanical Engineering

Stephen G. Azevedo
Laser Engineering Division
Electronics Engineering

The Computed Tomography (CT) project has continued to evolve to better meet the rapidly evolving needs of the Lawrence Livermore National Laboratory (LLNL). We remain committed to assembling the most comprehensive set of radiographic tools available anywhere, while at the same time solving the needs of an expanding customer base. Our continuing work in scientific and industrial CT and digital radiography (DR) has spanned all aspects of the theoretical and experimental practice. We continue to apply state-of-the-art methods of data acquisition, processing, analysis, and understanding, associated with penetrating radiation imaging. Our goal is to provide reliable and efficient nondestructive evaluation (NDE) techniques for use in probing the internal structure of fabricated objects and materials, associated with a broad spectrum of LLNL programs, industrial partnerships, and academic research. This report describes progress in our activities during FY-95, in the area of quantitative measurement tools for improved understanding of objects under inspection.

Introduction

In computed tomography (CT), a series of images is converted by substantial processing into higher dimensional data, revealing the internal structure of measured objects. CT continues to be an active area of research and development. LLNL's NDE section has made significant contributions to the field of non-destructive inspection of industrial parts, using x-ray CT. When many x-ray radiographs are combined using the proper CT algorithms, the internal structure of complex parts is revealed. Using today's fast computers and CCD camera technology, it is possible to rapidly acquire digital x-ray radiographs from many different angles around a part, and to process those radiographs into a full 3-D attenuation map of the part. Such maps provide information about materials properties (such as volume density, atomic number, porosity, and homogeneity), and about internal and external surfaces and component proximity.

In the past, industrial CT inspection was focused largely on the more qualitative aspects obtainable from the CT model, such as material porosity and homogeneity, or the relative size and location of components. More recently, there has been a rapidly increasing need for extracting quantitative information, such as actual component dimensions, angles, and curvatures, from the CT model. Starting with FY-94, we shifted our primary focus to extracting quantitative information from our CT data.[1] During the past year we have made significant progress in that direction.

One area of importance to both LLNL and industry, where quantitative CT can play an important role, is engineering. Computerized design, analysis, and manufacturing has become a necessity of modern engineering practice. A problem in common to both finite element modeling (FEM) and reverse engineering is 3-D geometry acquisition, necessary when no computer-aided drafting (CAD) drawings exist (e.g., in older systems or in medical

applications), or in cases where actual as-manufactured geometries are required. X-ray CT can provide a solution to this problem, as shown in **Fig. 1**. In contrast to laser range-finders, coordinate measurement machines, and other dimensional measurement devices that are sometimes used for 3-D geometry acquisition, x-ray CT allows determination of the spatial arrangement and dimensions of the components comprising a part, even when those components have internal boundaries. Recent progress in computerized data storage and computational speed makes it possible to rapidly acquire and process high spatial resolution x-ray CT data sets. The resulting voxelized 3-D x-ray attenuation model of the scanned part can be used for 3-D geometry acquisition. X-ray CT has thus become a cost-effective method for extracting 3-D geometries, especially in cases where internal structure and geometry are important.

Another area of importance to our quantitative CT work is increasing the accuracy of our CT image reconstructions. Commercially available x-ray sources produce polychromatic x-ray beams. This can introduce a source of error into CT reconstructions, if there is no accounting for the preferential absorption of lower energy ("soft") x rays over higher energy ("hard") x rays. Traditionally, correcting for this "beam hardening" has involved acquiring data using specially shaped phantoms. We have recently developed a new algorithm for determining beam hardening corrections that permits the use of nearly any phantom.

The functional requirements of some components and component assemblies demand a high degree of precision in the shape of the internal walls and boundaries. High pressure fuel injectors and single crystal turbine blades are two examples. For fuel injectors, the 3-D shape of internal fuel passages can account for significant differences in fuel efficiency. For turbine blades, the size and integrity of internal passages can affect the viability of the entire jet engine. It has been established that CT can dimension internal walls and component assemblies, but it is important to know at what accuracy. If CT-reconstructed volumes are to perform the task of detailed inspection of object internals, the error variance connected with measurements made should be reported with the measurement. Part of our effort this year was spent examining the error properties of dimensioning algorithms for CT data, and determining how the accuracy varies with different objects and conditions.

NDE often explores new applications of DR and CT. Over the past year, we investigated the application of CT to carbon-polymer composites in order to better understand behavior when static or fatigue loaded to failure. The main goal of this work was to demonstrate the feasibility of using CT to characterize the internal defects and microstructural damages for such materials. Our tasks included studying the typical load size of composite specimens (on the order of $20 \times 2 \times 200$ mm), and identifying the optimum scan parameter values for these samples, to correlate the defects with the corresponding loading situation. This allows us to evaluate the agreement of the real data with the mechanical models of failure proposed in previous studies.

Progress

Software and Algorithms

X-ray CT and CAD. The process of reverse engineering often requires or benefits from a CAD representation. X-ray CT can provide a cost-effective method of producing a CAD representation directly from a physical part, particularly when the part contains internal surfaces or structure.[2,3] The process of converting x-ray CT reconstruction into a CAD representation involves two major steps: 1) individual components within the part must be recognized and segmented; and 2) the exterior, and possibly the interior, surfaces of the components must be extracted and output in a form usable by CAD software (**Fig. 1**). We address the segmentation step below in the section describing the application of x-ray CT data to FEM analysis. For now, we will assume the segmentation step is complete.

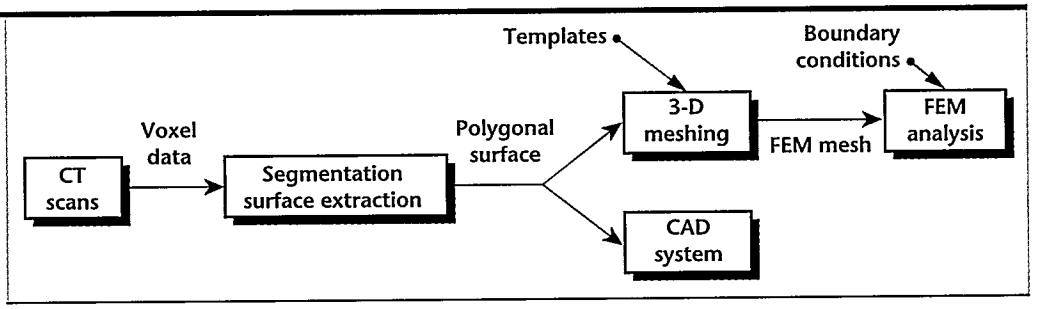


Figure 1. Schematic of data flow from CT to CAD and FEM.

The extraction of component surfaces can be done by the marching cubes algorithm.[1,4] It will extract surface polygons which are sorted and output in a form compatible with commercial CAD packages such as "Surfacer" from Imageware (Ann Arbor, MI). Such CAD packages reduce a set of surface data points (obtained, in this case, from the marching cubes algorithm) to non-uniform rational B-spline surfaces (NURBS) for CAD representation. However, this process is typically quite labor-intensive. A recent dissertation[5] on surface reconstruction from (possibly noise corrupted) surface points, and similar work ongoing at Stanford University, may provide an even more elegant, robust solution, and possibly less labor-intensive solution.

X-ray CT and FEM. The process of using x-ray CT data to support the geometry acquisition phase of FEM analysis (**Fig. 1**) requires meshing the interior region(s) of the components, based upon boundaries determined from the voxelized x-ray CT reconstruction of the part.[6] In simple cases (single component parts with good contrast, unlike air), the part boundary can be determined by proper selection of a threshold attenuation value. Unfortunately, many useful parts contain multiple components and simple thresholding is inadequate for proper segmentation. Accordingly, we have undertaken significant work in the area of 3-D segmentation of voxelized (x-ray CT) data. This work is being conducted in collaboration with other researchers at LLNL interested in FEM of human appendages, and in particular, the human hand.[7,8,9]

We began our collaboration by acquiring CT radiographic data on a medically amputated human hand. A CT reconstruction was performed, and segmentation work begun. The human hand provides a very difficult test case for the segmentation software we are developing. The low contrast between various soft tissues and the diffuse boundary between bone and ligament makes for a difficult segmentation environment (**Fig. 2a**). Additionally, the internal (trabecular) structure of bone provides an additional challenge to the segmentation software in this case, since at the current time, we are not interested in modeling the bone's internal structure.

Investigation of 3-D segmentation schemes began with the interactive segmentation package VOXELMAN (public domain software, University of Hamburg, Germany), which was developed for segmentation of 3-D medical datasets. VOXELMAN makes it possible to visualize three orthogonal views of the 3-D CT dataset, and to perform histogram-based thresholding and various other operations. After thresholding, several morphological operations, and a connected components analysis, the segmentation results can be rendered in 3-D and output as a bitmap volume. Errors in the segmentation can be detected from the rendering, and individual voxel segmentation assignments can be edited manually. Unfortunately, this type of segmentation is both tedious and labor-intensive. VOXELMAN's simplistic thresholding approach to the initial segmentation has significant troubles with both the internal structure of the bone, and the diffuse bone boundaries found in some locations.

From this experience, we determined that we needed both a more robust method of boundary detection and a scheme with less sensitivity to internal structure. We backtracked to the world of collections of 2-D images (slices), where work has been done, and visualization is significantly easier. Here, we first assessed the performance of edge detectors[10], which detect and localize sharp boundaries, in order to automate the segmentation process. Unfortunately, the internal structure of the bones and the texture of the surrounding soft tissue (**Fig. 2a**) created a large number of spurious edges, making this approach unreliable.

Moving on to an efficient region-based algorithm, based on advanced morphological operations, yielded much more accurate boundaries. A gray-scale dilation on a large neighborhood followed by a gray-scale reconstruction[11] significantly alleviated the problems (**Fig. 2b**). After the gray-scale reconstruction, the image histogram is bimodal so a threshold can be chosen to produce an initial, coarse segmentation. Using this coarse segmentation and the gradient image, watershed lines[12] can be extracted and used to produce the much more accurate solution shown in **Fig. 3**. The results in **Fig. 3b** clearly show the limitations of this approach when applied to the hand data set. In this

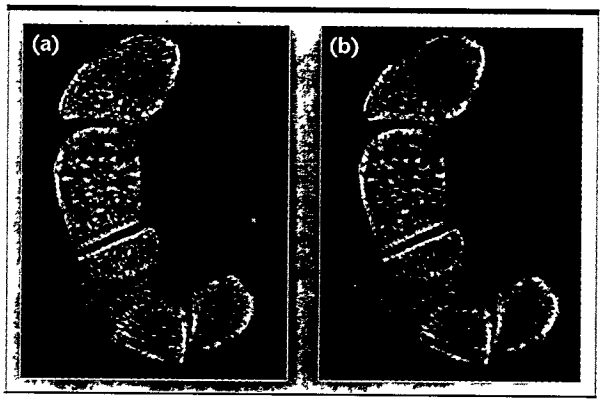


Figure 2. (a) Representative CT image slice of a hand; and (b) result after gray-scale mathematical morphology image reconstruction.

example, we cannot resolve two (bottom left) bones of the five finger bones during the thresholding step. After computation of the watershed lines, these two bones result in a single marker (as seen in **Fig. 3b**), demonstrating the need for some human interaction in some cases. We are hopeful that a 3-D extension of this algorithm will provide the extra information needed to produce an even more accurate segmentation result. Accordingly, we are currently working on the design of a fully 3-D segmentation algorithm. For a more detailed report on this segmentation work, see Bossart.[13]

Once the automatic segmentation is complete, the components can be meshed, or the segmentation can be further refined manually. We are preparing to use the 3DVIEWNIX package (public domain software, University of Pennsylvania, PA) for manual refinement, since it provides numerous advantages over the VOXELMAN package. With the segmentation complete, current automatic meshing technology can directly use the segmentation data and the octree algorithm to directly produce a tetrahedral mesh of each component.[6] However, tetrahedral meshes are not suited for the mechanical dynamics simulations required in our applications. Structural engineers strongly prefer hexahedral meshes for these problems. Although a tetrahedral mesh can be converted to a hexahedral mesh, the results are generally quite poor.

LLNL researchers have developed two methods of producing a high quality hexahedral mesh filling the segmented volume. Both involve the generation of a generic volumetric mesh that is registered against the segmented surface, and then projected onto that surface using the TrueGrid (XYZ Solutions Inc., Livermore, CA) meshing package. The resulting mesh is refined by smoothing and is then ready for input to the NIKE3D FEM Code developed at LLNL. **Figure 4** shows the mesh of a metacarpal bone produced by our work.

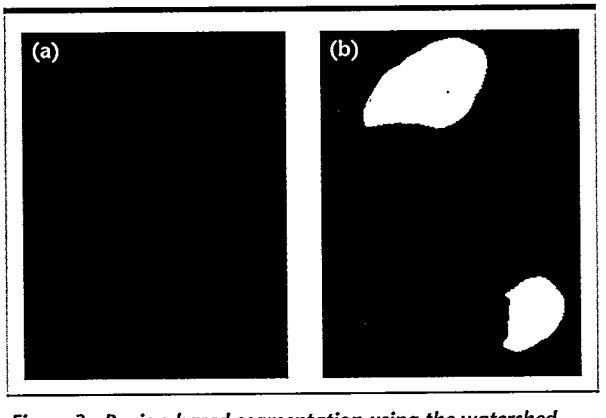


Figure 3. Region-based segmentation using the watershed concept: a) gradient image of the CT slice in Figure 2a; and b) results of region-growing algorithm.

Using these bone meshes derived from our CT data, our collaborators will analyze the joints of the hand under dynamic conditions. Eventually, we hope to segment and include various soft tissues into the models, including tendons and cartilage. Such analyses will not only greatly increase our understanding of joint mechanics, but will permit the computational exploration and evaluation of prosthesis.

Beam Hardening. Collaborative research with the Nuclear Physics Department at the University of Bologna, Italy, produced useful results in the last year. Researchers there are using LLNL software (the RECON package[14]) for image reconstruction and simulation of their CT data. They have built a high-energy system using a 2-MeV-linear accelerator source and solid-state cadmium-zinc-telluride detectors, as well as a neutron CT system using an image-intensifier and CCD camera detector. Recent images generated by these detector systems show reasonable speeds and good quality results.[15] However, beam hardening, a phenomenon where polychromatic x-ray beams are preferentially attenuated, is still a problem with the high-energy system.

Our collaboration resulted in a unique method of solving the beam-hardening problem. The common way to address beam hardening is to build a step-wedge of basis materials that gives accurate measures of attenuation vs part thickness for a given source/detector configuration and material. This implies that many step-wedges are needed for the various configurations of interest. Our approach

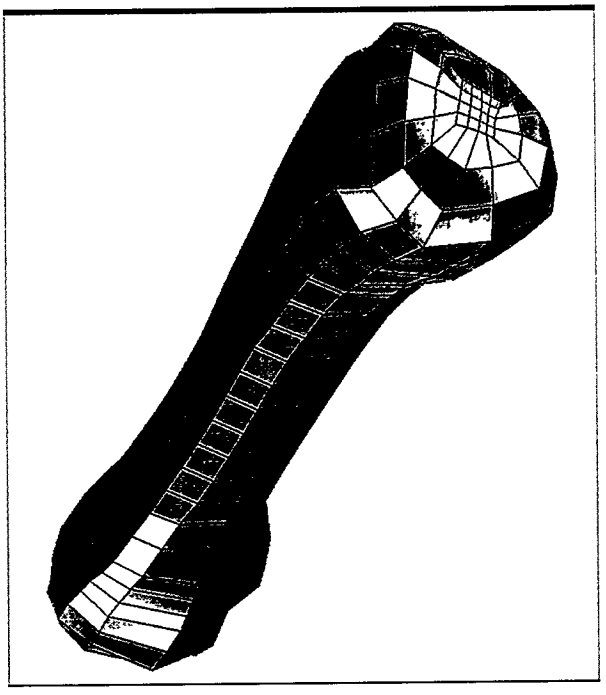


Figure 4. Volumetric mesh of a segmented metacarpal bone.

is to perform CT on a phantom of known material and size. It could be as simple as a cylinder with a diameter the size of the maximum expected material length in the unknown. Models of the phantom are generated in the RECON package and compared to the measured data over the entire sinogram. A polynomial fit of x-ray path length vs attenuation produces a model that is used to correct the sinogram of any part, at the same energy and similar material. **Figure 5** shows the results of the new beam-hardening method applied to a CT reconstruction of a steel phantom scanned on the Italian high-energy CT system. Notice that the line-out at the bottom shows much-reduced beam-hardening artifacts (constant values in the steel, near zero values in air). These techniques will be applied to LLNL scanners for improved beam-hardening correction.

Quantitative Measurements of Turbine Blades from CT. The USAF-funded study of turbine blade NDE, performed in collaboration with Southwest Research Institute, presented an ideal opportunity to examine dimensioning algorithms for CT data. This three-year study involved ultrasonic testing (UT), RT and CT inspections of some 100 blades, with companion destructive analysis of selected blades. Using the boundary identification approach common to most CT analysis software, we performed a series of simulations to establish the quantitative limits of the techniques for different levels of image blur, and different noise-to-signal levels. Second, we examined potential errors in the application of these algorithms for different types of scanning and different types of object constraints. Third, we assessed the calculated errors with the UT and destructive testing data.

The analysis of boundary identification algorithms with simulated data gave a number of results. With respect to spatial blur, CT wall thickness algorithms and boundary identification algorithms can

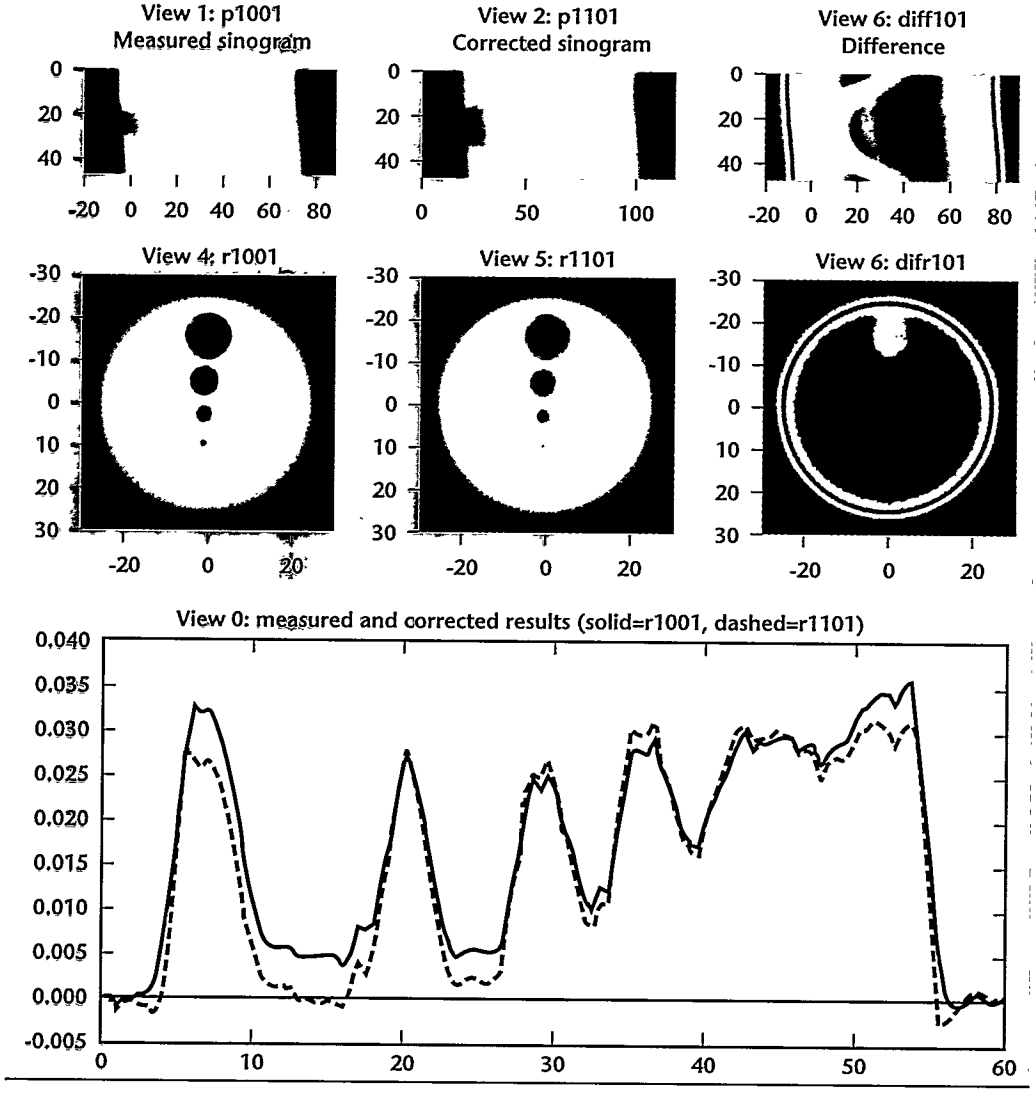


Figure 5. Beam-hardening correction. The two images at left show the sinogram and reconstruction of the raw data with beam-hardening. The corrected images are shown in the middle, with the difference at right. Vertical slices through the reconstructed images are shown at the bottom. The uncorrected slice is the solid line; the corrected slice is the dashed line.

be accurate down to 0.7 of the point spread function (PSF) of the scanner. The inner walls are rarely imaged as well as the outer walls, since the beam imaging the inner contents of the object includes a much larger scatter component. However, the necessity to scan around the circumference of the blade equilibrates the PSF across the entire image, enough to make the same algorithm efficacious for the entire image or volume. Provided the PSF is small enough, CT wall thickness algorithms can calculate dimensions to subpixel accuracy. CT wall thickness algorithms are particularly sensitive to noise content in the image, with the accuracy of the algorithm substantially decreasing for noise-to-signal greater than 1.5 %.

The application of CT wall thickness algorithms to CT-reconstructed data includes a number of operations that can make correlation with design documents, or destructive measurements difficult. In turbine blade design drawings, wall thickness is defined as the shortest distance across the wall, which is the path perpendicular to the two boundaries. Two circumstances can result in a CT wall thickness measurement not on the perpendicular path: 1) the operator, or the operator and software do not obtain a trace across the wall that is perpendicular to the boundaries, and 2) the blade geometry is not perpendicular to the beam in the vertical direction at the slice at which the operator acquires the line trace. Both deviations of the measured path and the perpendicular path result in measurement errors that increase with the difference angle, "γ." **Figure 6** is a pictorial of the difference between the two different paths.

The fraction error can be expressed quantitatively as:

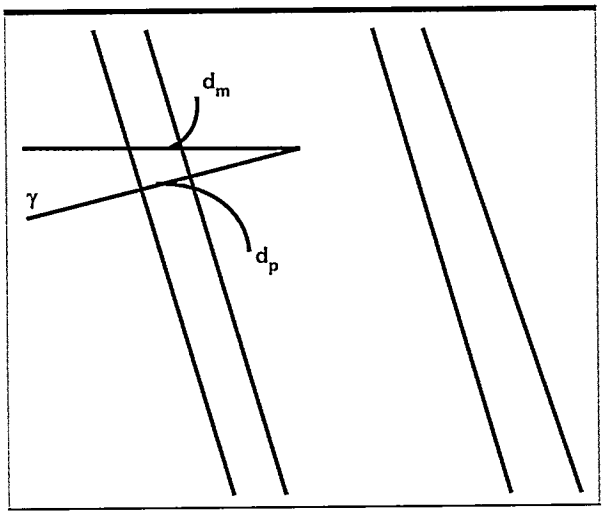


Figure 6. Illustration of perpendicular and level slice planes across a wall.

$$DE_{NP} = d_m - d_p = [(\tan^2(\gamma) + 1)^{0.5}] - 1$$

where:

d_m = the measured thickness of the wall,

d_p = the thickness of the wall measured along the perpendicular to the wall thickness boundary,

γ = the angle between the perpendicular path and the measured path.

Using this equation, **Fig. 7** is a plot of this function for angles up to 41°. The significance of the change in wall thickness with pitch angle is that many turbine blade designs include angles in excess of 20°. This is a particularly important result for correlation with UT data, since the UT measurement is always the perpendicular path (if done at all correctly). It is entirely possible for both the UT measurement and the CT measurement to be very accurate and produce different results for what seem to be scans in the same region.

The results from the measurements taken on the many different turbine blades, and the correlation with the UT and destructive analysis, follow the results from the analyses above. Overall, the UT and CT data correlate quite well, with a correlation of R = 0.93, and an average difference in reported thickness of less than 2 mils for turbine blades, and differences slightly larger for first and second stage vanes. For each blade or vane, six to eight different positions were identified for UT and CT measurement. In the positions where the pitch angle was low, the CT and UT measurements had average differences, on the order of one mil. However, for positions where the pitch angle was varying widely, the correlations were not as good. **Figure 8** contains a slice of a turbine blade volume that includes a varying pitch angle, and the wall thickness measurements from the different paths across the wall are identified. Turbine engine vanes involve greater mass, more pitch angle, and can be more difficult to penetrate with radiation (a condition which leads to greater noise content in the image). In our analysis we would expect less correlational accuracy for vanes with respect to UT data. A somewhat unexpected result from our study was the variation in wall thickness with the destructive results. We scanned the variety of turbine blades with a 0.97 µm pixel size and slice thickness. In areas with selected measurement sites close to surrounding pedestals, it was difficult to obtain a single-point destructive measurement (performed by hand with a caliper) that was at exactly the position of a single point CT measurement. Consequently, we developed the capability to calculate thicknesses for an entire

wall to perform the correlation. **Figure 9** includes a wall thickness calculation applied to an entire wall. Increased precision in CT and UT dimensioning has made correlation with destructive testing more difficult. It is not clear that the NDE community fully understands the significance of this result.

Applications

CT Inspection of Carbon-Polymer Composites. Carbon-polymer composites are becoming more and more common in engineering applications. They offer numerous advantages over conventional materials in many applications. Accordingly, understanding their properties and failure modes, and more crucial, predicting their failure and finding inspection techniques able to determine composite condition is increasingly important. Work at LLNL is focused on studying the lifetime and failure modes of some particular composites. We are applying different NDE techniques to help in investigations of polymer composites. Here we focus only on some of our x-ray investigations.

One composite application was x-ray CT inspection of a flywheel that had proven to be nearly impossible to evaluate using ultrasound techniques. The flywheel's annular structure of alternating fiberglass and carbon composite layers makes high-frequency ultrasonic inspection difficult, due to the large impedance mismatch between the fiberglass and composite boundaries. CT x-ray analysis is not affected by these boundaries, and therefore provides a clear view of the

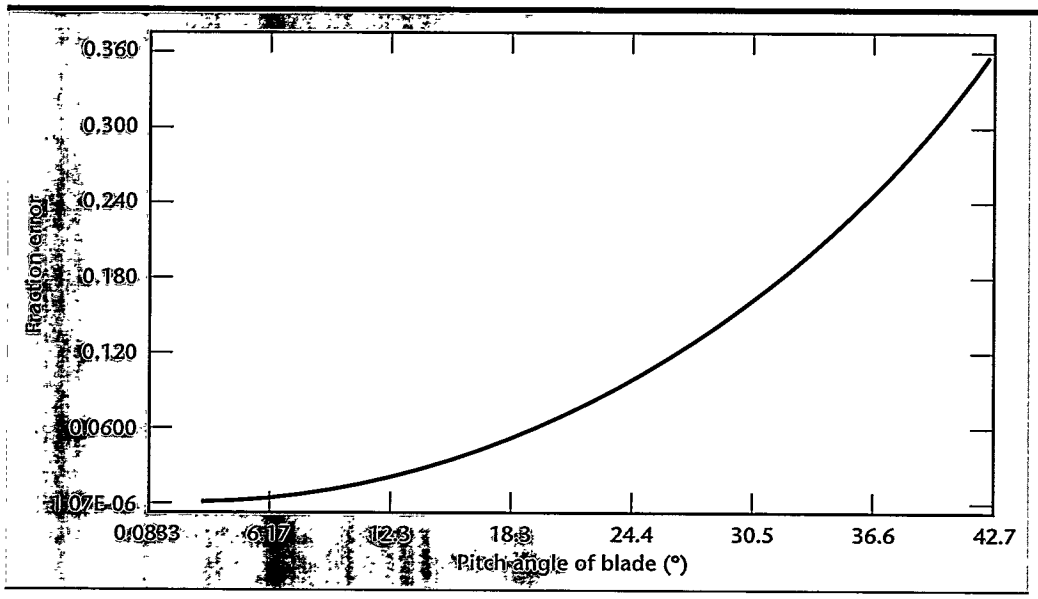


Figure 7. Plot of the relationship for the difference between level and perpendicular slice planes and pitch angle.

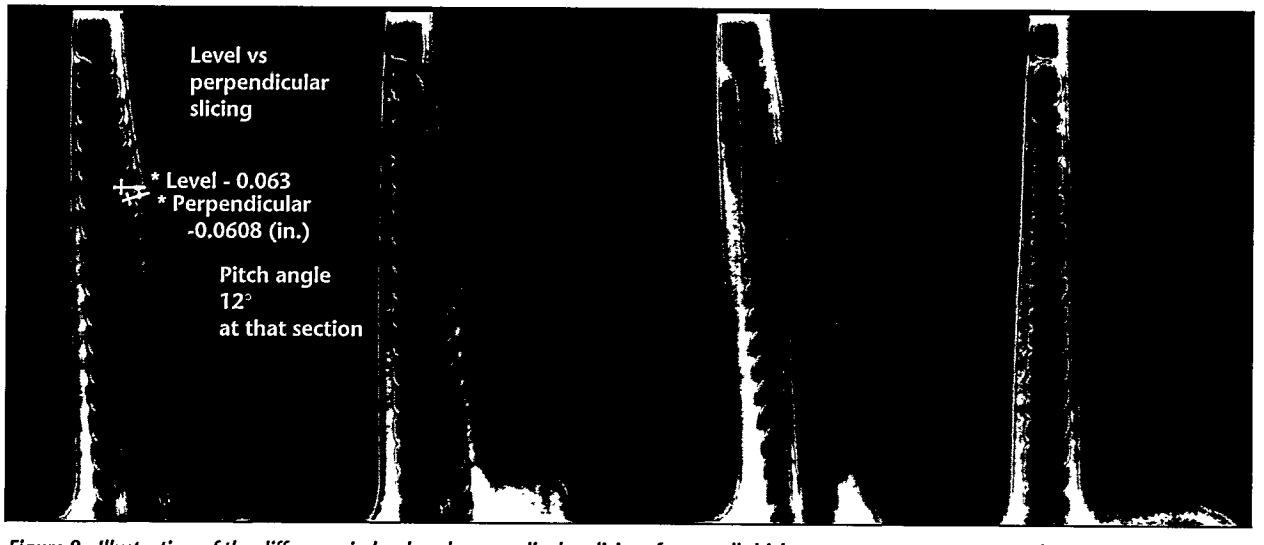


Figure 8. Illustration of the difference in level and perpendicular slicing, from wall thickness measurements on a turbine blade.

flywheel's internal structure, as shown in **Fig. 10**. This work was successful enough for us to continue to explore the possible application of this technique to inspect flywheels.

In addition to the flywheel investigation, this year we began another composite investigation in collaboration with the University of Calgliari, Italy.[16] This collaboration was initiated to explore the application of x-ray CT to carbon-polymer composites, in the hope of detecting the early signs of failure. We studied six different carbon-polymer composite samples with different matrix composition, geometric characteristics, and load profiles. Five of them were from Calgliari, and one from LLNL. The Calgliari samples were studied only after loading. Their previously unknown situation created uncertainty with respect to the origin of some observed features. The sixth sample was studied before and after four tensile-compression low-stress fatigue loading tests at LLNL. New CT data was acquired after every 0.2 Mcycles at 180° C. The fourth loading test was interrupted before completing its 0.2 Mcycles (total load of 961 401 cycles) because sample deformation, due to reduction in sample stiffness, exceeded 0.02 mm.

The study of the sixth, graphite-polyimide, sample allowed us to ascertain some interesting characteristics. First, we had the possibility of locating, from a spatial point of view, the defects inside the sample. We could establish the moment they appeared, their evolution along the loading test, and the relationship between such defects and the mechanical characteristics of the material itself. The use of different investigation techniques permitted us to look at several different aspects of the same phenomena. Representative results are shown in **Fig. 11**.

One of the main purposes of this study was to determine if x-ray CT can be used to evaluate carbon-polymer composite materials characterized by a low-contrast sensitivity between fibers and matrix components. We also wanted to determine if

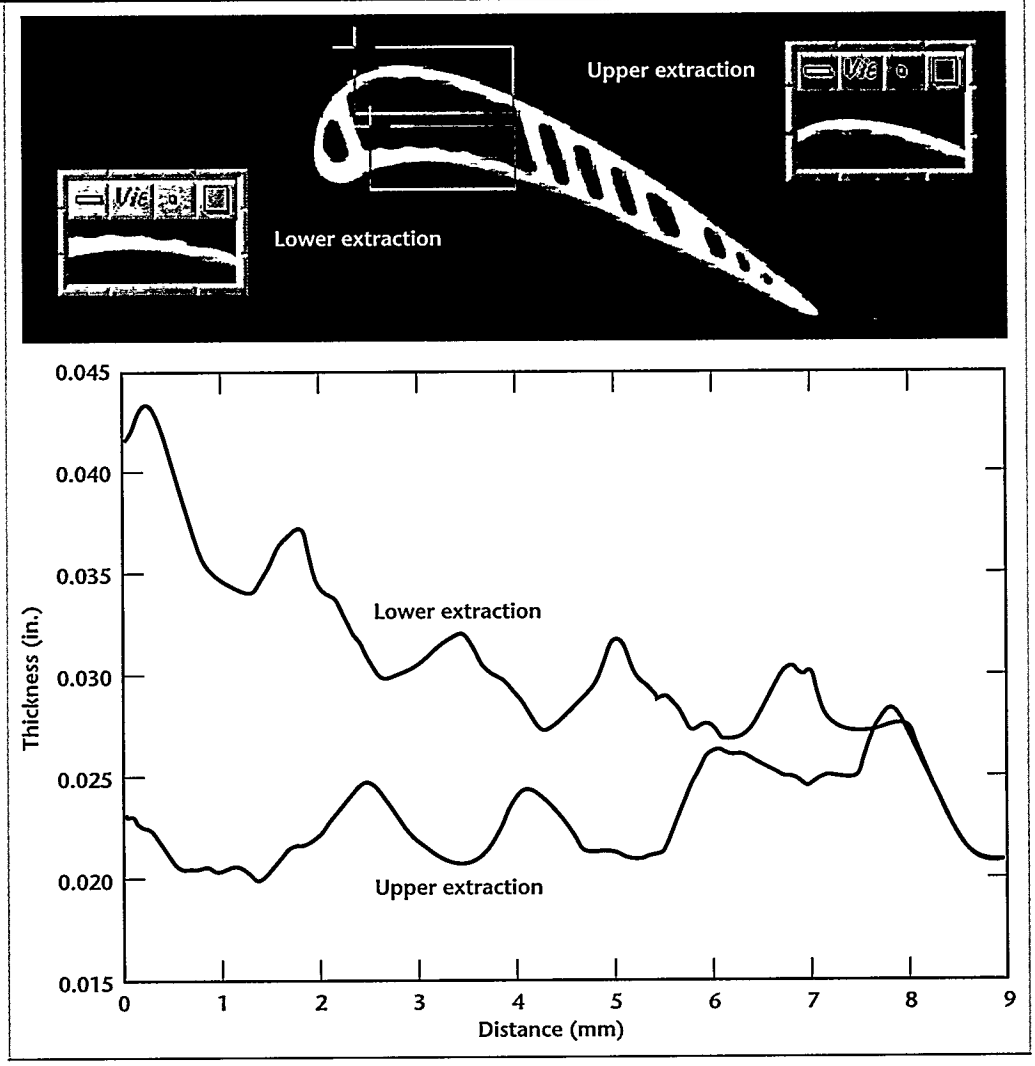


Figure 9. Variation of wall thickness for a turbine blade in an area close to supporting pedestals.

we could provide useful results without the use of any contrast enhancement dye, where the fiber dimensions are only a few microns in diameter.

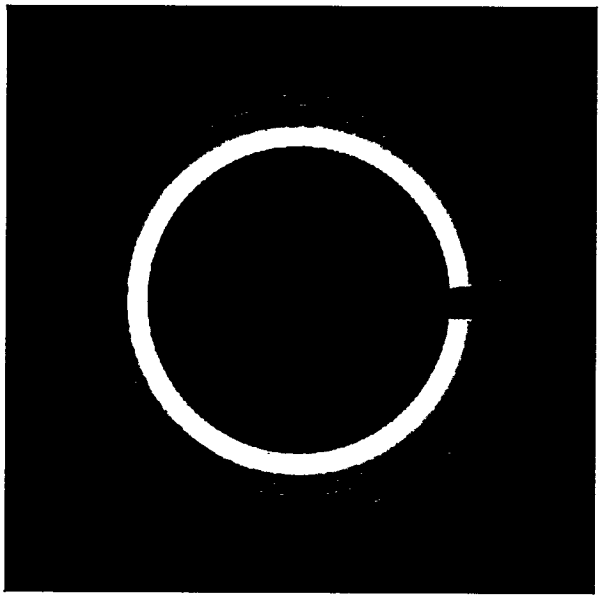


Figure 10. CT-reconstruction slice image of an experimental composite flywheel and a hexagonal fiducial. The fiberglass layers and the fiberglass inner hub are clearly visible. Notice the thinning of the third (from the inside) fiberglass layer in the left and upper left of the flywheel.

Preliminary results have shown that x-ray data are different in physical characteristics identified, and are complementary to other nondestructive methods. The x-ray CT results obtained are encouraging with respect to the study of thin carbon-polymer composite materials. However, there are some concerns as to whether we can see closed cracks and delaminations. We plan to investigate this further by scanning a part after it has been infiltrated by an x-ray contrasting agent

Postal Explosives Detection. LLNL has a long history of lending its expertise to problems facing various federal, state, and local agencies. Recently, the NDE thrust area was asked by the Forensic Science Center to provide assistance in postal bomb detection for the Livermore Police Department. We used our x-ray facility to obtain both film and digital radiographs of three mock postal bombs. Photographs of these mock bombs are shown in **Fig. 12** (top row). The corresponding DR results are shown in the middle row of **Fig. 12.** The film radiography results (not shown) are quite similar to the DR results; i.e., the internals details of the package are clearly identified. Details of the actual contents of two of the mock bombs are shown via a plexiglas window located underneath two of the bombs, as shown in the bottom row of

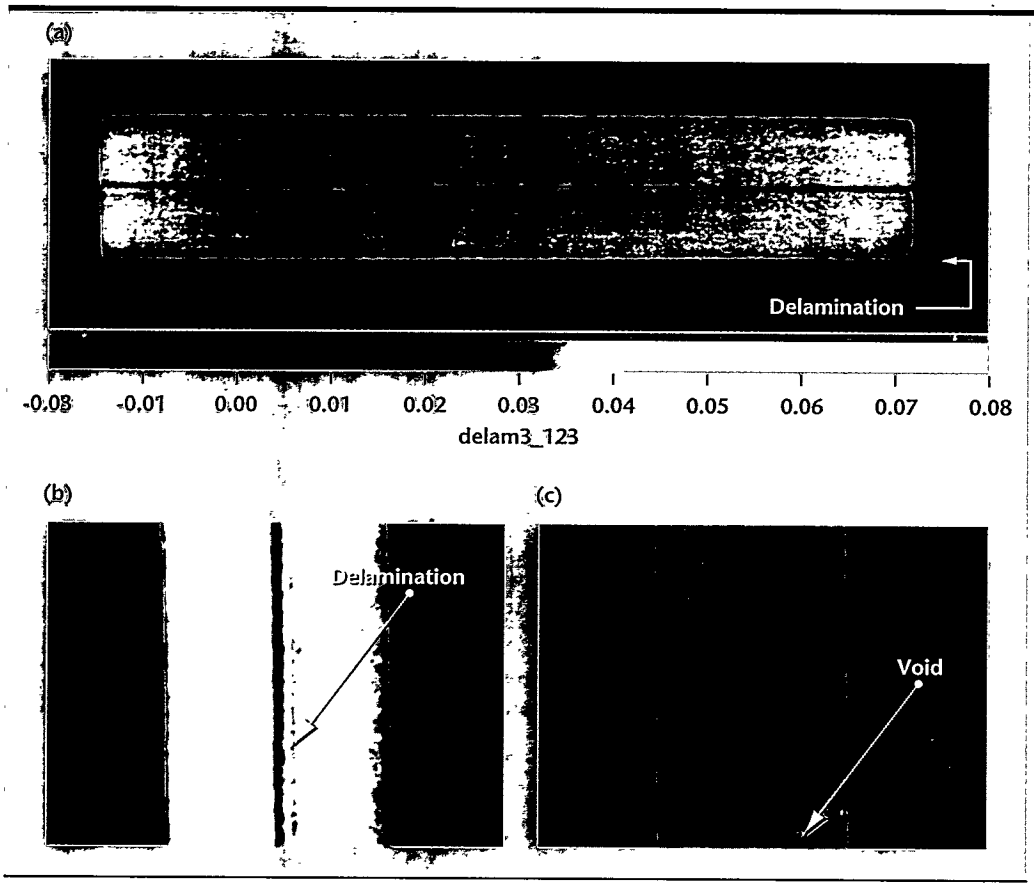


Figure 11. CT-reconstruction cross-sections of two composite tensile bars after loading at the University of Calgliari; (a) and (b) exhibit delaminations; (c) exhibits a void.

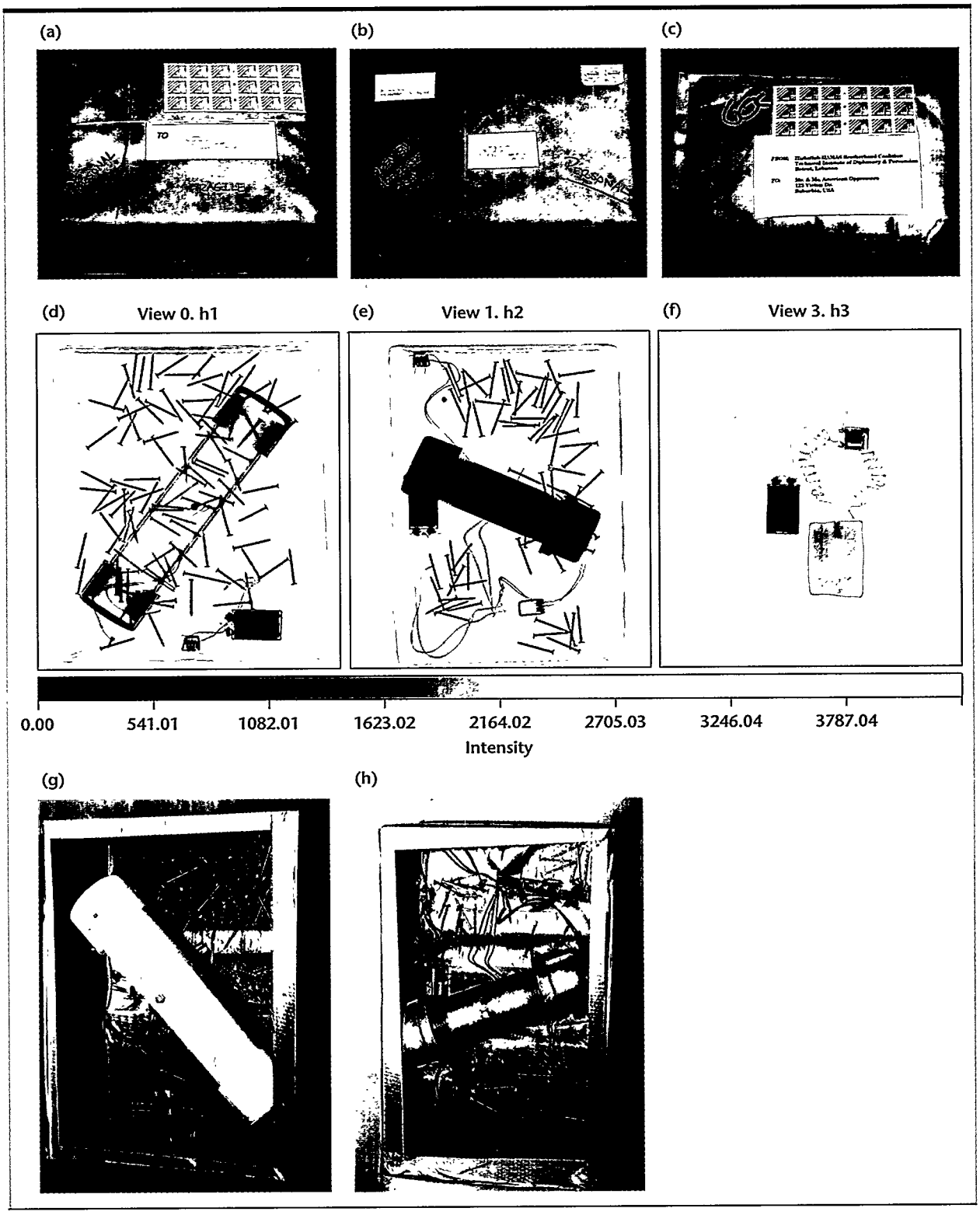


Figure 12. Three views of mock postal bombs. The top row is a series of photographs of the (unopened) mock bombs. The middle row shows the digital radiographs obtained for each, and the bottom row shows two of the mock bomb packages, opened to expose their contents.

Fig. 12. This work confirms that DR and film radiography can easily identify concealed postal bombs at the current threat level.

Summary

We have developed a suite of 2-D segmentation tools for segmenting 3-D x-ray CT data. These tools allow robust segmentation of even low contrast data in most cases. They also provide an excellent starting point for the development of robust 3-D segmentation algorithms and tools. We have also collaborated to make the results of our segmentation work directly available to various automated and semi-automated meshing codes, thereby helping to minimize the time required to acquire complex geometries for FEM analysis.

Developing ways to accurately measure inner boundaries and components in three dimensions has become more important as the functional requirements of certain objects has become more strategic. It is not enough to be able to measure the size with unknown accuracy, or with some *ad hoc* tolerance. In order for 3-D CT to be worth the time and effort to acquire and reconstruct a 3-D data set, the properties of that data must be well known. Automated inspections with CT data are becoming more and more tractable with faster scanning and reconstruction schemes. However, the reliability of these operations is built upon the accuracy of the reconstructed data. We will continue working to determine the error properties of CT data, as well as the error variance of analyses which can be applied to CT data in response to customer requests and to establish the viability of certain applications of CT scanning.

Future Work

We plan to continue our work on 3-D analysis of CT data along three major lines. First, we plan to significantly enhance our current CT scanner alignment capabilities. We will develop a robust set of alignment procedures and measurement. This will allow us to acquire better data with fewer artifacts. It will permit both the use of larger fan and cone angles (for higher magnification), and the use of more accurate 3-D reconstruction algorithms. We also plan to improve our software capabilities to deal with known, measured misalignments.

Second, we plan to continue development of 3-D analysis tools, including 3-D measurement techniques and segmentation methods. We will investigate the error properties of our existing dimensional measurement tools, and explore scale and sample interval issues. We will work to increase the robustness of the segmentation process and to more fully exploit the 3-D nature of the data. Promising 2-D segmentation algorithms need to be adapted and enhanced for 3-D application. We also need to decrease the time needed for human interaction during the segmentation process. This requires better human interfaces to the various segmentation algorithms, and improved visualization software.

Third, we plan to improve our cone beam reconstruction capabilities. We will deploy modern, "exact" cone beam algorithms to bring routine 3-D cone beam CT capability into everyday use.

Acknowledgments

The authors would like to acknowledge the valuable contributions to our FEM work by K. Hollerbach (Institute of Scientific Computing), and D. Schauer, E. Ashby, and Jeffery Weiss . We would like to thank P. Grant of the Forensic Science Center for referring the postal bomb inspection to us. Finally, we would like to thank E. Updike, D. Rikard, and G. Ford for lending their efforts and expertise in DR/CT data acquisition to all of our efforts.

References

1. Martz, H. E. (1995), *Nondestructive Evaluation*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-ID-119059).

2. Mohr, G. A., and F. Little (1994), "Effective 3D Geometry Extraction and Reverse CAD Modeling," *Review of Progress in QNDE*, vol. **14**, pp. 651–656.

3. Stanley, J. H., R. N. Yancey, Q. Cao, and N. J. Dusaussoy (1995), "Reverse Engineering and Rapid Prototyping for Solid Freeform Fabrication," *SPIE* vol. **2455**, pp. 306–311.

4. Lorensen, W. E., and H. E. Cline (1987), "Marching Cubes: A High Resolution Surface Extraction Algorithm," *Computer Graphics*, vol. **21**(2), pp. 163–169.

5. Hoppe, H. (1994), *Surface Reconstruction from Unorganized Points*, Ph.D. dissertation, University of Washington, Seattle, Wash.

6. Finnigan, P. M., A. F. Hathaway, W. E. Lorensen, I. J. Connell, V. N. Parthasarathy, and J. B. Ross (1991), "CATFEM - Computer Assisted Tomography and Finite Element Modeling," *Control and Dynamic Systems*, vol. **49**, pp. 298–337.

7. Weiss, J. A., B. N. Maker, and D. A. Schauer (1995), "Treatment of initial stress in hyperelastic finite element models of soft tissues," *ASME Bioengineering Conference, BED*-**29**, pp. 105–106.

8. Puso, M: A., J. A. Weiss, B. N. Maker, and D. A. Schauer (1995), "A transversely isotropic hyperelastic shell finite element," *Proc. ASME Bioengineering Conference*, BED-**29**, pp. 103–104.

9. Hollerbach, K., K. Underhill, and R. Rainsberger (1995), "Automated volumetric grid generation for finite element modeling of human hand joints," *Proc. ASME Bioengineering Conference*, BED-**29**, pp. 49–50.

10. Canny, J. (1986), "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. **8**(6), pp. 679–697.

11. Vincent, L. (1993), "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms," *IEEE Transactions on Image Processing*, vol. **2**(2), pp. 176–201.

12. Vincent, L., and P. Soille (1991), "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. **13**(6), pp. 583-598.

13. Bossart, P-L, H. E. Martz, H. R. Brand, and K. Hollerbach (1995), "Application of 3D X-ray CT Data Sets to Finite Element Analysis," Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-JC-120894), submitted to *QNDE Conference, Seattle, Wash.*, July 30 to August 4, 1995 (in press).

14. Brand, H. R., S. G. Azevedo, M. R. Gorvad, D. J. Schneberk, H. E. Martz, and G. P. Roberson (1995), *Computed Tomography Software and Standards: Revision 2.1*, Lawrence Livermore National Laboratory, Livermore, Calif., NDE Documentation, July 31, 1995.

15. Rossi, M., G. Baldazzi, E. Querzola, G. Guidi, M. G. Scannavini, P. Chirco, M. Zanarini, F. Casali, and S. Azevedo (1995), "High Energy Tomography using Cadmium Zinc Telluride Detectors," *Proceedings of 9th International Workshop on Room Temperature Semiconductor X- and Gamma Ray Detectors, Grenoble France*, September 18-22, 1995 (in press).

16. Savona, V., H. E. Martz, H. R. Brand, S. E. Groves, and S. J. DeTeresa (1995), "Characterization of Static- and Fatigue-Loaded Carbon Composites by X-ray CT," Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL-JC-121680), presented at *22nd QNDE Conference, Seattle, Wash.*, July 30-August 4, 1995 (in press).

omputed Tomography Software and Camera Upgrade

Hal R. Brand, Daniel J. Schneberk, and Michael R. Gorvad
Computer Applications Sciences and Engineering Division
Computations Organization

We have upgraded our core computed tomography (CT) reconstruction software to take advantage of our new SGI Onyx computer. We have used both the Symmetric MultiProcessing (SMP) and Reality Engine^2 (RE^2) graphics co-processor capabilities of the Onyx. Reconstruction times have been reduced by up to a factor of 80. We have also implemented a very powerful 2-D (area array) preprocessor for preparing raw CT data for image reconstruction. This software provides a fast, efficient, powerful, and user friendly upgrade to our previous preprocessing software.

We have enhanced our data acquisition capabilities in two ways. First, in support of the preprocessor work, the data acquisition software provides the preprocessing and reconstruction directives for the standard approaches to CT reconstruction. Second, two new cameras, AXIOM Research 12-bit and 16-bit cameras, have been added into the component hardware in the suite of scanners available for CRADA and customer work. We will use this new scanning capability to develop and evaluate scanning strategies for performing high energy digital radiography and CT in the coming year.

Introduction

The goal of x-ray computed tomography (CT) is the "reconstruction," via computational methods, of an object that has been scanned with x rays. The reconstructed object is not a physical entity, but instead is a digital representation of x-ray attenuation of the physical object, and is often referred to as an "image" in reconstruction literature. Details of computed tomographic reconstruction can be found in numerous references.[1]

CT of a physical object requires the acquisition of a collection of x-ray views or projections of the object. These projections are akin to medical or industrial radiographs familiar to most people. The theory of CT is based on the Radon Transform[2]. If it were possible to acquire an infinite collection of x-ray projections of an object, the Inverse Radon Transform could be used to obtain a representation of the physical object; i.e., a reconstruction. Though it is not possible to obtain an infinite collection of views, an approximate and serviceable reconstruction can still be computed.

Over the past year we have significantly upgraded both our preprocessing and our reconstruction CT software. Good preprocessing of CT data is required both to convert the measured radiographic intensity data into absorption data, and to remove or minimize various non-linearities and non-idealities introduced in the intensity measurement process. The new preprocessing software (known as PRECON) eliminates a number of deficiencies, both in capability and human interface design, in the old preprocessing system. The current implementation of PRECON supports only 2-D detector arrays, but extension to 1-D detector arrays and single detector systems is straightforward, and greatly simplified by reuse of existing code and algorithms already developed for the 2-D portion.

Once the CT intensity data has been properly preprocessed, CT reconstruction software computes a digital attenuation map of the object. While the calculation of object attenuation at any one small, discretized volume element (or voxel) within the object is not an extensive calculation, the process must be repeated for millions of voxels. In the past, large reconstructions required days or weeks of computational time. We have addressed this issue with a three-pronged attack. First, we have procured a new high-speed computer system using the latest 64-bit microprocessor. Second, we have made use of symmetric parallel processing available on the new computer. And third, we have coerced existing high-speed graphics hardware into solving

major portions of the CT reconstruction problem. All these efforts have resulted in a significant reduction in CT reconstruction time.

As mentioned in previous reports, CT scanners can be classified by detector architecture: 1) single detectors; 2) linear-array detectors; and 3) area array detectors. It is the last type of scanner that holds the greatest promise for fast 3-D scanning[3]. One type of area imager that we have pursued is a CCD camera-to-scintillator arrangement, where the x rays are converted to visible light by the scintillator, and the camera captures the visible light output of scintillation. Within the last two years, our research has documented improvements in spatial resolution and contrast performance of recently developed glass scintillators. Consequently, higher performance and faster scanning can be achieved with better cameras.

Prior to this year, our single high-performance camera was a PHOTOMETRICS 14-bit CH250, with a 1024 x 1024 THOMPSON chip. This camera is included in the PCAT scanner and has extended our range of customer work. However, the constant supply of customer work has made it difficult to use the larger dynamic range to evaluate new scanning techniques. CRADA work in digital radiography has resulted in the purchase of two new cameras from AXIOM research, both are 1536 x 1024 in size, but vary in pixel depth: one is 12-bits and the other 16-bits deep per pixel. We have upgraded our CT scanning software to include these two cameras. We are at present developing new fixturing arrangements for high-energy scanning (4 to 9 MeV).

Progress

Computed Tomography Reconstruction Software Upgrade

The Nondestructive Evaluation Thrust Area (NDE) recently acquired an SGI Power Onyx computer system for speeding CT reconstructions. This Onyx contains two symmetric 75 MHz R8000 processors and a Reality Engine (RE^2) graphics co-processor system. Very few software modifications were required to our CT image-reconstruction software, RECON, to operate (in single processor mode) on the Onyx. **Table 1** summaries the performance increases we obtained in moving our backprojection codes to the Onyx. Once RECON was functional on the Onyx, the (workhorse) backprojection codes within RECON were upgraded to take advantage of the Onyx's SMP and RE^2 capabilities.

SMP Backprojection

CT backprojection algorithms readily lend themselves to parallel processing. In theory, a processor could be assigned to compute and sum the backprojections for each cell (pixel or voxel) in the reconstructed image/volume. Ignoring such SMP hardware complications such as memory bandwidth and cache contention, a fully parallelized backprojection code should exhibit a direct proportionality to the number of processors (i.e., double the number of processors, double the performance). In the case of our two-processor Onyx, we have been able to demonstrate a 2X performance increase in the backprojection step, over the uniprocessor performance.

We have also introduced parallel processing into the "filtering" and (for the fan and cone beam cases) "weighting" steps. Combined, we have nearly achieved a 2X performance increase over the uniprocessor operation, with only the I/O portion of the backprojection codes not parallelized. **Table 2** summarizes the performance of our SMP codes. Comparing the run times of the Standard vs Parallel algorithms (**Table 2**) actually shows a performance increase ratio slightly greater than 2.0. This stems from other code enhancements that were introduced in the process of recoding the algorithm for SMP execution. When they are removed, the performance ratio decreases to 2.0.

Table 1. Relative performance of various NDE computers, for a fan beam reconstruction of a 256 × 256 image slice from 256 rays and 384 projections.

Computer	CPU	Time (s)	Relative performance
Sun4	Sun4c	153	1.0
Sun4	Sun4m	85	1.8
SGI Iris	33 MHz R3000	58	2.6
SGI Indigo	132 MHz R4600	40	3.8
SGI Onyx	75 MHz R8000	21	7.3

Backprojection via Texture Mapping

The SGI Reality Engine (RE^2) is very fast at generating 3-D texture-mapped polygons. Texture-mapping is basically the process of placing an image onto a 3-D polygon, and then rendering that "decaled polygon" via a 3-D viewing transform. The process is the digital graphics equivalent to the well-known physical process of pasting a decal onto a model. Rates up to 320 Mpixels/s are possible with a RE^2 containing four Raster Manager (RM) boards. In addition, the RE^2 is equipped with an "accumulation buffer" where multiple images can be summed. These two capabilities can be exploited to great advantage by the CT backprojection algorithm. For parallel and fan beam geometries, the backprojection process can be viewed as mapping a 1-D texture (the "filtered" absorption or "ray sum" values) onto the reconstruction slice pixels within a polygon. For the cone beam geometry, the backprojection process can be viewed as a 2-D texture-mapping process[4] where a portion of the digital radiograph forms the texture image. The summing of the backprojections can also be rapidly accomplished within the RE^2 using its accumulation buffer, which was intended for simulating motion blur and depth-of-field effects. Next, the relative rotation between each backprojection can easily and quickly be accomplished using the RE^2's 3-D geometry engine before the backprojection polygons are rendered. Finally, the "weighting" of the backprojections, as required in the fan and cone beam reconstruction algorithms, can be closely approximated using the RE^2's "depth queuing" or "fogging" capability, by assigning appropriate z coordinates to the polygon vertices describing the backprojection region.

We have written a parallel beam, a fan beam, and a cone beam reconstruction code that uses these capabilities of the RE^2 via the OpenGL library.[5] These codes also optionally use SMP filtering of the ray sum data, and thereby take full advantage of our Onyx. Performance of these codes varies with reconstruction size: the larger the reconstruction, the larger the performance increase. This is due to two factors: the filtering of the ray sums is not accelerated by the RE^2, and the setup and I/O (overhead) time required to use the RE^2 remains constant as the problem size increases. Thus, for small reconstruction sizes, the "filtering" and RE^2 overhead times make up a significant fraction of the total reconstruction time. For larger problems, their fractional contribution is reduced, and the performance factor increases accordingly. **Table 2** shows the performance of the RE^2 fan beam (convolution backprojection) algorithm as the reconstruction size is varied. Note that our Onyx system has only two R8000 processors, and its RE^2 has only one RM board. We could nearly quadruple the RE^2 backprojection speed by purchasing three additional RM boards. We could also increase the speed of the SMP steps by adding R8000 processors to our Onyx.

Table 2. Performance of fan beam reconstruction software, implemented on a uniprocessor (standard algorithm), two-processor SMP machine (parallel algorithm), and on a two-processor SMP machine with SGI Reality Engine^2 (RE^2 algorithm).

Case	Reconstruction size	Number of rays	Number of projections	Algorithm	Run Time (s)	% RMS difference
1	128 x 128	128	192	standard	3	
1	128 x 128	128	192	parallel	2	0.0017
1	128 x 128	128	192	RE^2	1	7.1
2	256 x 256	256	384	standard	21	
2	256 x 256	256	384	parallel	9	0.0055
2	256 x 256	256	384	RE^2	4	5.2
3	512 x 512	512	768	standard	152	
3	512 x 512	512	768	parallel	68	0.011
3	512 x 512	512	768	RE^2	18	3.8
4	1024 x 1024	1204	1536	standard	1163	
4	1024 x 1024	1024	1536	parallel	509	0.037
4	1024 x 1024	1024	1536	RE^2	107	3.5

One drawback to using the RE^2 is limited precision. The RE^2 was designed for graphics applications, where fixed point, 12-bit arithmetic is fully adequate. This is not the case for CT reconstructions. Our codes mitigate the effects of the RE^2's 12-bit fixed point arithmetic by pre-scaling the data before it is fed to the RE^2. Further inaccuracy is introduced by approximating the fan beam and cone beam "weighting" factors using "fogging." The last column of **Table 2** gives the Root-Mean-Square (RMS) difference over all reconstructed image pixels for the SMP and RE^2 algorithm results, referenced against the standard, single processor algorithm results. For most work, the image differences introduced by the limited precision of the RE^2 are small, and near the level of the data acquisition noise.

Computed Tomography Preprocessing Software Upgrade

Preprocessing the measured, raw x-ray intensity data into linear absorption data is a critical step in the process of x-ray CT inspection. Preprocessing must remove or mitigate all the non-idealities of the x-ray source and x-ray detection systems, and then convert the measured intensity data into the corresponding absorption values so that it can be operated upon by CT reconstruction codes to produce an attenuation map of the scanned part. Prior to this year's development of the PRECON preprocessing software, the preprocessing step was the least automated step in the x-ray CT inspection process, shown in **Fig. 1**. Largely, preprocessing was a manual operation, using the VIEW[6] signal- and image-processing package occasionally combined with some other highly specialized, but poorly documented codes.

The PRECON software package was designed to provide a powerful, integrated, well-documented, CT preprocessing specific solution to the preprocessing step. PRECON reads raw CT (intensity) data and the corresponding CT description file, as produced by the various NDE CT scanners, and produces ray sum (absorption) files that are ready for processing by the RECON reconstruction codes. It also produces an updated description file which now includes all preprocessing operations used on the data, as well as the parameters of those operations files (**Fig. 1**).

A full description of PRECON is given elsewhere.[7] Below we touch on the highlights of PRECON's capabilities.

Detector Subregion Selection. It is often desirable to eliminate some detector pixels from the top, bottom, left, and/or right of the detector array, since they contain only border information (i.e., no view of the part was ever projected into these pixels), or they contain slice information either above or below a specific region of interest (e.g., extra detector lines at the top and/or bottom). PRECON provides a mechanism for handling these cases. Parameters can be adjusted to eliminate border pixels at either the left and/or right of the detector array.

Bad Detector Element (Pixel) Replacement Algorithm. It often happens that detector arrays have elements which always give spurious measurements. One of the best methods for dealing with such elements is to replace their measurements with an average of the measurements of its neighbors. This is precisely what is done by the Bad Detector Element Replacement algorithm. A list of bad elements is located in a data file, and those elements that fall within the acquired data have their measurements replaced by the average of the eight adjacent neighboring elements.

Nearest Neighbor Outlier Rejection. Spurious measurements are sometimes acquired from properly working detector elements. To address this problem, the Nearest Neighbor Outlier Rejection (NNOR) algorithm is provided. It computes the average and standard deviation of the measurements in the neighborhood (up to eight adjacent detector elements) about each detector element. The measurement deviation is defined as the absolute difference between the detector element's measurement and the neighborhood average. The deviation is compared to three user-controlled parameters, and if the deviation exceeds all three, the measurement is replaced by the neighborhood average. The NNOR algorithm can be selectively used on the dark current image, the incident intensity (or I_0) image, the raw intensity radiographic images, and the ray sum radiographs.

Incident Intensity Exceedance Outlier Replacement. Raw intensity radiographic measurements that exceed the incident intensity (or I_0) measurement are theoretically impossible. In

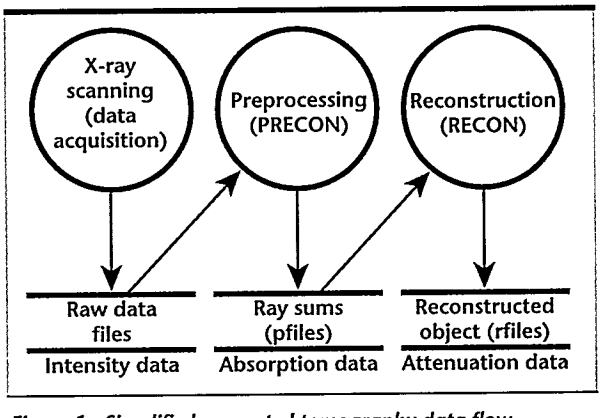


Figure 1. Simplified computed tomography data flow.

practice, this occurs largely for two reasons: 1) measurement noise, and 2) x-ray hits directly upon the light sensitive CCD detector.

The Incident Intensity Exceedance Outlier Replacement algorithm attempts to deal with the problem of x-ray hits on the detector. When a raw intensity radiographic measurement exceeds the incident intensity value by more than a user-controlled minimum percent, the measurement is considered an outlier, and replaced by the average of its adjacent neighbors, up to eight. If the measurement is very near the full scale value of the detector, then an additional algorithm is used. This algorithm looks for charge "spill" into the adjacent detector elements, as has been empirically observed on CCD detectors. Adjacent detectors that exhibit "spill" are also replaced by the average of their non-affected adjacent neighbors.

Thick Scintillator Blur Deconvolution. The conversion of x rays into visible light for detection using a CCD camera is accomplished by a specially formulated glass scintillation material, typically a doped glass. NDE often uses scintillation plates of significant thickness to obtain high sensitivity. However, these thick scintillators cause a roughly Lorentzian blurring of the radiographic signal transmitted to the CCD. To eliminate this blurring effect, it is necessary to deconvolve the CCD measured radiograph. It has been determined empirically that deconvolving with a Lorentzian works quite well. It has further been found that deconvolving in just the x dimension is adequate in most instances. The deconvolution process requires two parameters: the Lorentzian blur width in detectors (pixels), and an estimate of the noise-to-signal ratio. The values of these parameters will depend on factors such as the scintillator thickness, field of view, and noise characteristics of the CT system.

Radiograph Resampling. In some instances, it is desirable to resample that section selected via the preprocessing parameters of the radiographs. Accordingly, a full resampling algorithm is included in PRECON. This resampling algorithm preserves the full range of the data, unlike decimation algorithms which can drop the last few point(s), thereby altering the range. There are a number of resampling filters available including triangular, box, bell, cubic, B-spline, Lanczos, and Mitchell filters.

Beam-Hardening Correction. Beam-hardening is caused by the use of a polychromatic x-ray source and the higher attenuation coefficient of materials for lower energy x rays. These two effects result in a non-linear variation of absorption with x-ray path length. The ray sum radiographs may be corrected for this beam-hardening. The correction is implemented as a cubic polynomial applied to each ray sum value.

Ring Removal. Ring artifacts in reconstructed objects are caused by nonuniform, nonlinear detector-to-detector response. The best way to eliminate this problem is to calibrate each of the detectors in the 2-D array, i.e., each pixel of the CCD. However, this process is expensive, time consuming, and difficult. PRECON provides a heuristic alternative that works well in practice by correcting each detector pixel value, according to offsets determined directly from the radiographic ray sum data.

Radiograph Normalization. It can easily be shown that the total absorption, i.e., the sum of all the ray sums, in any radiograph of an object (where the whole object is within the radiographic region) is a constant, regardless of the viewing angle. This fact provides the basis for normalizing a series of radiographs for source and other temporal variation. Accordingly, each radiograph is scaled so that its total absorption summed over all detector pixels is the same.

CT Data Acquisition Camera Upgrade

Modern CCD area imagers can acquire an entire digital radiograph during a single integration period. As indicated above, improvements in high-density, high-Z, glass scintillators have created the situation where a better camera equates to a better scanning capability. Digital camera technology is a particularly vibrant area at this time and improvements in digitizing depth and read-out speed are occurring daily. To upgrade our CT data acquisition capabilities, we undertook two specific tasks: 1) to modify the VIDCT software, which controls the data acquisition for all of the area detector scanners, (MDCAT, HECAT, DCAT, and PCAT), to enable a better fit with automated CT processing and reconstruction tools, and 2) to integrate the newer cameras into our scanning software for use with any of our seven x-ray sources. The result is a faster CT scanning capability that will include the benefit of two new cameras.

The uncollimated beam from a typical x-ray machine source emerges from the source spot in the shape of a cone. CT reconstruction algorithms must account for the shape of the diverging beam in order to be quantitative. A number of additional features were added to the VIDCT software to measure and record the state of the beam divergence for the area detector based systems. These directives are used by PRECON and RECON to perform appropriate CT preprocessing and reconstruction in a more timely

fashion. **Figure 2** contains images of Geysers Rock scanned with a 2.5° cone, using the improvements in PCAT resulting from this upgrade work. These 4-in. core samples were inspected with radiography and CT within 8 h of extraction from the digging site. The papers and presentations connected to this data have received awards at conferences.

In the course of the CRADA work for Chrysler, a number of cameras were purchased for evaluating different digital radiographic techniques for inspecting raw aluminum castings. Two of these cameras were slow-scan, Peltier cooled, large format 1536 × 1024, CCD cameras with grade 0 CCD chips (no-defects) from AXIOM Research. One includes 12-bit digitization with fast full frame read-out; the other includes 16-bit digitization, but with a slower read-out for the full frame of data. The 12-bit unit is being extensively evaluated and used at Chrysler for inspecting transmission cases on a regular basis at the Kokomo plant. Both cameras are operated with the same low-level software from AXIOM. We can now control either camera from the VIDCT software, and are building the necessary fixturing and shielding for coupling these cameras with the 4- and 9-MeV LINACS. **Figure 3** is an acquired image of a test pattern with the AXIOM 16-bit camera.

Future Work

We would like to run our SMP codes on a SGI system using more than two processors, to verify the direct proportionality between processor count and performance. In addition, PC graphics hardware is beginning to become available that provides hardware acceleration of both 1-D and 2-D texture-mapping, and is packaged with an implementation of the OpenGL library. Such hardware could prove to be very cost-effective for medium-size reconstructions, and this possibility deserves exploration.

The PRECON software needs to be upgraded to include the preprocessing of 0D (single) detector system and 1-D (linear array) detector system data. In addition, early usage of PRECON has provided feedback on numerous minor human interface improvements that should have a high return on investment. Finally, a new code permitting interactive preprocessing of ray sum data, coupled with the parallel and fan beam reconstruction algorithms, would greatly speed the determination of some of the more critical preprocessing parameters. We expect to re-use a significant fraction of the PRECON code in this endeavor.

In the past two years we have established the value of higher-dynamic range cameras and higher performance scintillators for performing digital radiography and 3-D CT. A number of manufacturing inspection problems are emphasizing the importance of high-energy, high-spatial resolution 3-D imaging. This is also an important goal for fully automating the inspection of certain weapons components. We will use our enhanced capability in area detector imaging to extend the limits of spatial resolution for digital radiography and 3-D CT in the 4- and 9-MeV range.

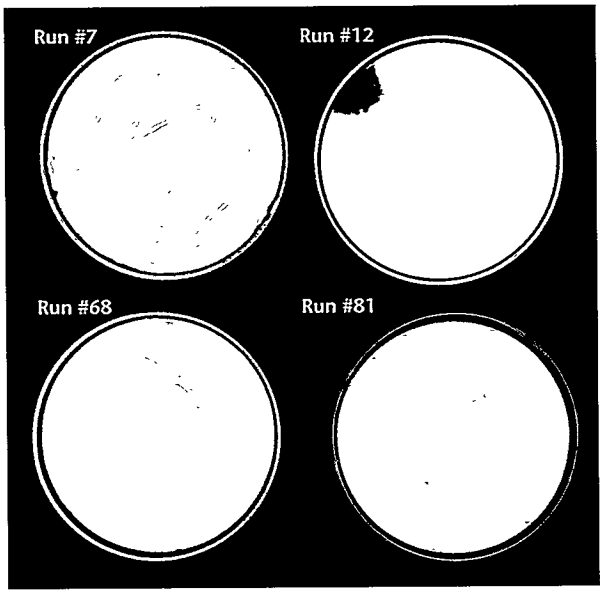


Figure 2. CT reconstructions of Geysers Rock core samples from UNOCAL Geothermal Plant, from different depths.

Acknowledgments

The authors would like to acknowledge the valuable contributions of S. Azevedo (Electronics Engineering Department), and H. Martz (Mechanical Engineering Department).

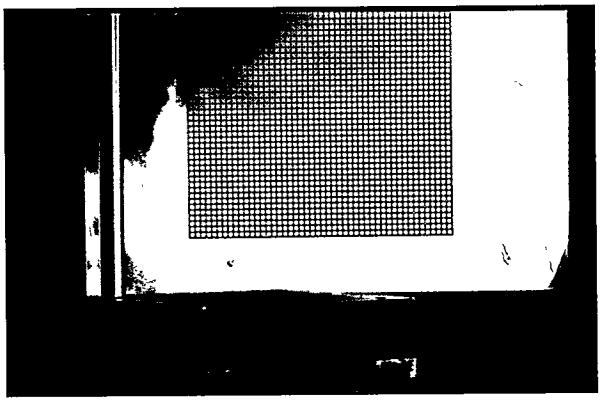


Figure 3. Visible light image of test pattern acquired with AXIOM 16-bit 1536 x 1024 camera, using Lawrence Livermore National Laboratory VIDCT software.

References

1. Herman, G. T. (1980), *Image Reconstruction From Projections*, Academic Press, New York, N.Y.

2. Radon, J. (1917), *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten (On the Determination of Functions from their Integrals Along Certain Manifolds). Berichte Sächisische Akademie der Wissenschaften, Leipzig, Germany, Math.–Phys. Kl.*, ***69***, pp. 262-267.

3. Georgeson, G. E., A. R. Crews, and R. H. Bossi (1992), *X-ray Computed Tomography for Casting Development*, Wright Patterson Air Force Base, Dayton, OH, (WL-TR-92-4032), September, 1992.

4. Azevedo, S. G., B. K. Cabral, and J. Foran (1994), "Tomographic Image Reconstruction and Rendering with Texture-Mapping Hardware", *Proc. SPIE, Mathematical Methods in Medical Imaging III Conf.* (San Diego, California, July 24–29); also Lawrence Livermore National Laboratory, Livermore, Calif., (UCRL-JC-117450).

5. Segal, M., and K. Akeley (1992), *The OpenGL Graphics Interface*, Silicon Graphics Computer Systems, 2011 N. Shoreline Blvd., Mountain View, Calif. 94039.

6. Brase, J., J. Phillips, H. Szoke, and M. Weiting (1994), *View User Manual*, Lawrence Livermore National Laboratory, Livermore, Calif., June 15, 1994.

7. Brand, H. R., S. G. Azevedo, M. R. Gorvad, D. J. Schneberk, H. E. Martz, and G. P. Roberson (1995), *Computed Tomography Software and Standards: Revision 2.1*, Lawrence Livermore National Laboratory, Livermore Calif., NDE internal documentation, July 31, 1995.

Applications of Micropower Impulse Radar to Nondestructive Evaluation

Jeffrey E. Mast
Defense Sciences Engineering Division
Electronics Engineering

Stephen G. Azevedo
Laser Engineering Division
Electronics Engineering

We have developed a nondestructive evaluation (NDE) tool, based on microwave impulse radar (MIR) technology and image processing codes, developed at Lawrence Livermore National Laboratory (LLNL), which provides an integrated platform for data acquisition and imaging. MIR technology provides a cost-effective and highly portable radar which is suitable for interrogating radiation in the microwave regime. The system is well suited for several applications, and we have demonstrated promising results for NDE applications in the areas of civil structures, buried utilities, and nonproliferation. In this report we discuss potential applications for which the system could be adapted through advancement of the imaging algorithms and modification of the MIR frequency range.

Introduction

NDE with microwaves has been attempted in the past for a small subset of materials and applications. Narrowband, i.e., single frequency or continuous wave (CW), radar can penetrate certain materials based on the dielectric constant and scatter from internal voids or inclusions. In the same way that ultrasound measurements of reflected acoustic energy provide information about the object, radar return values contain useful data about the dielectric properties of the object. In general, the imaging resolution of a device is on the order of the radiated wavelength in the material. **Figure 1** shows the electromagnetic (EM) spectrum and the wavelengths in air. Since microwaves have much longer wavelengths than x rays, for example, or even visible light, one would not expect to achieve the resolution available through those kinds of modalities. However, there is a class of materials, e.g., concrete or wood, for which microwaves may produce new information unavailable by other methods. Our work explores applications where this may be the case.

LLNL engineers have developed and patented an ultra-wideband radar device, called micropower impulse radar (MIR), that can be used for microwave NDE. It is a unique device that is small, safe and inexpensive to reproduce. It uses low power, yet produces clean microwave signals in the 1 to 4 GHz range. The wideband nature of the signals provides additional information that can improve the imaging and detection capability of microwave devices. The small size and low power make MIR suitable for array applications as well. It is this unique MIR device and a custom imaging system that we have used for our initial studies on microwave NDE. (Other MIR applications are described in the "Power Conversion Technologies" section of this book, and in numerous other reports.[1–7])

In our first year, we have assembled a prototype radar, a computer data acquisition system, and imaging software that can collect data in the field from a variety of samples. With this system, we have determined a class of NDE problems for which microwaves in the 1 to 4 GHz range are appropriate. In this report, we discuss our system, and propose new directions for microwave and millimeter-wave radar transceivers that will make them more widely usable in NDE.

Progress

In FY-95, we accomplished the first steps in integrating the data acquisition system and imaging algorithms into an imaging system. The system is

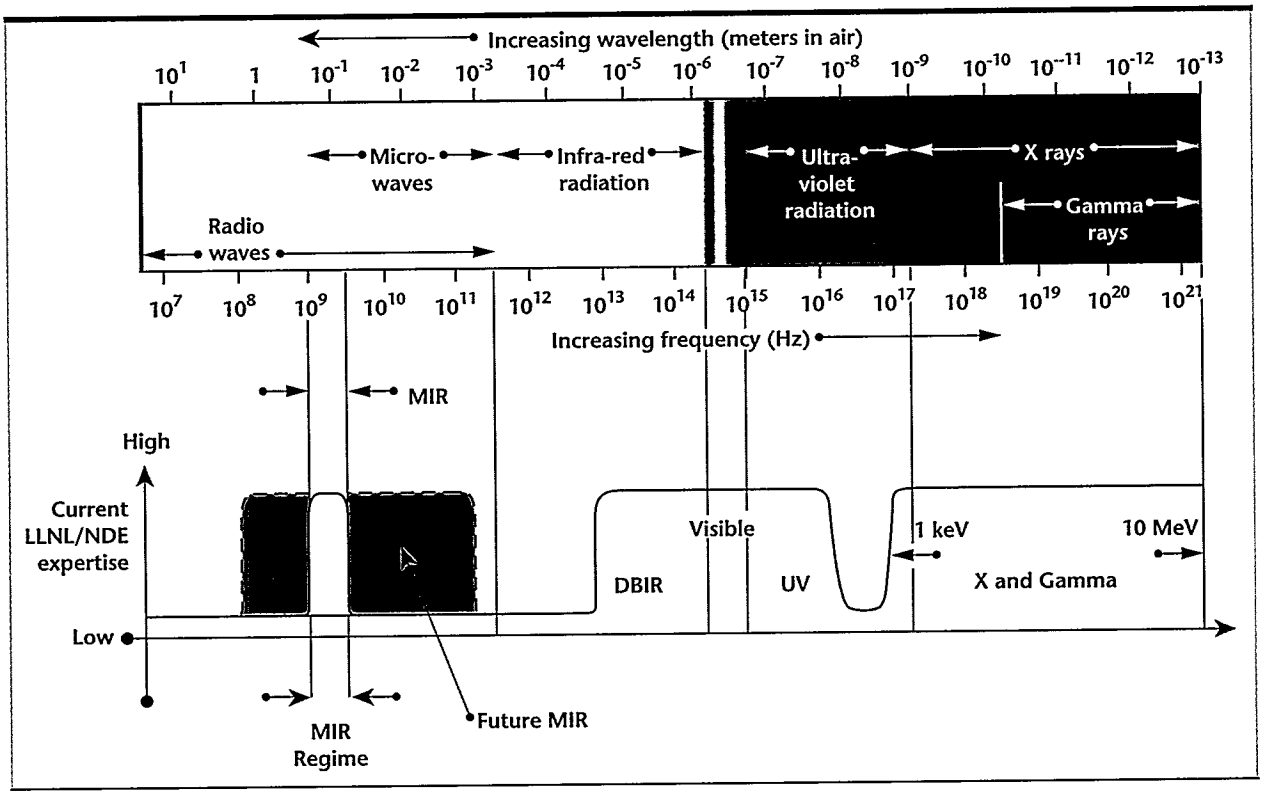


Figure 1. Electromagnetic spectrum, showing regions useful for NDE.

Table 1. Micropower Impulse Radar rangefinder imaging module characteristics.

General	
Module Configuration	Monostatic
Overall Dimensions	5 cm (height), 10 cm (width), 10 cm (depth)
Weight	< 0.23 kg
Input Power	6 to 12 V, 8 mA
Range (Maximum)	Up to 10 m in air; 0.2 to 0.5 m in soils (depends on soil conductivity and standoff range)
Transmitter	
Type	Ultra-wide bandwidth, impulse waveform
Pulse Repetition Frequency	2 MHz
Pulse Rise-time (Radiated)	~100 ps
Bandwidth	3.2 GHz (from ~0.8 to 4.0 Ghz)
Radiated Power (ERP)	< 100 mW, peak; < 10μW, average
Antenna:	Cavity-backed, resistively-loaded monopole
Beamwidth (-3 dB)	120°
Receiver	
Front End:	
Type	Equivalent-time sampler with sensitivity-time control
Bandwidth	<4.0 GHz
Noise Figure	25 dB
Antenna:	Cavity-backed, resistively-loaded monopole
Beamwidth (-3 dB)	120°

portable, with an integrated MIR positioning device providing a unique method for acquiring data suitable for imaging. Using this system, we were able to demonstrate the potential uses of MIR as a microwave NDE tool.

Data Acquisition and Imaging System

The integrated radar system consists of four components: a MIR rangefinder, a MIR positioner, a laptop computer, and imaging software. The MIR rangefinder is a swept-range gate impulse radar with specifications as shown in **Table 1**. The rangefinder operates in monostatic mode, transmitting a pulse into the medium under evaluation and receiving echoes from within. To perform 2-D synthetic aperture imaging, the MIR T/R unit is scanned in a linear sweep over the surface of the medium. Pulses must be transmitted and received at equally spaced intervals across this synthetic aperture. To facilitate this, another MIR unit, fixed with respect to the medium, acts as a positioning device, recording the location of the MIR T/R unit. This acquisition technique, (patent pending [IL-9719]), provides a flexible, highly portable, and accurate means for synthesizing the aperture required for imaging. Integrated software on the computer performs the data acquisition and the signal and image processing, and displays the computed image. The positioning device was also adapted to work with a previously developed mobile testbed.[8] The integrated system provided a 100-fold increase in data acquisition speed over our previous capability. This mobile system also provides a viable method for portable 3-D synthetic aperture imaging.

The acquired data is directly fed into the imaging software which was adapted from previously developed microwave imaging algorithms.[9,10] The entire

Table 2. Microwave NDE applications, showing materials, the need for evaluation, and the desired microwave parameters to perform the evaluation.

Application	Material(s)	Defect/ Evaluation	Frequency band(*)	Desired resolution	Desired penetration
Civil structures	dry concrete, granite, masonry	void, delamination, cracks, configura-, tion reinforcement	1 to 10 GHz	< 3 to 5 cm	> 20 cm
Buried utilities	earth, sand, soil, rock	location of buried utilities, (e.g., power lines, water/sewer pipes)	100 MHz 4 GHz	< 10 cm	> 20 cm
Shuttle tiles	aluminum, carbon composite, silica, silicone, felted nomex fiber	delaminations between tile bonding sheet, and aluminum or carbon composite	30 to 90 GHz	< 1 mm	< 10 cm
Utility poles	wood	detect water damage, rot, weaknesses	1 to 10 GHz	< 3 to 5 cm	30 to 40 cm
Non-proliferation	wood crate, other shipping materials	identify contraband, weapons in cargo	1 to 10 GHz	< 3 to 5 cm	> 30 cm
Waste management	barrels	determine rust deterioration beneath paint	30 to 90 GHz	< 1 mm	< 1 mm

(*) Frequency band offering most potential for desired penetration depths.

process, from data acquisition to display of the computed 2-D image takes only a couple of minutes, using a 486DX4 100 MHz laptop computer.

Potential Microwave NDE Applications

The current 1 to 4 GHz MIR rangefinder is capable of obtaining resolutions on the order of several centimeters, depending on the dielectric material under evaluation. **Table 2** shows several potential microwave NDE applications that we are evaluating. Note that the current MIR rangefinder is not suitable for some of the applications. However, a higher microwave frequency may provide resolutions that will provide adequate defect analysis. Note that the exact obtainable penetration depths ultimately depend on the power output of the radar and the conductivity of the materials.

Figure 2 shows a 3-D rendered image of a sewer pipe buried 25 cm beneath a reinforced concrete driveway. Using the data acquisition system, we acquired 108×126 spatial samples, and 512 temporal samples. The data was processed to produce a $108 \times 126 \times 80$ (1.4 m $\times$ 2 m $\times$ 0.3 m) spatial image. The entire data acquisition and imaging process of this size takes less than one hour using the current imaging system. To obtain greater penetration depths for locating deeply buried utilities requires modifying the MIR to launch frequencies below 1 GHz. However, the current MIR system provides adequate resolution and penetration capabilities for a wide variety of buried utility applications.

The current MIR is not well suited for several other potential microwave NDE applications. These applications require a frequency band different from 1 to 4 GHz, due to the nature of the defect and the materials under evaluation. For example, microwave NDE has the potential for detecting debonding between space shuttle tiles and the spacecraft structure. The debonds are on the order of 25 μm. Current MIR technology is not capable of identifying defects of this size. The MIR has a wavelength in air of about 7.5 cm, and positional stability in the radar pulse is better than 1 mm, but is certainly not in the micron range. A radar in the 30 GHz range or greater, with stability on the order of a micron (femtosecond), is required to locate these flaws.

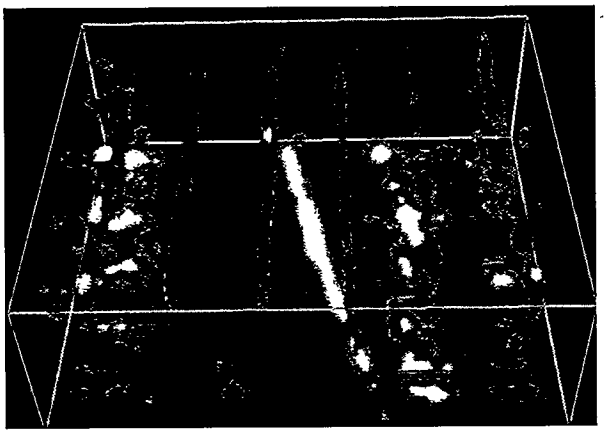


Figure 2. Microwave NDE image of subsurface beneath a concrete driveway, demonstrating the application of buried utility location. The rebar in the concrete is visible inside the concrete, and the sewer pipe is visible at a depth of ~25 cm.

Future Work

Future work mainly involves expanding our microwave NDE capability and extending the application base. As **Table 2** indicates, several identified applications are outside the current MIR frequency band of 1 to 4 GHz. Other work directly affecting the radar is antenna development. Different applications, imaging geometries, and frequency ranges require a variety of types of antennas for optimal pulse launching and coupling to the object under evaluation.

Other limitations of the current system preclude it from being used in many NDE scenarios, due mainly to geometry and apparatus constraints. The current system can only handle planar geometries imaged in monostatic reflection mode from a single surface. These constraints make it very difficult to image the interior of circular objects, such as utility poles, or take advantage of situations in which access from more than one surface or side of the object is available. Expanding our MIR imaging system to include the hardware and apparatus capabilities to do multistatic transmission mode geometries will greatly increase the scope and type of objects which we can evaluate.

Also, we currently lack the signal processing algorithms required to process the data from these expanded modes of operation. Additional work is required to develop, implement, and integrate multistatic and transmission mode imaging algorithms.

Acknowledgments

We would like to thank T. McEwan, E. Johansson, S. Nelson, and J. Warhus for their efforts in support of this project.

References

1. Bolomey, J.-Ch., and Ch. Pichot (1990), "Microwave Tomography: From Theory to Practical Imaging Systems," *International Journal of Imaging Systems and Technology*, Vol. **2**, pp. 144–156.

2. Devaney, A. J. (1983), " A computer simulation study of diffraction tomography, " *IEEE Trans. Biomed. Eng.* **30**, pp. 377–386.

3. Pichot, Ch., L. Jofre, G. Peronnet, and J.-Ch. Bolomey (1985), "Active microwave imaging of inhomogeneous bodies," *IEEE Trans. Antennas Propagat.* **33**, pp. 416–425.

4. Chommeloux, L., Ch. Pichot, and J.-Ch. Bolomey (1986), "Electromagnetic modeling for microwave imaging of cylindrical buried inhomogeneties," *IEEE Trans. Microwave Theory Tech.* **34**, pp. 1064–1076.

5. Caorsi, S., G. L. Gagnani, and M. Pastorino (1989), "Image Reconstruction by inverse scattering solution for multiple microwave views," *Proceedings URSI Int. Symp., Stockholm, Sweden*, August 1989, pp. 590–592.

6. Bolomey, J.-Ch. (1990), "Microwaves and their applications in analysis and in quality control of wooden pieces," *First European Wood Symposium, Paris, France*, May 1990.

7. Hasted, J. B., and M. A. Shah (1964), "Microwave absorption by water in building materials," *Brit. J. Appl. Phys.* **15**, pp. 825–836.

8. Warhus, J. P., S. D. Nelson, and J. E. Mast (1995), "Advanced Imaging Radar for Bridge Deck Inspection," *Engineering Research, Development and Technology FY94, Thrust Area Report*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-94).

9. Johansson, E. M., and J. E. Mast (1995), "Ultra-Wideband Radar Imaging for the Nondestructive Evaluation of Bridges," *Engineering Research, Development and Technology FY94, Thrust Area Report*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-94).

10. Johansson, E. M., and J. E. Mast (1994), "Imaging Algorithms for Synthetic Aperture Ultra-Wideband Radar," *Engineering Research, Development and Technology, FY93 Thrust Area Report*, Lawrence Livermore National Laboratory, Livermore, Calif. (UCRL 53868-93).

David N. Hunt, Thrust Area Leader

The Information Engineering thrust area develops information technology to support the programmatic needs of Lawrence Livermore National Laboratory (LLNL).

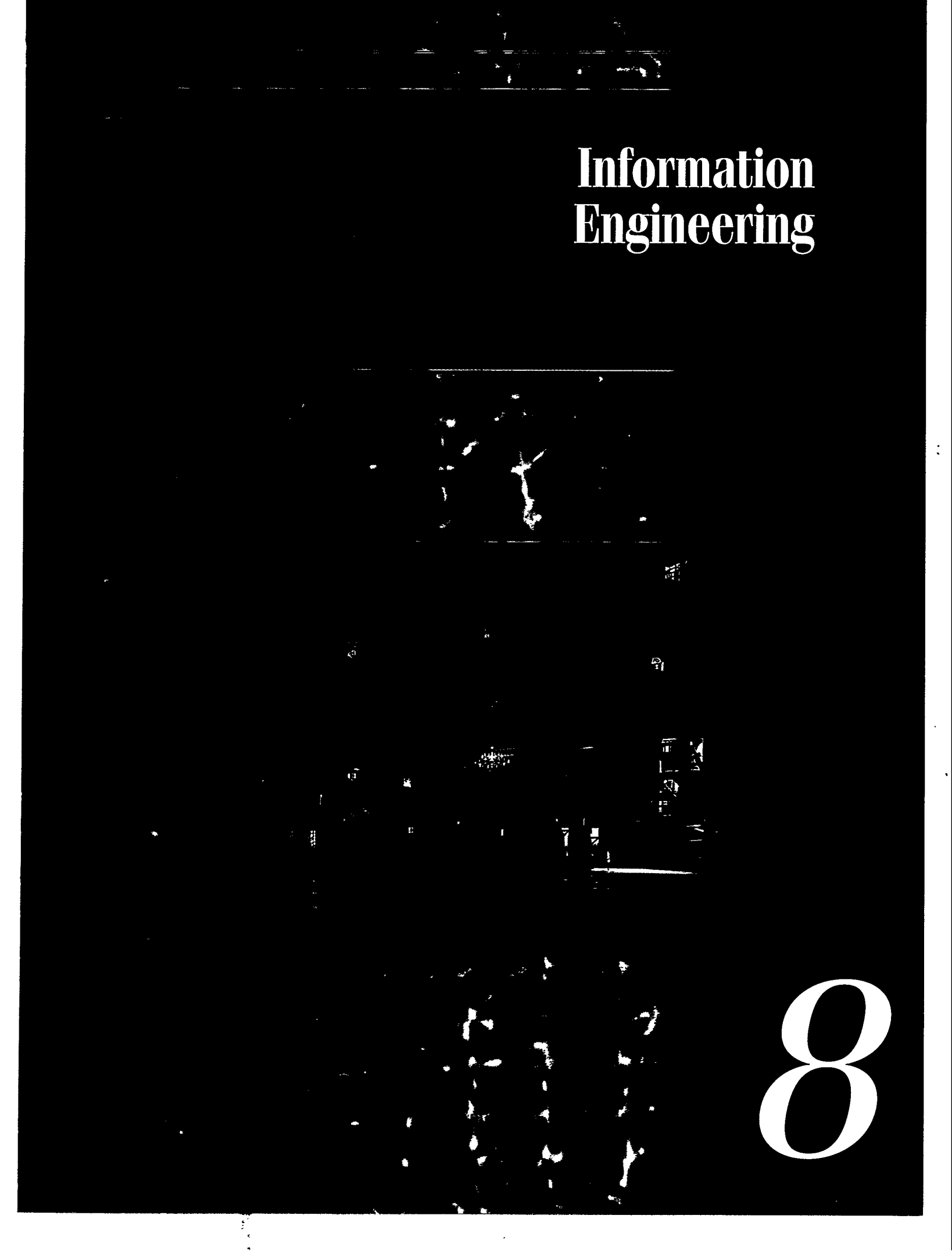
The mission of the thrust area is to maximize the information bandwidth between the system and its users. A major need to fulfill this mission is intelligent access for the user to the information served by the system. This intelligent access will require advances in compute platforms, networks, user 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. The thrust area is focused on the development and deployment of such information 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 thrust areas at LLNL, as well as with the industrial and academic sectors. The second major component of our 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 also increase the efficiency of day-to-day operations. In these days of fiscal austerity, the proper application of information technology can conserve precious programmatic funds. Savings of both time and money can be realized through investment in this area.

In FY-95, our first year of operation, we pursued projects that addressed both of the major components in the thrust area. Our activities for FY-95 are described in the following four sections: 1) Using the World-Wide Web as a Medium for Concurrent Engineering and Rapid Prototyping; 2) Subcarrier Multiplexing for Fiber-Optic Networks; 3) The World Wide Web As a Medium for Presenting LLNL Training Courses; and 4) Design and Implementation of 3-D MIME Data Access and Exchange.

Information Engineering



8

8. Information Engineering

Overview
David N. Hunt, Thrust Area Leader

Using the World-Wide Web as a Medium for Concurrent Engineering and Rapid Prototyping
Cecil W. Jordan, William A. Niven, Rudy E. Cavitt, George Pavel, Goffril Obegi, Fred E. Warren, C. James Buchanan, and Janine M. Taylor 8-1

Subcarrier Multiplexing for Fiber-Optic Networks
Paul D. Sargis, Ronald E. Haigh, and Kent G. McCammon 8-7

The World-Wide Web as a Medium For Presenting LLNL Training Courses
Darrel Lager and Sue Koopman 8-11

Design and Implementation of 3-D MIME Data Access and Exchange
Scott D. Nelson, Jeffrey S. Kallman, and Eric Brugger 8-17

Using the World-Wide Web as a Medium for Concurrent Engineering and Rapid Prototyping

Cecil W. Jordan, William A. Niven, and Rudy E. Cavitt
Laser Engineering Division
Electronics Engineering

George Pavel
Computer and Communication Engineering Division
Electronics Engineering

Goffril Obegi
Mechanical Engineering Department
Lawrence Berkeley National Laboratory
Berkeley, California

Fred E. Warren and C. James Buchanan
Computations Organization

Janine M. Taylor
Computer Sciences Department
University of California Davis
Davis, California

We are building a Concurrent Engineering Rapid Prototyping System (CERPS), which integrates engineering design, fabrication, procurement, and electronic commerce-linked vendors and suppliers for the rapid prototyping of electronic components and systems. Major benefits have accrued by speeding engineering development, facilitating multilaboratory and industry collaboration, reducing procurement time, and reducing overall costs. Programs at Lawrence Livermore National Laboratory (LLNL) and their industrial partnerships will benefit as well from the increased efficiencies by implementing engineering and commerce on the Internet.

Introduction

This project leverages three years of work funded by the Department of Energy (DOE) and a decade of work funded by the Department of Defense. Lawrence Berkeley Laboratory's WWW-based computer-aided design system, the Air Force's Computer Aided Logistics Program, the Air Force's Government Acquisition Through Electronic Commerce (GATEC) System, and the Technology Reinvestment Program's CommerceNet are incorporated in this work[1-9].

The CERP System integrates and leverages existing legacy computer-aided design systems and networks. The system makes use of commercial software and hardware (RSA security enhanced electronic mail, and Myriad workstation viewer), standard network technology (Internet, the World Wide Web, TCP/IP), and legacy computer-aided engineering design systems, such as Hewlett Packard EGS and Autocad.

Use of these technologies enabled rapid development of the CERP System, continuing return on the large investment in the existing design systems, and minimal connect cost for users. Users throughout the country can conveniently and cost effectively make use of this system.

Concept

The CERP System concept is depicted in **Fig. 1**. The system interconnects all users and providers to carry out the engineering development cycle

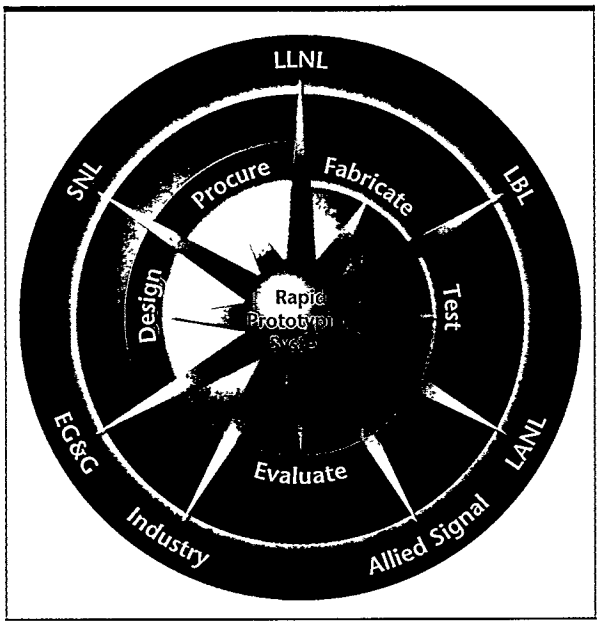
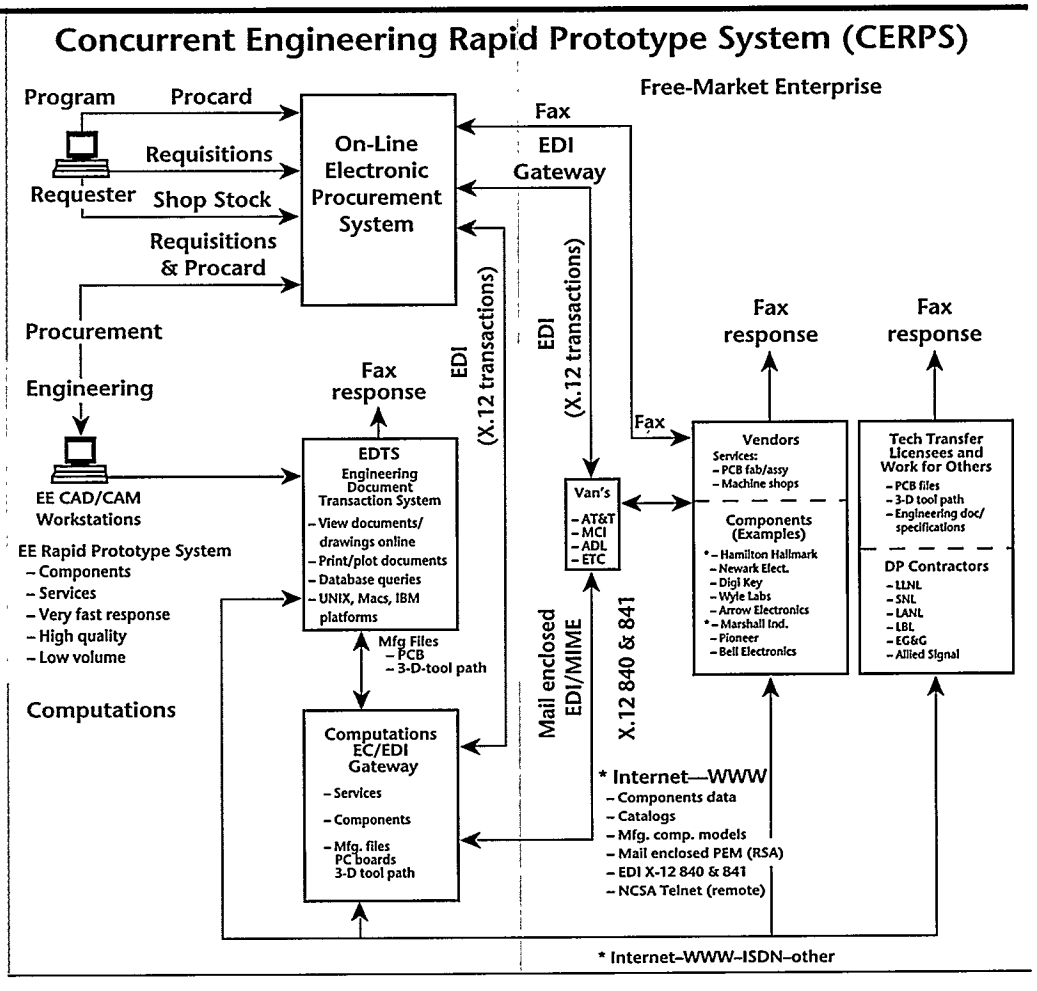


Figure 1. Illustration of Concurrent Engineering Rapid Prototyping System concept.

(design, procure, fabricate, test and evaluate) in a more highly integrated and efficient manner, independent of their geographic location.

Design, analysis, and simulation tools facilitate rapid, accurate and verifiable designs. Electronic commerce ties LLNL purchasing and financing, commercial vendors, and the LLNL end user together, to facilitate a very rapid procurement and delivery cycle which meets all DOE purchasing requirements. Computer-aided manufacturing tools and electronic commerce link industrial suppliers for rapid parts selection and prototype manufacturing. Electronic mail enclosures allow exchange of engineering drawings, specifications, project planning information and administrative information. Electronic privacy provisions assure confidentiality of transactions and protection of proprietary and contractual information. Electronic administration provides the record keeping and audit trails required by LLNL programs, industrial partners, and DOE.

Figure 2. Functional block diagram of Concurrent Engineering Rapid Prototype System.



With a centralized, sharable data source, multiple development partners are assured of accurate and up-to-date information and drawings.

Functional Description

The CERP System is a distributed and networked system. It is depicted in the functional block diagram shown in **Fig. 2**. Central to the system is the Engineering Document Transaction System (EDTS) which serves as a central focus for networked workstation client/server interaction. It provides the file conversion resources and database resources required to perform the multiplicity of engineering document transactions and data storage involved in designing, procuring, fabricating, testing and evaluating processes. It supports engineering documents of all types, such as drawings, printed wiring board, numerically controlled machine manufacturing files, specifications and spreadsheets. The Computations Commerce Gateway provides the interface with the Internet, linking designers with other designers, vendors, parts suppliers and fabrication sources, allowing rapid access and interaction for engineering information. The On-line Electronic Procurement System provides the procurement interface with vendors, parts suppliers and fabrication sources via the Internet.

EDTS is shown in the functional block diagram in **Fig. 3**. Central to EDTS are relational database resources such as ORACLE, INFORMIX, and ALLBASE, to store and retrieve engineering documents in the form of images, text, or CAD drawings.

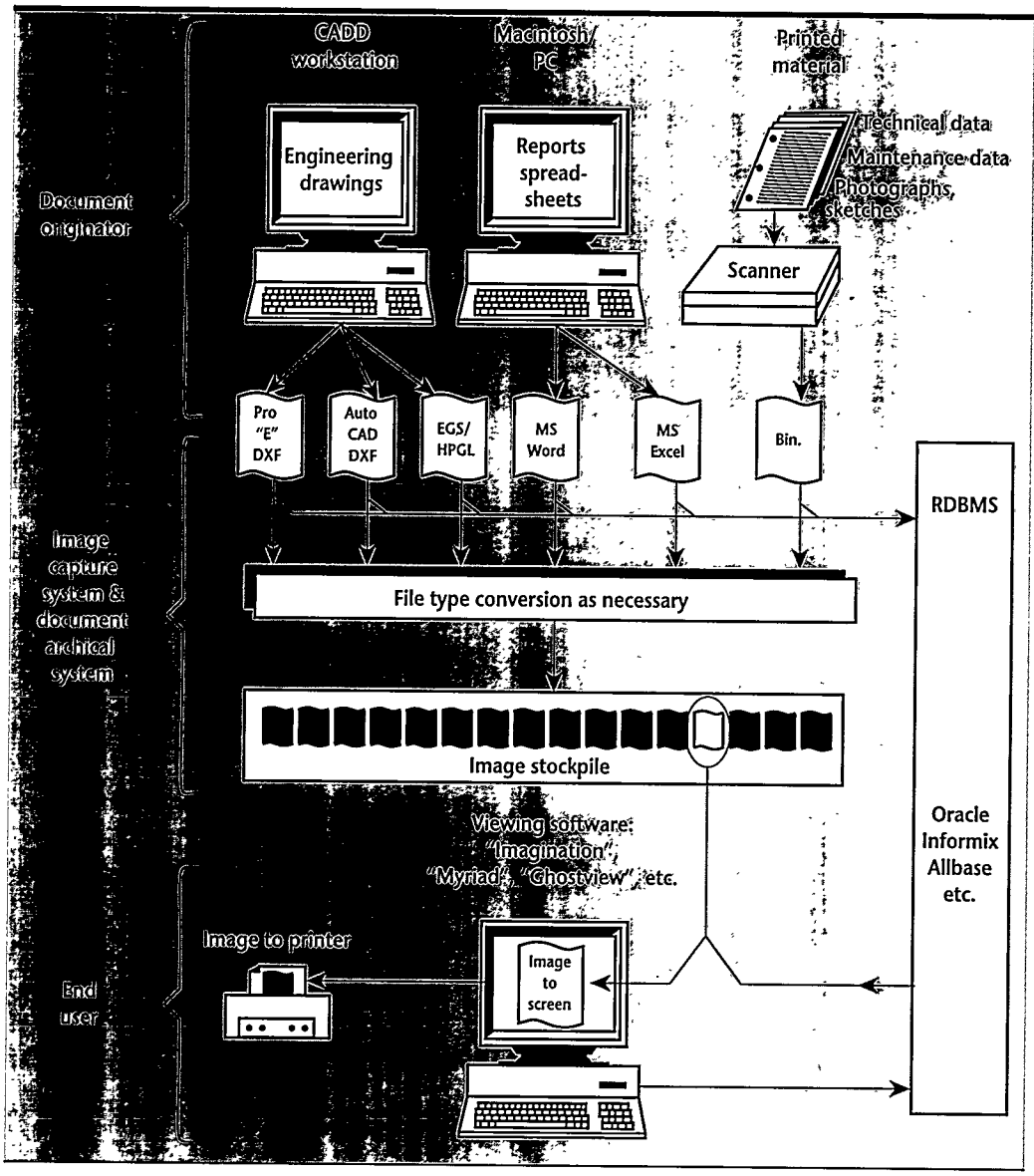


Figure 3. Functional description of Engineering Document Transaction System.

File type conversion software performs image, text and CAD file conversions, including PRO "E" DXF, AUTO CAD, EGS/HPGL, MS WORD, and MS EXCEL, all invisible to the user. Commercially available NetScape browser interfaces integrate file conversion and database transactions to greatly simplify use of the system.

Figure 4 shows the existing EE CAD/CAM workstation network that is supporting LLNL programs such as NOVA, ICF, and AVLIS; the Magnetic Fusion Program and the Environmental Program's Mixed Waste Management Facility. All areas are experiencing the benefits of being able to share files, level workload, increase productivity and being more highly integrated into the engineering development processes.

Progress

During this fiscal year the CERPS project implemented an electronic documentation distribution system which included the installation of the Electronic Document Retrieval System (EDRS) developed at Lawrence Berkeley Laboratory, a Xinet print engine that allows the printing and plotting of documents at any TCP/IP print device and the loading of the EDRS with documents drawn from various sectors of LLNL, LANL, and Sandia National Laboratory. This system was demonstrated to the DOE Office of Economic Competitiveness, and has been in constant operation since its inception, providing immediate access to a wide variety of electronic

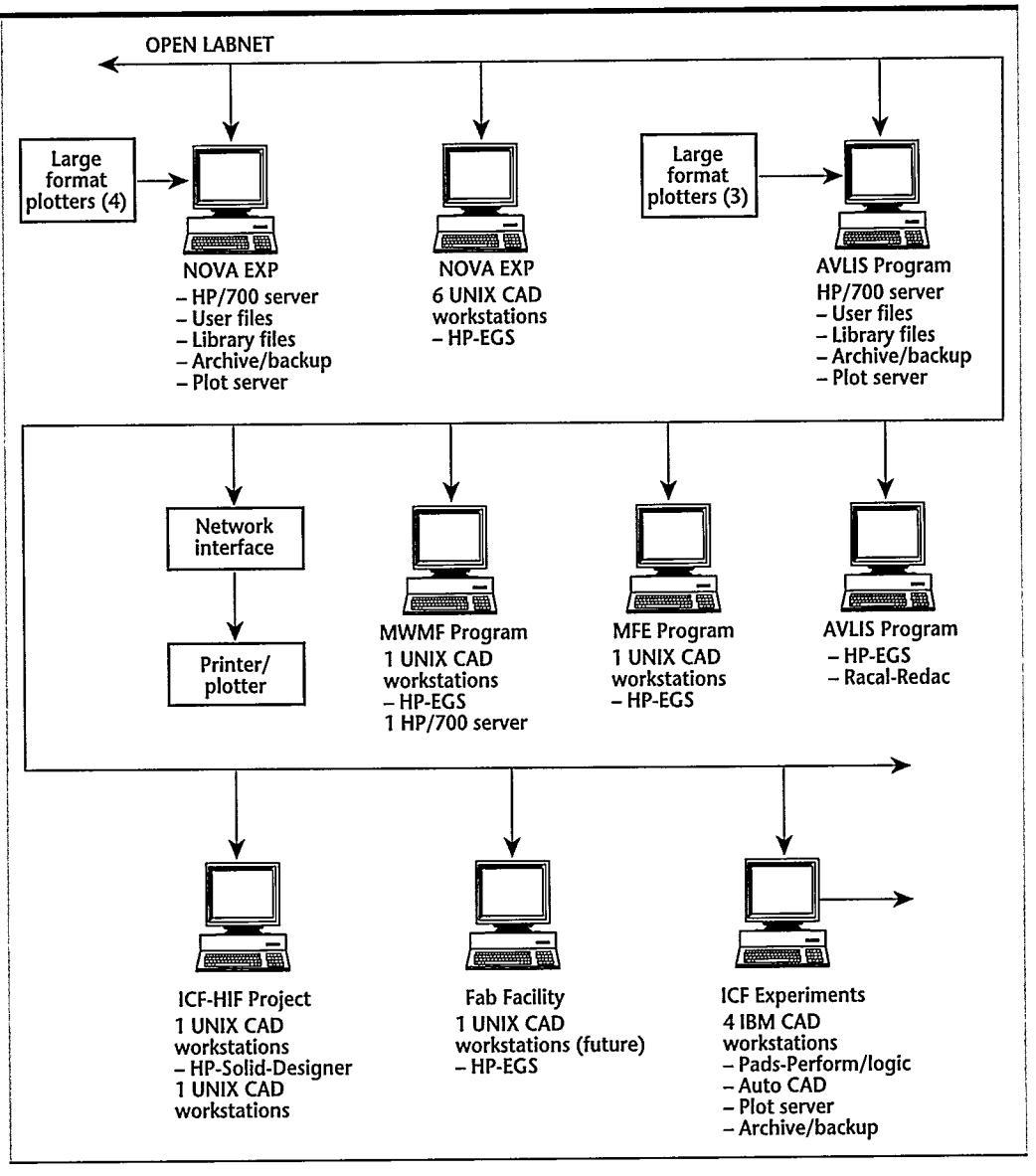


Figure 4. Existing EE CAD/CAM Workstation Network.

documents in many different formats. Viewer access to the documents is provided via a NetScape server/browser system.

A major accomplishment this year was the installation of the NetScape secure server. This World Wide Web server provides encrypted communications links to its client browsers, and provides a high degree of protection to the sensitive, but unclassified data stored in the host computer system's database. This secure communications scheme is critical for rapid procurement of prototype parts using the Internet.

The project has been able to take advantage of both the NetScape server and the ProCard procurement system to procure prototype printed circuit boards and mechanical components for the projects within LLNL's Laser Program. Part of this work involved the production of a part designed by EG&G (Nevada), giving us a chance to verify that rapid prototyping process works in collaboration with another DOE site. This CERPS effort eliminated paper for any step of the design or procurement process, and used engineering-developed data files to directly manufacture the parts. The process reduced procurement time from 25 working days or more, to five and one half days, in one typical example. The IPC charge was reduced from $174 to $25, which is very significant for low cost items.

CERPS has taken a lead role at LLNL in advocating using the Internet to lower the cost of doing business, particularly within the Engineering Departments. One of the major project successes in this area was the convening of the first meeting of the Engineering and Commerce on the Internet (ECI) Working Group. This group consisted of representatives from Procurement, Engineering, Computations, and Technical Information Department and resulted in the agreement to work together to promote Internet-based approaches to lowering the LLNL's cost of doing business. The CERPS team has also been invited to join the CALS/CommerceNet Working Group, which is taking the lead role for defining how government-developed electronic technical data standards (CALS) can be used to facilitate international enterprise integration. Participation in this working group will give LLNL access to a full suite of hardware and software tools, as well as visibility at the national level as a leader in the automation of design engineering.

The project has also focused on the use of a digital signature facility being tested by the Software Technology Center. This facility includes an X-500 directory service to provide for broad access to public keys, and the Entrust digital signature system which provides for the creation and use of digital signatures. The CERPS process used digital signatures this year to digitally sign manufacturing design files that were then released for manufacturing.

Future Work

Next fiscal year the project will focus more on the issue of digital signature blocks to facilitate the paperless engineering process. Engineering document processing depends upon authorization and certification of documents during the design cycle. Paper documents use the signature block to record these approvals, but electronic data files currently do not contain signature blocks. Digital signatures offer an excellent, and, in many ways, superior method to verify authorization and approval of digital files. The CERPS project will implement web-based secure digital signature blocks, which are fundamental for paperless engineering documents, and a graded approach to documentation. The availability of digital signature blocks will save time and money for everyone.

In addition to the work on digital signature blocks, CERPS will implement a pilot of the web-enabled engineering design and procurement process demonstrated this year. The project will focus on expanding the number of vendors involved to at least 15 local Small Manufacturing Enterprise (SME) vendors.

CERPS will implement its technology on the CALS/CommerceNet Intelligent Manufacturing Hub. This implementation will provide LLNL with experience in collaborative commerce relationships with other members of the CALS/CommerceNet group and provide a way to access an additional base of SME vendors that will be important to growing programs such as the National Ignition Facility.

Acknowledgments

The authors wish to thank M. Bell for graphic design and World Wide Web homepages; S. Hawkins for demonstrating solid modeling collaborative exchanges between LLNL, EG&G North Las Vegas, Allied Signal, Kansas City and commercial vendors; W. Edmonson for beta-testing ProCard and Unix Systems Management; B. Weber for ZAP file transfers to the Engineering Record 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. 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.

2. Functional Requirements Document for the Government Acquisition through Electronic Commerce (GATEC) System (1992), Lawrence Livermore National Laboratory, Livermore, Calif.

3. Functional Requirements Document for the Total On-Line Procurement System (in press), Lawrence Livermore National Laboratory, Livermore, Calif.

4. Ulery, D. L., and C. Wimple (1994), "The Future of CALS Interchanges with UN/EDIFACT Standards," *CALS/Enterprise Integration Journal*, Winter 1994.

5. Walton, R. E. (1989), *Up and Running: integrating information technology and the organization*, Harvard Business School Press, Boston, Mass., ISBN: 0-87584-218-6.

6. Gaffin, A. (1994), *Welcome to the Big Dummy's Guide to the Internet, version 2.2* , http://audrey.levels.unisa.edu/big-dummy/bdg_3.html, Electronic Frontier Foundation, Framingham, Mass., February 20, 1994.

7. Krol, E. (1992), *The Whole Internet Catalog & Users Guide*, O'Reilly and Associates Inc., Sebastopol, Calif.

8. Schwartz, L., and R. Schwartz (1993), *Learning Perl (the Llama Book), Nutshell Handbook Series*, O'Reilly and Associates Inc., Sebastopol, Calif., ISBN: 1-56592-042-2.

9. Schwartz, L., and R. Schwartz (1991), *Programming Perl (the Camel Book), Nutshell Handbook Series*, O'Reilly and Associates Inc., Sebastopol, Calif., ISBN: 0-937175-64-1.

Subcarrier Multiplexing for Fiber-Optic Networks

Paul D. Sargis, Ronald E. Haigh, and Kent G. McCammon
Defense Sciences Engineering Division, Electronics Engineering

We have developed a fiber-optic multiplexing technique which combines microwave and photonics technology. This system allows the simultaneous, long-haul transmission of multiple high-speed data streams over a conventional fiber-optic cable, while significantly reducing the effects of chromatic dispersion. Laboratory measurements on our prototype system confirmed the predictions we obtained using our computer model.

Introduction

To fully realize the enormous bandwidth offered by optical fiber, data must be multiplexed at several levels. Long-haul fiber systems typically operate at a wavelength of approximately 1550 nm to maintain compatibility with erbium-doped fiber amplifiers (EDFAs). Although wavelength-division multiplexing (WDM) is becoming a popular technique, the number of wavelengths is limited by the EDFA bandwidth. Subcarrier multiplexing (SCM) greatly increases the data capacity of a fiber by up-converting several baseband data streams to individual microwave frequencies, and transmitting them at a single wavelength. The number of subcarriers is limited by the data rate per subcarrier and the total modulation bandwidth of the SCM system. SCM, when combined with WDM, is especially effective at exploiting fiber bandwidth.

Experimental high-speed SCM systems presented in the literature to date have used coherent detection schemes[1] which are too complex to implement in a production telecommunications environment. Since long-haul SCM systems must operate near the 1550-nm wavelength, they are prone to chromatic dispersion in ordinary fiber. To circumvent this limitation, researchers have used special dispersion-shifted fiber to demonstrate their long-haul SCM systems.[2] Since existing fiber routes use ordinary fiber, practical optical networks of the future must function on what has already been installed.

Our approach allows baseband detection[3] and achieves dispersion reduction in the SCM receiver. Using commercial microwave mixers and a lithium niobate external modulator, sidebands are generated several gigahertz apart on the principal laser optical carrier. Digital data streams are impressed upon these sidebands for transmission over ordinary single-mode fiber.

Progress

In the past year, our efforts have focused on computer simulations, prototype development, and laboratory demonstrations.

Before we settled on a basic design, we developed some optical network models using National Instruments' LabVIEW. These models have enabled us to analyze the fiber dispersion problem and to evaluate the effectiveness of our SCM designs at reducing dispersion. As we will show later, system measurements have proven our models to be very accurate.

A block diagram of our prototype SCM system is shown in **Fig. 1**. Independent 2.5 Gbit/s data streams are up-converted to microwave frequencies f_1 and f_2, spaced several gigahertz apart. A diode-pumped, solid-state 1550-nm laser carrier is externally modulated by the microwave subcarriers using a Mach-Zehnder interferometric modulator having a bandwidth of 18 GHz. To suppress the principal laser optical carrier, the modulator is maintained at null bias. The resulting subcarrier multiplexed optical signal is capable of traversing hundreds of kilometers of fiber with the aid of appropriately-spaced EDFAs.

Prior to detection, the desired subcarrier is optically pre-selected using a fiber Fabry-Perot (FFP) filter having a bandwidth of 4 GHz and a

finesse of 125. This filtering process is illustrated in **Fig. 2(a)**. The ideal filtered optical spectrum is shown in **Fig. 2(b)**. A 2-GHz bandwidth optical detector converts the selected optical signal into a baseband electrical data stream, as in **Fig. 2(c)**.

One of our laboratory demonstrations featured SCM transmission through 220 km of spooled fiber. Four EDFAs were required to compensate for fiber attenuation on the order 0.3 dB/km. A pair of 2.5 Gbit/s data streams were placed on subcarriers at 7.3 and 14.6 GHz for this series of tests.

To demonstrate the dispersion reduction effect, and to verify the accuracy of our models, the 2.5 Gbit/s repeating pattern (11101010) shown in **Fig. 3** was impressed upon the 14.6 GHz subcarrier. (The 7.3 GHz subcarrier was turned off for this test.) With the FFP filter bypassed, the received pattern shown in **Fig. 4(a)** was severely distorted due to chromatic dispersion in the fiber. This result is in close agreement with the computer simulation in **Fig. 4(b)**. When the FFP filter was placed back in line and centered on the 14.6 GHz subcarrier, signal distortion due to chromatic dispersion was practically eliminated, as the received pattern shows in **Fig. 5(a)**. Results from the model show a similar outcome, as indicated by **Fig. 5(b)**.

Using eye diagrams, we evaluated system performance with and without spools of fiber. Uncorrelated pseudorandom bit streams having opposite polarities and different time delays were simultaneously impressed upon the 7.3 and 14.6 GHz channels. To remove dispersion from the experiment without eliminating noise sources, each spool of fiber between the EDFAs was replaced with an optical attenuator having equivalent loss. **Fig. 6 (a)** and **(b)** shows a comparison of eye diagrams that resulted from optically pre-selecting the 7.3 GHz subcarrier with optical attenuators and with fiber, respectively. **Fig. 6 (c)** and **(d)** shows the same comparison for the 14.6 GHz subcarrier. These results include crosstalk from the other subcarrier channel. With this design, 220 km of fiber presents no more of a challenge to the link than optical attenuation.

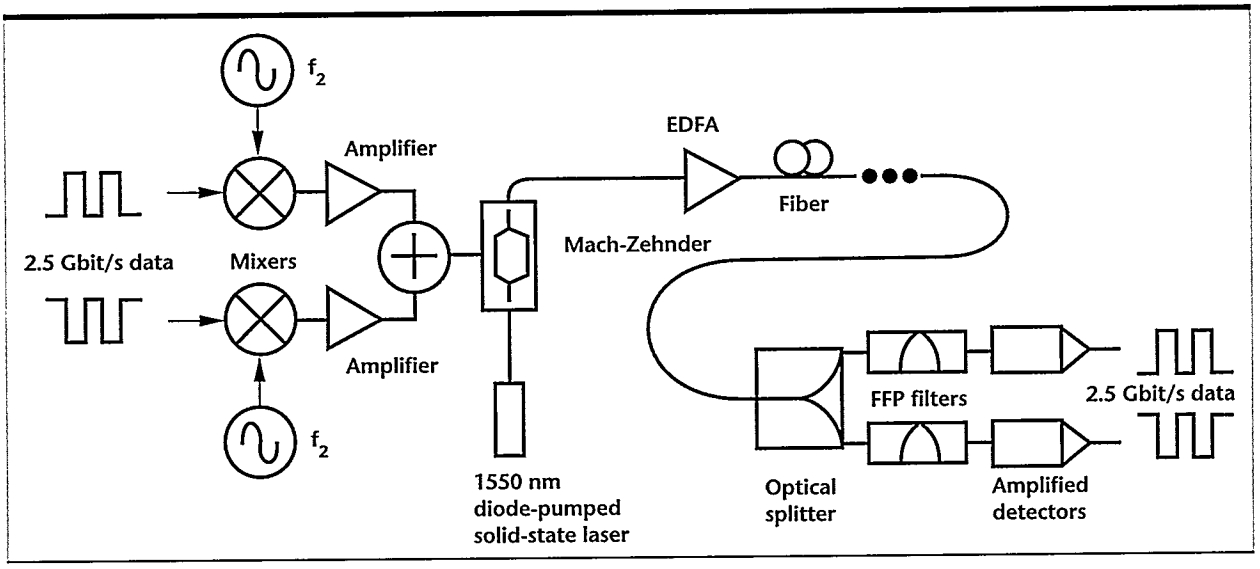
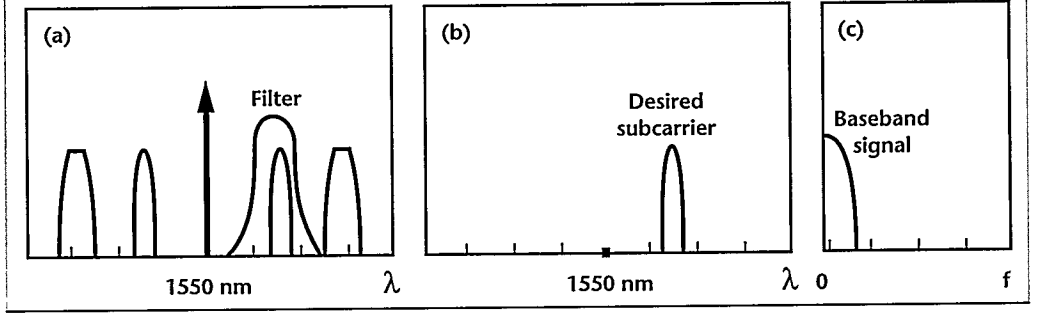


Figure 1. Block diagram of the subcarrier multiplexing system.

Figure 2. Optical spectrum (a) at the FFP filter, and (b) after the FFP filter. Frequency spectrum after the detector is shown in (c).



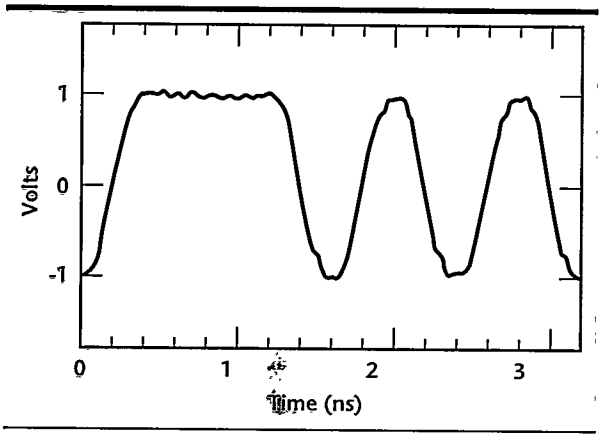


Figure 3. 2.5 Gbit/s repeating pattern at input of 14.6 GHz channel.

Future Work

We plan to conduct field demonstrations to prove the viability of SCM and to generate interest within the telecommunications industry. We will begin with a 90-km field trial using a Pacific Bell fiber route. This will give us an opportunity to test for compatibility between SCM and WDM. Later in the year, we expect to place SCM on one of the wavelengths of ARPA's 600-km NTONC network around the San Francisco Bay.

In parallel with these efforts, we will make hardware improvements to improve stability and reliability of the link. In addition, it will be

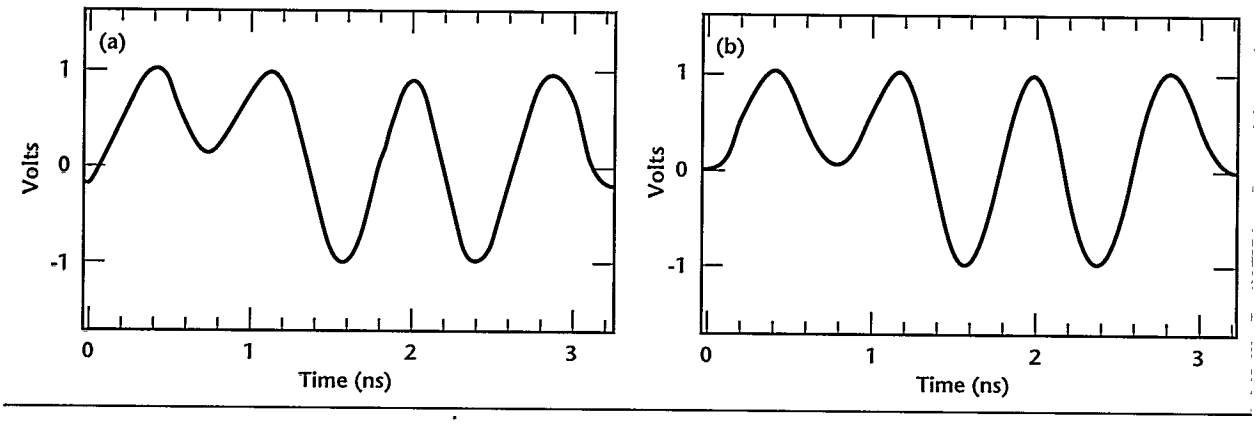


Figure 4. 2.5 Gbit/s pattern from 14.6 GHz channel output without dispersion reduction for (a) measured, and (b) modeled data.

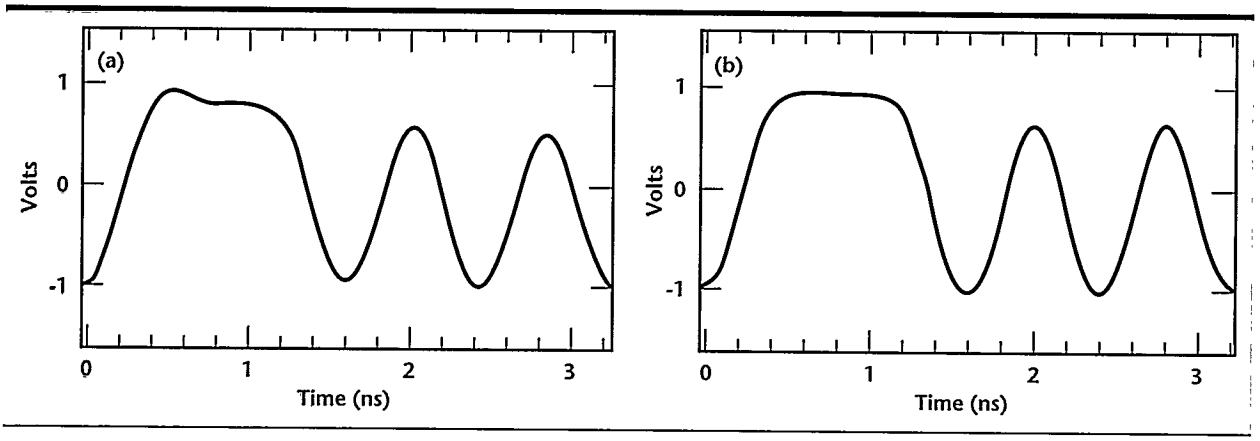


Figure 5. 2.5 Gbit/s pattern from 14.6 GHz channel output with dispersion reduction for (a) measured, and (b) modeled data.

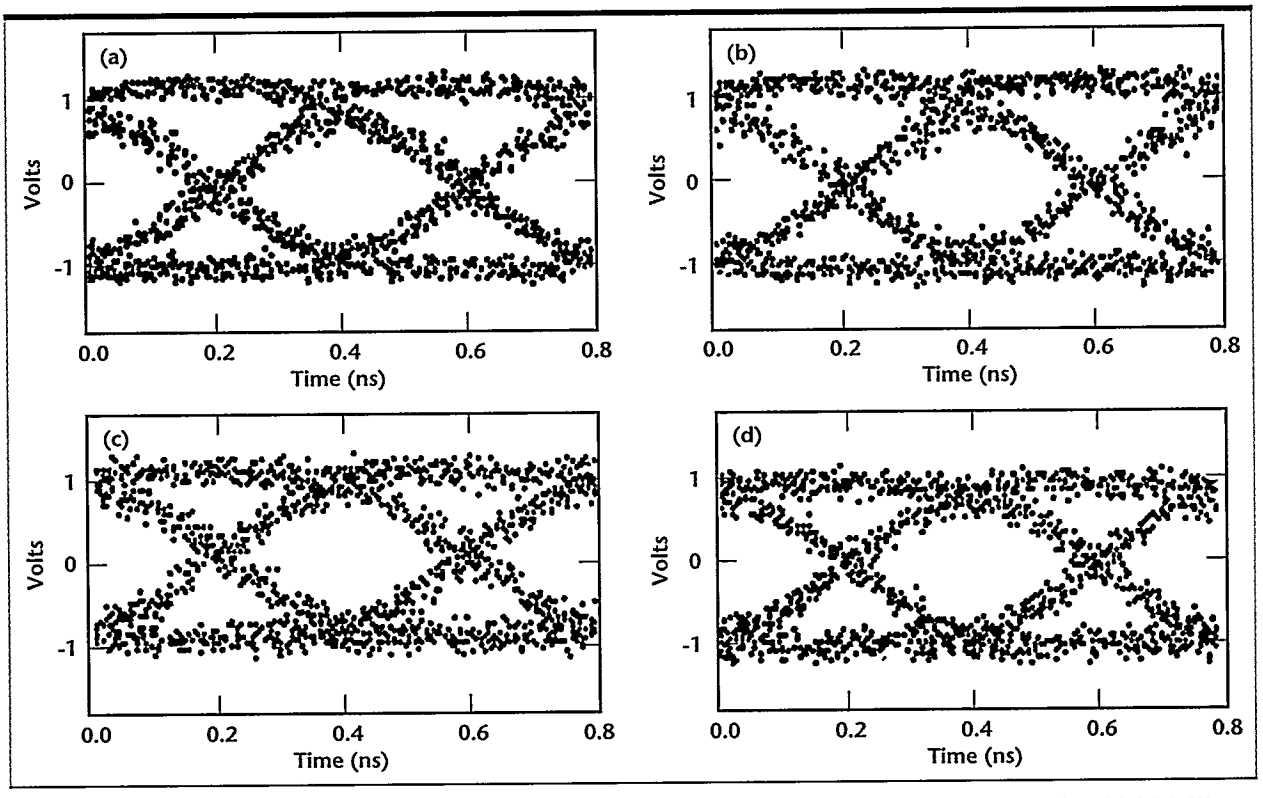


Figure 6. Eye diagrams for (a) 7.3 GHz channel with optical attenuators; (b) 7.3 GHz channel with 220 km of fiber; (c) 14.6 GHz channel with optical attenuators; and (d) 14.6 GHz channel with 220 km of fiber.

important to incorporate the behavior of EDFAs into the computer model.

One of our goals is to show that SCM will enable Lawrence Livermore National Laboratory, Los Alamos National Laboratory, and Sandia National Laboratory to be connected via a 40 Gbit/s wide-area network without regenerators. This network of the future is of vital interest to the ASCI and SuperLab projects. We plan to assemble a 2500-km recirculating loop experiment in the laboratory to demonstrate that this is a feasible proposal.

References

1. Hill, P. M., and R. Olshansky (1991), "8 Gb/s Subcarrier Multiplexed Coherent Lightwave System," *Photonics Tech. Lett.* 8, pp. 764–766.

2. Hill, P. M., and R. Olshansky (1990), "Bandwidth Efficient Transmission of 4 Gb/s on Two Microwave QPSK Subcarriers Over a 48 km Optical Link," *Photonics Tech. Lett.* 7, pp. 510–512.

3. Greenhalgh, P. A., R. D. Abel, and P. A. Davies (1992), "Optical prefiltering in subcarrier systems," *Proc. SPIE*, 1790, pp. 76–84.

The World-Wide Web as a Medium For Presenting LLNL Training Courses

Darrel Lager
Computer and Communication Engineering Division
Electronics Engineering

Sue Koopman
Education, Training and Safety Analysis
Hazards Control

We have put two prototype courses on the World-Wide Web: an introductory course for new personnel and an on-the-job-training course. This report describes our proof-of-concept work.

Introduction

Many large scientific organizations, such as Lawrence Livermore National Laboratory (LLNL), need to continuously train their employees to keep them knowledgeable and certified, for both safety and environmental/regulatory compliance. Also, new or re-assigned employees need introductory training. Often people need to take a course but must wait until instructors or facilities can be scheduled or until enough students sign up to justify its expense. The classical Lecture-Based Training (LBT) approach is to maintain a waiting list of prospective trainees until there are enough to schedule a class, then make arrangements for the students, the instructor, facilities, and course materials, all to be available at the same time and place.

Computer-Based Training (CBT) seeks to eliminate the scheduling problem by placing the course materials on a dedicated computer so students can cover the materials at their own rate. With this approach, students retain information better (70% for multimedia CBT, vs 10% for LBT). But, in general, CBT requires specialized hardware and software and a facility. Since it is often expensive to produce a course, many highly-specific training courses cannot be justified—for example, training three or four technicians in the safest procedure for working on a piece of high-voltage equipment.

Web-Based Training (WBT) is an evolution of the concept of CBT. WBT uses the technology available for the World-Wide Web for presenting course materials, providing capability for questions and feedback to the author, testing students to verify they have assimilated the information, and automatically notifying the training administrator when a student has taken a test. The existing Web browsers, e.g., Mosaic and Netscape, support presentation of text, images, movies, interactive forms, and interactive images, all of which are used in course presentation. We have added the capability for evaluating the student's progress with true/false multiple-choice tests administered via the Web.

There are several advantages of WBT over conventional LBT:

1) Students can view the course materials at any time and rate, since the Web is always available.
2) Students can view the materials in familiar surroundings, such as their normal workplace or home.
3) There is no need to schedule classes, instructors, or classrooms.
4) The instructor spends less time lecturing. For many on-the-job-training (OJT) courses, the instructor is a fellow worker or a supervisor, so less lecture time reduces time the instructor or supervisor is taken away from normal duties.

There are also several advantages of WBT over CBT:

1) Students don't need to learn to use an unfamiliar computer, since browsers are available for the most common computers at LLNL, i.e., Macintosh, PC, and UNIX workstations.
2) A large (and growing) number of people are already familiar and comfortable with the Web browsers so they are already familiar with the user interface.

3) The architecture of the system is determined by the needs of the organization, not by the software—centralized or decentralized, stand alone or networked.
4) Course materials become reference materials, with updates immediately available to everyone (vs the time and expense of distributing new hardcopy or new course CD-roms).
5) Refresher training for compliance is substantially faster, since students can do a quick review, then take the test.
6) With the Web, it is easy for the course author to include links to other materials, such as Department of Energy Orders, FSP's, OSP's, and prerequisite courses.

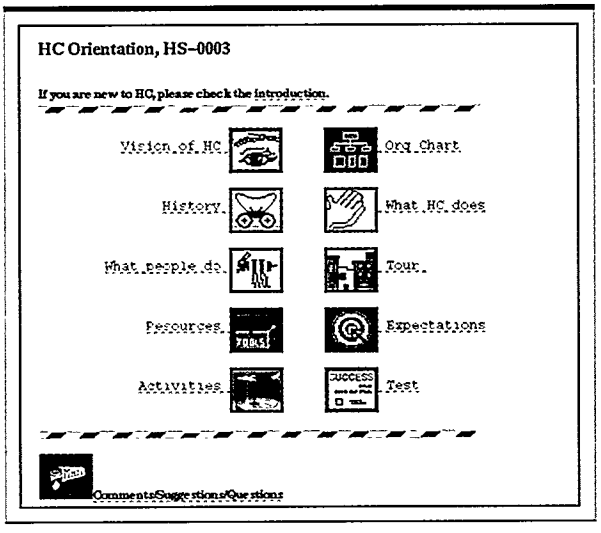


Figure 1. Topics for the HC Introductory Course, identified by icons. Standard icons identify the "Test" and "Opportunity for Questions and Feedback " sections.

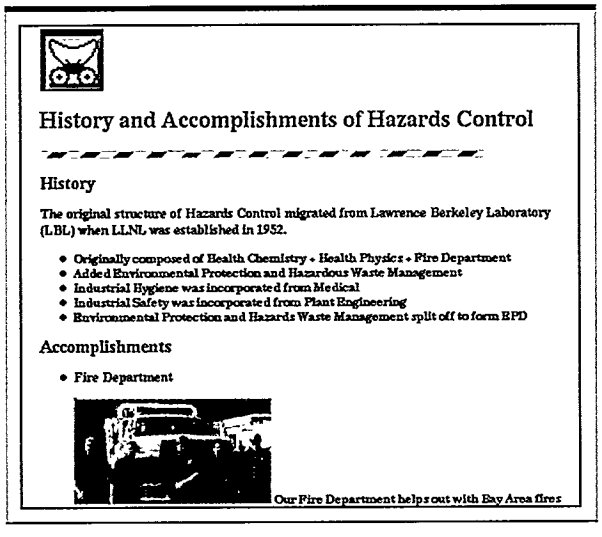


Figure 2. History Section for the HC Introductory Course, using text and photos.

7) The software for presenting the course is very inexpensive. Both browsers and servers are available free on the Internet. Commercial versions are nominally priced.
8) The language of the Web, html, is very easy to learn, so course authors need little training before producing course material.

We are exploiting the concept of course "templates," where the boilerplate and structure of a course are already in place, and the author simply inserts the relevant text and images. Courses can be designed as elaborately as necessary. For a simple course the author would gather existing electronic documents, take snapshots, and scan images. An elaborate course might use full Technical Information Department editing, professional camerawork, artwork, and scripting.

Progress

Two particular prototype courses have been put on the Web: an introductory course for new Hazards Control (HC) personnel (HS-0003) and an OJT course for Safing a Laser Amplifier (PSCL-WS-3213). These courses were chosen because they exemplified the characteristics most natural for presentation on the Web. The HC introductory course contains photo images of personnel and facilities, maps, organization charts, and short descriptions of facilities, all of which are readily assimilated without a great deal of interaction between student and teacher. This type of course tends to be a several-hour monologue on the part of

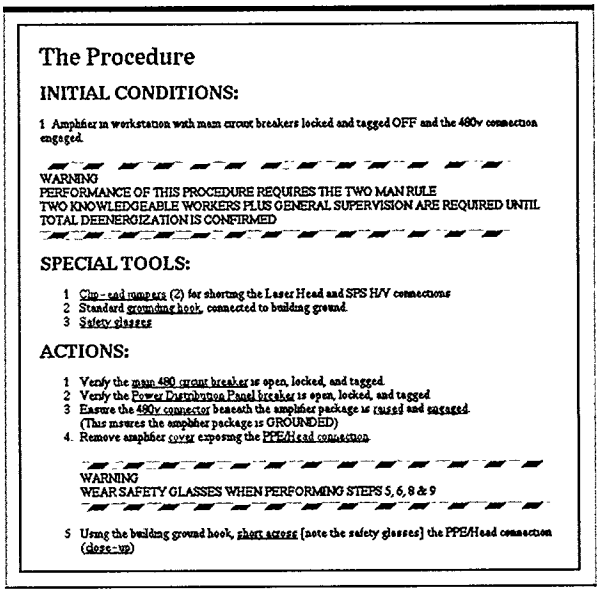


Figure 3. OJT Procedure, with the text linked to annotated photos.

the instructor. When presented on the Web, the course becomes very visual and easy to absorb. For example, the major topics of the course are identified by icons, (**Fig. 1**), and the history section is a mixture of text and photos, (**Fig. 2**).

The OJT course represents a type of course where the desire is to reduce the production costs. The text materials were taken from existing documents in electronic form and easily converted to html, the document language for the Web. For the Procedure section, (**Fig. 3**), we simply took the pre-existing text and made links to photographs of the apparatus, with annotations to indicate the object of interest, (**Fig. 4**). It is also easy to make the material more interesting by adding visual features, like the black and gray striped dividing lines in **Fig. 3**.

The combination of text and photographic images helps the student to become familiar with an apparatus, perhaps without having to disturb an on-going experiment.

Other courses are also at the prototype level, covering Electrical Safety Awareness, the Lock and Tag refresher, and the Hazard Communication portion of the Introduction to New LLNL Employees.

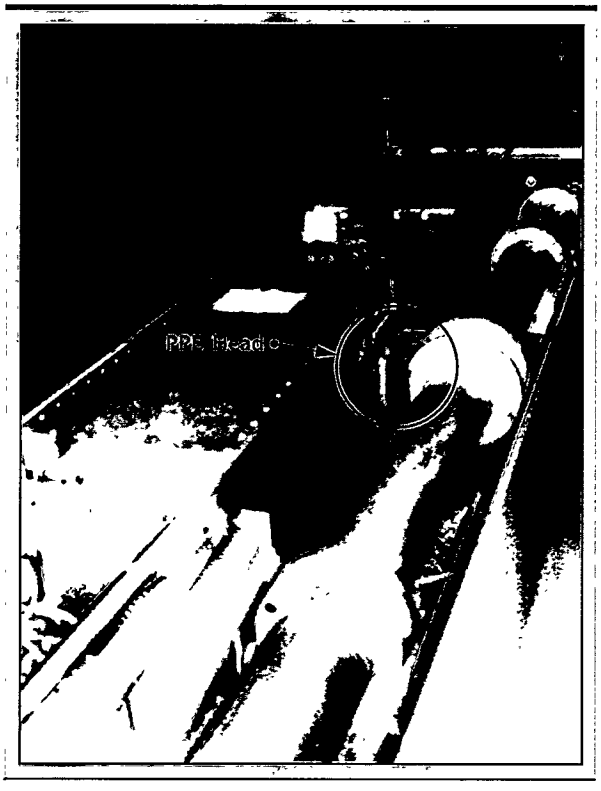


Figure 4. Photograph of the apparatus, annotated to show pulsed-power high voltage head.

Testing

For both types of courses we have provided a true/false, multiple choice test for evaluating how much of the information the student has absorbed. The student uses the Web browser to fill in forms for required information (**Fig. 5**), then moves to the test itself and selects the answers to the questions (**Fig. 6**). When the student is finished, the test is automatically scored (**Fig. 7**), and the results e-mailed to the appropriate training administrator.

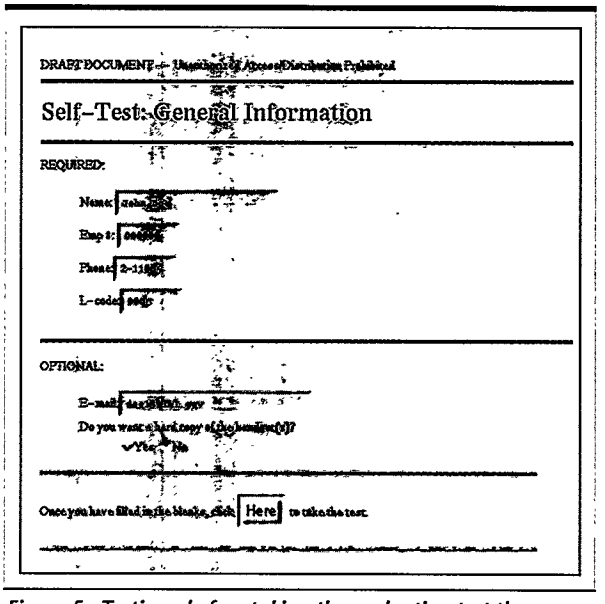


Figure 5. Testing: before taking the evaluation test the student enters information on a form.

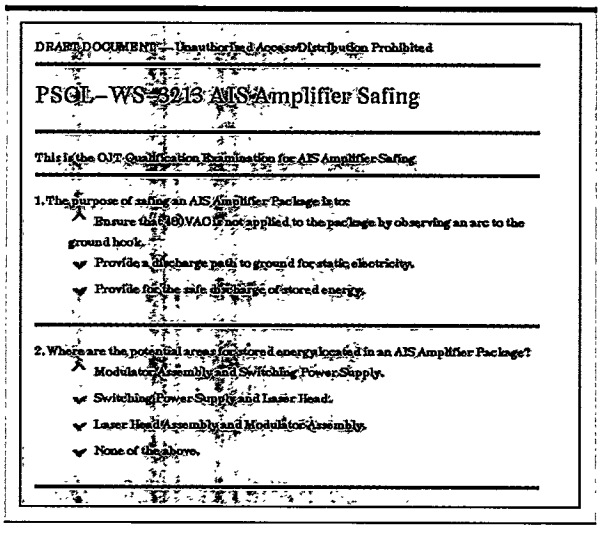


Figure 6. Test, with true/false multiple choice questions, with answers selected by mouse clicking over a button.

The actual text for the test is automatically formatted from a template file written by the course author and displayed by the Web browser. The author is freed from concerns of numbering the questions and learning the *html* commands for putting up the answer selection buttons, bars separating the questions, and the boilerplate at the beginning and end of the test. An example template is shown in **Fig. 8**, which produces the test questions shown in **Fig. 6**.

Due to limitations in the existing Web browsers and protocols, the administration of the test is on the "honor system" and is as much a self-assessment tool for the student as a mechanism for documenting the student's retention of the material. If it were really important to ensure the student was not cheating, the test portion of the course could be proctored. Future improvements in the Web technology, based on secure Web servers and "digital signatures," could reduce the need for proctoring. Also, one could always resort to the time-honored method of a paper hardcopy test administered in a closed room.

```
Test Results for --
 Name: John Doe
 Emp#: 000000
 Phone: 2-1100
 L-code: 000
 E-mail: doe1@llnl.gov

 Course Title: PSCL-WS-3213 AIS Amplifier Safing
   (from file: lasers/pscl-ws-3213.test)


 Date: Thu Sep 28 10:24:55 1995

1 was Wrong -- answer given: a, correct answer: c
2 was CORRECT!!
3 was Wrong -- answer given: , correct answer: True
4 was Wrong -- answer given: , correct answer: False
5 was Wrong -- answer given: , correct answer: c
6 was Wrong -- answer given: , correct answer: False
7 was Wrong -- answer given: , correct answer: c
8 was Wrong -- answer given: , correct answer: False
9 was Wrong -- answer given: , correct answer: False
10 was Wrong -- answer given: , correct answer: True
```

Figure 7. Scoring: the test is automatically scored; the results are shown to the student, and mailed to the training administrator.

```
PSCL-WS-3213 AIS Amplifier Safing

This is the OJT Qualification Examination for AIS Amplifier Safing


<QS>The purpose of safing an AIS Amplifier Package is to:
<A>Ensure that 480VAC is not applied to the package by observing an arc to the g
round hook.
<A>Provide a discharge path to ground for static electricity.
<AC>Provide for the safe discharge of stored energy.
</QS>

<QS>Where are the potential areas for stored energy located in an AIS Amplifier
Package?
<AC>Modulator Assembly and Switching Power Supply.
<A>Switching Power Supply and Laser Head.
<A>Laser Head Assembly and Modulator Assembly.
<A>None of the above.
</QS>
```

Figure 8. Template: the course author writes a test template, which is automatically processed to produce the test.

Future Work

The WBT courses presently available constitute a proof-of-concept prototype. We are working with training managers in the various LLNL programs to support WBT and develop the teams to put this technology to work. The technical issues to be addressed, include the following:

1) providing an explanation facility for the student to give feedback when an incorrect answer has been chosen;
2) randomly presenting questions from a large pool of questions, so tests are rarely identical;
3) automatic entry of test results into the LLNL training database software; and

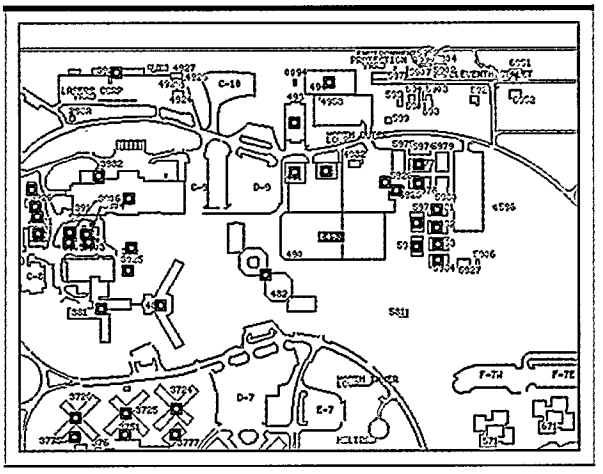


Figure 9. HotMap, indicating the facilities in the LLNL Laser Program as icons.

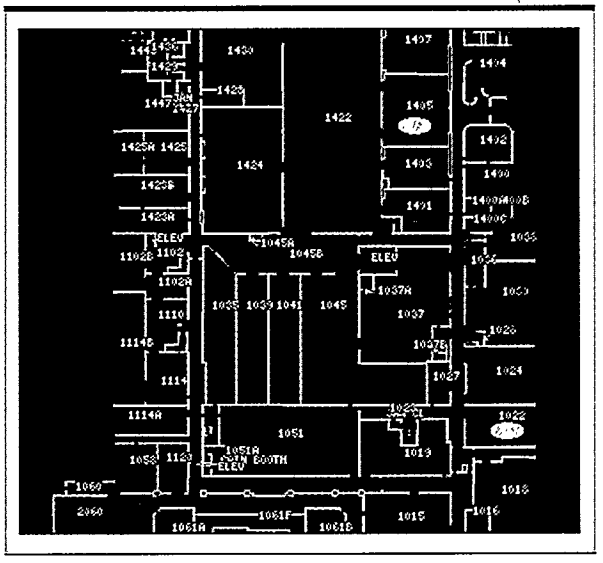


Figure 10. OJT Courses: clicking the Building 490 icon displays a drawing of the building, with icons in rooms where OJT courses are required.

4) management/maintenance issues, when a large number of courses from many organizations are in the system, i.e., keeping a database of information such as course authors and administrators.

None of these technical issues is insurmountable.

Another question we have addressed is how to find courses available for a particular facility. We have used HotMap by D. Ashkenas to display a map of LLNL with icons for facilities in the Laser Program (**Fig. 9**). Clicking on the icon for Building 490 will lead to a drawing of the building (**Fig. 10**), and clicking on the "OJT" icon in room 1022 will lead to the Web page for the OJT courses for that room (**Fig. 11**). HotMap is a tool for browsing course materials based on the geographical location of facilities containing the course subject.

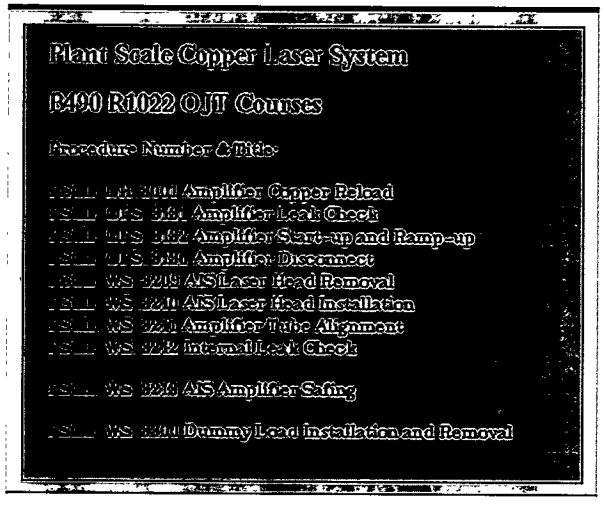


Figure 11. OJT Courses: clicking the OJT icon displays a Web Page, listing the courses for the laboratory. Clicking the Course Title begins the course.

Acknowledgments

The authors wish to acknowledge the funding contributions of the Electrical Engineering Department's Information Engineering Thrust Area (D. Hunt), the Concurrent Engineering Rapid Prototyping System-CERPS (C. Jordan), and Hazards Control (S. Hildum). C. Jordan also contributed several ideas, such as automatic mailing of test scores. We also wish to thank T. Thomas for producing the electronic documents for the OJT course, G. Schmere for acting as a guide (and model) for the snapshot photographs, and M. Bell for icon/logo development and photo scanning.

Design and Implementation of 3-D MIME Data Access and Exchange

Scott D. Nelson
Defense Sciences Engineering Division
Electronics Engineering

Jeffrey S. Kallman
Engineering Research Division
Electronics Engineering

Eric Brugger
Scientific Computing Applications Division
Computations Organization

This project created an infrastructure for the exchange of 3-D data objects through the network, by allowing users access to freely available viewers. The goals were as follows: 1) to combine existing software and to create new software; 2) to allow distribution from one location; 3) to allow users easy access to 3-D data; 4) to perform simple 3-D manipulations, including zoom, rotation, translation, and hide/show materials; and 5) to distribute the scientific and technical data of Lawrence Livermore National Laboratory (LLNL) in its native 3-D format. A table of acronyms is provided.

Introduction

Earlier world-wide web (WWW) data exchange services did not include facilities for moving 3-D, manipulable data objects. The previous practice was to generate a 2-D bitmap of the object in question, and to use that as the data object. This resulted in an increase in data size, loss of information, and limits in understandability. The WWW data exchange mechanism is becoming popular, both at LLNL and between LLNL and universities and companies.

The solution was to create a system to allow the portable sharing of data between different systems, through the existing network (**Figs. 1 and 2**). This approach is software-only and expands the existing capabilities of LLNL. Presently, there is not an ISI NWG[1] standard for 3-D solids exchange, although numerous proprietary standards, (such as AutoCAD), exist on some platforms. Our intent was not to duplicate existing tools. By creating the data type, the format, the transportation mechanism, and the visualization tools, we have created

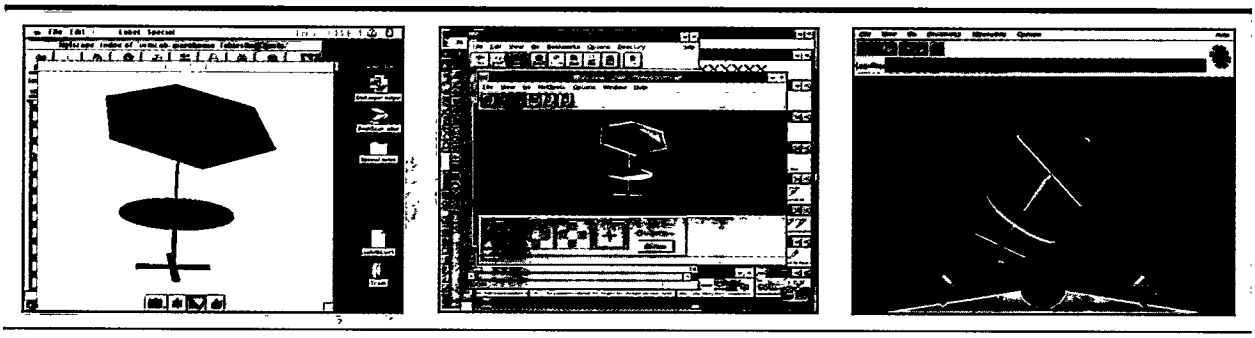


Figure 1. Multiplatform access to the same 3-D data on Macintoshes, PCs, and Unix systems. Access methods and operations are similar on the various platforms.

standards that can be used at LLNL, with universities, and with industry to communicate scientific and technical information.

The general functionality is accessed through a web browser, such as Netscape or Mosaic. The user selects an object; the object is transferred through the network to the user's computer; the computer automatically launches the appropriate 3-D viewer; and the user manipulates the 3-D object.

Progress

There are several kinds of 3-D vector and solid data types currently used on different computer systems for various purposes, such as AutoCAD, FIZ, BRLCAD, POV, Ingrid, and GRIZ/Taurus. Our approach was to select a simple subset of data objects common to these types. This subset can easily describe a 3-D vector object, or 3-D solid object, that meets the needs of LLNL in sharing scientific or vector 3-D data. For the scope of this project, the data objects of interest are limited to scalar and vector data, meshes, solids, chemicals, nodes, and wire frames.

The existing Mosaic-like WWW access tools are used as the data exchange engine. Our contribution was in the visualizer to access the 3-D data. The data subset selected defined generic objects that are of interest to LLNL. Part of this effort was getting the various data types registered with the Internet standards body,[1,2] such that users[3] can distribute 3-D data with a common mechanism. Two-dimensional data types[4] are already well-covered by numerous other data formats, and are outside of the scope of this project.

New Type Registration

The Internet uses a series of MIME data objects to exchange different kinds of data through the various pieces of hardware connecting the system. For new kinds of data, it is necessary to register the data objects such that hardware and software can recognize these data types. Existing data types[4] include text, video, audio, image, message, multipart, and application. As part of this project, a 3-D primary type called *model* is being defined,[5] and includes four pre-defined sub-types: vrml, mesh, iges, and chemical.[6–12] Note that IGES is already defined for 2-D as application/iges, but the discrimination issues between the 2-D type and the 3-D type are resolved in the RFC.

The type registration process for primary types differs somewhat from that of subtypes, in that the primary types require a more extensive review process, since the number of primary types is purposely limited. This RFC process involves several levels of peer review, and general working group approval. Primary types are reserved for transport issues, and deal with the type of data being sent and the ultimate hardware, which will articulate the data and does not reflect the content. The new *model* type is being created because hardware, such as stereo lithography machines, 3-D glasses, and peptide synthesizers, requires a new type, different from that required for the other data types. Thus, one would not send an audio stream to a stereo lithography machine.

The term "3-D," used above when describing the *model* type, is misleading. In actuality, the *model* type will support 4-D (x, y, z, t) or higher dimensions. This is the primary difference between the *model* type and all of the other types, which support two, or fewer, spatial dimensions, and one or no temporal dimension.

A *model,* in general, is defined by the following criteria which make it different from the other existing primary types:

1) three or more dimensions that are bases of the system;
2) orthogonal coordinate system (any orthogonal system is sufficient);
3) calibration factors to physical units (force, momentum, time, velocity, acceleration, size). Thus, an IGES file will specify a building of

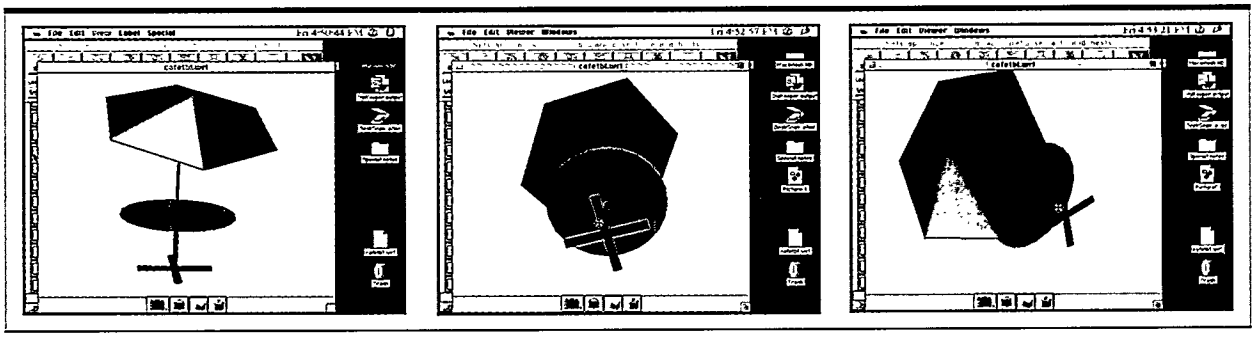


Figure 2. Illustration of data objects manipulated by the client in real time, not requiring further server operations. The burden is on the user's system; and parallel economies are realized and not server-limited.

non-arbitrary size; computational meshes and VRML models will have real spatial/temporal units; and chemical models will contain spatial/molecular/physical specifications.

With these assumptions:

1) The default dimensionality will be spatial and temporal (but any are allowed).
2) It is presumed that models will contain underlying structure (such as fluid dynamics vector fields, chemical bond structures, related IGES dimensional specifiers, and VRML materials and operators), which may or may not be immediately available to the user (**Figs. 3, 4, and 5**).
3) It is assumed that basis set conversion between model domains is lossless. The interpretation of the data may change, but the specification will not, e.g., 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, chemicals, and CAD files. The key is to identify the basis set conversion which makes sense. Thus, chemical bond strength may be a "dimension."
4) Models are grouped to assure less loss of information between the model sub-types than there would be between sub-types of other primary types, i.e., converting a chemical model into an image is more lossy than converting it into a VRML model.

Viewer Aggregation and New Viewer Development

As part of the organization effort, the WWW Viewer Test Page[2] was modified to serve as a distribution mechanism for the 3-D datatype viewers. This, in combination with the EMail MIME Test Page,

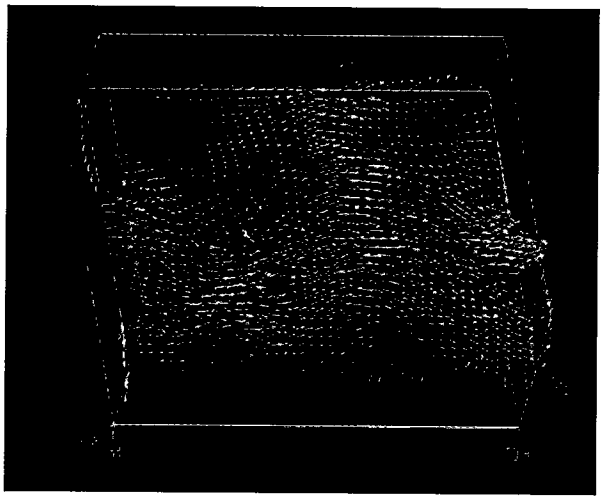


Figure 3. Illustration of how vector viewers allow users to integrate the display of vector data with other kinds of data, to select subsects of data to display, and to control the display format of the vector data.

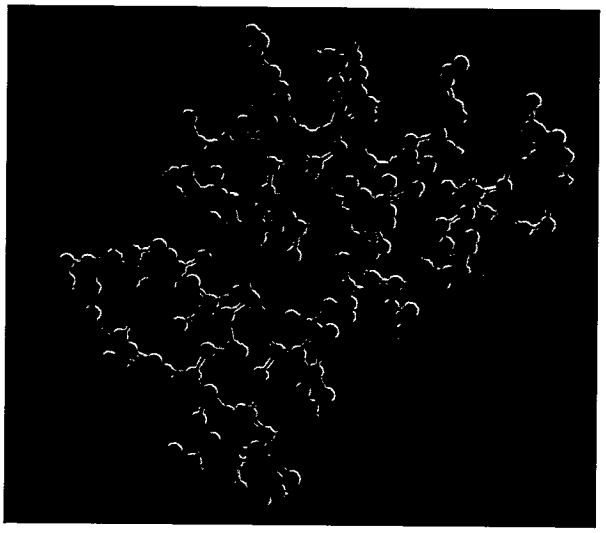


Figure 4. Illustration of how chemical viewers allow users to manipulate the chemical structure directly, change the type of displayed information, and change the level of detail.

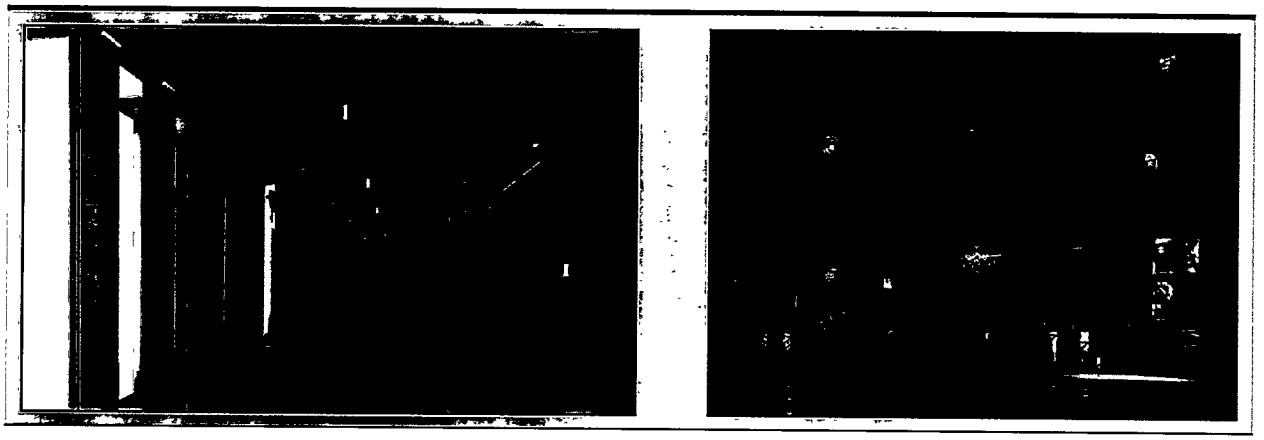


Figure 5. Illustration of how scene viewers allow users to walk through detailed scenes and to interact with various objects.

allows users to acquire, configure, and test the various pieces in their network system. Both clients and servers are provided with configuration information. Viewer software that has been developed on-site, and off-site by others, is freely available from the VTP. As of January 1, 1996, the present access rate to the VTP is about 2000 per day, with an access count of 400,000, and a unique machine count of 340,000 (#machines < #people < #accesses). These numbers do not include accesses to the VTP sub-tree structure. Information has been presented to on-site groups, and in meetings to outside groups. Additionally, these efforts have received national press attention,[13, 14] and the VRML effort as a whole has been well-publicized in the popular press.[15]

The primary focus of most of this effort was to consolidate existing viewers into a central location in an attempt to galvanize the 3-D data type efforts. The focus of this project was not to "reinvent the wheel," but to take advantage of as many existing pieces as possible. The model and VRML communities were unified under the *model* primary type that we created. The sub-types were extrapolated from that point.

The mesh viewer (**Fig. 6**) is an adaptation of the MeshTV viewer,[16–17] and was modified to create a simple and transportable viewer for a broader audience, while working within the framework of the web structure. These modifications included cross-platform compatability with current graphics standards, vector I/O and display, default display lists, command line file I/O, and synchronized multiplatform support.

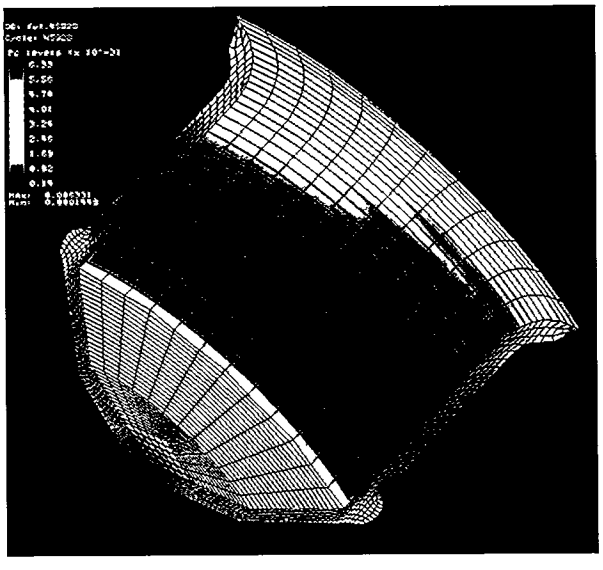


Figure 6. Illustration of how mesh viewers contain scalar/vector data in addition to the mesh, and allow users to manipulate the data and the mesh structure.

Converter Development

A series of data format converters was developed and acquired as part of this effort. These include filters for AutoCAD, IGES, TSAR, GRIZ(Dyna3D), FIZ, BOB, raw binary, and ASCII. There are numerous other formats that are convertible into one of these formats. This allows for cross-conversion to the desired format. Some of these formats are in a state of flux, notably IGES.

Future Work

The following work could be done at a point in the future when the technologies are more mature: Mac and PC 4-D mesh viewers; Mathematica object converters; DSI3D converters; newGriz format converters; better IGES support; and demonstrable "matter transport" to fabrication systems. Due to present funding limits, no future work is planned at this time.

Glossary

As for many acronyms, such as radar and laser, the origins of some in this list have been lost or are not commonly used. In the case of VRML, the acronym has remained the same while the originating text has changed: "markup" changed to "modeling."

Arena One of the many web browsers used to access HTML files.

CERN Conseil Europeen pour la Recherche Nucleaire (European Council (Center) for Nuclear Research); birthplace of the WWW, http://www.cern.ch/.

HTML Hyper Text Markup Language; the language that WWW pages are written in, derived from SGML.

IETF Internet Engineering Task Force; an international body to register new MIME types.

IGES Initial Graphics Exchange Specification; a common 2-D (non-standard) standard graphics format.

ISI NWG Information Sciences Institute Network Working Group; the discussion group that considers the new MIME type applications.

MIME Multipurpose Internet Mail Extensions; different data types are assigned a MIME type for transport/viewing purposes.

Mosaic One of the many web browsers used to access HTML files.

NCSA National Center for Supercomputer Applications; creators of Mosaic.

Netscape One of the many web browsers used to access HTML files.

RFC Request For Comments; an official application to the IETF to get new data types registered via peer review.

SDSC San Diego Supercomputer Center; home of the VRML repository.

TSV Tab Separated Values; one of the registered MIME types for databases.

VRML Virtual Reality Modeling Language; true 3-D interactive objects.

VTP WWW Viewer Test Page; repository for MIME viewers, http://www-dsed.llnl.gov/documents/WWWtest.html. In general, see the VTP, e.g., for more information and pointers to MIME types and RFC's, downloadable viewers and web browsers, pointers to VRML information, information on HTML sites and authoring, and the EMail MIME Test Page.

WWW World-Wide Web; the generic term describing the Internet when used with CERN compatible browsers (e.g., Netscape, Mosaic).

Acknowledgments

The authors would like to thank H. Rzepa for his efforts in the chemical community, M. Pesce for his VRML efforts, and K. Moore for his IETF primary type/sub-type comments.

References

1. Information Sciences Institute (1994), *ISI NWG RFC list*, University of Southern California, Los Angeles, Calif. (ftp://ftp.isi.edu/in-notes/iana/).

2. Nelson, S. D. (1995), *WWW Viewer Test Page*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (http://www-dsed.llnl.gov/documents/WWWtest.html).

3. Defense Sciences Engineering Division (1994), *DSED Home Page*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (http://www-dsed.llnl.gov/).

4. "Media Types" (1994), *ISI NWG*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (ftp://ftp.isi.edu/in-notes/iana/assignments/media-types/).

5. Nelson, S. D., C. Parks, and R. Mitra (1995), *IETA RFC list*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (ftp://ftp.isi.edu/in-notes/iana/assignments/media-types/media-types, "Model Primary Type registration").

6. Bell, G., A. Parisi, and M. Pesce (1995), *The Virtual Reality Modeling Language*, (http://sdsc.edu/SDSC/Partners/vrml/Archives/vrml10-3.html).

7. Nelson, S. D. (1995), *Registration of new Media Type model/mesh*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. [(will be) ftp://ftp.isi.edu/in-notes/iana/assignments/media-types/model/mesh].

8. Parks, C. (1995), *Registration of new Media Type application/iges*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (ftp://ftp.isi.edu/in-notes/iana/assignments/media-types/application/iges).

9. Rzepa, H. S., B. J. Whitaker, and M. J. Winter (1994), "Chemical Applications of the World-Wide-Web," *J. Chem. Soc.*, Chem. Commun. , p. 1907.

10. Casher, O., G. Chandramohan, M. Hargreaves, P. Murray-Rust, R. Sayle, H. S. Rzepa, and B. J. Whitaker (1995), "Hyperactive Molecules and the World-Wide-Web Information System," *J. Chem. Soc.*, Perkin Trans **2**, p. 7.

11. Baggott, J. (1995), "Biochemistry On The Web," *Chemical & Engineering News* **73**, p. 36.

12. Schwartz, A. T, D. M. Bunce, R. G. Silberman, C. L. Stanitski, W. J. Stratton, A. P. Zipp (1994), "Chemistry In Context—Weaving The Web," *Journal Of Chemical Education* **71**, p. 1041.

13. The Internet Comes of Age (1995), *PC Magazine*, May 30, p. 19.

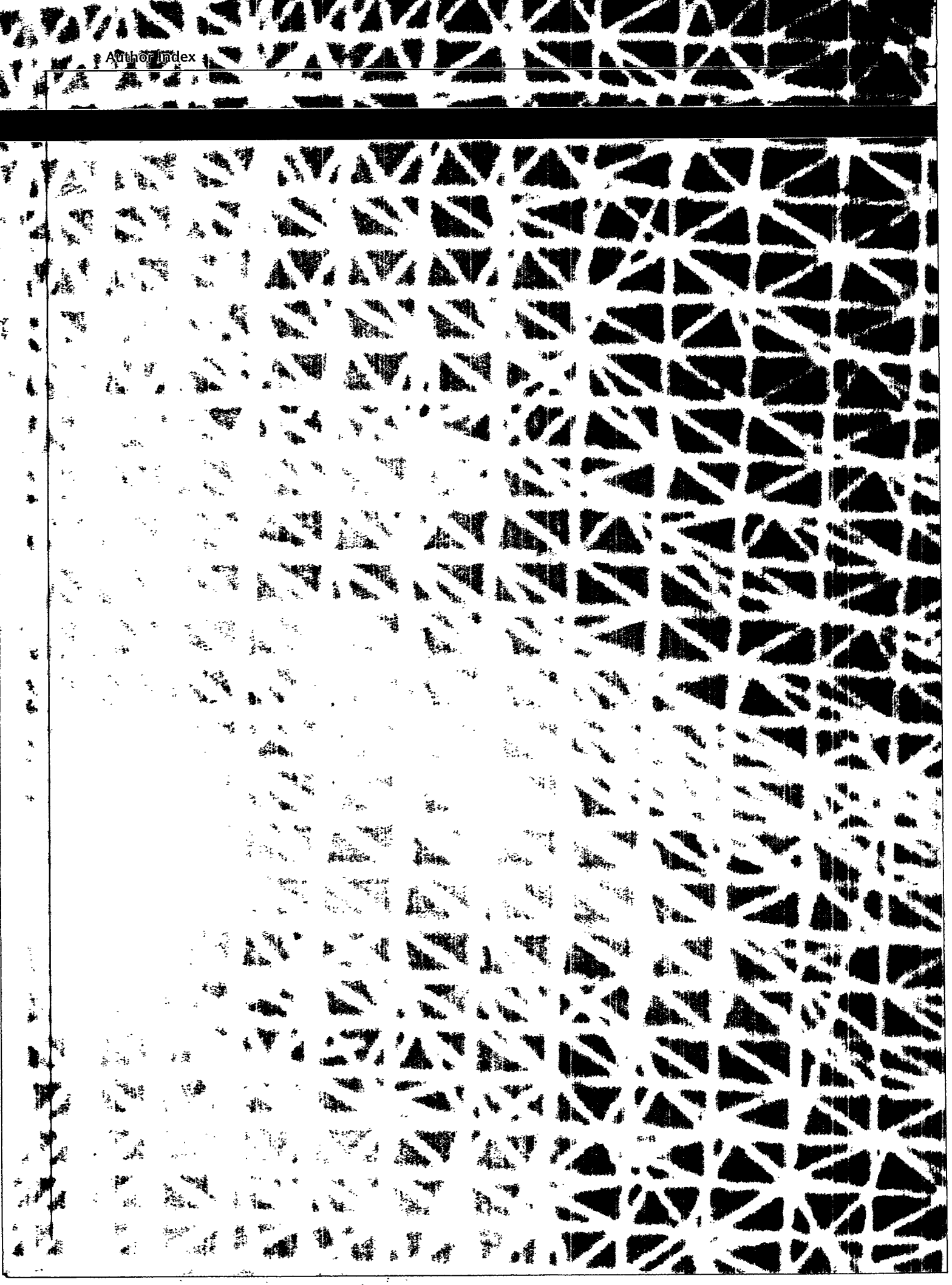
14. Holznagel, R.F. (1995), "The World Wide Web Top 1000," New Riders, Macmillan Co., Nov.

15. Meyer, M. (1995), "Surfing the Internet in 3-D," *Newsweek*, May 15, p. 68.

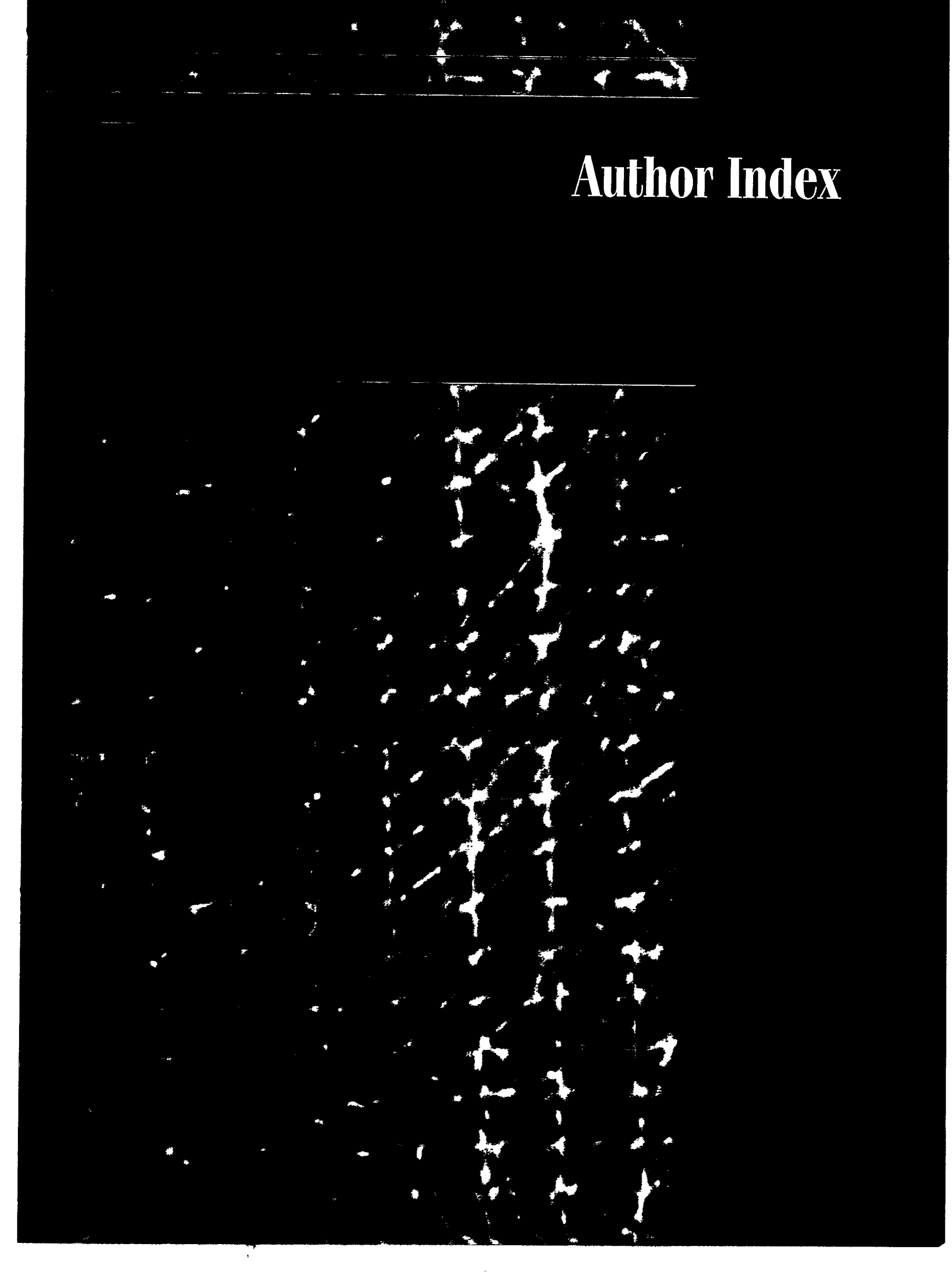
16. SILO User's Guide (1995), Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (UCRL-MA-118751).

17. Brugger, E. (1995), *Mesh-TV: a graphical analysis tool*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (UCRL-TB-115079-8).

18. Brown, S. (1995), "Portable Application Code Toolkit (PACT)," *PACT Home Page*, Lawrence Livermore National Laboratory, University of California, Livermore, Calif. (http://www.llnl.gov/def_sci/pact/overview.html).



Author Index



Azevedo, S.G. 6-17, 7-19, 7-39

Benett, W.J. 3-11
Bossart, P.-L. 7-19
Brand, H.R. 1-7, 7-19, 7-31
Brewer, L.R. 3-5
Brugger, E. 8-17
Buchanan, C.J. 8-1
Bumpas, S. 2-25

Campbell, G. 5-1
Candy, J. 7-13
Caporaso, G.J. 1-25
Cavitt, R.E. 8-1
Chinn, D. 7-13
Ciarlo, D.R. 3-7, 3-21, 3-23

Del Grande, N.K. 7-1
Deri, R.J. 3-1
DeTeresa, S.J. 5-7
DiJaili, S.P. 3-1
Dingus, C.W. 2-13, 5-15
Durbin, P.F. 7-1

Eveleth, G.G. 3-21

Fix, B.L. 4-1
Folta, J.A. 3-11

Genin, F.Y. 1-7
Gerhard, M.A. 2-25
Gorvad, M.R. 7-31
Grant, J.B. 1-33
Groves, S.E. 5-7

Haigh, R.E. 8-7
Hamilton, J. 3-19
Hanks, R.L. 6-1
Havstad, M.A. 5-15
Hayes, J.P. 4-5
Hiromoto, D.S. 5-1, 5-27

Jackson, D.R. 1-33
Jankowski, A.F. 4-5
Jiangning, Q. 2-17
Johansson, E. 1-19
Johnson, G.W. 6-7
Johnson, W.A. 1-33
Jordan, C.W. 8-1
Jorgenson, R.E. 1-33

Kallman, J.S. 1-1, 8-17
King, W. 5-1
Kirbie, H.C. 6-1
Koopman, S. 8-11
Kornblum, B.T. 2-13
Kozlowski, M.R. 1-7
Krulevitch, P.A. 3-15, 3-23

Lager, D. 8-11
Landram, C.S. 2-13
Langlois, R.G. 3-21
Lee, A.P. 3-15, 3-19, 3-23
Lehew, S. 3-23
Lesuer, D.R. 5-21, 5-27, 5-35
Logan, R.W. 2-31

Maker, B.N. 2-1, 2-17
Makowiecki, D.M. 4-5
Maltby, J.D. 2-13
Mariella Jr., R.P. 3-21
Martz, H.E. 7-19
Masquelier, D.A. 3-5, 3-21
Mast, J.E. 7-39
Mayhall, D.J. 1-13
McCammon, K.G. 8-7
McEwan, T.E. 6-17
McKernan, M.A. 4-5
Miller, D.E. 1-13
Molau, N.E. 1-7

Nelson, S.D. 1-19, 8-17
Newton, M.A. 6-1
Ng, W.C. 1-25, 6-7
Nguyen, D.T. 2-17
Nguyen, N. 5-1
Nieh, T.G. 5-21
Niven, W.A. 8-1
Northrup, M.A. 3-15, 3-23

Obegi, G. 8-1

Patterson, F.G. 3-1
Pavel, G. 8-1
Perkins, D.E. 7-1
Poole, B.R. 1-25
Puso, M.A. 2-1, 2-7

Raley, N.F. 3-7
Ramsey, P.B. 3-15
Ratowsky, R.P. 1-1
Riddle, R. 5-1
Rockway, J.W. 1-33

Sam, D.D. 2-39
Sargis, P.D. 8-7
Sawicki, R.H. 1-7
Schauer, D.A. 2-1
Schneberk, D.J. 7-19, 7-31
Shakeri, S. 3-21

Shang, C.C. 1-7, 1-25
Sharpe, R.M. 1-33
Sherby, O.D. 5-35
Sommer, S.C. 2-25
Spicer, J. 7-13
Syn, C.K. 5-27, 5-35

Taleff, E.M. 5-35
Taylor, J.M. 8-1
Thomas, G. 7-13
Trevino, J. 3-19, 3-23

Venkatesh, M.C. 3-21

Warhus, J.P. 6-17
Warren, F.E. 8-1
Weiss, J.A. 2-1
Wien, W. 5-1
Wilton, D.R. 1-33

Yee, J.H. 1-13
Yu, C.M. 3-3