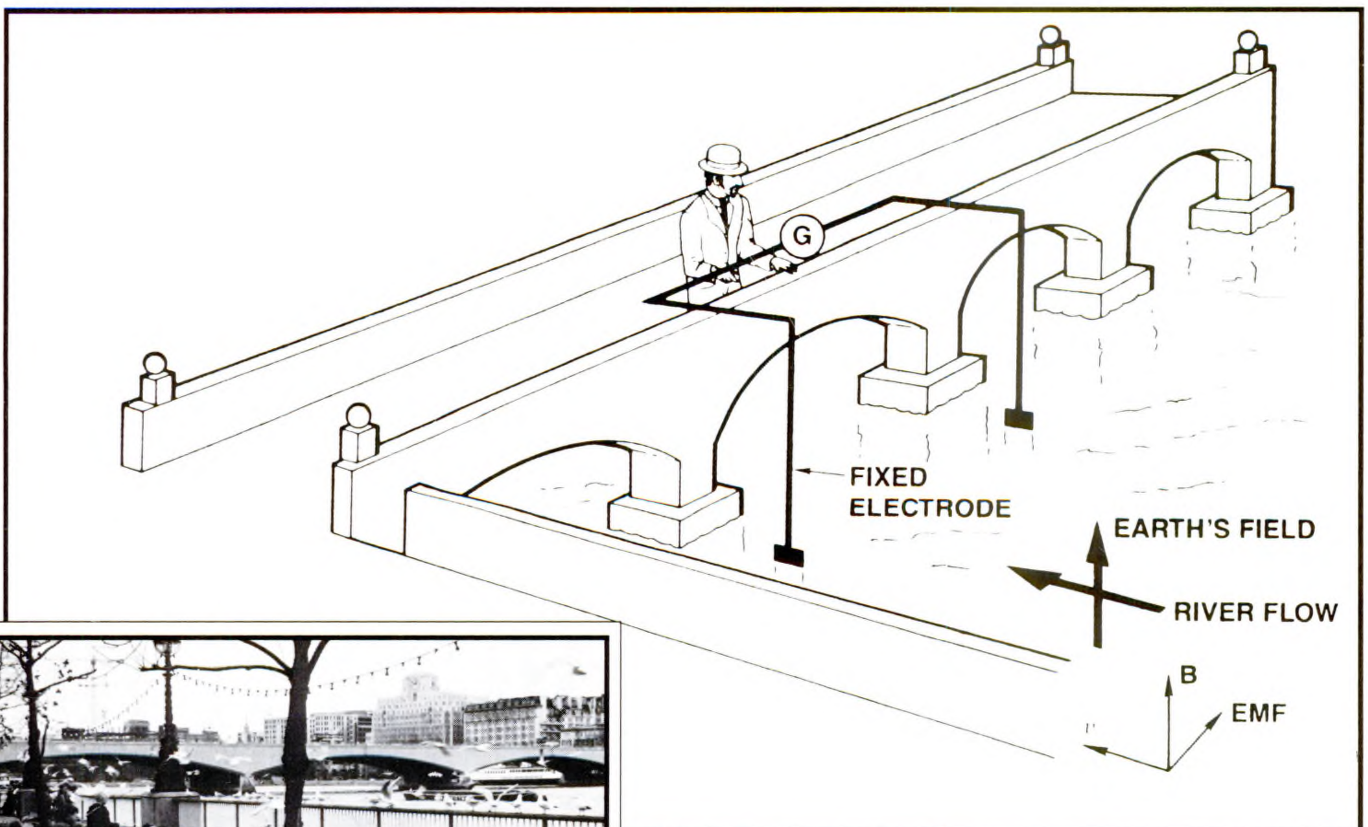
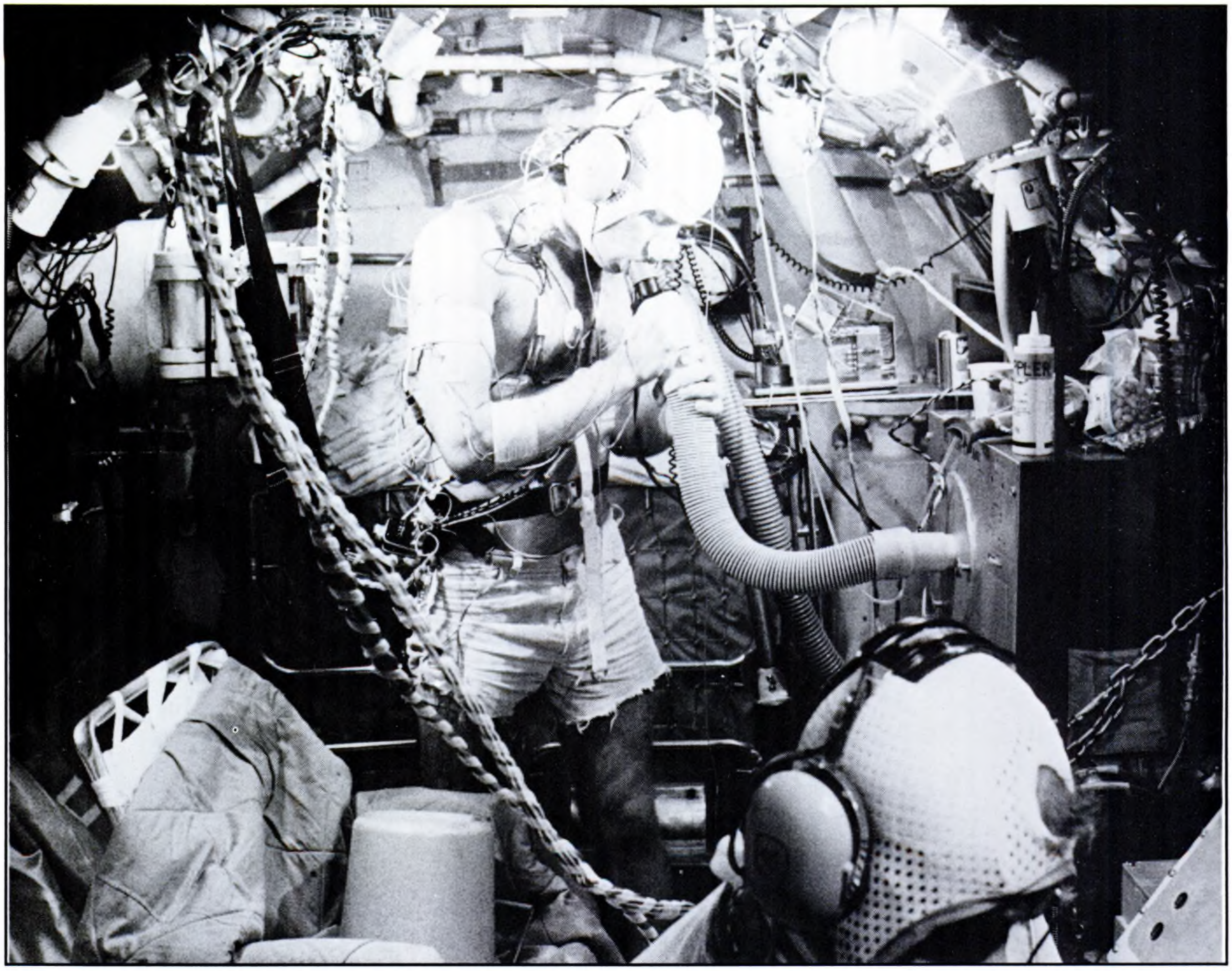
Naval Research Reviews

Office of Naval Research
Four / 1990
Vol XLII





Human Performance in the Underwater Environment

Diver Craig R. Carlson (background) performing measurements of flow resistance helium in the lung at 1600-foot depth at the University of Pennsylvania's Institute for Environmental Medicine in 1973. Michael R. Phillips (foreground) is engaged in mental, visual and auditory measurements. This was part of a long-term program at the Institute of Environmental Medicine, University of Pennsylvania, supported by the Office of Naval Research (ONR).

ONR research in the life sciences during the 1960's and 1970's provided fundamental knowledge concerning factors affecting man's ability to perform useful work in the underwater environment. The research focused on sensory, perceptual, psychomotor and cognitive processes as they are affected by hyperbaric pressure, cold, mixed breathing gases, optical properties of water, and a tractionless and viscous medium. This research led to: an understanding of changes in visual acuity, tactile and auditory thresholds, speech and hearing degradations, memory constraints and the relationships between environmental factors and complex task performance. Developments stemming from this research were design principles and guidelines for helium speech unscramblers, noise cancelling microphnones, underwater visual displays and training regimes for cold water exposure. In addition this research supported major elements of the Navy's Man-in-the-Sea Program including the Sea Lab and Tektite underwater habitat projects, which were sponsored by ONR.

Naval Research Reviews

Office of Naval Research
Four / 1990
Vol XLII

Articles

Departments

About the Cover

The insert photograph shows the Waterloo Bridge, London on a December morning. This is where Michael Faraday performed his pioneering experiment, over a century and a half ago (1832), to demonstrate the science of magnetohydrodynamics (MHD). Utilizing the Earth's magnetic field, and the flowing water of the River Thames as the conductor moving in the magnetic field, he perceived that an electric field would be induced. The sketch shows Faraday with the electrodes dipped in the water, trying to measure the electric field generated. (See article on page 2.)

Naval Research Reviews publishes articles about research conducted by the laboratories and contractors of the Office of Naval Research and describes important naval experimental activities. Manuscripts submitted for publication, correspondence concerning prospective articles, and changes of address, should be directed to Code OPARI, Office of Naval Research, Arlington, VA 22217-5000. Requests for subscriptions should be directed to the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20403. *Naval Research Reviews* is published from appropriated funds by authority of the Office of Naval Research in accordance with Navy Publications and Printing Regulations. NAVSO P-35.

Magnetohydrodynamics – Its Potential Defense Applications and Technical Issues

Gabriel D. Roy
Office of Naval Research

Introduction:

Magnetohydrodynamics or MHD is the science of conducting fluids interacting with electric and magnetic fields. This interdisciplinary science, though dates back to the early eighteen hundreds for its inception, however, has received attention from the research community only in the past two decades. The principle of MHD (power generation) is illustrated in Figure 1. The MHD generator transforms the internal energy of a gas into electric power in much the same way as a piston engine generator or a turbogenerator does, and the basic physical phenomena is the same. In the first two devices, the energy of the gas is converted into the motion of solid conductors in a magnetic field. In the MHD generator, the gas itself is the conductor. As this conducting gas moves (by expanding through a nozzle) through a magnetic field, an electromotive force is generated and a current flows according to Faraday's laws of induction.

On the other hand, if an electric field is applied to a conducting fluid moving in a magnetic field, mechanical energy is produced (as in a conventional electric motor), which could be realized as an increase in pressure, acceleration of the flow or thrust.

MHD devices have certain inherent advantages, general and specific to the applications. There are no moving parts, and hence the operation will be quieter. They are easy to start and stop, and are compact. A wide variety of fuels may be used. These qualify MHD as a potential candidate for several military applications. But there are several developmental issues as well. The purpose of this paper is to look at briefly the history of MHD, its potential military applications, and major technical issues in designing and building military MHD systems. It is not intended to address the governing equations and their solutions, nor the specific details to solve the issues.

History of MHD:

At the time of his discovery of the principle of electromagnetic induction, Michael Faraday, understood that it would apply to conducting fluids as well as to solids. Perceiving that the earth would produce the same effect as a magnet, he went on to experiment using earth's magnetic field. His first experiment (1832) was in Kensington lake to demonstrate electromagnetic induction with earth's magnetic field and water. But his pioneering experiment followed later that year in River Thames under the Warerloo Bridge, where he tried to demonstrate the interaction between flowing water and the earth's magnetic field (See figure in cover page). Faraday lacked the knowledge of the electrical properties of fluids. He wrote in the Bakerian lecture series, ".... although the electricity obtained by magnetoelectric induction in a few feet of wire is but small intensity and has not yet been observed except in metals and carbon in particular state, still it has the power to pass through brine; and as the increased length of the substance acted upon produces increased intensity, I hope to obtain effects from extensive masses of water, though quiescent water gave none." MHD was born!

Research in MHD proceeded at an intermittent pace. In 1860, Lord Kelvin experimented with liquid metal MHD. In the 1910's MHD patents began to appear, and addition of salt to water to increase the conductivity, was tried. In the 1940's Westinghouse was involved in MHD research (B. Karlovitz), and early hall generator was built, but still the knowledge of ionized gases was insufficient[1]. In the 1950s A. Kantrowitz (Cornell, AVCO) built the early generators, and also studied the electrical properties of thermally ionized gases and their interactions with a magnetic field[2]. S. Way (Westinghouse) in the early 1960's built an MHD generator and demonstrated generation of power from seeded combustion products[3]. The development was focussed on space and military applications. The MHD focus, however, shifted to power plant applications in the 1960's and R&D was supported by the Office of Coal Research (Department of the Interior), National Science Foundation, Department of Defense, Atomic Energy Commission, National Aeronautics and Space Administration, and the Tennessee Valley Authority. The MHD research efforts peaked in the 1970s and 80s by increased support from the Energy Research and Development Administration (later the U. S. Department of Energy), and the Strategic Defense Initiative Organization in support of MHD power plant demonstration, and space-based MHD power systems respectively. There is an international interest in MHD today (through the past decade), with the participation of USSR, Japan, Germany, India, Italy, Israel, France, Poland, Australia and China.

Development of MHD accelerators and pumps received attention in the last two decades. However, research of maritime MHD propulsion was carried out extensively during the 1960's. The concept of the MHD thruster, as marine vehicle power, was first evaluated in 1962 by O. M. Phillips. His analysis was based upon a 600 ft submarine utilizing a 0.6 Tesla (6000 gauss) magnetic field and setting the electrical conductivity of sea water at 4 mhos/m (1/m). The vehicle achieved a theoretical speed of 10 knots, with an efficiency of 8%. According to Phillips' calculations, to obtain a 30% effi-

Figure 1

Principle of MHD Power Generation

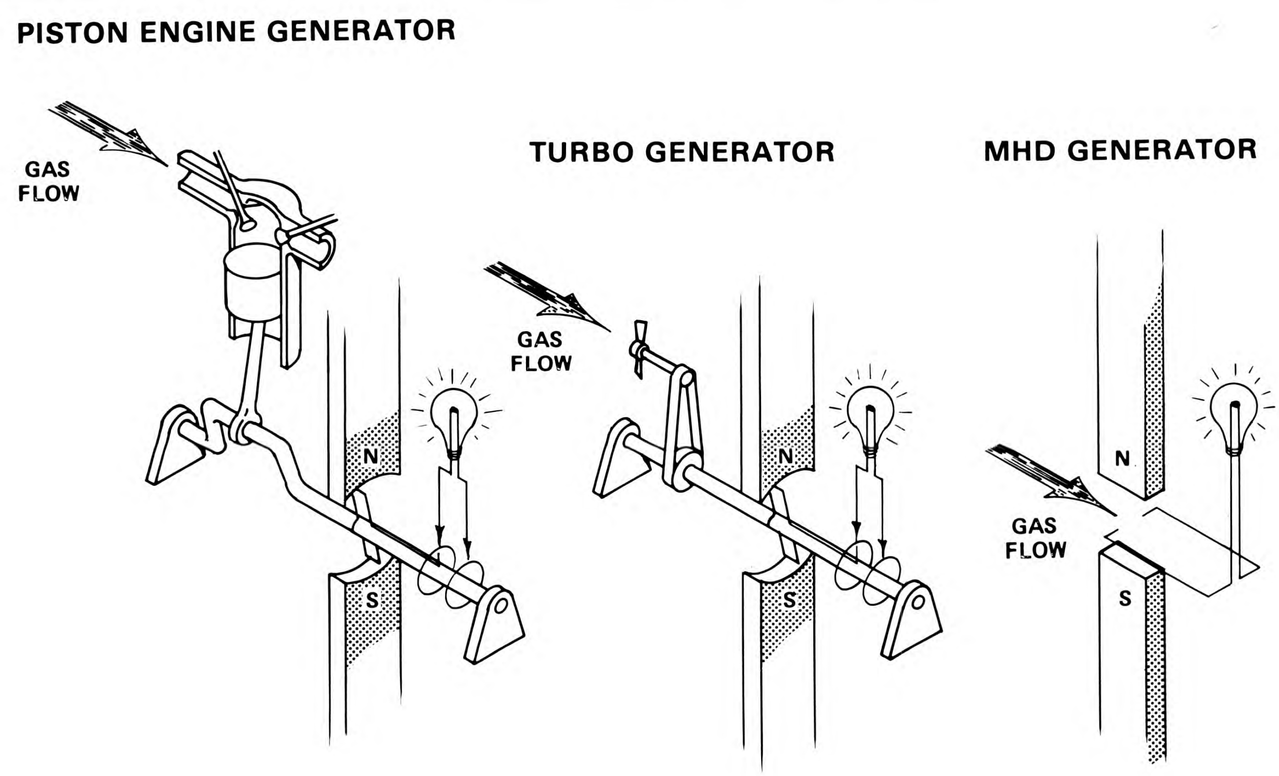
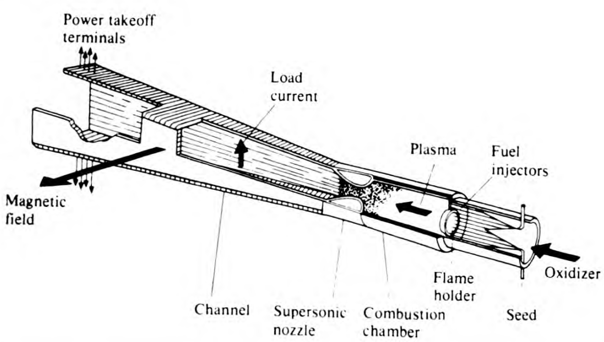


Figure 2

Schematic of a Linear Plasma MHD Generator (From Proc. Am. Power Conference, 1974)



cient system at 30 knots, a 2 Tesla magnet would be required[4]. The practical demonstration of this concept came about in 1966 when S. Way (University of California, Santa Barbara) constructed the EMS-1, a 10 ft. long, 900 lb, external field subscale submarine, with a magnetic field of 0.015 Tesla and 30 Volt internal power supply, and propelled this model at a velocity of 1 ft/sec for 20 minutes[5].

Recent advances in cryogenics have made moderate field superconducting magnets feasible for MHD propulsion. As early as 1979, A. Iwata and Y. Saji constructed the ST-500; a 12 ft long wooden model which was propelled by a 2.0 Tesla external field thruster, and obtained a velocity of 2 ft/sec[6]. The low speeds and efficiencies obtained in these demonstration devices are due to the low electrical conductivity of sea water and the lack of high field superconducting magnets.

Recently, there is a resurgence of interest in MHD propulsion, particularly due to the recent advancements in superconducting magnet technology and the development of high temperature superconductors. The Japan Foundation for Shipbuilding Advancement is actively pursuing the design, construction and testing of a prototype experimental ship using MHD thrusters. The recent movie, "The Hunt for Red October", which depicts an MHD propelled submarine used by the Soviets, has hit almost every leading news paper, with special attention given to MHD!

The Office of Naval Research (ONR) has provided open forums to the scientific community to evaluate the state-of-the art in MHD propulsion, to identify the challenging issues, and to provide the technical approach to realize MHD propulsion for undersea vehicles[7,8]. Research and development programs on MHD propulsion are currently sponsored by ONR, ONT and the Defense Advanced Research Projects Agency (DARPA).

Defense Applications and Technical Issues:

1. Plasma MHD

The MHD generator is a direct energy conversion device with practically no moving parts. It can be started and shut down quickly, and is relatively lighter and smaller compared to conventional power generators. A wide variety of fuels can be used to generate the combustion plasma. These advantages of MHD generators make it a favorable choice for space power applications. Figure 2 shows a schematic of a linear plasma MHD generator. The force and field vectors, and electrical connections are shown in Figure 3. Here the Lorentz force (J x B) acts opposite to the flow direction (whereas in a thruster, the Lorentz force provides the thrust). Considerable increase in the enthalpy extraction has been demonstrated in the past decade. Figure 4[9] shows the plot of specific energy extraction verses the product of interaction parameter and efficiency. (Interaction parameter is an important dimensionless group given by Iu or Ip = The ratio of the MHD forces to the fluid momentum or pressure force.) At very high interactions, MHD is suited for space power. The SDIO is actively pursuing MHD as a power source for space based systems. Both the linear

Figure 3

Linear Generator Geometry and Electrical Connections

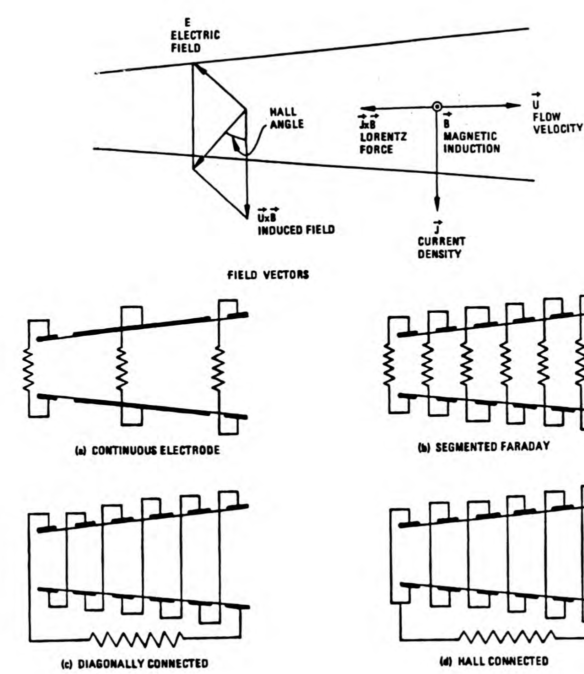
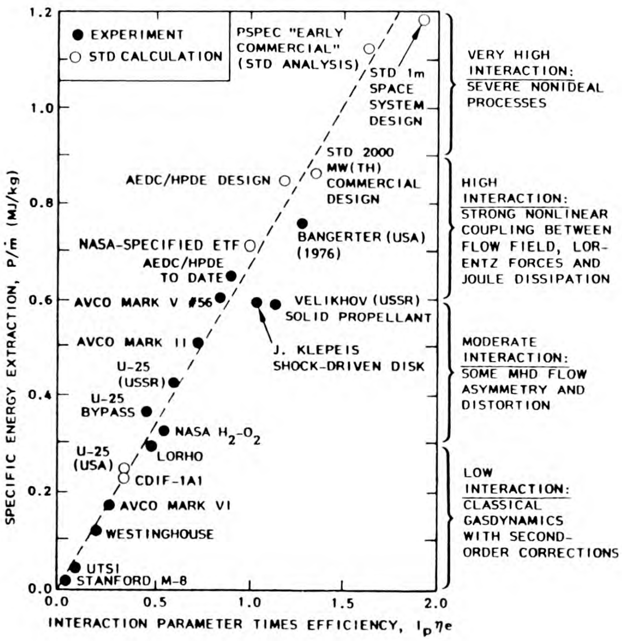


Figure 4

Specific Energy Extraction of Various MHD Devices



generator or the disk generator configuration can be used. In the linear channel (Figure.2), the exhaust gases will generate a thrust, and the diffuser design should take into account the elimination of this thrust. In the disk channel (Figure 5), the thrust is neutralized due to the circumferential exhaust. Under SDIO contracts, linear geometry is been pursued by STD Research Corporation, CA, whereas TRW Inc., CA, utilizes the disk design.

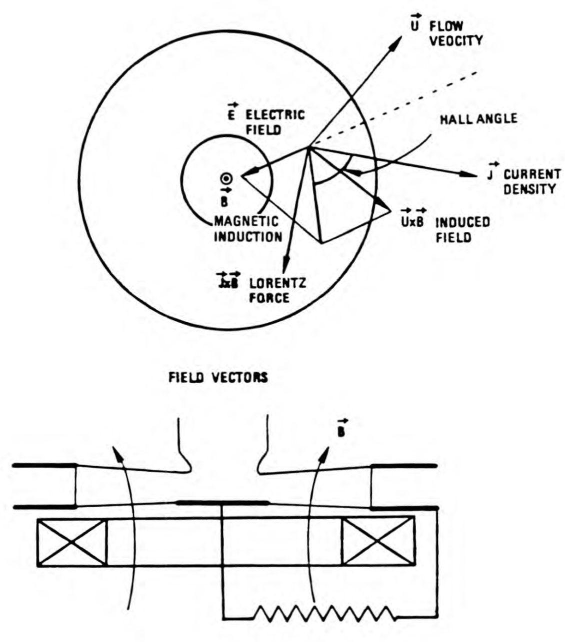
For directed energy systems, where high energy pulses are required for a relatively shorter periods of time, MHD offers potential. It has advantage over magnetic flux compression due to the flexibility of operation and repeated use. No mechanical deformation is required. For relatively long pulses combustion driven MHD system can be used together with a suitable pulse power switch to transfer the power at the required repetition rate to the weapon[10].

For extremely high power, but relatively short duration, detonation-driven MHD could be considered. Though the explosive driven linear MHD generator has been studied since the mid 1960's in the United States and in the Soviet Union, the last notable effort in the United States ceased in the early 1980's, whereas, this effort seems to continue as part of the national pulsed power program in the Soviet Union. The principle of this device is to cause a electrically conducting plasma confined between two electrodes to move across an externally applied magnetic field, as shown in Figure 6. The induced electric field will effect a current flow through the circuit, and a useful electrical energy pulse through the load results. Since energy is to be released very rapidly, explosives are detonated to produce the plasma. Pulse power of 10 giggawatts for durations on the order of 100 microseconds have been reported using linear explosive MHD generators[11].

The plasma phenomena in these devices are extremely complex and are not well understood. More detailed computational and simulation methodologies, and carefully controlled experimentation, with sophisticated diagnostics are needed to determine means for improving the inherently low efficiency of these devices. The plasma slug is finite rather than a continuous flow , and the process is transient and not a steady state process. Further, the explosive generator operates in the high magnetic Reynold's number regime, which means that the induced magnetic field accompanying the current in the plasma and electrodes is of the same order of magnitude as the applied field, and affects the resulting efficiency. A circuit analysis of explosively driven MHD was performed by H. J. Schmidt and associates[12].

Figure 5

Disk Channel Geometry and Electrical Connections



2. MHD Seawater Propulsion

As pointed out earlier, recently there has been a renewed interest in MHD seawater propulsion. The current interest in the U. S. is the application of MHD thrusters to submarines, torpedoes and autonomous underwater vehicles. The elimination of the propeller in the MHD thruster renders possible higher speeds that would be limited by cavitation in propeller-driven vehicles. The elimination of rotating parts and reduction gears can largely reduce the noise generation and acoustic signatures. Further multiple thrusters would increase the survivability in a wartime scenario, in addition to increasing the maneuverability. Recent advances in the superconducting magnet technology makes it possible to build reliable magnets in the 6-10 Tesla field range, offering the possibility of a descent efficiency. The success of the high temperature superconductor and magnet development, and light weight, high strength composite materials will have a beneficial effect on the efficiency, size and weight of the propulsion system.

Figure 6

Explosive MHD Generator Schematic

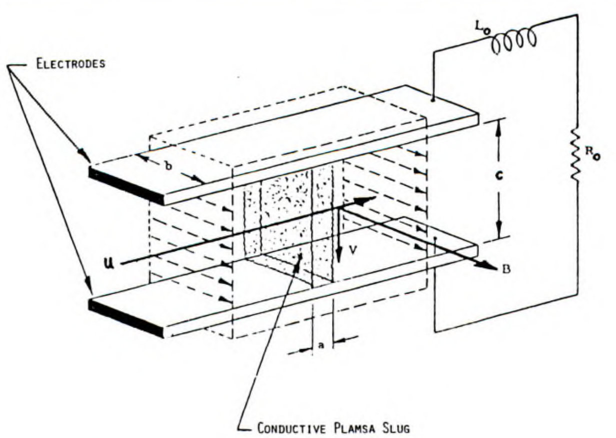


Figure 7

MHD Seawater Thruster Schematic

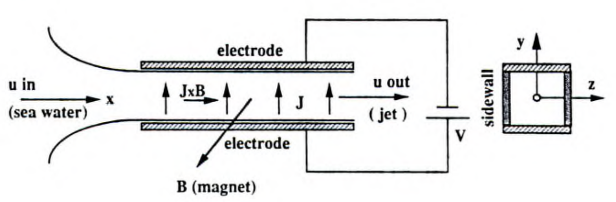


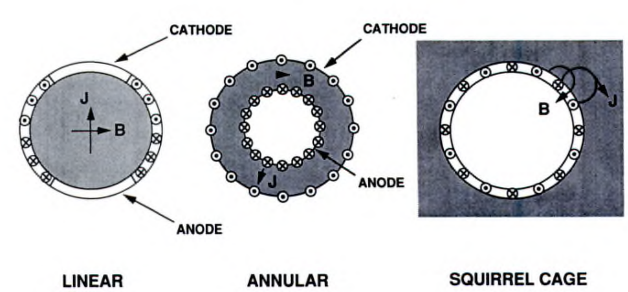
Figure 7 shows a schematic of a linear MHD seawater thruster. Its operation depends on the seawater conductivity to carry the electric currents. The seawater enters a rectangular duct (could be circular or annular), with continuous electrodes on the top and bottom (could be on the sides), placed in a magnetic field. When energy is supplied to the duct by an external electric field, the resulting electric current J, interacts with the magnetic field B, to produce a driving Lorentz force (JxB) that pushes the seawater through the duct. Therefore the MHD thruster can be viewed as an electromagnetic pump that ejects the seawater at a higher velocity than the intake velocity, thus providing the thrust.

Three major thruster configurations are possible. These are shown schematically in Figure 8[13]. The area where the seawater flows is shown shaded. The resulting JxB force is perpendicular to the plane of the paper. The linear thruster is the simplest, but has a relatively large magnetic leakage. The annular configuration minimizes this leakage, and also increases the volumetric flow rate (which increases the thrust), but however, the wetted area is almost doubled, and hence the drag losses. The third one, where the water surrounding the vessel is energized, and the thrust is obtained external to the vessel does not add any additional wetted area for the thruster (this type was used by S. Way in his demonstration experiment). The magnetic fringe fields and the energy losses to the surrounding water make this less attractive.

AVCO Research Laboratory chose the annular configuration, externally hull-mounted in particular (Figure 9), and performed extensive parametric analysis[13]. The internal flow (linear) MHD thruster proposed by J. C. S. Meng for torpedoes is shown in Figure 10. A porous shell in the intake area can serve for turbulent boundary layer control as well. Meng also performed simulation and performance characteristics studies

Figure 8

MHD Thruster Configurations



of clustered dipole and racetrack torroid magnet configurations for submarine propulsion[14].

However, there are several technical issues that have to be addressed in order that MHD thrusters can be developed for submarine or torpedo application. The major issues are addressed below:

(a) Conductivity:

The MHD thruster relies on the conductivity of seawater to carry the electric current. The thrust produced is proportional to the conductivity. The seawater conductivity is very low (about 4 mho/m), orders of magnitude lower than that of metal conductors. The salinity of seawater, as well as the temperature vary from ocean to ocean, and the conductivity varies with both. Figure 11 shows the average mean temperature, and the variation of conductivity with salinity and temperature is shown in Figure 12. It can be seen from these figures that a submarine can encounter over 50% change in conductivity, which could translate in similar changes in speed!

There is nothing one can do to increase the conductivity of seawater for submarine application. However, for a torpedo, where the mission duration is very small, and the electrode life beyond that is not of any significance, sacrificial electrodes that can chemically or electrically erode and mix with the thruster flow, offer potential to enhance the conductivity and speed (also efficiency). G. L. Holleck performed an experimental study of sacrificial electrodes, and found lithium to be an attractive candidate. The conductivity enhancement obtained in his experiment (doubling) is shown in Figure 13[15]. The signatures associated with this process need to be investigated.

(b) Electrolysis and Bubble Generation:

Seawater is an electrolyte, and hydrogen and chlorine gas will be generated at the cathode and the anode respectively. These gas bubbles could mix or be carried along with the flow stream if they get out of the boundary layer. The bubbles in the boundary layer can accumulate and decrease the current flow and impair the performance. Further, the noise generated by the bubbly turbulent flows could defeat the major advantage of MHD thruster, viz, quieter operation. Ideally one should look for a "gasless electrode", or minimize the bubble generation and its interference with current flow either by the choice of electrode materials to control the electrolysis, or by controlling the fluid mechanics of the flow.

Two electrode arrangements (top and bottom, or the sides) may be considered for a linear thruster (Figure 14). Experimental attempts have been made to quantify the bubble generation and its effect on current transport[16-18]. Tempelmeyer experimented with various electrodes and applied voltages to study the decay of current flow with bubble deposition[16]. See Figure 15. Koester used a prototype gated, intensified video camera to study the bubble generation and transport[15]. Figure 16 shows two video images taken at two current levels.

Figure 9

MHD Thruster Mounting Concepts

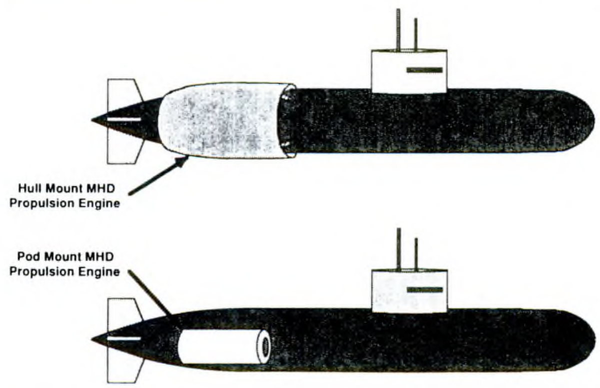


Figure 10

Torpedo Propulsion with Internal Flow Linear MHD Thruster

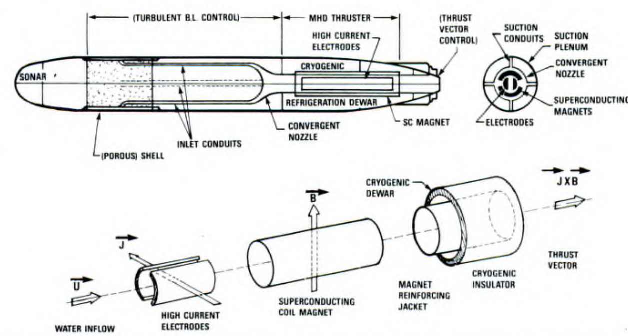
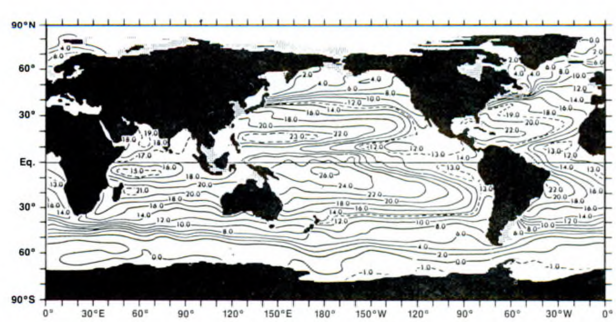


Figure 11

Average Mean Temperature of the Oceans



Annual mean potential temperature (°C) at 150 m depth.

The mechanics of bubble generation, its transport, and the acoustic implications (both transmitted in the wake, as well as structural), with the magnetic interaction need to be investigated. This highly complex interdisciplinary research will be a challenge to the scientific community. The fundamental understanding could lead electrode design that can alleviate these problems.

Figure 12

Variation of Seawater Conductivtiy with Salinity and Temperature

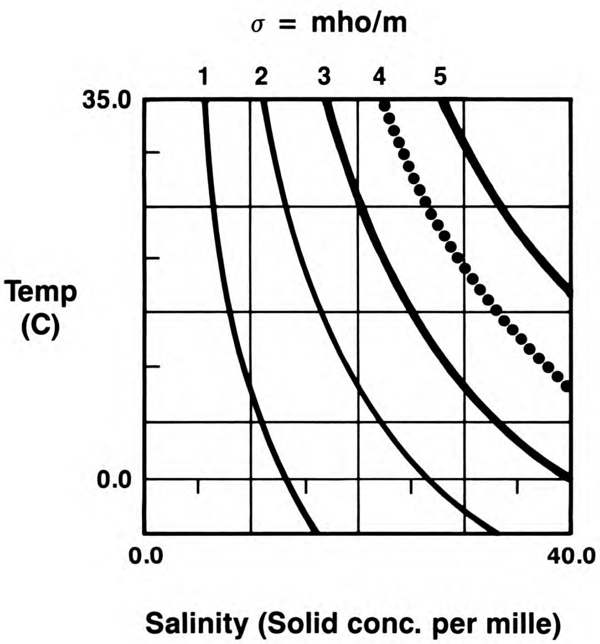
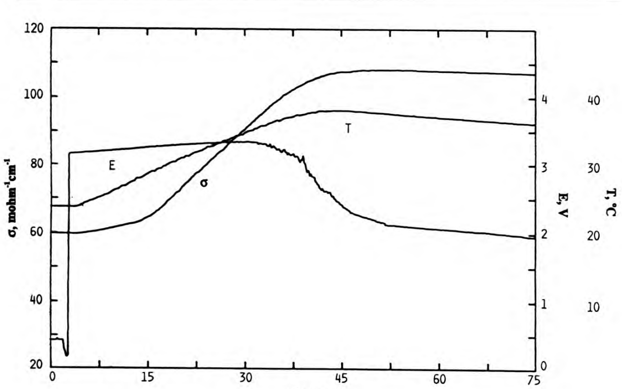


Figure 13

Conductivity Improvement Using Lithium in Simulated Seawater



(c) Magnet Size and Signatures:

As mentioned earlier, the performance of the MHD thruster can be improved by increasing either the conductivity of seawater or the magnetic field. For submarine propulsion, the latter is the only possible approach. Increasing the B field not only requires complicated cryogenic systems, but also extremely strong materials and heavy structure to withstand the high stresses caused by the strong magnetic field. Further there is a possibility of high fringe magnetic fields and signatures. With the exception of NMR magnets (truck mounted), the superconducting magnets built so far have been used for stationary applications, and not in a violent sea environment! Hence the issues are: light weight, survivability and signatures. Clever magnet designs with opposed magnetic field lines that cancel each other, without leaving fringe fields have been proposed as the approach by P. Marsten[19]. Aluminum or other light weight conductor stabilizer, supercritical-cooled or force-cooled helium as opposed to pool boiling, and designing the superconductor with high critical current density, are among the recommendations made by B. Wang[20].

(d) Efficiency:

Though reasonable efficiencies have been projected by analysis and computations, the demonstrated efficiency of MHD thrusters has been very low. To improve the efficiency of the thruster, one has to understand the various loss mechanisms that take place in and around the duct, and determine methods for improvement. Fluid mechanics plays a very important role for undersea vehicles as seawater is the environment. In addition, for an MHD thruster, the working fluid is also water. Thus the MHD thruster, in several aspects, is similar to an airbreathing engine. As a result, "fluid mechanics takes an added dimension to the traditional role which fluid mechanics has played in ships"[21]. The fluid mechanics of the duct flow is complicated by the electrolysis and MHD effects.

Figure 14

Linear Thruster Electrode Arrangements

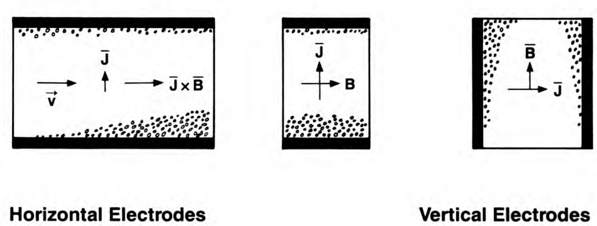


Figure 15

Conductivity Decay with Bubble Deposition

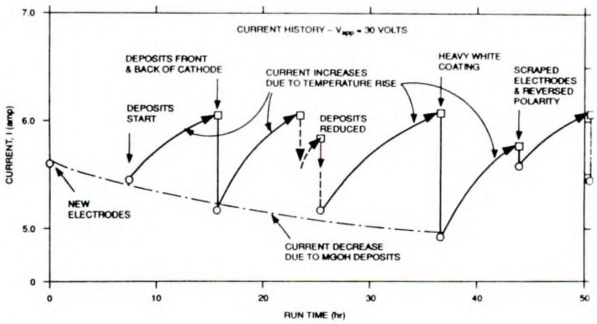
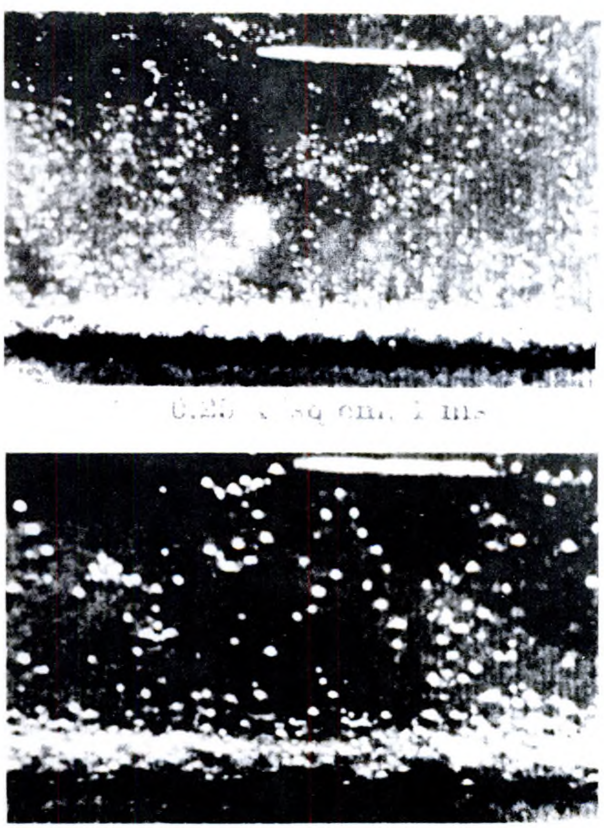


Figure 16

Video Images of Bubble Generation at Different Current Densities



J = 2.0 A/sq cm, 1 ms

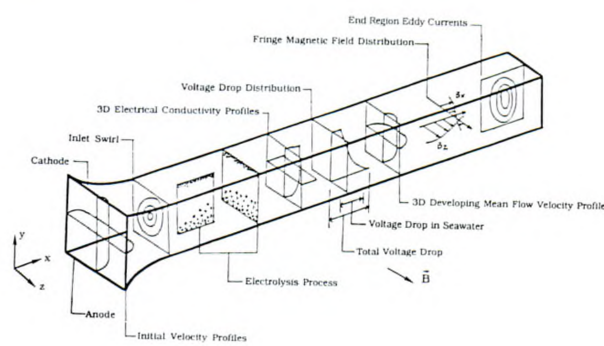
Though reasonable efficiencies have been projected by analysis and computations, the demonstrated efficiency of MHD thrusters has been very low. To improve the efficiency of the thruster, one has to understand the various loss mechanisms that take place in and around the duct, and determine methods for improvement. Fluid mechanics plays a very important role for undersea vehicles as seawater is the environment. In addition, for an MHD thruster, the working fluid is also water. Thus the MHD thruster, in several aspects, is similar to an airbreathing engine. As a result, "fluid mechanics takes an added dimension to the traditional role which fluid mechanics has played in ships"[21]. The fluid mechanics of the duct flow is complicated by the electrolysis and MHD effects. Even in a linear seawater thruster, an inherently three-dimensional phenomena occurs. This is schematically shown in Figure 17[21]. Needless to say that a full three-dimensional analysis and numerical modelling of MHD thruster flow with electrolysis effects, structural interaction, and inlet and wake effects is in order. E. Doss performed preliminary 3-D numerical computations[22], but a full computation is still to come.

Conclusions

Magnetohydrodynamics technology offers significant potential device applications related to defense. However, there are a number of technical challenges that need to be addressed in the research and development levels. Even with the international attention over the past two decades, it is still not a mature technology. This is mainly due to the fact that the efforts were not focussed. A clear change in trend is seen. Understanding at microscopic levels to control macroscopic processes are still needed in some areas. A well coordinated parallel research and development effort could make the application of MHD to practical working systems a reality.

Figure 17

Three-Dimensional Phenomena in a Linear Thruster



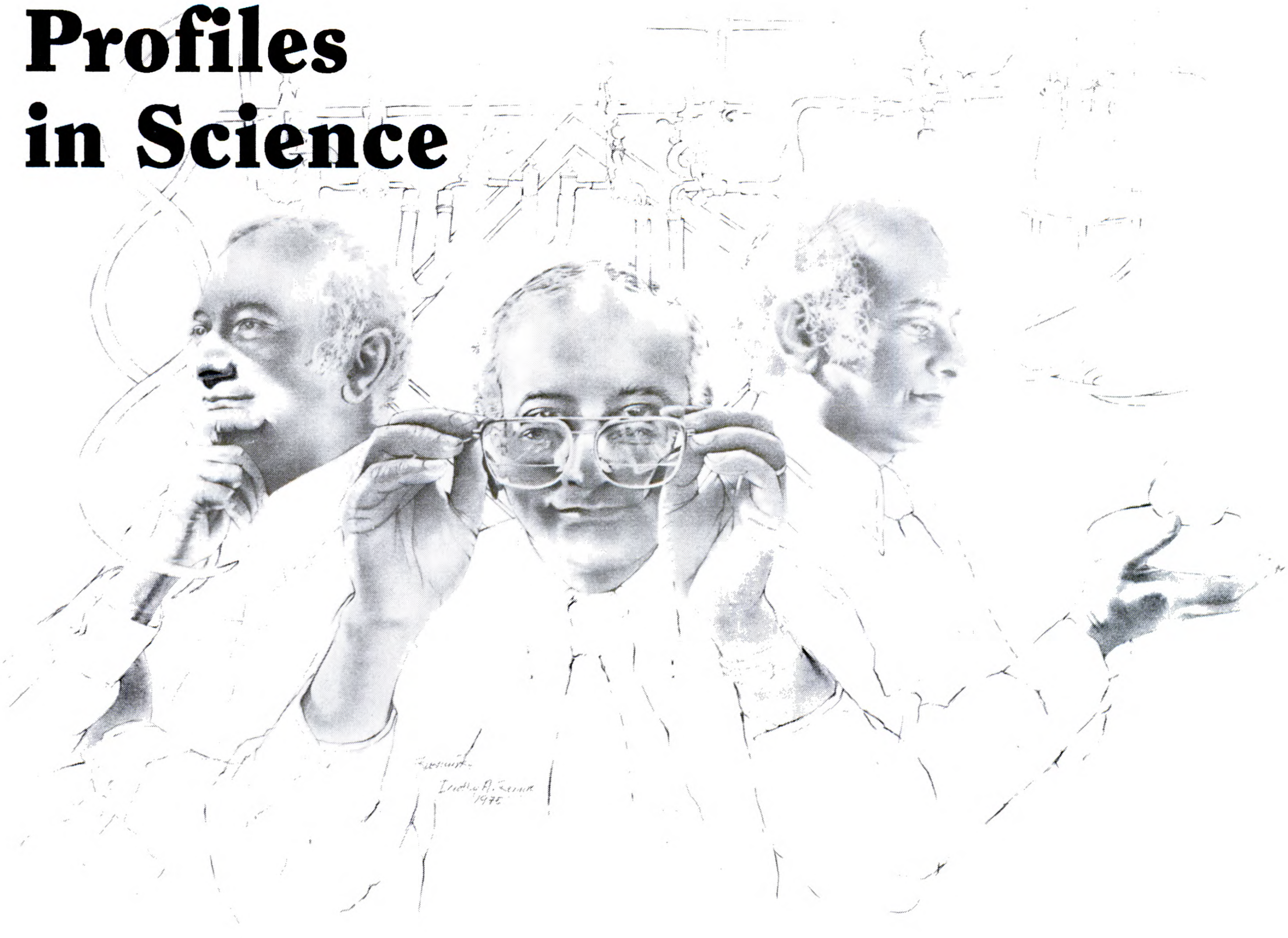
Biography

Gabriel D. Roy Received his Ph.D in Engineering Science from the University of Tennessee. He served as a faculty member for 12 years and as a researcher and research manager in university and industry for 15 years. At the Office of Naval Research Dr. Roy manages the propulsion and pulse power physics programs. He has over 30 publications and two patents, and his research interests include plasma dynamics, MHD, pulse power and propulsion. He is a recipient of the Roll of Honor award from TRW Inc. Dr. Roy is a member of ASME and Sigma Xi Honor Society, and an associate fellow of AIAA. He is also an associate editor of the AIAA Journal of Propulsion and Power.

References:

1. Karlovitz, B., and D. Halasaz, "History of the K & H Generator -Conclusions Drawn from Experimental Results", Proc. 3rd Symposium in Engineering Aspects of Magnetohydrodynamics, 1964
2. Lin, S. C., E. L. Resler, and A. R. Kantrowitz. "Electrical Conductivity of Highly Ionized Argon Produced by Shock Waves", J. App. Phy., 26 (1): pp 83-95, 1955
3. Way, S. et. al., "Experiments with MHD Power Generation", J. Eng. Power, 83A:397, 1961
4. Phillips, O. M., "The Prospects of Magnetohydrodynamic Ship Propulsion", J. of Ship Research, pp 43-51, 1962
5. Way, S, "Propulsion of Submarines by Lorentz Forces in the Surrounding Sea," ASME Paper 64 WA/ENER7, 1964
6. Iwata, A., Y. Saji, and S. Sato, "Construction of Model Ship ST-500 with Superconducting Electromagnetic Thrust System," Proc. IECEC 8, pp 775-784, 1980
7. Roy, G. D., W. Jackson, "First ONR Workshop on Magnetohydrodynamic Underwater Propulsion, Washington, D. C, 1988
8. Roy, G. D. and L. Parnell, "Second ONR Workshop on Magnetohydrodynamic Submarine Propulsion", San Diego, CA, 1989
9. Maxwell, C and S. T. Demetriades, "MHD Pulsed Power", STD Report, 1988
10. Roy, G. D., "High Power, High Repetition Rate Switches: An Overview", IEE Colloquium on Pulsed Power, Digest No: 1988/134, London, UK, 1988
11. Gill, S. P., J. D. Shimmin, and J. D. Watson, "Principles of Efficiency in High Power MHD Experiments" Artec Report FR-165, Hayward, CA, 1984
12. Schmidt, H., L. W. Crawford, and J. N. Chapama, "A Study of Explosive MHD Generators", The University of Tennessee Space Institute Report, Tullahoam, TN, 1987
13. Swallom, D. W., "Magnetohydrodynamic Submarine Propulsion", Proc. Second ONR Workshop on Magnetohydrodynamic Submarine Propulsion, San Diego, CA, 1989
14. Meng, J. C. S.,"Superconducting Electromagnetic Thruster for Seawarwe Propulsion ," ibid
15. Holleck, G. L and G. S. Jones, "Sacrificial Electrodes for the Enhancement of Seawater Conductivity", Proc. Third ONR Propulsion Meeting, Middletown, RI, 1990
16. Tempelmeyer, K. E., "Electrical Characteristics of a Seawater MHD Thruster", Proc. Second ONR Workshop on Magnetohydrodynamic Submarine Propulsion, San Diego, CA, 1989
17. Brouillette, et. al, "Bubble Phenomena in Seawater Thrusters," ibid
18. Koester, J. K., "Low Noise Electrode Systems for MHD Propulsion ," ibid
19. Marsten, P. "On Magnet Design for MHD Propulsors" Presented at the Fist ONR Workshop on MHD Underwater Propulsion, Washington D. C., 1988
20. Wang, B., "Report on MHD Propulsion Magnet," Proc. Second ONR Workshop on Magnetohydrodynamic Submarine Propulsion", San Diego, CA, 1989
21. Wu, Y. C. L., "Some Fundamental Fluid Mechanics and Electrochemical Issues of MHD Seawater Thruster", ibid
22. Doss, E. D. and G. D. Roy, "Comparative Three-Dimensional Flow Analysis of MHD Generators and Seawater Thrusters" Proc. 28th Symp. Eng, Aspects of MHD, Chicago, Illinois, 1990

Profiles in Science



Rustum Roy

Rustum Roy has been an Office of Naval Research Principal Investigator for 40 years. His work spans the field of both new materials and processes. His most renowned research is in the areas of hydrothermal and high pressure anvil **processes**; perovskites, synthetic clays and zeolites; low expansion ceramics, and diamond films.

Professor Roy started his career as a postdoctoral student under Professor E.F. Osborn who was then funded by the first ONR grant to The Pennsylvania State University in 1948. This grant from ONR's Geophysics Branch was to prove very significant to geosciences and materials science. It provided the support for the discovery of the sol-gel process which is now a substantial branch of ceramic science. It also led to the development of hydrothermal phase equilibria and processing science and the design of vital research equipment in that field. The equipment developed under this grant became the world's standard for laboratory experimentation in geosciences.

In 1957 when Professors Roy and Tuttle extended hydrothermal high pressure range to high pressure Bridgman anvils from 5 kbar to 200 kbars, they became pioneers in diamond research, again with ONR support. Many of Penn State's achievements in high pressure physics and chemistry came from ONR and National Science Foundation support. During this early phase of high pressure work, Roy and his students synthesized more new high pressure phases than all other laboratories combined.

Recently, it was an ONR grant to Roy which brought diamond film research to the U.S. In subsequent ONR research, he confirmed the then unbelievable Soviet reports on a process of laser-solid interaction to yield high pressure phases in the ambient atmosphere.

Statistical Analysis on a Distributed System

Mark J. Schervish
Carnegie-Mellon University

Introduction

A network of computers can be made to act together as a parallel processor using data-flow algorithms. Many statistical methods can be implemented via data-flow algorithms. One requirement is that the computation can be broken into *independent* parts, i.e., parts which can be performed in any order. If an application can be broken into independent parts, and one has p processors available to perform the computations, then each of the p processors can work on one of the parts at a time. When one processor finishes, it can work on the next unfinished part of the computation. If the division into parts is done well, this can allow a speed of the computation to increase by almost a factor of p. This paper briefly describes a distributed system and discusses several applications of distributed data-flow computation in statistics.

1 Parallel/Distributed Computation and Statistics

Chernoff, *et al.* report that "Concurrent computers in a variety of architectures are already upon us and will be the dominant supercomputers, and possibly the ordinary computers, for science and technology in 10 years or less."[1] In statistics, the time is now. There are several common types of statistical computations which can make effective use of a distributed computation system. For example, if one must analyze a large data set, and the processing of one observation does not depend on the processing of the others, then distributed computation can be used. But this is not the only type of statistical computation which can easily be distributed. This paper presents some examples of statistical computations which have made use of distributed computation. Some of the examples appear to have benefited more from distributed computation than others.

The examples all have a key feature in common. They can each be decomposed into a large number of independent tasks. That is, the computation consists of many parts which can be performed in any order, regardless of the results of the other parts. Even this condition is not completely necessary. If only some of the parts of the computation can be performed in any order, but then there are some other parts which depend on the results of these first ones, the computation can still be distributed, but perhaps not as effectively as when there is no dependence at all. The most common examples of computations which can be distributed are various summations. There are summations of large numbers of terms, and there are summations of terms whose computation is time-consuming. Both types of summations appear in the examples. In some of the examples, there is more than one way to divide the computation into independent parts. We take a brief look at how different choices for the division into independent parts affect the efficiency of the computation.

In Section 2, there is a description of the distributed computation system which was used for the examples. Section 3 describes the examples in more detail.

2 A Distributed System

Eddy and Schervish (ES86) developed a parallel/distributed system which is being used in actual applications.[2] The system hardware architecture is a network of processors. The parallel system appears to be a data-flow system. A data-flow system is one in which the arrival of data at a processor signals that processor to begin a computation. Section 2.2 gives a brief description of the system. For an excellent description of several types of parallel architectures, see Treleaven, Brownbridge, and Hopkins.[3]

2.1 Distributed Data-flow Computation

Any network of processors which supports some simple communication primitives can be used to implement a data-flow system. Since nearly every computer installation has a network of processors, such systems can be made widely available. It is not *necessary* to have any special-purpose languages nor any special understanding in order to program a data-flow system. The complex synchronization issues that are usually attendant to parallel programs do not arise in a data-flow system. The major feature that distinguishes the data-flow system of ES86 from other systems is the *granularity* of the problems that are appropriate for it, that is, the size of the computational tasks which will be treated as indivisible units. A vector or array processor treats vector operations as indivisible; these are very fine grains compared with those of ES86[2]. For example, O'Leary and Stewart discuss a system (named ZMOB) with architecture similar to ES86;[4] however, the communication bandwidth of the ES86 system is substantially smaller than the bandwidth of ZMOB, the communications latency of the ES86 system is substantially greater than AMOB, and the processor speeds in the ES86 system are substantially larger than those of ZMOB.

In order for a problem to run efficiently on the ES86 system, the interprocess communication cost must be a small fraction of the process computation cost. The individual processors in the ES86 system are VAXes, which have substantial computational capability. Hence, it should often be the case that the communication cost is small relative to the computation cost. Most problems for which parallel computation is appropriate can be adapted to the system in a reasonably efficient manner, so long as one can determine the correct size grains (indivisible computational tasks) for the particular application.

2.2 A Particular Data-Flow System

The distributed system described by ES86[2] runs on VAX computers under the VMS operating system using the DECnet communication protocol. The mode of operation is called *master-slave*. The master process communicates with the slaves by writing to and reading from *network devices* (DECnet's way of defining communication channels). An alternative system, described by Gardner, *et al.* uses stream and datagram sockets based on TCP/IP under Berkeley Unix 4.2.[5]

Data-flow is implemented via *asynchronous communication*. For example, the master begins by assigning a task to each slave. Each slave has a network device associated with it. The master process writes the appropriate data and/or instructions to the network device associated with each slave, and then reads from the network device, but does not wait for a response. Figuratively speaking, the master says "Let me know as soon as something arrives." Then the master goes on to the next slave. When "something arrives" from a slave, the master deals with the response and sends another task (if any remain) in the same way as before.

The slaves begin by reading from the network device and waiting for a task to arrive from the master. Each slave does its work, writes its response to the network device, and waits for another task. When the work is finished, the master can release the slaves or keep them waiting for another set of tasks.

All of the examples in Section 3 proceed in the fashion described here. However, a master-slave data-flow system is slightly more versatile than this process suggests. For example, if several processors are all searching for a particular result, as soon as one processor finds it, the master can signal all of the others to stop looking. This can be particularly useful if the overall computation consists of several stages, each of which contains such a search.

3 Statistical Applications

In this section, I give brief descriptions of some of the statistical applications which have used the system of ES86. Comparisons are made with a single processor running the same application. Timings for the distributed system are measured in *wall clock time*, since the goal of distributed processing is usually to get the job done sooner. The total amount of CPU time on all of the processors is often much smaller due to communications overhead (e.g. a slave must wait for the master to respond). For a single processor, CPU time and wall clock time are nearly identical.

3.1 Discrete-Finite Inference

The first application which made use of the system of ES86 is one which could not even be attempted without a great deal of computational power. Even the distributed system goes only a small way toward attacking the problem. Suppose X is a random variable which has d possible values, namely $x_1,\ldots,x_d$. Let p_j stand for Pr(X = x_j for $j = 1,\ldots,d$, where $\Sigma_{j=1}^{d}\ p_j = 1$. In addition, assume that each p_j has only m possible values, $v_1,\ldots,v_m$, each between 0 and 1. The vector $\underset{\sim}{p} = (p_1, \ldots, p_d)$ is called the *model vector* and it gives the distribution of the quantity X. The set of possible model vectors is called the *model space*. For convenience, begin by assuming that each possible model vector is equally likely *a priori*.

This model can be applied to several observables $X_1,...,X_N$ if we model them as independent and identically distributed with distribution p conditional on the value of $\underset{\sim}{p}$. If we observe some of these X_i values, we may wish to make inference about some of the others. Suppose we observe $X_1,...,X_n$. We can first calculate the posterior distribution of $\underset{\sim}{p}$, and then calculate the conditional distribution of future observations given the observed data. For example, suppose that the observed values are $X_1 = x_1,...,X_{14} = x_{14}$, and we are interested in X_{15}. For each $\underset{\sim}{p}$, we must use Bayes' Theorem to calculate the posterior probability

$$\Pr(\underset{\sim}{p} \mid X_1,\ldots,X_{14}) = \frac{\prod_{i=1}^{14} p_i}{\Sigma_{All\ \underset{\sim}{p}} \prod_{i=1}^{14} p_i} \qquad (1)$$

Then, we use the Law of Total Probability to calculate the predictive distribution of the next observation

$$\Pr(X_{15} = x_j) = \sum_{All\ \underset{\sim}{p}} p_j \Pr(\underset{\sim}{p} \mid X_1,\ldots,X_{14}). \qquad (2)$$

The sums in these formulas are over all the possible values of $\underset{\sim}{p}$. If, for example, $m = 20$ and $d = 10$, we get 6,906,900 possible model vectors. Some restrictions on which model vectors are possible can reduce the number of terms over which we must sum. For example, one could require that no two adjacent coordinates of $\underset{\sim}{p}$ can be more than k/m apart, when k is some small integer. This requirement says that the model vectors cannot be very rough.

ES86[2] report on a particular example in which there were $n = 14$ observations. The discrete-finite model used had $d = 19$ and $m = 22$. In addition, to insure that the model vectors were somewhat smooth, they used $k = 2$, as described above. In order to calculate the predictive distribution of the next observation X_{15}, they needed to calculate the sums in (1) and (2). Each of these sums has 38,226,040 non-zero terms. That is, there were 38,266,040 model vectors which did not have zero posterior probability. Using a single MicroVAX II computer to perform the calculations, it took 11 hours and 8 minutes (40,100 seconds). It took only 4,303 seconds on a distributed system consisting of eight MicroVAX II's, six MicroVAX I's and one VAX 11/750. Comparisons were made of the three types of machines working on the same problem. It was estimated that a VAX 11/750 ran approximately 75% as fast as a MicroVAX II and a MicroVAX I ran approximately 25% as fast as a MicroVAX II. Hence, the distributed system was approximately the equivalent of 10 MicroVAX II computers. Surprisingly, the speedup was nearly a factor of 10.

In addition to the faster execution on a distributed system, there is another benefit of distributed computation. A distributed system can provide more accurate calculation of a sum than can a single processor. In the example, the computation was divided into 6129 tasks of approximately 6237 model vectors each. This was done to maximize the accuracy of the partitioned summation. Single precision distributed computations agreed with double precision serial and distributed computations to five significant digits. On the other hand some of the sums differed by as much as 5% in a comparison of single precision and double precision serial computations.

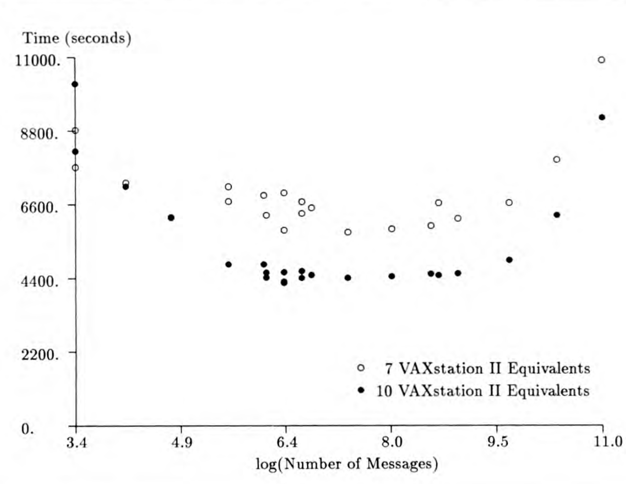
The choice of task size approximately equal to number of tasks helped to maximize numerical accuracy. It might also be interesting to see what type of division of the total work would minimize total time to completion. ES86[2] report the results of using two different configurations of distributed system and several different task sizes. One configuration was the equivalent of 10 VAXstation II's as described earlier. The other was the equivalent of 7 VAX station II's. The total time to complete the computation described earlier is plotted against the natural logarithm of the number of tasks in Figure 1. Although 6129 tasks does not seem to provide the absolute minimum time, it is not far from optimal.

3.2 Large Data Sets

The need to analyze very large data sets is quite common in statistical practice. As an example, consider the 1979 survey of Inmates in State Correctional Facilities conducted by the Bureau of the Census for the Bureau of Justice Statistics. Kim and Schervish analyzed a part of this data set consisting of responses to interview questions from 9566 inmates.[6] The questions concerned their incarceration careers, that is their movements into and out of prisons and the crimes for which they were sentenced. A hierarchical model was used, which had a separate incarceration rate θ for each inmate. The θ values were assumed to be selected independently from a distribution F. The distribution F had two parameters which

Figure 1

Time (seconds) vs log (Number of Messages)



were left unspecified (to be estimated). Both Bayesian and likelihood based methods of inference about the parameters of F require many calculations of the likelihood function. Each calculation of the likelihood function involved an integral over θ for each inmate separately. So each evaluation of the likelihood function required the calculation of 9566 integrals. Using a single DEC VAXstation II, it took 57.5 minutes to evaluate the likelihood function once.

To run this application on the distributed system, the calculation of the likelihood function was broken up into 95 tasks consisting of 100 inmates each and one task with 66 inmates. Every method of analyzing the data was able to make use of this same division of the work. Three different methods were used:

- the simplex method as programmed by Olsson to find the posterior mode[7]
- the Laplace approximations of Tierney, Kass, and Kadane to find posterior moments[8]
- the adaptive quadrature method of Smith *et al.* to find posterior moments.[9]

Each method made use of the distributed system in precisely the same way. Each time that a value of the likelihood function was needed, the master would create tasks to distribute to the slaves. There was no need to modify any of the algorithms for methods themselves. On 10 VAXstation II's, each log-likelihood evaluation took only about 7.1 minutes (wall clock time).

An application of this type is perhaps the most suitable for a parallel/distributed system. The general computational methods do not need to be altered in order to adapt the application to parallel processing. All that was needed was a division of the work which was peculiar to the particular application. Since so much of the computation time is spent performing nearly identical calculations for each observation in the data set, applications involving large data sets are just the sort for which a distributed system should work well. The division of the data set into tasks, one for each processor, is easy to program and can provide nearly optimal speedup.

3.3 Multiprocess Models

In Section 3.2 I described an application in which a single calculation on a large data set required a great deal of time. In this section, I describe a case which requires a multitude of calculations on a small data set. The application involves Bayesian models for autoregressive processes which allow for changes in the model to occur. For example, a particular observation may be an outlier, that is it may have a distribution other than what the autoregressive model predicts. But one can rarely, if ever, be certain that an observation is an outlier, or that it is not. For this reason, one might wish to introduce two different models, one which says that the observation is an outlier, and another which says that it is not. One can use the posterior probabilities of these two models to decide how likely it is that the observation is indeed an outlier. Schervish and Tsay developed the theory for models of this sort.[10]

The number of computations begins to increase when one considers the possibility that several different observations are outliers. For two observations, perhaps both of them are outliers, or perhaps neither of them is an outlier. There are four possibilities. When one has n observations to consider, there are 2^n possibilities. For even moderate sized samples, 2^n is too many possibilities to consider. Schervish and Tsay ignore all but a fixed number of the possibilities with the highest posterior probabilities.[10] Possibilities which are highly unlikely after a few more observations are taken will be eliminated from consideration. In order to keep possibilities open for as long as possible available, Schervish and Tsay make one task out of each possibility whenever a new observation arrives and divide the tasks among several processors.[10] Bayes' theorem is applied after each observation and the possibilities with the highest posterior probability are maintained.

Schervish and Tsay report results using twenty different data sets.[10] Their timing results deal only with the prediction of the last 20 observations from each data set. This was done so that comparisons of the predictions could be made with other statistical methods which require large initial segments of the series in order to choose a single model to represent the process. After each prediction is made, the 20 models with the highest posterior probability are retained. Each of these models is further refined according to whether the next observation is an outlier, a level shift, or neither. This creates 60 models for the next observation, and each processor is assigned one model to update using Bayes' theorem. Updating a model consists of finding the posterior distributions of all of the parameters for the autoregressive models of order 0 through 20 and finding the predictive distribution of the next observation. On a single MicroVAX II, it took 1391 seconds for a typical data set, while a distributed system of six MicroVAX II's needed only 360 seconds (wall clock time).

Note that a six processor parallel system did not run even four times as fast as a single processor. There are several possible reasons. First, the master process must do a substantial amount of work keeping track of which of the sixty models are the 20 best. Because of storage limitations, it is not possible to store all information on all sixty models at the same time. Hence, the master must spend time sorting through the models as they are completed so that we never store more than the 20 best. Some or all of the slaves sit idle while the master is doing this bookkeeping. Secondly, since the posterior distributions of the parameters do not change when the current observation is an outlier, such models take very little time to fit. Hence, 20 of the 60 models fit are done in virtually no time. For one data set, only the AR models from other 0 to order 10 were fit, and in this case, parallel and serial computation both took about the same time (280 seconds), since the slaves were simply not doing enough work to justify the use of a distributed system.

3.4 Hierarchical Models

The National Crime Survey (NCS) is conducted by the Bureau of the Census to assess patterns in victimization (in contrast to reported crimes). The ultimate sampling unit is a household. Suppose that there corresponds a Bernoulli random variable X to each household, where X indicates whether or not the household was victimized during a particular time period. In the sample, the households are clustered, and the clusters are subsets of enumeration districts (ED), which in turn lie in larger primary sampling units (PSU). The particular sample dealt with had 10,117 households in 411 PSUs. Lehoczky and Schervish used a hierarchical model in which households were at the deepest level (the data level) of the hierarchy, preceded by EDs (parameter level), PSUs (hyperparameter level), and the entire nation (hyper-hyperparameter level).[11]

The model can be described briefly as follows. A separate parameter θ_{ij} was introduced for each ED j in each PSU i. We suppose that an observation X from a household in the aforementioned ED has conditional probability θ_{ij} of being victimized given all of the parameters of the model. The distribution of these parameters was taken to be beta with (hyper)parameters P_iQ and $(1\text{-}P_i)Q$. That is, for each PSU i, the θ_{ij} were assumed independent random variables with $Beta(P_iQ, [1\text{-}P_i]Q)$ distribution given the hyperparameters $Q, P_1, P_2, \ldots$ The P_i were assumed to have $Beta(PR,[1\text{-}P]R)$ distribution given the hyper-hyperparameters P and R. Finally Q and R were given convenient distributions consistent with the amount of variation expected between EDs and between PSUs respectively, while P had a beta distribution with parameters consistent with prior opinions about national crime rates.

The goal was to estimate the proportion of households victimized. To do this, it was necessary to calculate posterior means for the P_i. This, in turn, required many four dimensional integrations, one for each i. The forms of these integrals were

$$\int_0^\infty \int_0^\infty \int_0^1 \int_0^1 p_i f(p_i, p, q, r) dp_i \; dp \; dq \; dr. \tag{3}$$

One option for distributing the computation would have been to divide the PSUs into small groups and calculate the integrals in parallel. This, however, would have required a great deal of duplicated effort. A second option was to split the iterated integral (3) into two parts. The outer integrations over q and r were done by the master, while the inner integrations over p_i and p were done by the slaves. These inner integrations were each 15 point Gauss-Legendre quadratures. Each task to a slave was to perform the integration over p_i for fixed values of p, q, and r.

Serial computation on a single VAXstation II required 3,969 minutes to fit the model. A distributed system of 11 VAXstation IIs took 660 minutes (wall clock time). Once again, we see that the gain in speed is far less than a factor of how many processors were used. There really were not enough tasks to go around (only 15 for each fixed value of q, r, and p.) The slaves spent a lot of time all working on the same task at the end of the computation. It might have been more efficient to fix several values of q, r, and p simultaneously and ask each slave to perform several integrals. This, however, would have required that the serial computation be reprogrammed to a much greater degree than in any of the other applications.

3.5 Asynchronous Iteration

Many statistical methods are iterative by nature. At first, it might appear that distributed computation would be hindered because each iteration would need to finish before the next could begin. This is not true, and several authors have shown how iterations can be performed asynchronously while maintaining the convergence properties of the iterative method (see Baudet for example.)[12] Eddy and Schervish illustrate how a typical iterative method can be implemented on a distributed system.[13]

The idea behind asynchronous iteration is that partial results from ongoing iterations can be used as starting points for other partial iterations. For example, suppose that we are trying to solve for the eigenvector corresponding to the largest eigenvalue of a very large matrix A using the power method. That is, we let x_o be an arbitrary vector and calculate $x_{n+1} = Ax_n/c_n$, where c_n is chosen so that the coordinates of x_{n+1} are roughly of the same size as those of x_n. Suppose the dimension of A is k and k is large. Then the calculation of Ax_n can be time-consuming. We can split this calculation into several tasks, say by partitioning A into submatrices with fewer than k rows each. For example, if $k = 500$, then we might partition A into 25 20 x 500 submatrices $A_1, \ldots, A_{25}$. Then each slave could be assigned the calculation of $A_i x_n$ for some i and some n. The value of $A_i x_n$ is only part of an iteration of the power method. Rather than issue the 25 tasks (in this example) and wait for them all to finish before issuing the next task, an asynchronous iteration scheme would proceed as follows. Each time a new task is issued, n is increased by 1. When a previously issued task, say calculation $A_i x_m$, completes, the coordinates in x_n corresponding to the rows in A_i are changed to the new values, and the resulting vector is x_{n+1} for the next iteration. The other coordinates of x_{n+1} are identical to those of x_n. This fact makes it slightly more complicated to monitor the convergence of the algorithm. One would not want to compare x_n to x_{n+1}, but rather, for each coordinate, compare its most recent value to its previous value.

Eddy and Schervish give an example of the calculation described above with $k = 499$ and 10 tasks, 9 of them consisting of 50 coordinates and one consisting of 49 coordinates.[13] On a single VAXstation 2000, it took 8 hours and 39 minutes of CPU time to reach a prespecified convergence criterion based on a sum of squared distances between successive iterates. On a system of 10 VAXstation 2000s, it took 1hr 9min of wall clock time to reach the same level of convergence. The synchronous calculation (one computer) required 2332 iterations of the power method to converge. The asynchronous calculation required 28818 partial iterations (1/10 each). Even though more calculation was required for the asynchronous method, it still took less wall clock time due to the parallel computation.

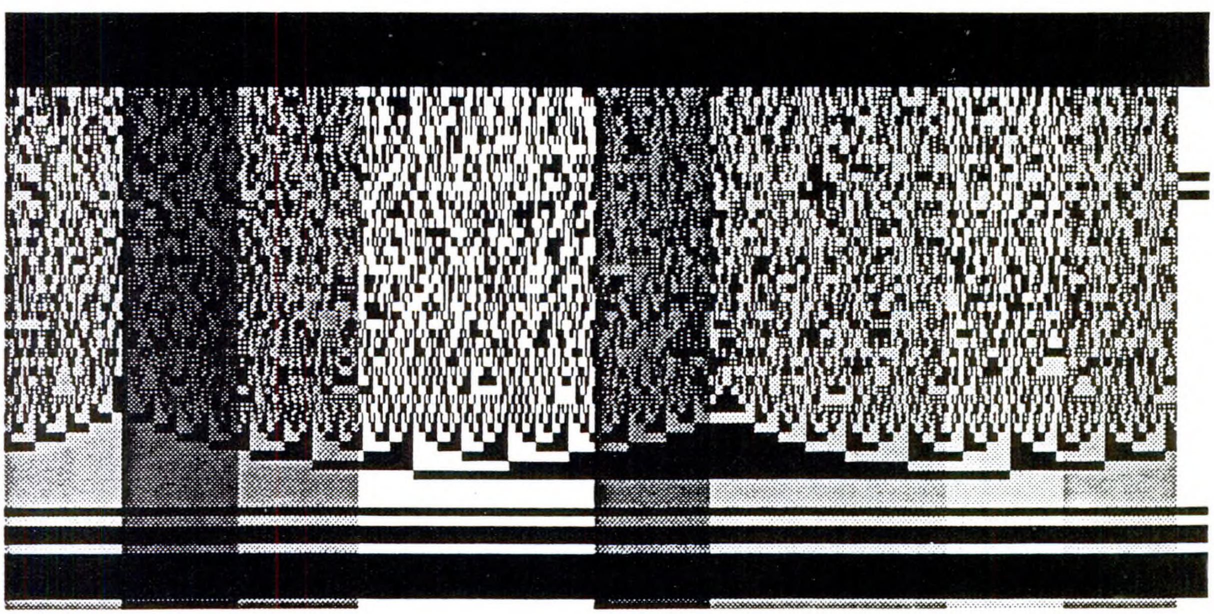
In order to see how the asynchronous calculation affected the convergence of the algorithm, Eddy and Schervish constructed an animated video display of the convergence.[13] The eigenvector corresponding to the largest eigenvalue was known to have all coordinates identical and the eigenvalue was 18.99999999992728. For each iteration, a 512 x 256 pixel bitmap was created. Each of the first 499 columns of the bitmap were used to represent one of the coordinates of the vector x_n after iteration n. The last 13 columns were used to represent the value 18.999..., since this was the value to which the algorithm forced all of the coordinates of x_n to converge. This allowed comparison of the coordinates to see to what extent convergence was not uniform across the vector. To represent a number, stored as a 56 bit mantissa and 8 bit exponent, we added a constant to every number to force the exponents to be the same and divided 224 of the vertical pixels into 56 groups of 4 each to represent each bit of the floating point representation of each coordinate. The most significant bits were at the bottom of the graph. Figure 2 is an example of one frame from the video tape. The different shades (colors in the video tape) represent the different processors which were used. (Which coordinates are shaded which way might change from one iteration to the next depending on which processor most recently updated those coordinates. Some of the sets of coordinates are the same shade because the same processor was the most recent to update them all.) We see some notable features in Figure 2. About 10% of the way from the left, there is a cusp where the convergence is more advanced than elsewhere. Also, about 60% of the way from the left, there is the reverse phenomenon where convergence is much slower. This pattern proceeds throughout the video tape. It appears to be related to the second and third largest eigenvalues which are equal and not substantially smaller than the largest.

4 Distributing an Application

In order to use a system like ES86 effectively, an application must be decomposable into several parts which can be performed independently of each other. For example, the sum of a large number of terms can be decomposed into sums of smaller numbers of terms, such as in the discrete-finite inference example. In that case, each of the many terms to add was easy to calculate, but there were a very large number of them. In the example involving the large data set, there were not as many terms to add, but each of them required significant computation time. In both of these cases, it was straightforward to break the sum into tasks of approximately equal size.

Figure 2

One Frame from Video Tape of Asynchronous Iteration



Numerical integration is very much like summation. One approach to multidimensional integration is to do iterated quadrature. This is a simple, but often inefficient method. More sensible approaches include multidimensional quadrature as described in Chapter 7 of Isaacson and Keller.[14] In the hierarchical model application, we needed an appropriate set of integration grid points to call a task for the slaves. The choice of which grid points to include in each task appears to have been made poorly. The less time it takes too compute the integrand, the more grid points that should be included in each subtask in order to reduce the communication costs. A better choice might have been to divide the grid points into arbitrary subsets, not necessarily with fixed values of the outer integration variables. This would have required a more serious programming effort than what was done in the example in Section 3.4.

Another type of application which can make use of a distributed system is a computation in which several different types of calculations can be performed concurrently. In the multiprocess time series application, the 60 different models being fit concurrently were of three different types. Each slave was capable of performing the calculations for each of the three different types of models. The master needed only to inform each slave of which type of model its task involved. This concept clearly generalizes to other types of computations which are made up of several parts that do not need to be performed in a fixed sequence. Each slave can either be capable of performing every type of task or it can be dedicated to a particular type of task.

5 Conclusion

This paper has described a parallel/distributed system and its use in statistical analyses of computationally intensive applications. The applications were similar in some respects, but each had its unique features. The most important feature common to all of them was that a large portion of the computation consisted of sets of calculations, the elements of which could be performed independently of each other. This allowed the slaves to calculate the elements of each set asynchronously. The master would perform whatever computation was needed to merge the results from the slaves.

In four of the five applications, the set of calculations was some collection of summations. Almost the entire computation was calculating the sum in the discrete-finite inference example. There were a very large number of terms in the sum. In this case, distributed processing effected a speedup almost equal to the increase in CPU power. In two of the other summation applications (inmate survey and NCS) the computation consisted of several summations. In the inmate survey, the likelihood function needed to be calculated many times. Each time that the parameter values changed, the slaves would start over, summing the logarithm of the likelihood contributions for different sets of inmates. In the NCS application, there was a separate numerical integration over P and the P_i for each pair of values for R and Q. The multiprocess time series application was different in that the computation consisted of doing 60 different calculations (one for each model) with the same data, and then repeating the procedure for each data value.

The criminal careers, National Crime Survey, and time series applications had another feature in common. This was that the distributed computation was needed, in essentially the same form, several times during the application. The entire computation need not be broken up into parts which can be performed independently. A system of distributed processors can be used to perform calculations only parts of which are parallelizable. Of course, a large portion of the computation must be distributed before the application will make effective use of the system. The time series application is a somewhat disappointing example, showing how time savings can be eaten up by overhead. The NCS example shows how some applications may require a significant reprogramming effort in order to realize significant time savings. If such an application will be run many times, perhaps with many data sets each time, it may be worth the effort to fine-tune the division of work into tasks in order to make better use of the distributed system. Two final notes are in order, despite the less than optimal speedup in some of the applications. First, the ES86 system runs at low priority on time-sharing machines, so that other users can still make use of the computers in the system. Second, even though the total amount of time on all of the processors used by an application may be larger than that used by a single processor, such as in the asynchronous iteration, real savings in wall clock time are still possible with distributed computation.

Biography

Mark J. Schervish is Professor in the Department of Statistics at Carnegie Mellon University, Pittsburgh, Pennsylvania. He earned the Ph.D. degree in Statistics in 1979 at the University of Illinios at Urbana-Champaign. His research interests include statistical computation, foundations of inference, and multivariate analysis.

Acknowledgments

This work was supported in part by the Office of Naval Research and National Science Foundation. The computer equipment used to perform the computations discussed herein was acquired with partial financial support from the National Science Foundation and the Office of Naval Research. The author would like to thank W. F. Eddy for help in developing an understanding of distributed computation and W. F. Eddy and Joel Welling for producing the video tape.

References:

1. Chernoff, H., Bickel, P.J., Brockett, R. W., Dantzig, G. B., de Boor, C., Eppen, G. D., Fefferman, C.L., Glimm, J.G., Hamming, R. W., Keller, H.B., and McDonald, G.C. (1990). *Selected Opportunities for Mathematical Sciences Research Related to the Navy Mission–An Update*. National Academy Press, Washington.

2. Eddy, W.F. and Schervish, M.J. (1986), "Discrete Finite Inference on a Network of Vaxes," *Computer Science and Statistics: 18th Symposium on the Interface*, 30-36.

3. Treleaven, P.C., Brownbridge, D.R. and R.P. Hopkins (1982), "Data-driven and Demand-driven Computer Architecture," *Computing Surveys*, 14, 93-143.

4. O'Leary, D.P. and G.W. Stewart (1985), "Data-flow Algorithms for Parallel Matrix Computations," *Communications of the ACM*, 28, 840-853.

5. Gardner, T.J., Gerard, I.M., Mowers, C.R., Nemeth, E., and Schnabel, R.B. (1986), "DPUP: A Distributed Processing Utilities Package," *SIGNUM Newsletter, ACM*, 21:4, 5-19.

6. Kim, C.E. and Schervish, M.J. (1988), "Stochastic Models of Criminal Careers," *Bayesian Statistics 3*, Bernardo, J., *et al.* (eds.) University of Valencia, (in press).

7. Olsson, D.M. (1974), "A Sequential Simplex Program for Solving Minimization Problems," *Journal of Quality Technology*, 6, 53-57.

8. Tierney, L., Kass, R.E., and Kadane, J.B. (1986), "Approximation of Posterior Expectations and Variances Using Laplace's Method," Technical report #385, Department of Statistics, Carnegie-Mellon University.

9. Smith, A.F.M., Skene, A.M., Shaw, J.E.H., Naylor, J.C., Dransfield, M. (1984), "The Implementation of the Bayesian Paradigm," *Communications in Statistics, Theory and Methods* 14(5), 1079-1102.

10. Schervish, M.J. and R.S. Tsay (1988), "Bayesian Modeling and Forecasting in Large Scale Time Series," *Bayesian Analysis of Time Series and Dynamic Models*, J.C. Spall (ed.), Marcel Dekker, New York.

11. Lehoczky, J.P. and Schervish, M.J. (1987), "Hierarchical Modelling and Multi-level Analysis Applied to the National Crime Survey," Proceedings of Workshop on the Design and Use of the National Crime Survey, (unpublished manuscript).

12. Baudet, G.M. (1975), "Asynchronous Iterative Methods for Multiprocessors." *J. Assoc. Comput. Mach.*, **25**, 226-244.

13. Eddy, W.F. and Schervish, M.J. (1988). "Asynchronous Iteration." In *Computing Science and Statistics: Proceedings of the 20th Symposium on the Interface*, E. Wegman, D. Gantz, J. Miller, Eds. 165-173.

14. Isaacson, E. and Keller, H.B. (1966), *Analysis of Numerical Methods*, John Wiley and Sons, New York.

Fiber-Interphase-Matrix Interactions in Ceramic Matrix Composites

B. A. Bender, T. L. Jessen and D. Lewis III
Naval Research Laboratory

Many advanced Navy systems require new materials for high temperature structural applications. This is particularly true for those applications requiring material temperatures in excess of 1000°C. For these temperatures, current high temperature structural materials, e.g., the superalloys, are operating at or above their intrinsic limits and are either totally unsuitable or suffer from extremely short service lives. For these very high temperature structural applications one of the most promising class of new materials is that of ceramic fiber-ceramic matrix composites (CMC's). These composites also may prove competitive with other structural materials operating at intermediate temperatures, e.g., titanium alloys, intermetallics, and metal matrix composites which may be used in the 200-1000°C temperature range. The CMC's typically offer unique combinations of low weight, high strength, high toughness, and the ability to perform at temperatures as high as 1500°C.

These ceramic composites may also offer significant advantages over the other possible high temperature structural materials, particularly intermetallic matrix-ceramic fiber composites and carbon-carbon composites, in a greatly diminished need for protection against the operating environment. In the case of carbon-carbon composites the problem is extremely severe, and despite very extensive work on the problem, oxidative degradation of carbon-carbons still precludes their use as critical load-bearing structural materials operating at high temperatures in air. Presumably, a CMC system based on suitable fibers and matrix would show greater intrinsic stability under high temperature oxidizing conditions, and thus would prove suitable for such applications. The present difficulty, however, with such use of CMC's is in the need to control the nature of the fiber-matrix interfacial region in the composite, both during processing and during subsequent service at elevated temperature. A major part of the research described here is devoted to dealing with this critical problem in CMC's.

Fiber-Matrix Interreactions

One of the major limitations currently in employing CMC's in high temperature structural (and other) applications is that of the deleterious interactions between the various phases in the composite. Typically, the matrix, fiber, and fiber-matrix interphases interdiffuse and react, altering and typically degrading the properties of the composite. The constituents also react with the environment, inasmuch as the environment is free to penetrate the interior of the composite, either by diffusion or by penetration down interfaces or cracks in the material. The reactions with the environment also tend to degrade the composite, either by altering the properties of the constituents (e.g., weakening of the fibers), or by altering the nature of the fiber-matrix interface (increasing the degree of bonding in a CMC).

These deleterious reactions are particularly likely to occur during processing, assuming that processing and fabrication takes place at higher temperatures than service. Some composite processing currently takes place at temperatures ranging from 1200-1700°C vs. typical maximum service temperatures of <1200°C, as would be found in gas turbine hot sections and hypersonic vehicle skins. Note that efforts to reduce the processing temperatures significantly may be doomed to failure, since a material processed at temperatures far below those of service is almost certainly destined to be thermodynamically unstable in service and subject to further reactions and change. Appropriate processing temperature should presumably be somewhat higher than the projected service temperatures, the difference depending somewhat on the required service lifetime. Thus, composites to be used at temperatures of <1200°C will probably be processed at temperatures of at least 1300-1400°C, and the processing and interfacial control system must be such that the composite survives such processing and use temperatures with little degradation or deleterious reactions.

Modelling

The chemistry and physics governing the matrix-interphase-fiber-environment interreactions are exceedingly complex, and the resultant mathematical problems nearly intractable. The mathematical problems involved are highly nonlinear, since the material properties in these problems are spatially heterogeneous and time-dependent, as the local composition and structure depend on previous diffusion and reaction. The attempts to model the processes occurring are also hindered by the lack of data on high temperature diffusion, chemical kinetics and thermodynamic properties, and also by the unknown character of some of the phases present, e.g., amorphous BN coatings on fibers in the present study. It is thus necessary to base any efforts to model the processes occurring at the fiber-matrix interface in these ceramic matrix composites on a firm base of detailed experimental observations. About the best that can be expected at present from such modelling efforts are parametric studies which qualitatively duplicate the nature of the interfacial region evolved during processing of the composite or during service.

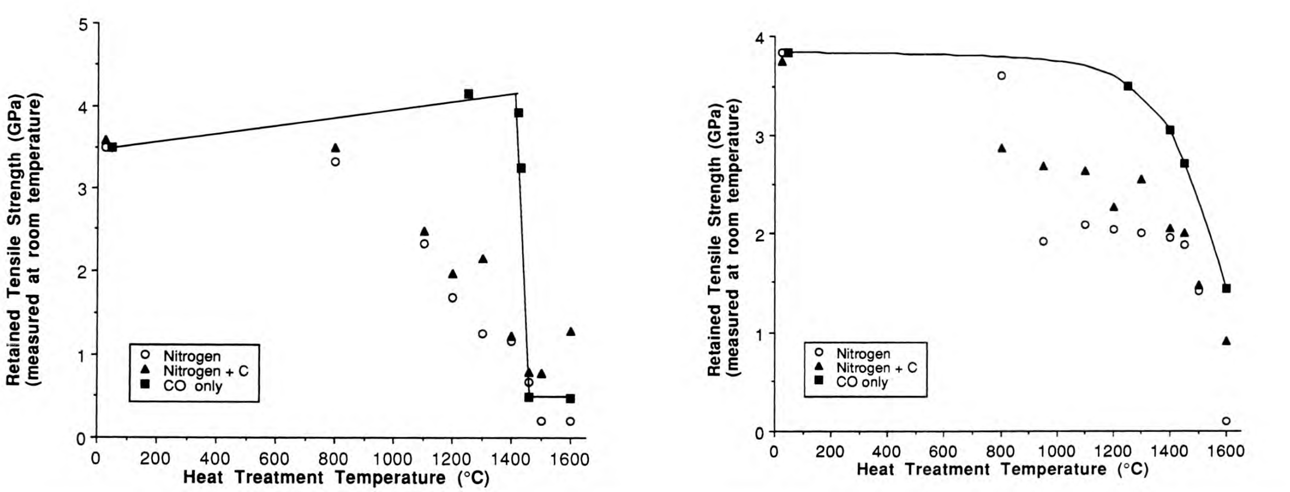
Ongoing Research

Current efforts in the Composites and Ceramics Branch at NRL have focussed on detailed observations of interactions of individual fibers with various representative environments (temperature and atmosphere), and on such observations in a relatively well-developed, model ceramic matrix-ceramic fiber composite system. The former includes both characterization of fiber-environment interactions and of fiber-coating-environment interactions[1-3]. The coatings are here and elsewhere applied to the ceramic fibers to control the nature, properties and stability of the fiber-matrix interface.

Previous efforts at NRL have resulted in the development of single layer BN and bilayer BN/SiC coatings for polymer-derived SiC fibers to be incorporated into various ceramic matrices. The BN is used in these cases to minimize the extent of fiber-matrix bonding (to produce a relatively weak interface), and the SiC top layer in the two layer coating is used to protect the BN from interaction with oxide matrices and/or with an oxidizing service environment, both resulting in the undesired formation of B_2O_3. Other investigators have ex-

Figure 1

Effect of heat treatment temperature on ambient temperature tensile strength of polymer-derived SiC Fibers: (a) Ti-doped SiC Tyranno™ fibers, and (b) Nicalon™ SiC fibers



plored other possible fiber coatings, including pyrolytic carbon which is also used as a weak interface material, though the carbon is even more susceptible to oxidative loss than the BN. Parallel efforts in the Composites and Ceramics Branch on advanced metal matrix composites have also explored the use of multilayer coatings for interfacial control, in this case as diffusion and reaction barriers, compliant layers, and to effect strong bonding of fiber to matrix.

With regard to ceramic matrix composites, the earlier work at NRL focused on the change in the character of the reinforcing fibers with heat treatments in various atmospheres, intended to be representative of those present during composite processing, and those found during service in applications such as advanced gas turbine engines. These previous studies documented the changes in and degradation of the polymer-derived SiC fibers used, both coated and uncoated, in a variety of atmospheres, including vacuum, argon, nitrogen, and carbon monoxide[1-3]. In the last case, very interesting results were obtained which have been employed in the present work on the model composite system. The CO treatment was found to increase the strength of the heat-treated fibers, and greatly extended the possible range of processing temperatures. Another interesting finding here was the possibility of in-situ formation of fiber coatings by suitable heat treatments in SiO and nitrogen atmospheres, resulting in the formation, variously, of dense SiC layers, SiC whiskers and Si_3N_4 whiskers on the surface of the nominally SiC fibers. The SiC coating would be useful in minimizing the reaction of these fibers with the matrix and the processing environment; the whisker deposits should prove useful in enhancing the bonding of the fibers to the matrix in materials such as carbon-carbon and MMC's, where strong bonding is desired.

The CMC system currently being used as a model material at NRL to study interface development and degradation consists of commercial (Nicalon™) polymer-derived 'SiC' fibers, with various coatings (BN, BN/SiC, SiC/BN/SiC) to control the interfacial character, in a $ZrTiO_4$ matrix derived from alkoxide precursors, using processing described previously[4]. The matrix here is predominantly zirconium titanate, with small amounts of residual reduced zirconia and titania, and was chosen for an expansion match to the fibers, a relatively low elastic modulus, and low processing temperatures (1200-1300°C). The current effort focuses on this more complex composite system in which a much greater variety of coupled, interactive diffusion, reaction, and recrystallization processes are possible. The initial constituents in this composite system (ignoring possible impurities such as SiO_2 which are known to be present) include:

$SiC+C+O+BN$ or $BN+SiC$ or $SiC+BN+SiC+ZrTiO_4+ZrO_{2-x}+...$
(Fiber) (Coating) (Matrix)

+ Vacuum or Argon or N_2 or CO
(Processing Atmosphere)

with possible products:

$SiC, C, CO, CO_2, SiO, SiO_2, BN, B_2O_3, Si_3N_4,$
$ZrO_2, TiO_2, ZrTiO_4, ZrSiO_4, TiC, TiN,$ etc.

Figure 2

Effect of fiber coatings on ambient temperature flexural strength and fracture toughness of model CMC.

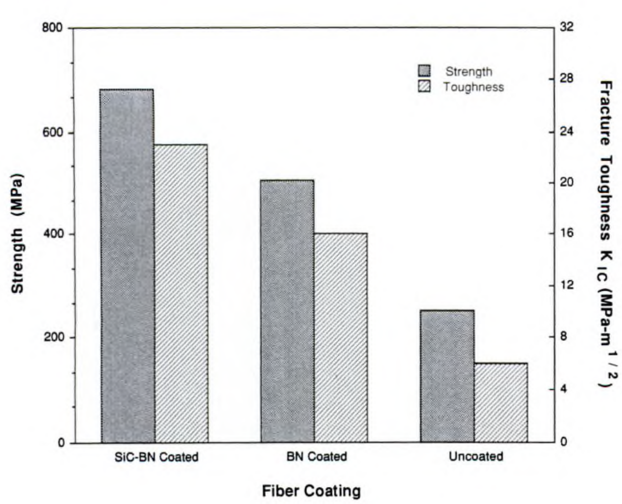


Figure 3

Effect of hot pressing environment on strength and toughness of CMC.

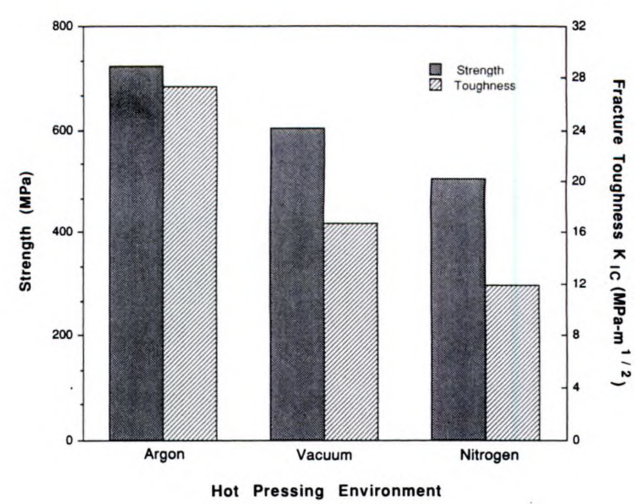
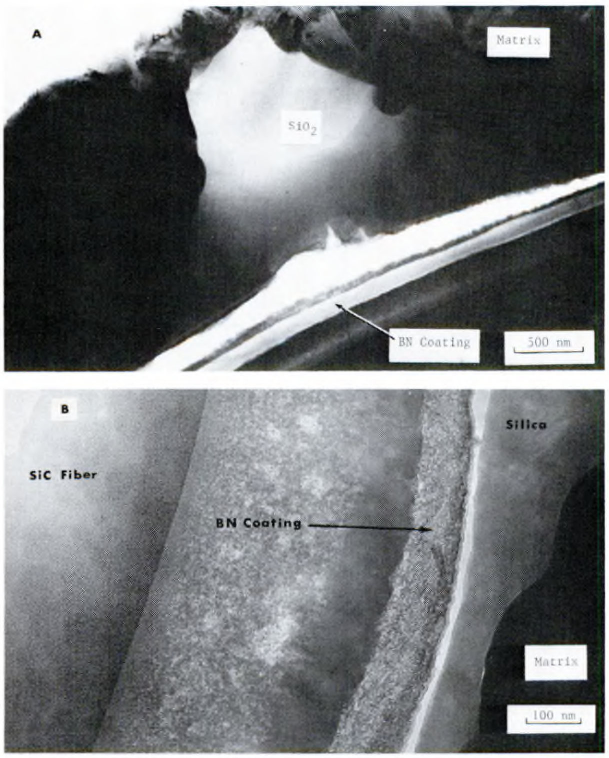


Figure 4(a) and (b)

Transmission electron micrographs of the interfacial region in a $ZrTiO_4$ matrix/SiC fiber composite incorporating coated fibers and processed at 1250°C.



Note that some of the possible reaction products in this system are gases at the processing temperatures of 1200-1700°C (CO, CO_2) and others would show very high vapor pressures at these temperatures (B_2O_3, SiO). Some of these compounds are liquids and possible glass-formers (SiO_2 and B_2O_3), and others are expected to be present as well-defined crystalline compounds (SiC, TiC). Many of the possible compounds here can exist in various forms, depending on the process of their formation, e.g., amorphous, hexagonal and cubic BN, SiO_2 as amorphous, cristobalite or quartz, etc. Other compounds exhibit variable stoichiometry depending on local conditions, TiC_{1-x}, ZrO_{2-y}, etc. The material in this relatively simple case, and the processes involved in production of the composite and its degradation during service are clearly very complex.

The current efforts with regard to CMC's at NRL have attempted to investigate the details of the processes using the best analytical tools available. Currently these tools include transmission electron microscopy (TEM), analytical electron microscopy (AEM), scanning Auger microscopy (SAM), and scanning electron microscopy with microanalytical capabilities (SEM/EDS). These techniques, while requiring lengthy and difficult sample preparation, with possible effects on the structures being observed, provide detailed information on the physical and chemical nature of the interface region. At present, there are **no** analytical tools available which can provide in-situ characterization of the chemical and physical nature of the interfacial region in composites of this type, despite the obvious need for such tools. Thus current research efforts stress the use of such invasive analytical techniques as TEM/AEM where the extensive sample preparation (sectioning, mechanical polishing, dimpling, ion milling) very probably causes extensive changes in the nature of the material being characterized. Corollary efforts on the model composite system here include measurement of the key mechanical properties of the composite (presumably a structural material): flexural strength, fracture toughness, and failure modes, as affected by the various material and processing parameters. Failure modes are assessed here through high magnification video imaging of a flexural specimen which simultaneously incorporates tensile, shear and compressive stresses, the relative magnitudes of which may be manipulated through specimen and loading geometry.

Current Findings

Fibers Heat-Treated in CO

Preliminary work in this program on the effects of various atmospheres and processing temperatures on properties of the fibers produced one very interesting result which has led to significant further improvements in composite properties. Figure 1 shows the effect of heat treatment temperature on the ambient temperature tensile strength of two polymer derived SiC fibers: Tyranno™ in (a), and Nicalon™ in (b). In the case of the widely used Nicalon™ fibers, the presence of the CO atmosphere preserves the initial tensile strength of the fibers to significantly higher temperatures (400-600°C than other atmospheres). In the case of the Tyranno_ fibers, the CO heat treatment appears to significantly increase the strength of the fibers, in addition to greatly extending the useful range of processing temperatures for composites made with these fibers. Recently, experiments have been conducted where the fibers have been heat treated in CO prior to application of the CVD multilayer barrier coatings. These heat treatments permit higher CVD processing temperatures, more rapid deposition of the coatings, and deposition of coatings which are more crystalline and hence more stable and possibly more resistant to oxidation. Composites made with such pre-heat treated fibers have already been shown to be over 10% stronger than composites made by the conventional route, with qualitative indications (fracture modes) of large toughness increases.

Effect of Process Variables on Strength and Toughness

Figures 2 and 3 show examples of the strength (tensile strength as observed in a 4-point flexural test where failure occurs in tension) and fracture toughness (single edge notched beam in flexure) of the CMC model material. Note here that the fracture toughness numbers should not be taken too literally, since the fracture proceeds in the nominal direction only in the case of the low toughness specimens; for these the numbers should be relatively accurate. For the higher toughness materials, failure does not normally occur by self-similar propagation of the starting notch, thus the K_{Ic} values are likely to be **underestimates** of the true values. In addition to the complications above, it is well-known that the fracture toughness values in these types of fiber composites exhibit R-curve behavior, with K_{Ic} increasing with increasing crack length, thus the fracture toughness values shown are appropriate only for the crack size used in these tests (ca. 1-2 mm). In the program described here, the fracture toughness values are used primarily in directing the development of suitable processing parameters and optimal fiber coating systems, not as engineering design parameters.

The results in Figure 2 show the rather dramatic effects on both strength and toughness of the coatings on the fibers in the CMC. The single layer BN coating, a patented NRL development[5], provides on the order of 100% improvement over the results obtained with uncoated fibers. This is primarily the results of the reduction in matrix-fiber bonding afforded by the presence of the coating, as well as a reduction in the degradation of the fiber properties associated with reaction with the matrix. Note however that recent experiments with composites with uncoated fibers, but processed in a CO atmosphere have quite different results, and the use of such processing atmosphere, or variations on them, may minimize or

Figure 5

TEM micrograph showing development of TiC crystallites within BN coating in NRL CMC material.

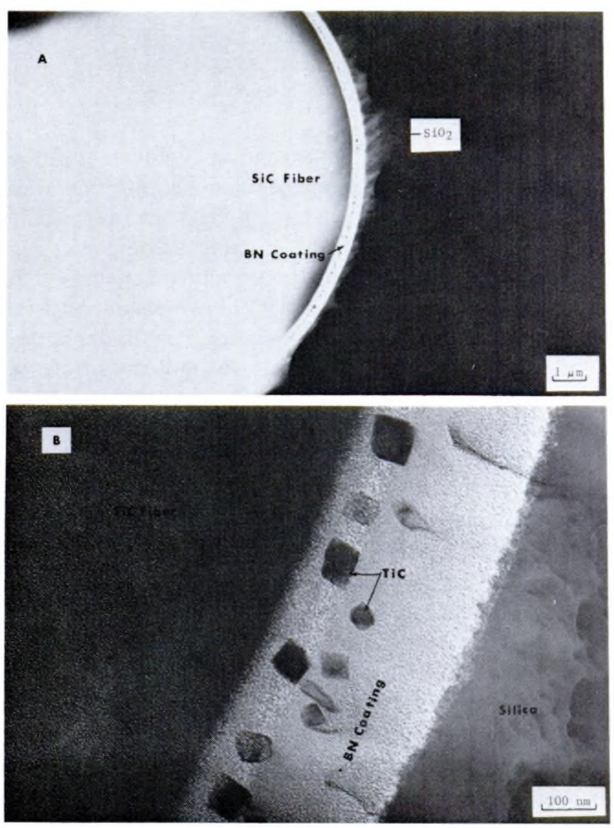
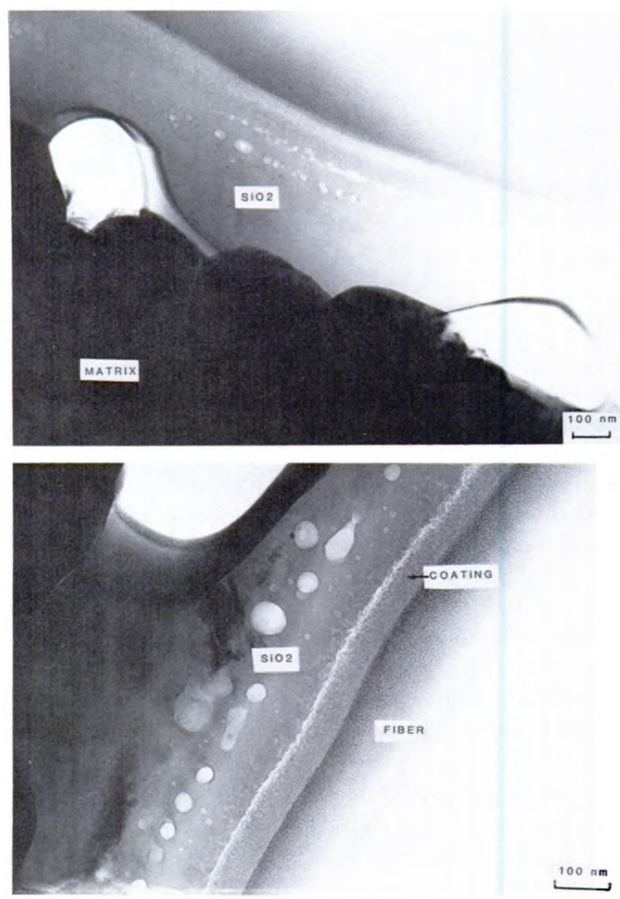


Figure 6

Bright field TEM micrograph showing good bond resulting from the formation of a silica layer at the fiber-matrix interface, and crystalline areas within SiO_2 layer.

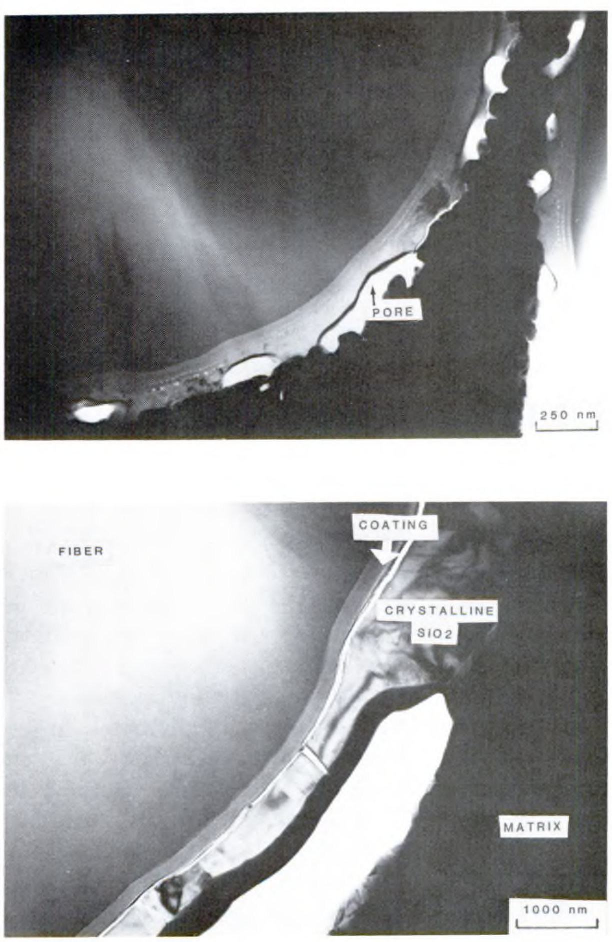


eliminate the need for such barrier coatings on fibers, though for many systems and applications coating systems such as discussed below may still be necessary.

The two layer BN/SiC coating in which the SiC overlayer inhibits the reaction of the BN with the oxygen in the matrix, provides even more dramatic improvements, to approximately three times that obtained with uncoated fibers. The bilayer coating also has other benefits in applications, since the BN coating is susceptible to degradation by oxidation with service in high temperature oxidizing environments (gas turbines). The SiC overlayer tends to protect the underlying BN through formation of a relatively impermeable SiO_2 layer, through which oxygen diffuses very slowly. (Note that SiC coatings are also used for oxidation protection of carbon-carbon composites for similar reasons, but in that case they are macroscopic coatings with thicknesses of ca. 1 mm, compared to the mm thickness of the fiber coatings used here.) The CMC's prepared at NRL with the two-layer coating exhibit values for strength and fracture toughness comparable to those of the common engineering alloys used for aerospace structures, e.g., aluminum and titanium alloys, and should maintain those properties to relatively high temperatures.

Figure 7

Bright field TEM micrograph of interfacial region in NRL CMC with SiC/BN-coated SiC fibers in $ZrTiO_4$ matrix, after oxidation at 1000°C in air for 16 hours. Note crystalline silica layer and separation of silica layer from fiber coating.



CO Processing Atmosphere

With recent improvements in composite fabrication techniques (these laboratory samples are essentially hand-made), and with the change to a CO processing atmosphere, strengths have been improved even further. Current samples, with the bilayer coating and approx. 50 v/o Nicalon fibers, show tensile strengths ranging from approx. 750-1250 MPa, averaging about 1000 MPa (145 ksi), twice the strength of the aircraft aluminum alloys and comparable to the strength of conventional titanium alloys, with densities (approx. 3500 kg/m^3) intermediate between those of aluminum (2700) and titanium (4500) and much less than those of the ferrous alloys and superalloys (ca. 8000 kg/m^3). The fracture toughness values typical of these CMC's, e.g., 25-50 MPa-$m^{1/2}$, are comparable also to those of the light alloys currently used for aerospace structure, and the combination of high strength, high fracture toughness and possible high use temperature make these composite materials very appealing for use in future high temperature structural applications, supplanting the MMC's and superalloys which now are used.

Other Processing Atmospheres

Figure 3 shows the effects of the processing atmosphere on the bulk mechanical properties of the model composite (here with BN/SiC coated fibers). Not shown here are the more recent results cited above for composites processed in a CO atmosphere which gives substantially higher toughness and strength. In the case of nitrogen, argon and vacuum environments, deleterious effects would be expected, based on possible reaction of the constituents with the atmospheres, as noted elsewhere[3], and based on the gaseous nature of some of the reaction products (see above).

The nitrogen atmosphere presumably degrades the composite by some reactions with the constituents (fiber, matrix, coating). These reactions would be expected to be those which would produce nitride or oxynitride products such as Si_3N_4, SiAlON, or TiN. However, despite the obvious effects on the strength and toughness of the composite, no such products have ever been detected by TEM or AEM analysis, indicating that either different, unknown processes are responsible for the degradation, or that the nitride products formed are unstable under the conditions in the composite during processing.

The vacuum atmosphere, typical of much consolidation via hot pressing, tends to promote reactions which produce gaseous products such as that which is known to occur in the fibers between the SiC, free C and free O, producing SiO and CO:

SiC + C + O Æ SiO≠ + CO≠

and thus tends to encourage degradation of the fibers and composite. This degradation results from the formation of porosity in the fiber and voids at the fiber-matrix interface as discussed below.

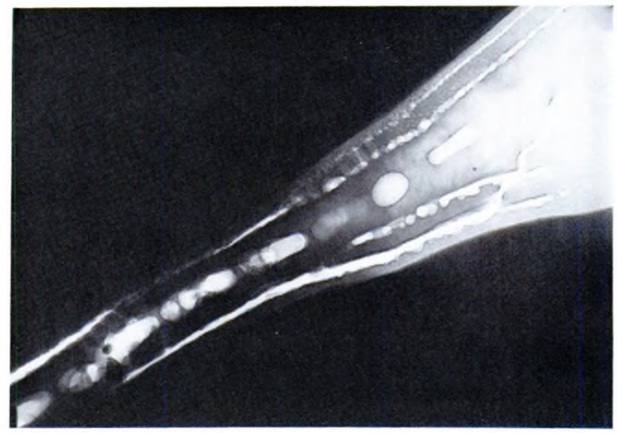
The argon atmosphere, which is nonreactive with the composite, but which provides a small overpressure to inhibit gas-producing reactions provides the best results, next to those produced by the more recently developed CO atmosphere, the trials of which were delayed by the need to assure adequate safety in handling this very toxic gas. With the argon atmosphere, there is still some opportunity for degradation by evolution of gaseous species, especially if the vapor presssures are higher than the limited argon overpressure. The results obtained by the CO atmosphere, which are the best to date, might also be further improved by a substantial increase in the CO overpressure as might be obtained in a hot isostatic press. The current experiments in a conventional hot press are limited to overpressures of approx. 0.1 MPa, constrained by the seals in the system.

Complexity of Fiber-Matrix Interfacial Region

Figure 4 shows some of the details of the fiber-matrix interface in a composite made with BN/SiC fibers. Figure 4b shows the Nicalon™ SiC fiber with its microcrystalline nature, the amorphous BN coating and the crystalline $ZrTiO_4$ matrix. Note also the pocket of amorphous SiO_2 in Figure 4a and the nearly continuous layer of SiO_2 in Figure 4b. These result either from oxidation of the SiC coating on the fibers (by reduction of the surrounding matrix), or from SiO_2 impurities in the matrix precursors (note that most ZrO_2 sources contain significant amounts of SiO_2 since the starting material is typically $ZrSiO_4$ sand). In Figure 4b, note also the fine structure of the interfacial region which suggests a number of reaction and diffusion processes are occurring there, e.g., the fine layers at the boundaries of the DBN coating and the SiO_2 reaction layer. These effects result from processes such as SiO escaping from the fiber and condensing at the fiber-coating-matrix boundaries. Note also the gap where the fiber has physically debonded from the matrix at the silica/BN coating interface, preventing the bonding of the silica to the fiber and leading to a significantly tougher composite. Various theoretical calculations suggest that optimum toughness in these brittle matrix composites requires a relatively weak interfacial bond, typically about 1-10 MPa[6].

Figure 8

Voids by production of SiO or CO gas at interface in NRL CMC during prolonged high temperature oxidation.



The TEM micrograph (Figure 4.) shows the complexity of this particular interfacial region, which spans a distance of less than one micrometer, and how it consists of numerous discrete layers, suggesting that a number of complex reaction and diffusion processes are occurring at the silica/BN coating interface. Another example, indicating the complexity of the processes occurring at the interface is shown in Figure 5. In this case, Ti diffusing in from the matrix and C diffusing out from the fiber react **within** the BN coating to form a nearly continuous layer of TiC crystallites in approximately the middle of the BN 'barrier' layer. Experiments here have shown that it may be possible to form a continuous layer of TiC in this manner, which itself may be useful as a barrier or interfacial control layer.

Effects of Oxidation of Interfaces and Properties

One of the key problems in the potential use of CMC's as high temperature structural applications has been the deleterious effects of oxidation on the character of the interfacial region and on the resultant mechanical properties. In composites where C or BN are used to form an weak interfacial layer between fiber and matrix, this layer is subject to oxidative attack. In the case of C, the result is the loss of the material as CO or CO_2, possibly forming voids in the composite in the process. In the case of the BN coating, the coating oxidizes to B_2O_3, which may also volatilize at sufficiently high temperatures, but which also has the potential for forming low viscosity glasses with the silica and other impurities in the composite. Formation of such glassy phases can produce strong fiber-matrix bonding, as shown in Figure 6. Such strong bonding is undesirable in these brittle matrix composites and can have other negative effects. Vaporization of the B_2O_3 or SiO (under locally reducing conditions) can form substantial voids (cf. Figures 7 and 8) in the material, thus weakening it, and the SiO_2 can crystallize as well to either christobalite or quartz (note the crystalline areas within the SiO_2 layer in Figure 6 and throughout the SiO_2 layer in Figure 7). Both of these crystalline phases undergo very destructive phase transitions. The typical effect of such changes in the nature of the fiber-matrix interface is both a loss in strength in the composite, and a substantial decrease in fracture toughness, both of which are unacceptable changes in a critical structural material. The problems here associated with the oxidative degradation of these ceramic matrix composites are probably the most important to be solved if these materials are to achieve any degree of utility as high temperature structural materials. The solution is very likely to result from a combination of appropriate selection of fiber and matrix, and through the use of appropriate multilayer fiber coating systems, to minimize the effects of such exposure.

Summary

Work at the Naval Research Laboratory on a model ceramic fiber-ceramic matrix composite system has shown the strong effects of the processing environment on the nature of the fiber-matrix interface and the resultant bulk mechanical properties of the composite. Particular results indicate the efficacy of the bilayer BN/SiC coating on the SiC fibers, and the significant benefits of the use of a CO atmosphere during processing to minimize fiber degradation. Analysis of the details of the interfacial region in this composite suggest very complex interactions between the various phases present and the processing conditions and indicate the need for sophisticated modelling to describe the various diffusion and reaction processes occurring. The unsolved problems associated with the effects of high temperature oxidation on the interfacial region and the composite properties remain the most critical for the future use of these composite materials.

Biographies

Barry A. Bender is a materials research engineer with the Naval Research Laboratory, Ceramics Branch. He earned a B.S. in metallurgical engineering from Virginia Polytechnic Institute in 1977 and an M.S. in materials science engineering from Lehigh University in 1979. Mr. Bender was at Western Electric Co. before joining NRL in 1980. Since then his research has been involved in the microstructural analysis of ceramics and ceramic composites using transmission electron microscopy and analytical electron microscopy.

Todd L. Jessen is a ceramic engineer in the Ceramic and Composites Branch at the Naval Research Laboratory. He earned a B.S. in ceramic engineering from Alfred University in 1982 and a M.S. in Ceramic Science from Pennsylvania State University in 1987. He joined NRL in 1987 after graduation. His current research interests are metal particulate-reinforced glass composites and the mechanical properties of ceramic fiber/ceramic matrix composites.

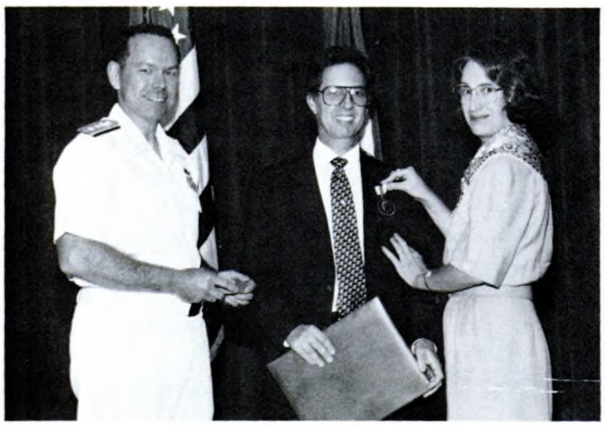
Dr. Lewis joined NRL as a research ceramic engineer in 1979. He is currently the associate branch head of the Composites Branch and is the head of the Structural Ceramics Section. His areas of expertise include processing and characterization of ceramics and ceramic composites, fracture mechanics, mechanical properties of ceramic and ceramic composites, theoretical modeling of the fracture and toughening processes in ceramics and ceramic composites, and the statistical analyses of mechanical properties of ceramics. He has authored approximately 75 publications and made over 200 technical presentations.

References:

1. B. A. Bender, D. J. Schrodt, and J. S. Wallace, "Thermochemical Treatment of Tyranno Fibers," in *Metal Matrix, Carbon and Ceramic Matrix Composites*, J. D. Buckley, ed., NASA Conf. Publ. 2482, pp 201-212 (1987).
2. B. A. Bender, J. S., Wallace, and D. J. Schrodt, "Effect of Thermochemical Treatments on the Strength and Microstructure of SiC Fibers," to be publ. in *Metal Matrix, Carbon and Ceramic Matrix Composites*, J. D. Buckley, ed., (1989).
3. D. Schrodt, B. A. Bender and D. Lewis III, "Thermochemical Degradation of Polymer-Derived Silicon Carbide Fibers," in *1987 NRL Review*, NRL Publ. 112-2630, pp 91-92 (1988).
4. B. A. Bender, D. Shadwell, C. Bulik, L. Incorvati and D. Lewis III, "Effect of Fiber Coatings and Composite Processing on Properties of Zirconia-Based Matrix SiC Fiber Composites," *Bull. Am. Ceram. Soc.* 65 (2), 363-369 (1986).
5. R. W. Rice, "BN Coating of Ceramic Fibers for Ceramic Fiber Composites," U. S. Patent No. 4,642,271, 1987.
6. D. Lewis, B. A. Bender, T. L. Jessen and S. H. Lawrence, "Processing and Environmental Effects on Interfaces in a Polycrystalline Matrix-Ceramic Fiber Composite," to be publ. in *Proc. Electrochem. Soc.*, 1990.

Research Notes

Dr. Boris Wins Captain Robert Dexter Conrad Award



Dr. Boris is congratulated by his wife Elizabeth and RADM Miller.

Dr. Jay P. Boris, chief scientist and director of the Laboratory for computational Physics and Fluid Dynamics (LCP&FD), at the Naval Research Laboratory (NRL), is the recipient of the Captain Robert Dexter Conrad Award for 1990. He was presented the award by RADM Miller, Chief of Naval Research.

The award is granted to "an individual who has made an outstanding contribution in the field of research and development for the Department of the Navy." Dr. Boris' citation commends his leadership that has "shown the way nationally and internationally toward practical use of the latest generation of supercomputers for commercial applications in aerospace, weapons system and engine development." According to the nomination, he and his interdisciplinary research teams he has assembled, "tackle difficult scientific and engineering problems important to the Navy, the DoD and the nation."

Dr. Boris holds a PhD and MA in astrophysical sciences and a BA in physics, all from Princeton University. He joined the NRL staff in 1970 as a research physicist in the Plasma Physics Division (PPD). He served as consultant in computational physics to the PPD from 1972 to 1976, when he became head of the Plasma Physics Branch. He was appointed to the senior executive service and became Chief Scientist and Director of the Laboratory for Computational Physics in 1979. In 1986, Dr. Boris' laboratory was expanded, becoming the Laboratory for Computational Physics and Fluid Dynamics.

He is internationally known for his solutions to difficult fluid, reactive flow, plasma, and many-body dynamics problems, which have been published in more than 300 papers, books, reports and articles.

Dr. Boris has been the recipient of numerous awards for outstanding scientific achievement, including the SES Meritorious Senior Executive Award. He holds the NRL Chair of Science in Computational Physics and is a fellow of the American Physical Society and member of the American Astronomical Society and the American Institute in Aeronautics and Astronautics.

Studies Light Detection by High Tc Superconducting Films

Scientists from the Naval Research Laboratory's (NRL's) Electronics Science and Technology Division have developed a technique for the study of the detection of light by thin films of high-Tc metal oxide superconductors.

Scientists state that the possible short-term application of this work will be optical detectors operating at accessible temperatures. However, operation at over a wide spectral range is anticipated, including ultraviolet, visible, and infrared.

This method relies on the change in surface resistance of a sample positioned in a microwave bridge spectrometer when light is absorbed. Since neither contacts nor a complete conductive path are needed, the microwave approach is more flexible, and in some instances more informative than the conventional dc-biased photoconductivity experiments commonly used to study the light detection.

The NRL scientists say that when studying the detection of light by superconducting thin films using conventional transport methods, homogeneous samples and good contact are necessary. Both requirements are often hard to satisfy in the case of the recently discovered metal oxide high Tc superconductors. This hinders obtaining an understanding of the detector mechanisms and improving the detector performance.

The detection of light by high-Tc films using contactless microwave techniques has been demonstrated. The sample is placed in the cavity of a microwave spectrometer and illuminated through a slotted cavity wall. Light absorption leads to a change in the surface resistance, thus unbalancing the microwave bridge and yielding a signal. By this means, the detection of light has been studied for a variety of different superconducting compounds and structures, with optical wavelength, light intensity, temperature, and magnetic field a variable, the team reports.

Progress in both the physics of these new materials and their possible application as detectors requires obtaining an understanding of the interaction of light with carrier pairs and quasiparticles, their excitation, and subsequent relaxation or recombination, the team notes. The microwave experiments may provide more direct and unencumbered information on these questions than transport studies.

In addition, the microwave method provides a way of spatially mapping the properties of an extended film, since only the illuminated region generates a signal. Thus a focussed laser beam could be rastered over a film, yielding maps of the properties of interest, Dr. Kaplan concludes.

Metal Fluoride Glass Windows

Dr. Ishwar D. Aggarwal, a Naval Research Laboratory (NRL) scientist, reports the development of a technique for the preparation of 20 kilogram (kg) heavy metal fluoride glass windows, the largest fluoride glass windows that have been made to date.

"Heavy Metal Fluoride (HMF) glass offers a number of advantages such as zero optical path difference relating negligible phase distortion. These windows may be used for shipboard laser tracing and weapon systems."

The 20 kg windows, which are greater than one inch thick and nearly 18 inches in diameter, are produced under contract to the Owens-Corning Fiberglass Company and are reported large enough to be used for ground-based and shipboard high energy laser systems. The window program at NRL is sponsored and being carried out in association with Space and Naval Warfare Systems (SPAWARS) Command, PMW-145 offices.

Although Heavy Metal Fluoride (HMF) glasses offer a number of advantages over other materials, until now, they could not be produced in large enough size to be of practical use as a high energy laser (FEL) optical component. However, the new windows are large enough and can compete with, and surpass some of the current HEL optical components.

In the development, the new preparation techniques were combined with traditional optical glass forming techniques to produce these windows. Small batches of glass were premelted and used as cullet. The cullet was then remelted in a large platinum crucible, and the molten glass was cast from this crucible into a brass mold preheated to an appropriate temperature to prevent cracking and crystallization of the window.

Since the discovery of HMF glasses in 1975, research and development efforts have concentrated on the production of long-length, low-loss, optical communications fibers. HMF glasses are desirable for high energy laser optical components for three primary reasons; 1) they have very low optical loss at longer wavelengths, 2) they have a broad transmission range, and 3) they offer the possibility of producing windows with zero optical path distortion.

DEPARTMENT OF THE NAVY
OFFICE OF NAVAL RESEARCH
ARLINGTON, VA 22217

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300