Laboratory and Space Plasma Studies

SAIC Final Report 97/1167

August 1, 1996

Laboratory and Space Plasma Studies

Final Report

SAIC 97/1167

Prepared Under:

Contract No. N00014-93-C-2178
Project No. 01-0157-03-6984

Submitted to:

Dr. Victor Serlin
Laser Plasma Branch - Code 6730
Plasma Physics Division
Naval Research Laboratory
Washington, DC 20375-5000

Prepared by

Ellis Hyman, Principal Investigator

Contributing Authors Include:
Yefim Aglitsky, Orville Barr (Pharos Technical Enterprises), Avraham Bar-Shalom, Chia-Lie Chang, Joseph Czarnaski, Alban Deniz, Adam Drobot, Shmuel Eidelman, Gabriel Font (Cornell University), Amnon Fruchtman, David Garren, Bahman Hafizi, John Hardgrove, Barton Lane (Plasma Dynamics), Thomas Lehecka, Malcolm McGeogh (Plex Corporation), Nathan Metzler, Dennis Papadopoulos (University of Maryland), Philip Serafim (Northeastern University), Ian Smith (Pulse Sciences, Inc.), Ravi Sudan (Cornell University), Steve Swanekamp, Ioana Triandaf, Kang Tsang, Kwok Tsang, and Xialong Yang

August 1, 1996

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

TABLE OF CONTENTS

I. INTRODUCTION

The work performed by Science Applications International Corporation (SAIC) on this contract, "Laboratory and Space Plasma Studies," Contract Number N00014-93-C-2178, SAIC Project Number 01-0157-03-6984, encompasses a wide range of topics in experimental, computational, and analytical laboratory and space plasma physics. The accomplishments described in this report have been in support of the programs of the Laser Plasma Branch (Code 6730) and other segments of the Plasma Physics Division at the Naval Research Laboratory (NRL) and cover the period 27 September 1993 to August 1, 1996. SAIC's efforts have been supported by subcontracts or consulting agreements with Pulse Sciences, Inc., Clark Richardson, and Biskup Consulting Engineers, Pharos Technical Enterprises, Plex Corporation, Cornell University, Stevens Institute of Technology, the University of Connecticut, Plasma Materials and Technologies, Inc., and GaSonics International, Inc.

In the following discussions section we will describe each of the topics investigated and the results obtained. Much of the research work has resulted in journal publications and NRL Memorandum Reports in which the investigation is described in detail. These reports are included as Appendices to this Final Report.

II. DISCUSSION

A. NIKE KrF Laser Support

A.1 Introduction

NIKE is a 5 kJ KrF laser facility at NRL designed and built to perform laser plasma experiments. The primary goal of the system is to produce very uniform (less than 2% RMS nonuniformity in a four nanosecond pulse) high intensity (greater than 10^{14}W/cm^2) illumination on flat targets to study hydrodynamic instabilities under conditions similar to those expected for inertial confinement fusion (ICF) targets.

The laser system can be broken down into five functional subsystems: the laser front end, the 20 cm electron beam (e-beam) amplifier, the 60 cm e-beam amplifier, the propagation bay optical system, and the target area systems. A drawing of the laser system showing the location of these various systems is shown in Fig. A.1-1. In addition to these functional subsystems there are associated support systems crucial to laser development and operation. These include: laser and target diagnostics, data acquisition and archival, laser alignment systems, and laser control systems. Finally, a substantial effort has gone in to modeling and numerical simulation of the laser system and target plasma, and SAIC has had a strong role in contributing to this effort as well.

NIKE Floor Plan

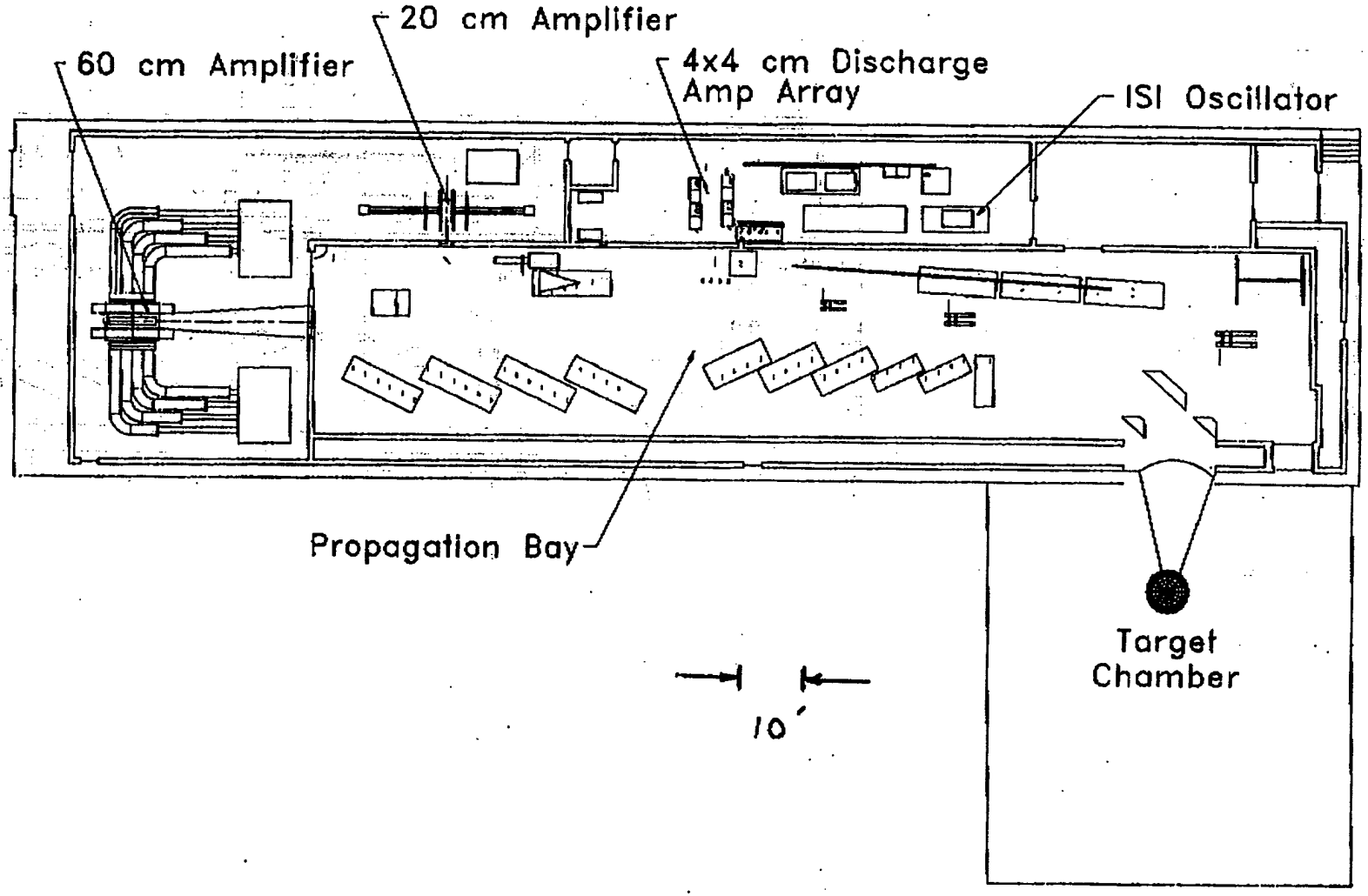


Figure A.1-1

During the past three years under this contract SAIC has been involved in development and operation of all of the above subsystems. In addition SAIC has performed work to integrate these subsystems into an operational laser system which has, to date, achieved virtually all of the original design goals.

In the following subsections of Section A we will summarize the work performed under this contract on the NIKE system. Section A is organized as follows. We devote a subsection to each of the above mentioned functional and support subsystems and to the different modeling efforts. Within each subsection the work performed is presented in an approximately chronological order. For brevity, the reader is referred for details to publications that have resulted from this work if available; they are reproduced in Appendices.

A.2 The laser front end

This portion of the laser includes a KrF oscillator, two commercial amplifiers, a custom built discharge preamplifier and an array of four custom built discharge amplifiers. Echelon free induced spatial incoherence (ISI) beam smoothing is introduced in this area to produce a uniform focal profile along with optical pulse forming to produce the desired four nanosecond laser pulse. The purpose of the front end is to produce a uniform focal profile and amplify this beam to the one to two Joule level in a 4 ns pulse.

SAIC personnel were involved in the design, fabrication and testing of the majority of components in this area. A novel oscillator scheme was developed to produce the desired uniform focal profile. This work is summarized in a publication in Optics Communications, "KrF oscillator system with uniform profiles" and is included in this report in Appendix A. Additionally, results obtained from the laser system were presented as an invited talk at the Conference on Lasers and Electro-Optics (CLEO), May 1993, titled: "Nike KrF Laser Project-Status." The published write-up for this talk is reproduced in Appendix B.

In collaboration with Mark Pronko of NRL and Julius Goldhar of the University of Maryland we developed a Pockels cell with a thin KD*P crystal and plasma-discharge electrodes. We demonstrated that the Pockels cell caused very large non-uniformities in the focal profile during the rise of the laser pulse, and we provided a method to nearly eliminate these non-uniformities. The non-uniformities occur because the Pockels cell is at the Fourier transform plane of the target plane. Polarized light passing through the Pockels cell at 0° will have its polarization rotated by 90° and will therefore be blocked by the crossed polarizers (except for leakage by the polarizers, which is acceptably small). However, light passing through at a non-zero angle will be in another polarization state, which depends on the angle; this results in an angle-dependent leakage and transforms into a position dependent leakage at the target plane. The relative non-uniformity is negligible when the Pockels cell is completely turned on, but is very large if the transmission is

low, such as during the foot or the rising edge of the laser pulse. It remains to be seen if these non-uniformities are too large for performing target experiments. They can be nearly eliminated by having the light pass through a MgF_2 crystal of the proper thickness, which to a very good approximation undoes the effect of propagating through the KD*P crystal at a non-zero angle. Preliminary tests indicated that the MgF_2 compensated for the KD*P crystal, but also caused non-uniformities because of its internal strain, as expected. This can be compensated for by imposing an external strain onto the MgF_2 crystal.

During 1994, SAIC directed efforts to further characterize laser performance and increase system reproducibility, resulting in improved performance of the Nike front end. This portion of the system routinely achieved the desired performance, allowing us to improve the reproducibility of the remainder of the laser system.

A.3 20 cm amplifier development

During 1993 and 1994 the laser front end was integrated with the 20 cm amplifier. SAIC personnel directed the effort to characterize and improve the performance of this system. This culminated in the demonstration of the required laser uniformity at high energy, an accomplishment that was one of the key issues at a DOE review of the Nike program in late 1993. The results of this work were published in an article titled "Production of high energy, uniform focal profiles with the Nike laser," Optics Communications, **117**, 485 (1995). The paper is reproduced in this report as Appendix C. Also, these results were presented at the 1994 CLEO meeting in a talk entitled "Optical beam smoothing on the Nike KrF laser," and in an APS poster entitled "Production of Uniform Laser Illumination with the Nike Laser." The CLEO article is reproduced in Appendix D and the abstract for the APO poster follows.

> Production of Uniform Laser Illumination with the Nike Laser.* T. LEHECKA, A.V. DENIZ, J. HARDGROVE, Science Applications International Corp., S.E. BODNER, K.A. GERBER R.H. LEHMBERG, E.A. McLEAN, S.P. OBENSCHAIN, C.J. PAWLEY, M.S. PRONKO, J.D. SETHIAN, J.A. STAMPER, Plasma Physics Division, Naval Research Laboratory: Nike is a KrF laser at the Naval Research Laboratory designed to produce uniform illumination on flat targets for hydrodynamic and laser plasma interaction experiments. The goals for Nike include $> 2 \times 10^{14} W/cm^2$ intensity on target (> 2kJ, 4 ns, 600 μm diameter) and less than 2% RMS fluence nonuniformities. Approximately one-half of the final system is operational and is currently being tested. To date we have produced more than 120 J in a 4 ns pulse with 4% peak to valley linear tilts and 2% RMS deviation from a flat top profile. Modifications are being made to reduce this below the desired 2% peak to valley tilt level. Experimental results and future plans will be presented.
> **This work is sponsored by the US Department of Energy.*

A.4 60 cm amplifier development

SAIC contributed in several areas to the development and operation of the Nike 60 cm amplifier. These included: rewiring the low voltage control signals for diagnostics, vacuum pumps, and solenoids to reduce EMI noise in the control system generated by the 60 cm amplifier's 3 kilogauss magnet; rewiring all high voltage lines for controlling gate valves and cryo pumps to comply with electrical code and reduce personal shock hazard; installing and modifying the 60 cm amplifier control panel to accommodate full remote control capability via fiber optic link to the central control system for the 60 cm amplifier's gas fill and vent evacuation, valve controls, and diagnostics; installing two cryopumps, dry star vacuum pumps, air dryer, coolers, control valves, gate valves, various gauges for diagnostics, and all necessary electronic controls for the cryo overnight regeneration system in order to reduce down time; building and installing the necessary controls for operating the 60 cm amplifier's diode gate valves to prevent opening in case of vacuum loss during firing of the amplifier. This latter is accomplished by closing both north and south gate valves seconds before firing the amplifier and reopening them if no vacuum loss is detected in the diodes, so as to prevent time consuming regeneration of cryo's and to prevent damage during foil breakage. Currently an SAIC employee is involved in day to day operation of the 60 cm amplifier including repair and modification. He is a qualified operator for running the Nike laser system during the experiment, and is a back up operator for trigger laser alignment and operation for the 60 cm amplifier.

SAIC participated in the integration of the 60 cm amplifier with the Nike system, and operation of the system commenced in October 1994. This result was a major milestone for a DOE review of the program. It involved development of diagnostics for the laser, integration of the existing system with the 60 cm amplifier, and debugging of the 60 cm amplifier to achieve improved reliability.

A.5 Propagation bay and optical system

SAIC personnel were involved in the design, fabrication, and optimization of the Nike optical system throughout the contract period. This optical system includes over 500 transmissive and reflective optics to angularly multiplex 56 laser beams from the front end, through the e-beam amplifiers, and onto target. In addition to the optical components themselves SAIC managed the design and testing of all the optomechanical components to mount these components without excessive distortion or vibration. The alignment system which controls these optics is discussed in section A.9.

Construction of the optical system was completed in early 1994. SAIC performed tests of atmospheric propagation and optical quality to assess the optical system performance. Working with NRL scientists adjustments were made to correct system astigmatism, and point spread

functions (PSF) of 15 times the diffraction limited value (XDL) were demonstrated for the entire laser system at 248 nm, slightly above the design goal of 12 XDL.

In collaboration with Carl Pawley of NRL SAIC measured the non-linear n_2 coefficient of the index of refraction of Xe at 248 nm. SAIC provided assistance in acquiring and analyzing the data. We determined that the n_2 coefficient is indeed negative, and calculations by R.H. Lehmberg of NRL indicate that the effects of propagating the intense Nike beams through air (self-focusing and self-phase modulation) might be compensated for by propagating them through Xe. This would allow a higher useful KrF power density. The paper titled "Two-photon resonantly-enhanced negative nonlinear refractive index in xenon at 248 nm" published in Optics Communications described this effort in detail and is included in this report as Appendix E.

A.6 Laser diagnostics

With Ed McLean of NRL, SAIC chose and installed a spectrometer to measure the laser beam bandwidth in the front end. This system monitors the bandwidth on every target shot. The RMS deviation of the focal profile due to its incoherence is $(\tau/T)^{1/2}$ where τ is the laser coherence time, and T is the pulse width. Because beam uniformity is very important for direct-drive laser fusion, it is important to know the bandwidth when optimizing the laser uniformity. Decreasing the bandwidth of the laser is one way of increasing the short-scale-length non-uniformity. For this to be done in target experiments, the bandwidth must be monitored.

With Rob Morse of CTI, SAIC designed and installed a system to monitor the time-integrated focal profile of a single beam after amplification by the 60cm laser. CTI provided mechanical design and fabrication services, and SAIC designed the optical system and the attenuation method. The attenuation necessary is approximately nine orders of magnitude, and is done by a combination of transmission through a 95% reflective mirror, reflection by 4% uncoated surfaces, and transmission through high-quality neutral-density filters. The attenuation was tested to make sure that the profile was not unacceptably distorted.

SAIC developed software to calculate the energy of the Nike laser beams from existing calorimeter or pyrometer measurements. The software determines the energy measured, even in the presence of expected noise and glitches. It also checks the accuracy of the fit.

A.7 Nike target area systems

After completion of the laser system, initial experiments were performed to demonstrate target acceleration uniformity and characterize X-ray emission from aluminum targets. These experiments, along with laser beam diagnostics at the target plane, indicated that the beam uniformity of the overlapped beams was unacceptable.

SAIC personnel determined that the alignment procedures that had been used for these initial experiments were inadequate. A modified target viewing system was developed and

implemented to improve beam focusing and overlap at the target. These improvements allowed the demonstration of uniform focal profile production with multiple beams overlapped on target, meeting the requirements for laser target interactions.

SAIC then participated in single and double foil experiments to investigate the acceleration pressure nonuniformities created with these ultra-uniform focal profiles. SAIC in collaboration with Ed McLean of NRL set up, calibrated, and tested photomultiplier tubes and an optical streak camera to image the rear surface of the target. The PM tubes were absolutely calibrated, and their signals are used to calculate the rear-surface temperature as a function of time. The streak camera was used to measure the time uniformity of the shock breakout, from which the pressure uniformity of the laser-produced plasma is determined. The results were presented as an invited paper at the 1995 APS Plasma Physics Meeting. The paper describes this effort. It was published in Physics of Plasmas entitled "The Nike KrF laser facility: Performance and initial target experiments" and appears here in Appendix F. The results were also reported at the 1995 CLEO meeting in a talk entitled "Initial results from the Nike KrF Laser facility," which is included in this report in Appendix G.

SAIC personnel routinely support ongoing laser target experiments. Alignment schemes for various diagnostic arrangements are developed as needed. A variety of inexpensive commercially available components have been acquired to allow a great deal of flexibility in target viewing and alignment at low cost. Additionally, on-site target fabrication is performed to support experiments. This includes alignment targets to check diagnostic alignment and determine diagnostic resolution, and fabrication of backlighting targets with various materials.

During 1995 SAIC supervised and implemented the design and fabrication of a computerized target control system for the Nike facility to remotely position and load laser targets without breaking vacuum in the target chamber. The system was completed in October 1995 and has been in continuous operation since then. The system hardware consists of a 4-axis (x, y, z, rot) target positioner and a 4-stage robotic target inserter that transfers targets between the positioner and a vacuum-interlocked loading chamber that holds up to 5 targets simultaneously. A standard (STD) bus computer with motion boards is used to control the stepper motors and translation stages that move the hardware. The control program is written using LabWindows/CVI and runs under MS Windows. It provides a mouse-enabled graphical user interface to operate the target control system. Target positioning is done via a joystick-like multiple button widget interface. Target loading and unloading is simplified to single-button controls. The program also continuously monitors for motion errors and event sensors to minimize hardware accidents, e.g., running into a wall barrier or knocking over a detector mount.

SAIC is developing a laser diagnostic system to support Nike cryogenic target fabrication development. Laser interferometric measurements will be used to check target uniformity and

surface finish on a submicron scale in situ under cryogenic conditions. In addition, computerized control and data acquisition, using a general purpose interface bus (GPIB) and analogue to digital converter (A/D) boards, is being implemented to improve control and monitoring of fabrication processes. The diagnostics system currently consists of a 1mW stabilized HeNe laser at 632.8 nm that is electronically shuttered to produce millisecond pulses. Interferometry fringes are recorded by a CCD camera onto a video recorder. Then single frames are video captured to PC for fringe analysis. Photodiodes are used for transmission and reflection measurements. Pump vibrations from the cryogenic fabrication chamber currently prevent measurement of interferometry fringes. The addition of a turbopump vibration isolator and the use of pulsed interferograms should eliminate this problem in the future.

SAIC was responsible for fabrication and installation of a target control system for a fully automated vacuum system for a 600 cubic foot vacuum chamber, with two turbo vacuum pumps and controllers, vacuum gauges for diagnostics, two 10 inch cryo pumps, several roughing vacuum pumps, a 40 inch gate valve, and several smaller gate valves, all controlled and operated from a remote control panel with built in safety features.

A.8 Target diagnostics and x-ray spectroscopy

In addition to laser diagnostic development SAIC is involved in diagnostic support and development for laser target experiments. This has included involvement with optical diagnostics such as shadowgraphy and optical imaging, extreme ultraviolet imaging with multilayer optics, and X-ray diagnostics such as gated and time integrated pinhole cameras, photodiode arrays, and crystal spectrometry.

SAIC characterized the spatial resolution, RMS signal-to-noise, and linearity of a Hamamatsu 2 ps UV streak camera and a Grant Applied Physics 120 ps UV framing camera. KrF light was used for the streak camera and the characteristics for different camera gains, streak speeds, and light levels were measured. SAIC constructed a model that explains the measured RMS signal-to-noise. We determined that this camera is not suitable for measuring the focal profile, although it is suitable for laser-plasma experiments. For the UV framing camera we used visible light, and measured the characteristics for different camera gains, exposure times, and light levels.

The NRL Nike KrF laser was designed to produce uniform irradiation and acceleration of planar foils for direct-drive fusion studies. The laser produces an energy of up to 3 kJ on target in a 4 nsec pulse at a wavelength of 0.248 μm. Excellent beam uniformity is achieved by induced spatial incoherence optical smoothing and by overlapping of up to 44 beams in the focal spot. The intensity distribution of the laser radiation that is incident on targets has a full width at half maximum of 750 μm and a flat central region that is 400 μm in diameter.

Some of the initial experiments carried out using the Nike laser have focused on characterizing the x-ray emission from the target plasma. These studies are necessary to determine the uniformity and brightness of the x-ray emission from plastic and metal foil targets. The x-ray emission from plasmas created by the laser was characterized using several spectroscopic instruments. The targets were thin foils of aluminum and titanium and were irradiated by laser energies in the range 100 J - 1500 J. The focused laser intensity ranged from 5×10^{12} to 8×10^{13} W/cm^2. These laser conditions are relevant to a number of direct-drive and backlighter applications. Using a spherical-crystal imaging spectrometer operating in the 1-2 keV x-ray region the density, temperature, and opacity of aluminum plasmas were determined with a spatial resolution of 10 μm in the direction perpendicular to the target surface. Spectral line ratios indicated that the aluminum plasmas were relatively dense, cool, and optically thick near the target surface. The titanium plasmas, on the other hand, have an optical depth for a resonance line close to unity.

This work is described in detail in Appendix H in a paper entitled "X-Ray Emission from Plasmas Created by Smoothed KrF Laser Irradiation" and has been accepted for publication in Physics of Plasmas.

A.9 Alignment system

Nike's unique feature is its excellent beam uniformity. The Nike laser system uses 56 separate laser beams that are first multiplexed, then pass through two large amplifiers before they are recombined on to a target. Forty-four of the beams are used for target acceleration and 12 as part of a backlighter for target diagnostics.

During the 93/94 Nike project calendar year the primary function of the laser system was the ongoing effort to improve the beam profile. One of SAIC's primary responsibilities was the alignment of the laser beams. This required operating the Nike preamplifiers and 4x4 discharge amplifiers.

While improvements and testing of the Nike beam profile were taking place, SAIC was installing and testing the alignment system. The automation of the recollimation, intermediate de-multiplexing, and turning arrays was under design and fabrication. Each mirror array required 56 mirrors and 112 motors to be operational. Several thousand feet of cable had to be cut to length, and more than 2,000 connectors crimped onto motor wire cables. Then the cables of each array had to be installed and tested. Nearly 3,000 connections were tested on each array, a total of over 10,000 connections for the 4 arrays under fabrication.

At the same time the power and communication systems for each of the array's card cages were being developed and installed. Improvements in the alignment's system communications also were made involving converting the system from a point to point RS232 system to multi drop RS485. This improvement reduced the number of cables in the propagation bay and allowed for a

multiprocessor bi-directional communication system. It involved changing the Nike alignment software communication modules in the PC, in the control room, and in the embedded processor systems out on the propagation floor. Software changes were required for the 45 degree rotation of the beam on the intermediate array. Other improvements and debugging of the alignment software were also an on-going effort at this time.

In addition, 224 mirrors were placed into X-Y mounts each containing two customer linear actuators. The custom fabrication and assembly of 300 actuators also took place as we prepared to have all 44 beams on target by August 1994.

At the same time an alignment/vibration damping system for the 60 cm mirror was developed and installed. The goal was to reduce vibration of the beams and provide an alignment point to which the 60 cm beam would be directed.

During the 95/96 project year Nike had become a fully operational laser system. Its mission was to conduct experiments on the interaction of plasmas for direct drive fusion target materials. During this time SAIC continued to help in operating the laser during experiments, with the specific responsibility for the alignment of the 56 beams onto target. As the system transitioned into a routinely operated laser system the primary goals changed to upgrading and repairing the alignment system. We have also been involved in the development of new alignment systems to aid the researchers in setting up experiments and diagnostic equipment.

Over the course of the year many changes have been made to improve the existing alignment system. The demultiplexer array shutters have been integrated into the alignment software to aid in the final target alignment process. The main beam lens system has been motorized, providing the user with the ability to focus the beams on target while in the control room. While a position feedback system has been designed for the lens focus system, it has not been determined whether this system need should be implemented due to its cost. Additional mirrors, lenses, and shutters have been added and motorized to bring the backlighter system on line and under automated control.

Software changes have also been implemented as operational patterns have become better defined. Software upgrades have been made to correct various performance and alignment problems. Beam alignment of Nike's intermediate and recollimation mirror arrays became fully automated in time for the dedication ceremony in June.

The Nike alignment system is a large multi axis beam positioning system. It consists of hundreds of motors and mirrors and miles of cable all controlled by thousands of lines of software code and electronic hardware including cameras, microprocessors, and motor drivers. Its function is to facilitate the alignment of the Nike laser. The alignment process is still being adapted as it grows to meet the expanding needs of its users.

A.10 Data acquisition and archival

SAIC has been developing interactive graphical analysis programs using interactive data language (IDL) widgets for analyzing Nike experimental data. Since Nike is routinely obtaining daily multiple shots, there is an urgent need for efficient and automated data analysis. General purpose widget programs were written to facilitate rapid and thorough inspection of line and image data using cursors and sliders. Specialized programs for analyzing laser beam shapes and profiles and accelerated target foils have also been developed.

SAIC designed and implemented a system to archive computer files containing experimental data from Nike to off-line media. The system mirrors (on magnetic disk) the data taken at several locations, and, when the disk containing the data becomes full, transfers the data, with verification, to removable optical media, deletes the data on disk, and creates a file containing the names of the files transferred. This system has been routinely archiving data from Nike target experiments.

A.11 Design of a 2 MJ KrF laser fusion facility

A major effort during this contract period has gone into the development of a detailed model for an advanced 2 MJ KrF laser fusion facility. An amplifier model was first benchmarked against Nike laser data and then used to design higher energy modules with segmented pumping. It was found that the optimum multiplexed beamline energy as it impacts facility size, and hence cost, is in the range 100-200 KJ. For that reason two 68 KJ modules were proposed to be combined into a 136 KJ multiplexed beamline and a total of 16 beams to target was shown to result in a net energy at the target of 2 MJ. This major effort was a collaboration of Dr. Malcolm McGeoch of PLEX Corporation, Dr. Ian Smith and staff at Pulse Sciences, Inc., and NRL scientists. Dr. McGeoch was a consultant to SAIC for this effort and Pulse Sciences performed the work under subcontract to SAIC. The result of this work entitled "Conceptual Design of a 2 MJ KrF Laser Fusion Facility" has been submitted for publication in Fusion Technology. It is included in this report in Appendix I.

A.12 Calculation of plasma radiation from the Nike laser target

A calculational method known as the Super Transition Array (STA) model has been developed and used for opacity calculations to simulate the complex bound-bound and bound-free transitions that constitute the Nike laser target emission. A review of this method is described in the following:

The most complex contribution to the spectrum emitted from hot and dense plasmas arises from the huge number of bound-bound and bound-free transitions. Unresolved clusters of many neighboring partially overlapping transition lines are created, constructing a complex intensity profile. The STA model reveals this complicated structure by a convergence procedure that increases the resolution of the calculated spectrum until the required accuracy is achieved.

Particularly, in each step the entire bulk of transitions is divided into groups G of neighboring lines and each group is described as a Gaussian having the exact group moments i.e. total intensity, average energy, and variance. The resolution is thus increased with the number of groups. Specifically, the total spectrum can be written as a sum over spectral groups

$$S(E) = \sum_G S_G(E) \tag{1}$$

where

$$S_G(E) = \sum_{i,j \in G} N_i \; w_{ij} \; P_{ij}(E - E_{ij}) \tag{2}$$

In eq. (2) the summation is over all the transitions $i \rightarrow j$ in G where i, j indicate the corresponding initial and final levels. N_i is the population of the initial level, w_{ij} is the transition probability, and P_{ij} is the corresponding line shape centered on the transition average energy, E_{ij}. For normalized symmetric line profiles we obtain for the group moments:

Intensity:

$$I_G \equiv \int S_G(E)dE = \sum_{i,j \in G} N_i \; w_{ij} \tag{3}$$

Average energy:

$$E_G \equiv \frac{\int S_G(E)EdE}{I_G} = \frac{\sum_{i,j \in G} N_i \; w_{ij} \; E_{ij}}{I_G} \tag{4}$$

and Variance:

$$(\Delta E_G)^2 \equiv \frac{\int S_G(E)(E - E_G)^2 dE}{I_G} = \Delta_G^2 + \Delta_P^2 \tag{5}$$

where

$$\Delta_G^2 = \frac{\sum_{i,j \in G} N_i \; w_{ij} \; (E_{ij} - E_G)^2}{I_G} \tag{6}$$

and

$$\Delta_P^2 = \int P(E - \overline{E})(E - \overline{E})^2 dE \tag{7}$$

is the variance of the individual line shape assumed equal for all the lines in the group G. The central achievement of the STA theory is the ability to obtain, under certain conditions detailed below, analytical formulae for the moments, bypassing the impractical need to account for the huge number of transitions one by one. Only two assumptions are made:

1) The plasma is in local thermodynamic equilibrium (LTE), yielding Boltzmann populations N_i.

2) The configuration widths are smaller than kT.

The definition of STA groups described below has two advantages:

1) It enables us, using these assumptions, to derive analytic expressions for the group moments

2) It facilitates a group splitting strategy that speeds up the convergence.

For bound-free transitions the final level j belongs to the continuum and the moments of G are obtained by integration over the continuum.

In order to account for the non Gaussian nature of P we first construct a Gaussian from the moments I_G, E_G and Δ_G^2 :

$$\Gamma(E-E_G)=\frac{I_G}{\sqrt{2\pi}\Delta_G}EXP\left[-\frac{1}{2}\left(\frac{E-E_G}{\Delta_G}\right)^2\right] \tag{8}$$

and then construct the spectrum $S_G(E)$ by convolution with the individual line shape

$$\mathbf{S}_G(E)\equiv\int\Gamma(E'-E_G)P(E-E')dE' \tag{9}$$

having the same moments as the original spectrum of G defined by eq. (2).

In order to complete this brief description of the theory we now define the STA groups. An STA group (termed 'STA') is the collection of all the transitions between two super-configurations. A super-configuration is a collection of ordinary configurations defined symbolically by the product over super-shells, s

$$\Xi\equiv\prod_{\sigma}\sigma^{\,Q_\sigma} \tag{10}$$

A super-shell in turn, is the union of energetically adjacent ordinary atomic sub-shells $s\equiv\mathbf{j}_s\equiv n_s l_s j_s$. In eq. (10), the super-configurations are constructed by distributing the Q_σ electrons occupying the super-shell s among the sub-shells s in all possible ways subject to $\{\sum_{s\in\sigma}q_s=Q_\sigma\}$:

$$\sigma^{\,Q_\sigma}\equiv\sum_{\{\sum_{s\in\sigma}q_s=Q_\sigma\}}\prod_s\mathbf{j}_s^{q_s} \tag{11}$$

Clearly each partition of Q_σ is an ordinary configuration. The transition between two configurations constitutes an unresolved transition array (UTA) and an STA is thus a collection of energetically near UTAs.

The convergence procedure mentioned above splits super-shells to smaller super-shells according to their energy spread. For each super-configuration in its turn, at each step, super-shells that give rise to relatively well-separated configurations, are preferentially split. The detailed structure of the spectrum is thus gradually revealed, yielding a converging spectrum. This procedure converges to the UTA spectra where each UTA is completely unresolved.

Over the last several years the STA model has been continually upgraded and the results compared with experimental measurements. Details of this work can be found in a series of papers relevant to the model and to Nike laser plasma radiation. We list the papers below and describe each briefly.

1. W.H. Goldstein, A. Zigler, P.G. Burkhalter, D.J. Nagel, A. Bar-Shalom, J. Oreg, T.S. Luk, A. McPherson, and C.K. Rhodes, "X-Ray Emission from a 650 fs Laser-produced Barium Plasma," Phys. Rev. E **47**, 4349 (1993).

In this work we studied the x-rays emitted from a BaF_2 target irradiated by a KrF laser pulse of 120 mJ and 650 fs length. The BaF_2 target was chosen because it is sufficiently transparent to low intensities of the uv laser so that no prepulse plasma is created. We showed, using both hydrodynamic codes and spectral analysis, that a cold homogeneous optically thin plasma core, close to solid density, is created at early times before hydrodynamic changes occur. This core, localized in both space and time is therefore an optically thin LTE source of radiation with well defined density and a relatively small temperature spread of no more that 150 ev. Very good agreement between the STA theory and the experimental results were obtained. The investigated spectral range included the Dn=1 transition arrays, where the occurrence of a later time hot corona was not apparent. This paper appears in this report as Appendix J.

2. A. Bar Shalom, J. Oreg and W.H. Goldstein, "The Effect of Configuration Interaction on LTE Spectra," J. Quant. Spectrosc. Radiat. Transfer **51**, 27 (1994).

In this work we presented a method for including the effect of configuration interaction (CI) on the STA spectra. We obtained analytic expressions for correcting the UTA and STA intensities due to CI. These expressions, though quite complex, were easily incorporated into the STA code. A few examples are presented in the paper to show conditions for which CI affects the spectrum.

3. A. Bar-Shalom and J. Oreg, "Case studies by the STA model" in 'Third International Opacity Workshop & Code Comparison Study" edited by A. Rickert, K. Eidmann, J. Meyer-ter-Vehn, F.J.D. Serduke and C.A. Iglesias, Garching (1994).

Thirty two case studies were calculated for comparison with results of other codes and experiments. The comparison showed very good agreement between the STA spectra and the few

experimental results for Fe, Ge, and Nb. Also very good agreement was found in most of the cases between the STA results and the detailed OPAL spectra where detailed OPAL calculations are still possible.

4. A. Bar Shalom, J. Oreg and W. H. Goldstein, "The Effect of Configuration Widths on LTE Spectra," Phys. Rev. E **51**, May, 4882 (1995).

The analytic expressions for the STA moments were extended to include without additional approximation the effect of level spread within configurations. The extended expressions were included in the computer code and results showing the importance of this effect on various spectra, in comparison with experiments, were presented. This paper appears in Appendix K.

5. G. Winhart, K. Eidmann, C.A. Iglesias, A. Bar Shalom, E. Minguez, A. Rickert, and S. Rose, "XUV Opacity Measurements and Comparison with Models," J. Quant. Spectrosc. Radiat. Transfer, **54**, 437 (1995).

The opacities of Aluminum, Iron and Holmium were measured and compared with theoretical results of the OPAL, STA, IMP, JIMENA and SAPHIR codes.

6. A. Bar Shalom, M. Klapisch and J. Oreg, "Phase-Amplitude Algorithms for Atomic Continuum Orbitals and Radial Integrals," Comp. Phys. Commun. **93**, 21 (1996).

We have presented a new, fast, and accurate phase amplitude algorithm for the calculation of atomic continuum orbitals *needed for the bound-free part of the STA model* as well as for cross section computations of various atomic processes in plasmas. A coarse, energy independent, mesh is sufficient to achieve high accuracy. A straightforward application of a predictor corrector procedure to the non linear differential equations would fail, in particular, for high energy free electrons in any atomic potential. The present algorithm overcomes this problem. In addition, we describe a novel method for calculating the radial integrals by integration over the phases instead of r. With the use of Gaussian trigonometric formulas over half periods, the integrals are transformed into alternating series. Levin's transform for convergence acceleration then provides the sum of the series with only a few terms. These methods are applicable in a relativistic as well as non relativistic framework. This paper is included as Appendix L.

7. C. A. Iglesias, B. G. Wilson, F. J. Rogers, W. H. Goldstein, A. Bar Shalom, and J. Oreg, "Effect of Heavy Metals on Astrophysical Opacities," The Astrophys. J. **445**, 855 (1995).

The STA model was used to calculate the opacities of heavy elements in astrophysical plasmas. Complete tabulation of the opacities, over the relevant grid of density and temperature was presented. This paper is included in Appendix M.

8. A. Bar-Shalom and J. Oreg, J.F. Seely, U. Feldman and C.M. Brown, B.A. Hammel, R.W. Lee and C A. Back, "Interpretation of Hot and Dense Absorption Spectra of Near-LTE Plasma by the STA Method," Phys. Rev. E **52**, 6686 (1996).

The Super-Transition-Array (STA) model is shown to be a very convenient tool for the interpretation of near-LTE hot and dense plasmas. Specifically, we interpret here the absorption spectra of the CH/Ni/CH foil experiment performed at Livermore using the backlighter technique. In this experiment a laminar foil composed of 200 A Ni with 1000 A CH on both sides was radiatively heated by the x-ray continuum from a nearby gold plasma and was backlit by the x-ray continuum from a distant gold plasma that could be time-delayed with respect to the heating pulse. This setup was designed to achieve a uniform density and heating of the Ni middle layer. It is found that the Ni absorption features depend very weakly on the density of the foil but are quite sensitive to the foil temperature. Remarkably good agreement between the theory and the experiment is obtained for the Ni 2p-3d spectrum. The detailed features indicate that the plasma temperature is confined to a narrow range between 14-18 eV, demonstrating that the foil design, with the goal of creating an homogenous Ni plasma, was successful. These results represent a new temperature diagnostic for high Z plasmas. This paper is included in this report as Appendix N.

9. A. Bar-Shalom and J. Oreg, "The Photo-Electric Effect in the STA Model," accepted for publication in Phys. Rev. E (1996).

In this work we present the STA approach for calculating the detailed photo-electric spectra under LTE conditions. We define the bound-free super transition arrays (STA's) and obtain analytic expressions for their moments (total intensity, average energy, and variance). It is shown that the various STA's connected with a specific photo-electron can be combined first, and then integration over the continuum is carried out only once. The various initial super-configurations give rise to the structure of the photo-electric spectrum near the ionization edges. The details of this structure are gradually revealed by the convergence procedure, inherent in the STA model. The efficiency of the method is discussed. Results of a few examples are given and compared with the average atom method and with detailed term accounting in cases where this approach is possible.

Finally, the following essential points should be emphasized:

a. The convergence procedure permits the use of first order energies in the Boltzmann factor for the level configurations.

b. Orbital relaxation is applied by taking different potentials for different super-configurations. This relaxation is carried out for different STA's as well as for the lower and

upper super-configurations of the same STA. It improves significantly the agreement between the calculated and experimental spectra.

c. The STA model has been extended to include configuration widths (reference 2) i.e. the interaction between j-j configurations belonging to the same LS configuration. These developments improve significantly the agreement with experimental results (references 1 and 3) and with the OPAL code (reference 3). The STA code is a very convenient tool for spectral diagnostics (reference 7). Given the material density and temperature it produces the entire spectrum in a single run. It allows fast identification of the various arrays and lines.

d. The STA model has been extended to account for all the bound-free transitions (reference 9).

The theory and code development described above has been applied specifically for application to the hydrocodes used for studying the Nike laser plasma. This included the calculation of the Planck and Rosseland means for use in the hydrodynamics codes. It involved first running the code to calculate a set of detailed spectra for a particular element (e.g. Hydrogen or Carbon). After establishing a database for the detailed opacities corresponding to the desired elements at the desired temperatures and densities, the MIX code was run to compute the Rosseland and Planck spectral means for the frequency groups required for a hydrodynamic simulation of a laser-irradiated plasma. The resulting opacity tables for aluminum at high density and low temperature were used in the paper "Opacity of Dense, Cold, and Strongly Coupled Plasmas" by Mostovych, et al, published in Physical Review Letters **75**, 1530 (1995). The paper is included in this report as Appendix O. It describes tests of the opacity models in the regime where the probing photon energies are of the same order as the average interparticle interaction energies in the plasma. Overall agreement between experiment and theory within a factor of 2 has been obtained.

In a related effort, SAIC developed an improved two-moment radiation transport theory. The theory was useful particularly in spatial regions that are either optically thick or optically thin. It was incorporated into a one-dimensional hydrodynamics code. The results were presented at the APS:DPP meeting in November 1994 and at the European Conference on Laser Interaction with Matter (ECLIM) at Oxford, England in September 1994. The latter resulted in publication in the Conference Proceedings of a paper titled, "Improved Two-Moment Radiation Transport," that appears in this report as Appendix P.

In addition, SAIC developed a non-LTE model for plasma simulations based upon Busquet's non-LTE theory [Phys. Fluids B **5**, 4191 (1993)]. The development of this theory was necessary because traditional LTE does not yield accurate conversion efficiencies (ratio of energy radiated to input laser energy) and coronal temperatures in laser-produced plasmas. The fundamentals of Busquet's radiation-dependent ionization model were used to transform NRL's 2-

D Eulerian multigroup hydrodynamics code (that used LTE opacities and an equation of state) into a non-LTE code. This development was difficult because Busquet's code is designed to be incorporated into a hydrodynamics code using temperature and mass density as independent variables. In contrast, NRL's code uses mass density and internal energy as independent variables in order to most fully utilize the aspects of flux-corrected transport. This complication was not fully appreciated when the decision was made to incorporate Busquet's model for non-LTE computations into NRL's hydrodynamics codes. SAIC devised a work-around involving an iterative calculation of the average charge state, the electronic temperature, and the so-called "ionization temperature." Unfortunately, convergence appears to be linear rather than quadratic as desired. Various methods were tried to improve the convergence properties, but none were better than the simple choice of a constant "accelerating factor" of one-half. More precisely, the algorithm appears to converge most quickly when the average of the current and the previous values (for the average charge state, the electronic temperature, and the "ionization temperature") are taken as the input to the next iteration. A paper entitled "Variable Eddington Radiative Transport" was presented at the Anomalous Absorption Conference in Aspen in 1995. A set of viewgraphs from the presentation are included here in Appendix Q.

A.13 Calculations of laser target instability growth

(1) Rayleigh-Taylor growth: comparison between experiment and computation

The effect of Rayleigh-Taylor (RT) instability growth for different targets was computed using the FAST2D code. Code results were compared with experimental measurements. The results are described as follows. In order to compare computational results to experiment, the following CH (plastic) target was used:

47 micron thick

Rippled with a 60 micron wavelength, 0.5 micron amplitude.

Experimental conditions were as follows:

No side lighter beams. Installed calorimeter at target chamber on beam 19 to be included in the energy calculation. 50.3 J/beam, 300 J in 6 beams, 1.8kJ in 36 beams, (2.2 in 44, 2.8 in 56).

The experiment showed RT growth, but did not break through the target. The actual laser pulse shape (Fig. A.13-1) was used as an input in the computer simulation. The computation was performed over a width of one wavelength. The computed mass variations in Fig. A.13-2,

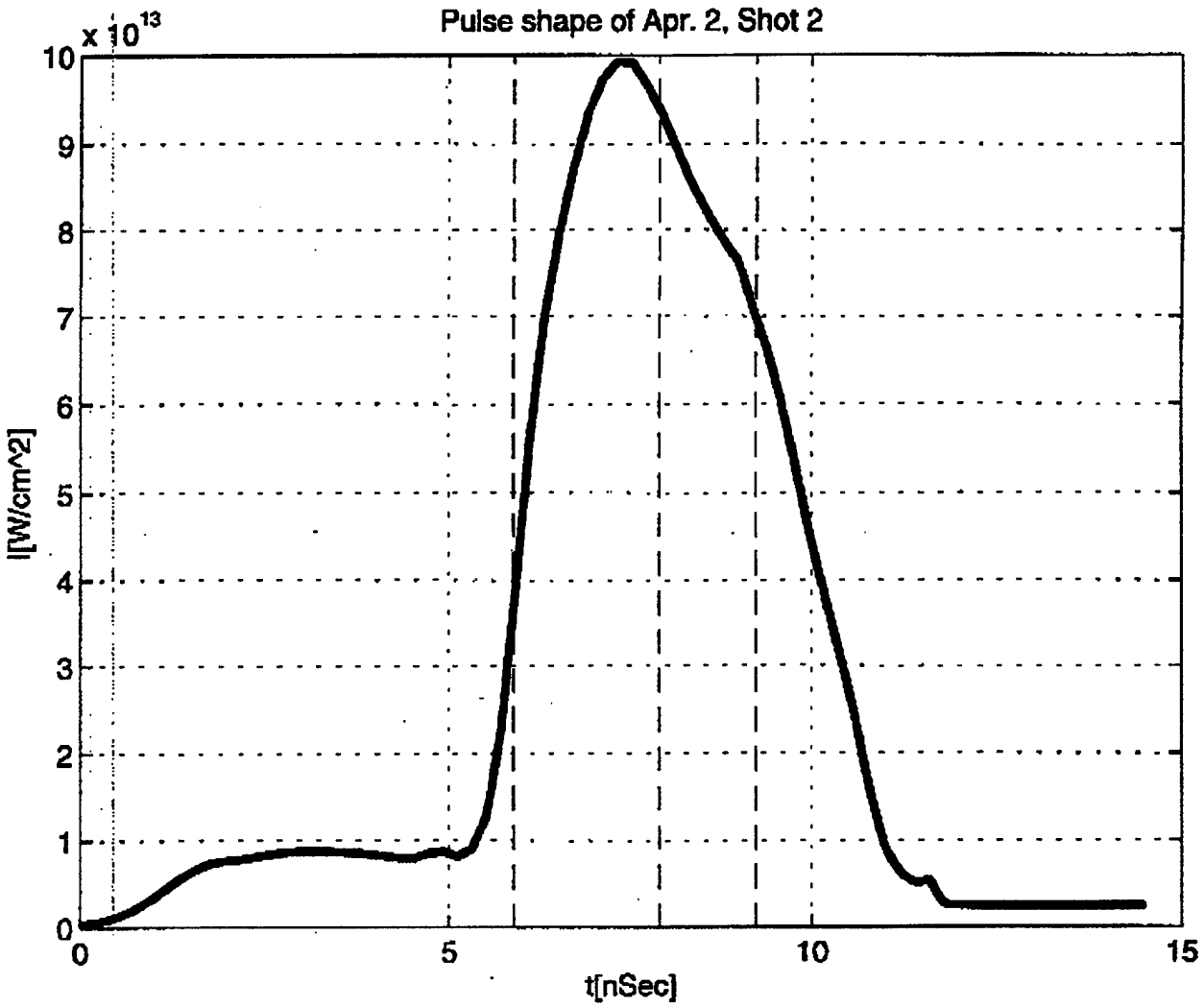

Pulse shape of Apr. 2, Shot 2
x 10^13
I[W/cm^2]
t[nSec]
0
1
2
3
4
5
6
7
8
9
10
0
5
10
15

Figure A.13-1

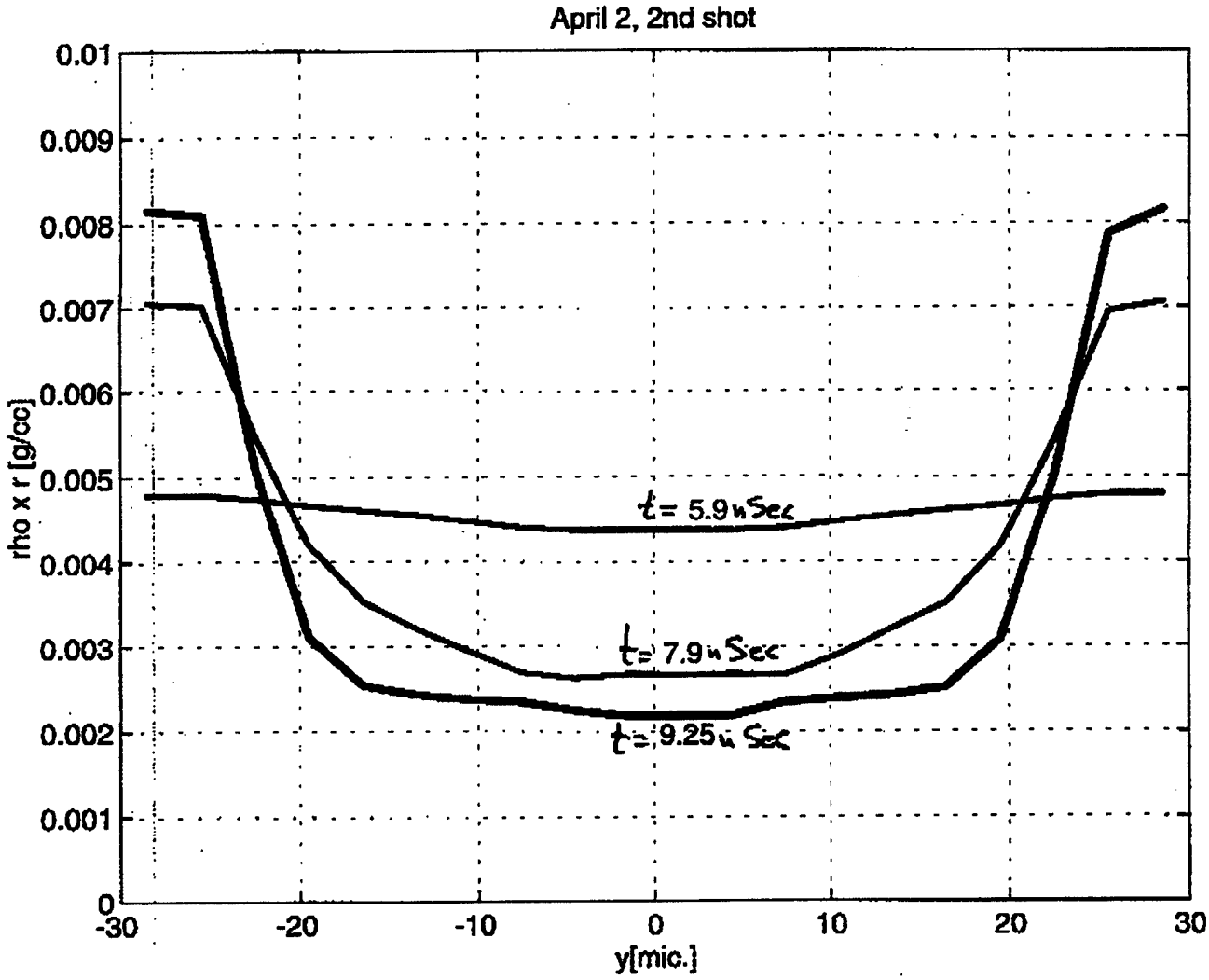

April 2, 2nd shot
rho x r [g/cc]
y[mic.]
t= 5.9 u Sec
t= 7.9 u Sec
t= 9.25 u Sec
0.01
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
-30
-20
-10
0
10
20
30

Figure A.13-2

are comparable to the experimental ones, shown in Fig. A.13-3. Note that in Fig. A.13-2 "-x" is the laser direction and "y" is the transverse direction (over one wave length), while in Fig. A.13-3 "x" is the transverse direction (over the spot size) and "-y" is in the laser direction. Thus, in both the longitudinal origin is at an arbitrary position. The transverse mass variation obtained from experiment (Fig. A.13-3b) is qualitatively comparable to the computed one at intermediate time (t=7.9 sec in Fig. A.13-2). Even at later time, both the computation and the experiment predict no burn through of the target due to RT instability. To be sure that the laser power input taken from the experiment is scaled properly, a comparison between experiment and computation was made for the target acceleration. The actual laser pulse (Fig. A.13-4) was used in the computation. The comparison between the experimental and computed target trajectory is shown in Fig. A.13-5.

(2) Reduction of the laser imprint

Computer simulations predicted that single mode laser beam nonuniformity will grow significantly slower if the target is made of a plastic foam with gradually varying density, instead of a solid plastic. Experimental results were obtained using a few constant density discrete foam layers in place of the gradually varying foam density and they verified the prediction. Even if ISI-smoothed beams are used, the computation predicts significant reduction of the growth rate if foam is used. The advantage of foam disappears, according to the simulations, if the radiation transport is "turned off." Currently, computer simulations using multiple materials but spherical targets rather than flat foils, are in progress.

B. SPACE PLASMA STUDIES

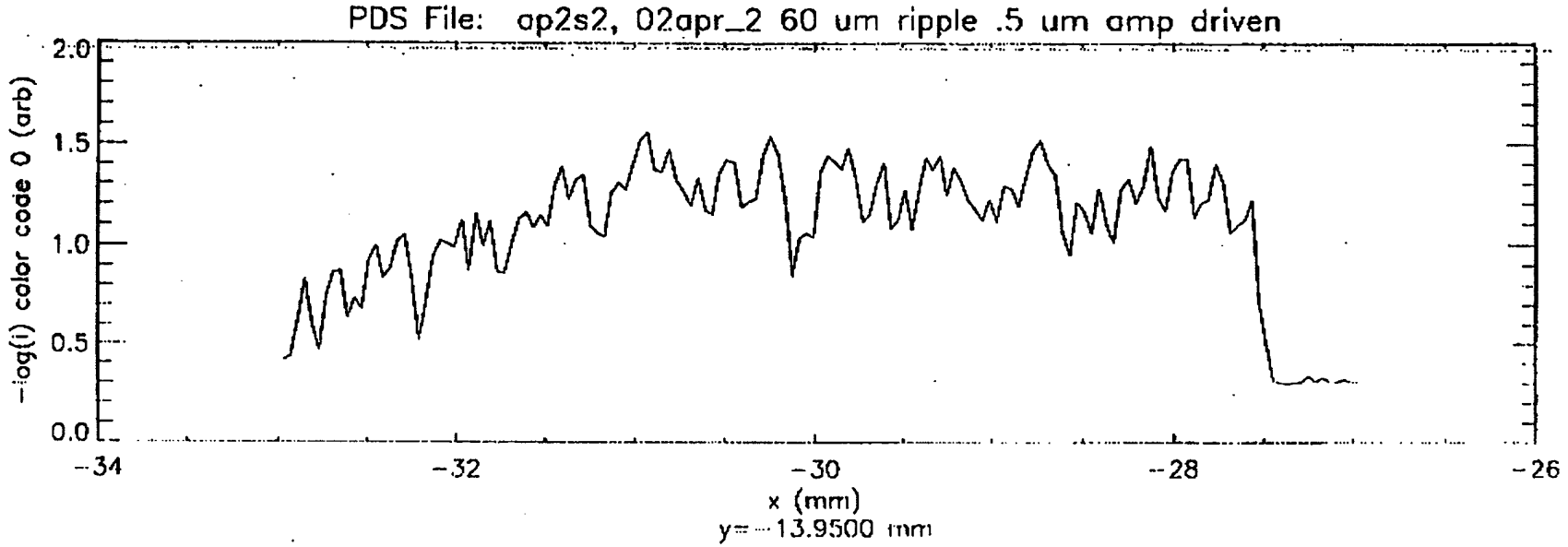
During the current contract period our work focused on using the active properties of the ionospheric plasma for several Navy related tasks such as communications, detection, and countermeasures. More specifically we examined the following:

B.1 Ionospheric response to conductivity transients

A combined analytic and numerical technique was developed to address the current closure problem. The set of the cold electron plasma equations and Maxwell's equations were first solved analytically in ($\underline{k}$, ω) space. Inverse Laplace and 3-D complex Fast Fourier Transform (FFT) techniques were subsequently used to numerically transform the radiation fields and plasma currents from the ($\underline{k}$, ω) space to the ($\underline{k}$, t) space. The results show that the electron plasma responds to a time-varying current source imposed across the magnetic field by exciting whistler/helicon waves and forming an expanding local current loop, driven by field aligned plasma currents. The current loop consists of two anti-parallel field-aligned current channels concentrated at the ends of the imposed current and a cross-field Hall current region connecting these channels. The characteristics of the current closure region are determined by the background

Growth of mass perturbation of sinwave modulated target

(lineouts near beginning of main pulse (top) and 2.6 ns later)



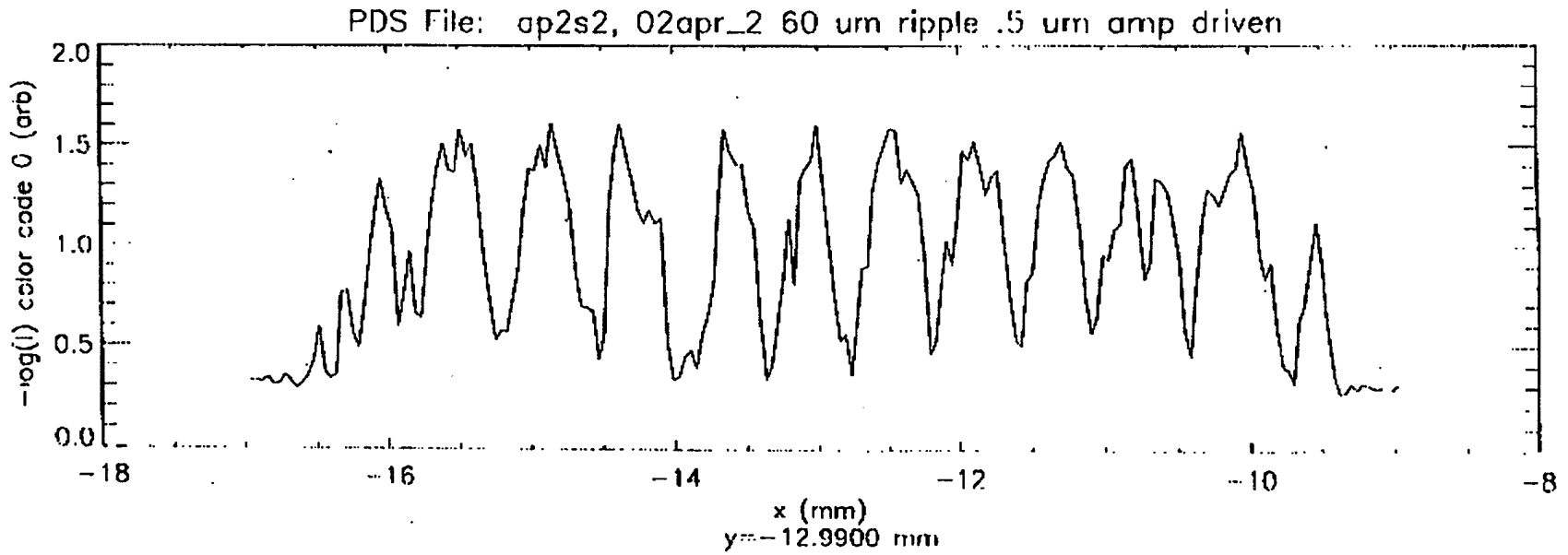


Figure A.13-3

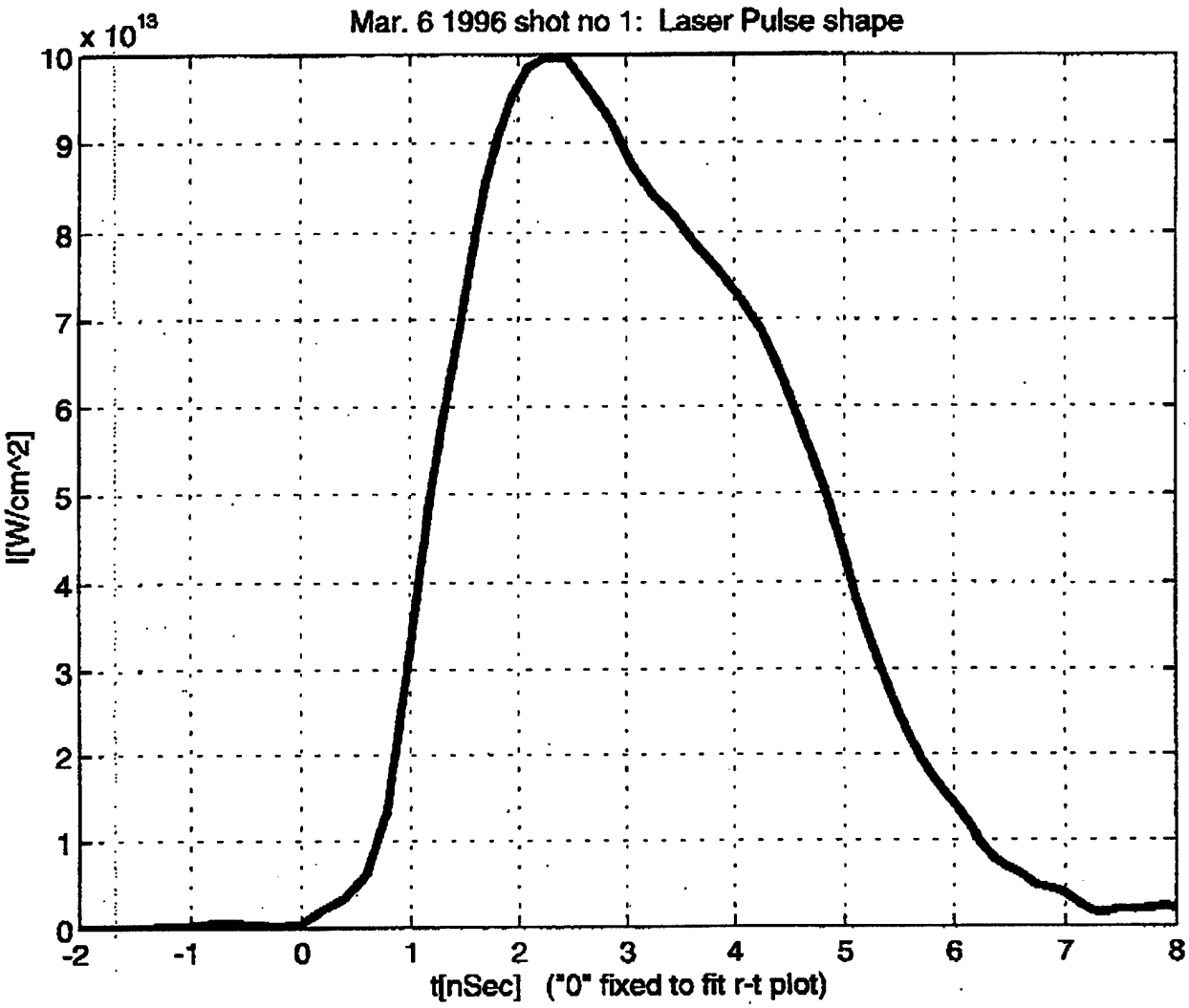

Mar. 6 1996 shot no 1: Laser Pulse shape
x 10^13
I[W/cm^2]
t[nSec] ("0" fixed to fit r-t plot)
10
9
8
7
6
5
4
3
2
1
0
-2
-1
0
1
2
3
4
5
6
7
8

Figure A.13-4

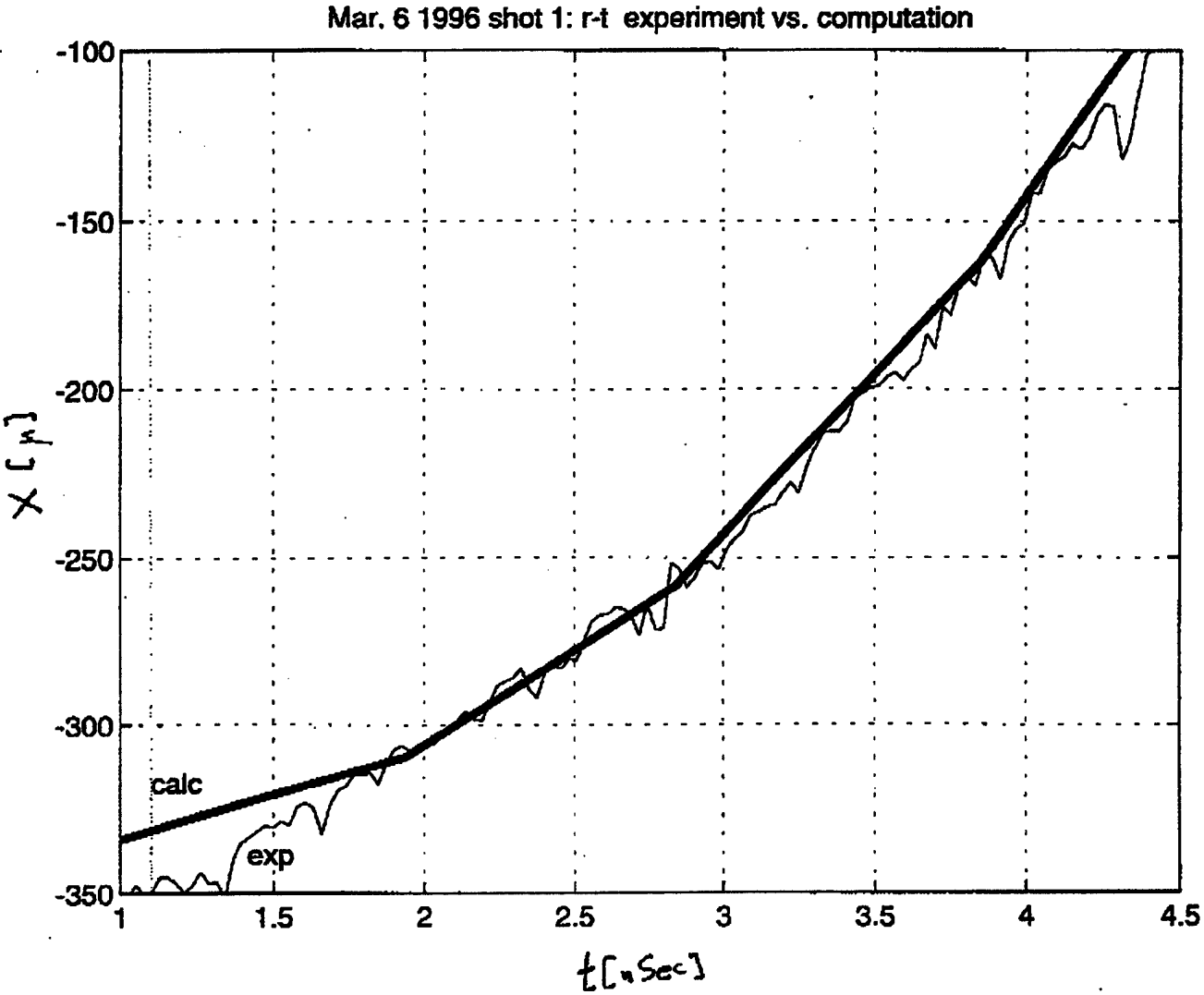

Mar. 6 1996 shot 1: r-t experiment vs. computation
-100
-150
-200
-250
-300
-350
X [μ]
calc
exp
1
1.5
2
2.5
3
3.5
4
4.5
t[μSec]

Figure A.13-5

plasma density, the magnetic field and the time scale of the current source. The results were applied to the ionospheric generation of ELF/VLF radiation using amplitude modulated HF heating. Details of this work appear in a paper published in Physics of Plasmas entitled "Electromagnetohydrodynamics response of a plasma to an external current pulse" and appears in this report as Appendix R.

B.2 Cerenkov excitation of the earth ionosphere waveguide

A comprehensive theoretical analysis of direct Cerenkov excitation of the earth-ionosphere waveguide using ionospheric heating was performed. The model relies on transient ionospheric heating with a heater spot moving horizontally at the bottom of the waveguide with speed close to the speed of light. The cases of isotropic ionospheric conductivity, corresponding to heating altitudes below 70 km, and of anisotropic conductivity, corresponding to heating at higher altitudes were examined separately. We found that enhanced radiation coupling requires that the speed of the heater approach the speed of light. For the anisotropic case such enhancement occurs independently of the direction of motion, while for the isotropic case motion parallel to the ambient electric field is required. Details of this work have been submitted for publication in Radio Science in a paper entitled "Direct Cerenkov Excitation of Waveguide Modes by a Mobile Ionospheric Heater" and it appears here in Appendix S.

B.3 Excitation of the TEM mode by a moving source

The excitation of the long range TEM mode in the earth-ionosphere waveguide by a current source moving horizontally in the lower ionosphere was studied. The moving current source was generated by a novel HF modification scheme that requires the horizontal sweeping of the HF heater beam. Motion of the current source in the ionosphere can excite plasma waves similar to the traditional Cerenkov excitation by a charged particle moving through a media. The excited waves may propagate into the earth-ionosphere waveguide and couple to the TEM waveguide mode at ELF/VLF frequency. The resulting radiation pattern of the TEM wave at the ground level peaks in the direction of the motion. The frequency spectrum, wave amplitudes, and power of injection were derived. Our analysis included realistic effects such as anisotropic plasma and vertical density profile in the ionosphere. A paper describing this effort in detail has been written but not yet submitted for publication. Titled "Excitation of ELF/VLF Waves in the Earth-Ionosphere Waveguide by a Moving Current Source" it is included in this report as Appendix T.

B.4 ULF excitation by ionospheric heating

Excitation of ULF waves by HF heating including chemistry effects was examined. We showed that in the ULF range electron heating reduces the electron recombination rate with NO^+ and O_2^+, leading to electron density enhancement and concomident modification of the conductivity at ULF. Strong Pc1 and Pc2 signals can be produced. Applications to underground

imaging were considered. Details of this work appear in Appendix U. This work was presented at the Santa Fe meeting of the HF Active Auroral Research Program (HAARP) in April 1996.

B.5 Ionospheric focused heating experiment

SAIC was a participant in analyzing measurements taken in support of the Ionospheric Focused Heating (IFH) Experiment. This experiment employed a rocket that released 30 kg of CF_3Br into the F region ionosphere, resulting in an ionospheric hole (electron depletion). Two radio beams from Arecibo were focused on that region of the ionosphere and showed substantial enhancements in power density. Simultaneously, in situ ionospheric measurements were made and airglow measurements were obtained. This work has been described in detail in a publication in the Journal of Geophysical Research titled "The Ionospheric Focused Heating Experiment" and appears in this report in Appendix V.

B.6 Plasma sheet boundary layer

SAIC, in collaboration with Dr. Gurudas Ganguli of NRL, studied effects on the plasma sheet boundary layer (PSBL) during aurorally active periods, at which time it becomes stressed. Simulations performed showed that the electron-ion-hybrid (EIH) instability plays a key role in the nonlinear relaxation of the stressed PSBL. This work is described in detail in a publication in Geophysical Research Letters titled "Relaxation of the Stressed Plasma Sheet Boundary Layer." It appears in this report in Appendix W.

C. NOVEL LASER INVESTIGATIONS

C.1 Generation of far infrared radiation

In this work a device is proposed and analyzed that combines features of photo-switched semiconductors and of radiation conversion using a periodic capacitor array. The device is referred to as a Photo-switched Periodically-biased Semiconductor (PPS). Radiation is generated by applying a static, periodic electric field across the surface of a planar semiconductor. A laser pulse is injected into the device propagating along and nearly parallel to the surface and perpendicular to the applied static field. The propagating pulse generates photocarriers within the semiconductor and a transient current develops in response to the applied periodic electric field; the currents generate electromagnetic radiation. A planar conductor, placed parallel to the photoconductor will confine and enhance the generated radiation. For typical parameters tunable electromagnetic radiation can be generated with wavelengths in the 50-500 μm range, pulse durations in the pico - or subpicosecond range, and peak powers on the order of 100 W.

A paper describing this work in detail has been submitted and accepted by Physical Review E. It is titled "Radiation Generation by Photo-Switched, Periodically-Biased Semiconductors" and is included in this Report in Appendix X.

C.2 Intense laser pulse solitons in a plasma

An analysis was performed that uncovered one-dimensional soliton behavior associated with intense laser pulses propagating through a uniform plasma. The solitons are classified by the number, N, of nodes of the envelope of the laser vector potential. The properties of these solitons are investigated for N=0 and N=1 and for the N>>1 case. This work has been submitted for publication in Physical Review Letters and is included in this report in Appendix Y.

D. PLASMA OPENING SWITCH SIMULATIONS

The plasma opening switch (POS) is the key component for compact terawatt inductive-energy-store pulsed-power systems. In this effort the detailed physics associated with high density POS's has been investigated. The numerical tools include two fluid MHD and particle-in-cell (PIC) codes. In the closed state, the POS must conduct megampere currents for times approaching 1 μs. During this time electrical energy is converted into magnetic energy and stored in the primary storage inductor. To achieve high output powers, the POS must open on a time scale short compared to the conduction time and deliver the magnetic energy to a particle-beam-diode or imploding-plasma load. The power flow as the switch opens, out of the POS and into the magnetically insulated transmission line (MITL), that connects the POS and load, has not been well understood. Simulations were performed with the Mission Research Corporation PIC code, MAGIC, to analyze POS experiments on the Decade Prototype Module 1 (DPM1) at Physics International. Details of these simulations and results have been published in the Journal of Applied Physics, titled "Power flow between a plasma-opening switch and a load separated by a high-inductance magnetically insulated transmission line." This paper is included as Appendix Z.

In a related effort we focused on electrostatic gap formation processes near a cathode in the context of long-conduction-time (~ 1 μs) POS. The work extends previous PIC simulations to plasma densities up to 5×10^{15} cm^{-3}, more than two orders of magnitude higher than earlier simulations. A new model was proposed and results of simulations based on it were presented. The model gives significantly different predictions than previous models; differences that the validity of will be determined by future experiments. This work has been published in Physics of Plasmas under the title "Gap formation processes in a high-density plasma opening switch" and is included in this report as Appendix AA.

Finally, recent theoretical studies predict that the magnetic field can penetrate into initially magnetized plasma on a time scale faster than either the Alfven speed or resistive diffusion; we performed PIC code simulations to provide the first verification of Hall penetration, the mechanism of this fast penetration. These results were presented in the context of the POS but have much wider applicability. The simulations used the MAGIC PIC code. In addition to the fast penetration, the PIC simulations show that vortices in the electron flow accompany the penetration

and are a natural consequence of electron inertia. Details of this work appear in a paper entitled "Particle-In-Cell Simulations of Fast Magnetic Field Penetration Into Plasmas Due to the Hall Electric Field" that has been accepted for publication in Physics of Plasmas. It appears in this report as Appendix BB.

E. INVESTIGATIONS IN NONLINEAR DYNAMICS

During this contract period SAIC has performed several different investigations in nonlinear dynamics. The topics are described as follows.

E.1 Control methods for dynamical systems

We have introduced a new control algorithm for spatio-temporal processes and applied it to a reaction-diffusion system modeling chemical kinetics laws in the Couette flow reactor. This algorithm stabilizes unstable states by using the natural dynamics of the system modeled as a time series. The technique involves classical linear control combined with nonlinear analysis embedding techniques. The algorithm was presented in the paper "Controlling Unstable States in Reaction-Diffusion Systems Modeled by Time Series," published in Physical Review E, and appears in this report in Appendix CC.

The control ideas were further pursued in the context of wall bounded turbulent shear flow. Embedding methods were extended to a Karhunen-Loeve model of wall-bounded turbulent flow to explore nonlinear structure around unstable orbits having nonlinearly unstable directions. Control was achieved in a special case that is reported in "Detecting Motion on Center Manifolds from a Time Series-An Example" that appeared in the Proceedings of the IEEE International Symposium on Circuits and Systems and which is attached in Appendix DD.

To further extend these techniques, control methods for time series were also applied to a simple model of coupled pendulums, and an array of diffusively coupled maps that exhibit rich dynamical behavior was created.

Another research direction we pursued was to use Karhunen-Loeve decomposition to represent data with a minimal number of degrees of freedom and create equivalent lower dimensional models for spatio-temporal processes. Initially designed for fluid dynamics experiments, the Karhunen-Loeve decomposition was applied to generate a low-dimensional model for a reaction-diffusion process. The dynamics of the Karhunen-Loeve model was fully investigated over a wide parameter range. Insight was gained into how the dynamics in a high-dimensional model carries over to the low-dimensional model. Also the dynamics of the modes can explain some of the phenomena in the high-dimensional model. The results were reported in a paper titled "Chaos and Intermittent Bursting in a Reaction-Diffusion Process" in the journal Chaos, and it is included here as Appendix EE.

Finally, an algorithm for preserving chaotic transients was developed and applied to a laser model, and a paper is currently in preparation on the subject.

E.2 Pattern recognition applied to analyzing oil debris particles

Debris particles accumulating in machine oil are monitored and analyzed in order to determine the state of wear of the machine. The aim is to assess, by the shape of the particles, the state of wear of the machine and time the necessary parts replacements. The results of this project are described in a paper entitled "Advances in Optical Debris Monitoring Technology". This paper appeared in the Proceedings of the Integrated Monitoring, Diagnostics and Failure Prevention Conference organized by the Joint Oil Analysis Program (JOAP) and the Mechanical Failure Prevention Technologies (MFPT). The meeting was held in Mobile, Alabama, April 1996. One of the main goals of this project is to obtain a classification of the particles according to wear source and wear state. To achieve this aim we have determined features of the boundary of the particles such as aspect ratio, circularity, external compactness, kurtosis of the distribution of angles in the edge contour, and variations of curvature in the edge contour, which allow us to distinguish wear particles at different stages of wear and to match the experimental classification with a quantitative description of the particles. This work involved using inversion methods to analyze scattering data. It will be followed by artificially generating fractal particles to analyze their boundary.

E.3 Analyzing second harmonic scattering data to determine electro-chemical interface structure

We proposed a general approach for establishing correlations between the optical second harmonic response generated from a metal-electrolyte interface and the interface structure. In specific case studies the response function describes the dependence of the dipolar second harmonic generation signal on the local applied electric field, which in turn can be connected via electrochemical models with the molecular structure of the interface. The novelty of the approach consists in that the analysis and interpretation of influences of electrochemical potential on the second harmonic response are made in terms of quantities at the molecular level. An abstract of the paper in preparation follows.

Case Study Analysis of Second Harmonic Scattering
from a Metal-Electrolyte Interface Based on Construction
of a System Response Function

S.G. Lambrakos, P.P. Paulette, and Ioana Triandaf

We present a case study analysis of second-harmonic scattering data for scattering of polarized light from a metal-electrolyte interface. Our analysis is based on the construction of a response function whose form is deduced according to the variation of the second-harmonic-scattering intensity and

certain general and consistent assumptions concerning the molecular-level character of second-harmonic-scattering sites at the metal surface.

E.4 Application of the coupling model to polymer properties

To study cosmic ray acceleration of charged particles under moving magnetic structures, Fermi proposed to use a map [E. Fermi, Phys. Rev. **75**, 1169 (1949)], in which a particle bounces back and forth between a fixed and an oscillating wall. The Fermi map determines the velocity and phase of discrete time n+1 from information at time n.

It is well known also that in billard systems boundaries partly made up by arcs of circles will cause chaotic behavior in the motion of a particle [M.V. Berry, Ann. Phys. **131**, 163 (1981); Proc. Roy. Soc. London A**423**, 219 (1989)]. The simplest of such billard systems is the well known stadium system in which the sharp corners of a rectangular boundary are replaced by arcs of radii R.

We studied the combination of the Fermi map system and half a stadium to determine the effect of additional nonlinearity in the well known Fermi acceleration problem. We compared the relaxation in the Fermi-stadium map with different R's to that in the Fermi map and found the relaxation retarded for different values of R. After a crossover time, the Fermi relaxation can be approximated by an exponential function, while the Fermi-stadium relaxation can be approximated by a stretched-exponential function. The fractional exponent beta decreases further from unity with increasing nonlinearity. The result bears strong similarity to the basic features suggested by the Coupling Model [K.L. Ngai, Comments Solid State Phys. **9**, 127 (1979)] and seen in neutron scattering. Details and results were published in a paper entitled "From Chemical Structure to Viscoelastic Properties of Polymers" and appears in this report in Appendix FF.

We also analyzed a system of interacting arrays of globally coupled nonlinear oscillators. We compared the relaxation in the interacting arrays with different interaction strengths to that in an array not subject to interaction with others and found the relaxation of the latter to be an exponential function of time. On the other hand the relaxation of the interacting arrays is slowed down and departs from an exponential of time. There exists a cross-over time, t_c, before which relaxation of the interacting arrays is still an exponential function. However, beyond t_c, relaxation is no longer exponential but is well approximated by a stretched exponential, $\exp-(t/T)^{\beta}$. The fractional exponent beta decreases further from unity with increasing interaction strength. The result again bears strong similarity to the basic features suggested by the coupling model and seen experimentally by neutron scattering for relaxation in densely packed interacting molecules in glass forming liquids. Details and results will be published in a paper entitled "Relaxation in Interacting Arrays of Oscillators" and will appear in Phys. Rev. E as a rapid communication. It is included here as Appendix GG.

F. COMBUSTION MODELING

Three distinct studies have been carried out for the numerical simulation of combustion processes. They are described in the following subsections.

F.1 Modeling combustion in fire plumes

The goal of this research was to develop a combustion model appropriate for buoyancy driven fire plumes and thus to study the structure and energetics of hydrocarbons burning in a pool fire configuration. We made use of finite chemical kinetic rate equations to numerically simulate a laminar diffusion flame. The code was constructed to consider the viscous effects in a mixing layer, heat conduction, the multi-component diffusion and convection of important species, the finite rate reactions of these species, and the resulting interaction between the fluid mechanics and the chemistry. The numerical model was used to obtain a detailed description of laminar diffusion flames obtained above 2-D methane/air burners. Results were compared with experimental thermocouple measurements for methane-air burners with similar geometries and flow configurations. This research is an ongoing study and is described in more detail in Appendix HH. We plan to submit a paper describing this work for publication to the Journal of Fire Research. It will be titled "Numerical Simulation of Combustion in Fire Plumes."

F.2 Burning rates of RDX propellants

A mathematical model for a three-tiered system consisting of solid, liquid, and gas was derived for studying the combustion of RDX propellants. The resulting nonlinear two-point boundary value problem was solved by Newton's method with adaptive gridding techniques. In this study the burning rate is computed as an eigenvalue, which can remove the uncertainty associated with employing evaporation and condensation rate laws in its evaluation. Results were obtained for laser-assisted and self-deflagration of RDX monopropellants and were compared with experimental results. The burning rates were computed over a wide range of ambient pressures and compared well with experimental results from one to ninety atmospheres. The burning rate was found to be proportional to the pressure raised to the 0.76 power. Sensitivity of the burning rate to initial propellant temperature was calculated and found to be extremely low, in agreement with past theoretical predictions and experimental data.

Results for laser-assisted combustion showed a distinct primary and secondary flame separated by a dark zone, the length of which is dependent upon the incident laser flux intensity. A detailed description of this work is provided in Appendix II. A paper describing this work has been accepted for publication in the Journal of Combustion Science and Technology and will appear shortly. The title is, "An Eigenvalue Method for Computing the Burning Rates of RDX Propellants."

F.3 Burning rates of HMX propellants

A mathematical model for a three-tiered system consisting of solid, liquid, and gas was derived for studying the combustion of HMX propellants. The resulting nonlinear two-point boundary value problem is solved by Newton's method with adaptive gridding techniques. In this study the burning rate was computed as an eigenvalue, which removes the uncertainty associated with employing evaporation and condensation rate laws in its evaluation. Results were obtained for laser-assisted and self-deflagration of HMX monopropellants and were compared with experimental results. The burning rates were computed over a wide range of ambient pressures and compare well with experimental results from one to ninety atmospheres. The burning rate was found to be proportional to the pressure raised to the 0.82 power. Sensitivity of the burning rate to initial propellant temperature is calculated and found to be extremely low, in agreement with past theoretical predictions and experimental data.

Results for laser-assisted combustion showed a distinct primary and secondary flame separated by a dark zone, the length of which is dependent upon the incident laser flux intensity. A detailed description of this work has been provided in Appendix JJ and has been submitted to the Journal of Combustion and Flame for archival publication. The title is, "An Eigenvalue Method for Computing the Burning Rates of HMX Propellants."

G. MODELING AND SIMULATION OF VAPOR PHASE PROCESSING AND MANUFACTURING

The objective of this SAIC research program is to develop a numerical simulation capability that will aid in understanding, development, and improvement of advanced vapor phase processing and manufacturing technologies. The program is divided into focus areas that address three particular material processing technologies: Thermal Spray Coatings; Metalorganic Chemical Vapor (MOCVD) deposition; and Flame Deposition and Synthesis of Nanoscale Materials. To date significant progress have been made with simulations in all three areas. The following address our specific achievements and the current status of research in the focus areas.

G.1 Thermal spray processing

SAIC has developed the first comprehensive simulation capability for thermal spray deposition based on a High Velocity Oxy/Fuel system. The model includes accurate simulations of both gas and particle flow, and allows analysis of the effects of particle/gas turbulence interaction on particle parameter distribution. The numerical simulations were benchmarked against experimental results and showed excellent agreement for both gas and particle flow predictions. SAIC used the numerical simulations to analyze the key parameters that affect the quality and efficiency of thermal spray systems. The methodology and results of analysis were presented to industry at a number of national conferences, and were included in five conference papers (four

proceedings publications), and two journal articles. The journal articles are titled "Numerical Analysis of a High-Velocity Oxygen-Fuel Thermal Spray System," published in the Journal of Thermal Spray Technology and included here as Appendix KK and "Use of Thermal Spray Methods for Coating Nanoscale Materials," here in Appendix LL.

SAIC signed a CRADA with Idaho National Engineering Laboratory (INEL) that is dedicated to further validating the developed numerical model, as well as to experimentally verifying the coating regimes that were developed using numerical simulation. INEL/SAIC experiments under the CRADA began in March 1996.

The University of Connecticut, which is subcontracted to develop TS coatings of nanostructured material, has demonstrated single phase, high density WC/Co coatings, that have the same parameters as high pressure, high temperature sintered WC/Co. These coatings were facilitated through the understanding of the TS process gained through our numerical simulations.

SAIC will use the developed numerical capability to improve thermal barrier coatings on turbine blades for ABB Power Plant Laboratories of Windsor, Connecticut (under a separate contract). We are also investigating the use of numerical optimization of coating systems for the WPAFB program, currently being performed by SAIC, to develop environmentally compliant coatings.

Status of the Thermal Spray Focus Area

Under this focus area a unique simulation technology was developed and validated. Currently the main effort is directed towards technology transfer and its application to the development of improved coatings for government and industry.

G.2 Metalorganic Chemical Vapor Deposition (MOCVD)

Stevens Institute of Technology, under subcontract to SAIC, has demonstrated the feasibility of a nanoscale yttria-stabilized zirconia (YSZ) coating with columnar structure and very good thermal properties in an MOCVD reactor. This demonstration has enormous technological importance, since similar materials are now produced in an e-beam PVD reactor that costs about 100 times more than a MOCVD reactor. To date, this type of coating has only been demonstrated in an experimental reactor on a small substrate. It will be a significant challenge to scale up the process from a research reactor to a full scale turbine blade or other engine part that needs thermal barrier coating. The challenge is to assure uniform exposure of the coated surface to the metalorganic precursors.

SAIC performed full three dimensional simulations of the MOCVD reactor with a detailed geometry of the turbine blade (obtained from Howmet Corporation). The simulations were done using tetrahedral grids that allow efficient and accurate description of complex geometries. These simulations are the first example of using unstructured grid methodology for CVD reactor design for a truly complex three dimensional substrate. We performed parametric analysis of the flow

regimes in the reactor. As a result of this analysis, we changed the blade orientation in the reactor, which resulted in to low vorticity uniform flow over the blade even for very high precursor velocities.

Status of the MOCVD Focus Area

For this focus area, SAIC developed a simulation tool that facilitates the design of a MOCVD reactor for a complex three-dimensional substrate. This tool is currently benchmarked against experimental results and some classical problems. Stevens Institute developed a unique YSZ thermal barrier coating as well as diagnostics in a MOCVD reactor. Work is underway to use CFD simulations for reactor design.

G.3 Flame deposition and synthesis of nanoscale materials

SAIC has developed a simulation model for a low pressure Flat Flame (FF) reactor used by our subcontractor, Rutgers University. The FF reactor is an implementation of Chemical Vapor Condensation methods that allow high rate production of nanoscale powders of ceramic and other materials. Rutgers University has demonstrated the use of the FF reactor for the synthesis of nanosize SiO_2, TiO_2, and other materials. The FF reactor is a scaleable device that can be used for the production and deposition of nanoscale materials with unique properties. Rutgers University has equipped their FF reactor with extensive nonintrusive diagnostics that allow measurement of gas temperature and composition. The numerical simulation of the processes in the reactor will lower the cost of reactor design and will facilitate its rapid scale-up from a laboratory scale reactor to an industrial process.

SAIC has simulated flow and flame conditions in the Rutgers FF reactor. We have developed a five-reaction chemical kinetics model, which is able to simulate a steady, self-sustained flame in the reactor. The simulation results were benchmarked against experimental data and found to be in good agreement. SAIC is working on further validation of the numerical model through detailed diagnostics and development of the reduced chemical kinetics models for the simulation of a range of materials synthesis processes.

Status of the Flame Deposition and Synthesis of Nanoscale Materials Focus Area

SAIC developed a model for the simulation of the FF reactor. Simulations are currently being benchmarked against experiments. Preliminary results show good agreement. Rutgers University developed and installed diagnostics for the reactor and demonstrated synthesis of nanoscale materials in the reactor.

Technology Transfer Status

We have continued interaction with our industrial team members at TAFA, and Engelhard, and have developed contacts with other players in the thermal spray industry. We have also developed a working relationship with government labs that are active in the area of thermal spray. Our effort has been recognized by the industry, which resulted in an assignment to use

developed CFD tools for improving the quality of an industrial TS coating process. We are collaborating fully with our university team members for the experimental validation effort, and for defining the simulation technology requirements for materials science and technology needs.

H. A COMPREHENSIVE MODEL OF PLASMA ETCH REACTORS

This SAIC modeling effort involves a collaboration with Cornell University and with two etch reactor manufacturers, GaSonics International, Inc., manufacturer of a high pressure chemical downstream etch (CDE) reactor, and Plasma & Materials Technologies, Inc. (PMT) manufacturer of a low pressure helicon reactor. The objective is to (1) develop a detailed computational model of plasma etch reactors valid over a broad range of operating parameters (pressure, input power, flow rate, and geometry); (2) identify important physical mechanisms that govern reactor performance and validate model assumptions and algorithms with experimental data; and (3) construct simplified codes for use by industry for design and control of reactor performance.

Among the elements required in a plasma etch reactor are the capability of determining the (a) electromagnetic field including the impact due to the response of a high density plasma, (b) spatial dependence of the electron energy distribution, (c) ion dynamics, (d) neutral flow, (e) plasma and neutral chemistry, and (f) surface effects at the reactor walls and substrate.

Because of the wide range in time scales in these reactors we have developed a frequency domain (FD) electromagnetic solver (fastest time scale in the reactor) to determine the microwave modes and provide the self consistent electric field in the presence of plasma. This required the development of a new algorithm capable of generating a converging solution for the non-positive-definite and non-diagonally dominant operator that results on inclusion of plasma response terms in Maxwell's equations.

The current status of code development for the PMT reactor is as follows: We have

(a) calculated the RF E field including the effect of a high density plasma ~ $5x10^{12}cm^{-3}$.

(b) calculated typical (low energy) electron orbits in the reactor using a particle code. Here the electrons are constrained by the static B field and energized by the RF E field. The calculation was performed for three cases: (1) no collisions, no RF field, (2) collisions, no RF field, and (3) collisions plus RF field.

(c) used a Direct Simulation Monte Carlo (DSMC) code including chemistry to calculate neutral and ion densities with ions constrained to move along the static B field. Density contours determined include Cl_2, Cl, Cl_2^+, Cl^+, and Cl^-.

The status of GaSonics reactor code development:

(a) local Boltzmann code for electron energy distribution extensively tested.

(b) module for oxygen-nitrogen chemistry developed and tested.

(c) modules for thermal conductivity, diffusion, and viscosity developed and tested.

(d) a fluid code using a finite analytic Navier-Stokes (FANS) algorithm has been implemented for the GaSonics reactor's plasma discharge tube.

(e) simulation of flow in a plasma discharge tube using a high aspect ratio simplification has been performed in presence of a model plasma discharge distribution.

In addition to the above, two separate studies were carried out: (1) a special numerical investigation for the PMT reactor of the effect of input nozzle placement on the efficiency and uniformity of flux to the substrate and (2) for the GaSonics reactor an experimental investigation of the role of impurities in the oxygen CDE resist etching process.

A set of viewgraphs describing the above in more detail and presenting results was presented to our DARPA sponsor in June 1996 and is included herein as Appendix MM, entitled "A Comprehensive Model of Plasma Etch Reactors." In addition, an invited article in the new journal The Industrial Physicist was written and published. Entitled "Computer Modeling of Deposition and Etching," it is included here as Appendix NN. Also, an article was written in connection with a presentation at the 33rd Aerospace Sciences Meeting of the American Institute of Aeronautics and Astronautics (AIAA) at Reno, Nevada, in January 1995. That article, entitled "Modeling Chemical Vapor Deposition and Etching Processes" is included in this report as Appendix OO. Finally, in Appendix PP we include a paper titled "Role of Impurities in O_2 Chemical Dry Etching." This paper details an experimental investigation carried out at SUNY Albany on the effect of the impurities N_2 and H_2 in an oxygen resist CDE reactor on parameters including the etch rate, oxygen atom concentration in the discharge region, and surface modifications on the resist.

Appendix A

A KrF Oscillator System with Uniform Profiles

Optics Communications 106 (1994) 113–122
North-Holland

OPTICS
COMMUNICATIONS

Full length article

A KrF oscillator system with uniform profiles

A.V. Deniz
Science Applications International Corporation, McLean, VA 22102, USA

and

S.P. Obenschain
Plasma Physics Division, U.S. Naval Research Laboratory, Washington, DC 20375, USA

Received 1 June 1993; revised manuscript received 19 August 1993

A KrF oscillator system that has produced highly uniform flat-top focal distributions is described. The oscillator system is part of a large laser system that will utilize the echelon-free induced spatial incoherence technique to obtain uniform illumination of planar targets for fusion research. With this system, focal profiles with small long scale length nonuniformities have been obtained. The nonuniformity was determined by performing a least-squares fit to a series of profiles, and calculating the deviation of each fit from a flat-top profile. With a linear fit, the deviation averaged over the series is $\pm 0.5\%$, and with a quadratic fit, it is $\pm 1.4\%$. Details of the oscillator system configuration, focal uniformity measurement techniques, and resulting focal profiles are presented.

1. Introduction

One of the requirements for high-gain direct-drive inertial confinement fusion is a highly symmetric implosion of the spherical fuel pellet. Ablation pressure nonuniformities less than a few percent are thought to be required. If an ultraviolet wavelength is used (which couples more efficiently to the target than longer wavelengths), then there is only modest lateral smoothing of the ablation pressure [1]. Therefore, success with direct drive laser fusion requires the development of techniques for highly uniform illumination of fuel pellets. It is perhaps impossible with today's technology to have a uniform focal illumination with a nearly diffraction limited beam. Progress has instead been made by methods that employ controlled spatial and temporal incoherence with focal profiles that are smooth when averaged over many temporal coherence times [2–11]. These smoothing techniques have been shown to reduce laser plasma instabilities [12–18]. However, with existing high energy glass lasers, peak-to-valley focal nonuniformities still are typically $\sim 10\%$ with these techniques [9–11].

The NIKE KrF laser system [19–21] currently under construction at the Naval Research Laboratory, is being built to have focal profiles that are uniform enough to produce ablation pressures flat to within 2% on planar targets. NIKE will use the echelon-free induced spatial incoherence (EFISI) technique [3] to produce 60 time diffraction limited flat-top focal profiles with at least two kilojoules in 4 ns on target.

The EFISI technique is illustrated in fig. 1. An oscillator with spatially incoherent output illuminates an object aperture, whose image is relayed by the two lenses to the image plane. Light from each point in the object aperture illuminates an aperture at the Fourier plane with the same intensity. The light passes through the object aperture, is amplified at the Fourier plane, and is then focused at the image

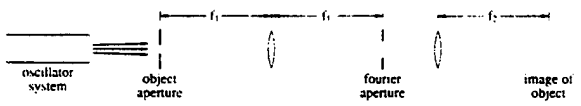


Fig. 1. A simplified schematic of the echelon-free induced spatial incoherence technique.

SSDI 0030-4018(93)E0422-C

Volume 106, number 1,2,3 OPTICS COMMUNICATIONS 1 March 1994

plane. With the EFISI technique, the image profile would not be strongly affected by the spatial nonuniformity of the amplifier gain because light from each point in the object aperture is amplified by the same amount. Also, because the image is already many times diffraction limited, it will not be strongly affected by the phase aberrations of the laser system.

There are several requirements that the oscillator must meet:

(i) The light at the object aperture must have sufficient spatial incoherence (divergence) to fill the Fourier aperture.

(ii) The oscillator output must have sufficient temporal incoherence to produce a time-averaged smooth profile.

(iii) The light at the object and Fourier apertures must produce a uniform flat-top profile at the image plane.

(iv) The light from each point in the object aperture must illuminate the Fourier aperture with the same profile.

Requirements (i) through (iii) are perhaps straightforward. Requirement (iv) is necessary so that the focal profile be insensitive to gain nonuniformities of an amplifier located at the Fourier plane. These requirements are discussed in the next section.

Here we report on a KrF oscillator system with unconventional resonator optics that has come close to meeting the above requirements. This oscillator, with a two stage Pockels cell pulse slicer, produces 4 ns flat-top focal profiles (using an $f/130$ lens) with tilts on the order of 1% and the temporal coherence time of 0.6 ps. When a nonuniform amplifier gain was simulated by blocking half the Fourier aperture, the tilts were still less than 3%. This system approached the focal uniformity goals for the NIKE laser.

In this oscillator, the laser medium is imaged back onto itself by the resonator optics. Thus a photon that makes a large angle with the longitudinal axis will still pass through the laser medium, even after many transits (unless the angle is so large that the finite size of the laser chamber windows blocks it). The output of this oscillator therefore has greater angular divergence than the output of an oscillator with conventional stable resonator optics. In addition, the Fourier aperture is illuminated more nearly uniformly by each point in the object aperture, as required by the EFISI technique.

The next section describes the oscillator setup in greater detail and presents its measured characteristics. Section 3 presents details of the pulse slicing system. Section 4 presents details of the imager used to measure the profiles. Section 5 presents the focal profiles and the algorithms used to evaluate them. Section 6 summarizes the results and presents the conclusions.

2. The oscillator system

As mentioned in the introduction, the oscillator output must be spatially incoherent so that it can produce a beam that is many times diffraction limited. The focal profile does not then depend strongly on the phase errors encountered during propagation. The focal profiles presented are 60 times diffraction limited, although the NIKE system has been designed to propagate beams which are up to 120 times diffraction limited without vignetting. If both the object and Fourier apertures are circular, then the beam is N_d times diffraction limited with

$$N_d = Dd/\lambda f, \tag{1}$$

where D is the Fourier aperture diameter, d is the object aperture diameter, λ is the wavelength of the light (248 nm), and f is the focal length of the lens between the object and Fourier apertures.

The second requirement is that the oscillator output be temporally incoherent. At any given time, the focal profile is a complicated speckle pattern. After one coherence time, it is a different pattern. The measured profiles are time-integrated, so that if the measurement is made over more coherence times, then both the shot-to-shot variation due to the speckle and the random variation from one coherence zone to the next are decreased. The coherence time must be short enough so that these variations are not too large.

The third requirement is that light from each part of the object aperture illuminate the Fourier aperture with the same angular energy distribution. This requirement can be stated more precisely as follows: Let $\bar{I}(\boldsymbol{x}, \boldsymbol{k})$ be the time-averaged intensity profile at point $\boldsymbol{k}$ within the Fourier aperture due to light from several coherence zones (which have a size of

$\sim d/60$) around point x in the object aperture. The requirement becomes

$$\nabla_x \bar{I}(x, k) = 0 . \quad (2)$$

Note that $\bar{I}(x, k)$ need not also be uniform across the Fourier aperture. If eq. (2) is satisfied, then the focal profile will be flat, even if the light is subjected to nonuniform amplification at the Fourier aperture.

Equation (2) can be checked by placing a small pinhole in the object aperture and measuring the time-averaged intensity profile at the Fourier aperture. The resulting Fourier intensity profile should be the same regardless of where the pinhole is placed within the object aperture. Adherence to eq. (2) can also be checked by blocking various regions of the Fourier aperture, and observing whether or not the image maintains its flat-top shape. The latter method was used in this work.

Figure 2 shows the optical setup for the oscillator used for the results presented here. The oscillator optics image the laser medium back onto itself. All the oscillator configurations investigated used a 1 cm × 2 cm × 80 cm discharge pumped KrF laser medium. The rear optics consist of a positive lens and a high-reflectivity flat mirror with an aperture, and the front optic is a 50% reflectivity flat mirror. The distance between each mirror and the lens is equal to the focal length of the lens (1 m). Light from point a inside the laser medium will be imaged, after reflection by both the front and rear optics, at point b. This property of imaging the laser medium back onto itself produces a large angular divergence.

Conventional discharge oscillator optics did not satisfy these requirements. A stable resonator setup consisting of a flat front mirror and a rear mirror with a large radius of curvature (5 m, 10 m, and ∞) was tested. While the image of the object aperture was flat, the angular divergence was too small to produce the required 60 times diffraction limited beam.

The front of the laser medium (which is approximately 10 cm from the chamber window) was relayed with a telescope to the object aperture. This geometry was found empirically to produce the flattest images. The image was not as flat when the center of the discharge was imaged onto the object aperture, or when the object aperture was ~ 50 cm from the front oscillator mirror outside the oscillator (no telescope was used in the latter case).

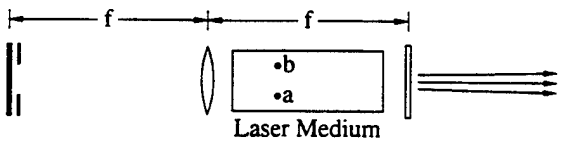


Fig. 2. The optical setup for the oscillator.

The images produced by this setup had flat-top focal profiles, but with residual tilts on the order of 10% in the vertical direction. This was most likely due to irregularities in the laser cell and/or electrodes. The tilts also varied during the oscillator pulse. Finally, there was a shot-to-shot variation of the tilts, presumably caused by irreproducibilities of the discharge.

All profile tilts were substantially reduced by adding the telescope system shown in fig. 3. The oscillator output was split into two beams and then recombined. One beam passed through a single telescope, and the image was inverted. The other beam traveled the same distance but passed through two telescopes, and the image was not inverted. Combining the beams at the object aperture resulted in images with a very small tilt ($\sim 1\%$), independent of when the oscillator pulse was sliced. The tilts of the profiles also had a very small shot-to-shot variation.

The rear reflector of the oscillator, which is located close to a Fourier plane, was apertured to limit the angular divergence of the output at the object aperture. This affected the concavity of the profile at the image plane (after passing through the Fourier aperture). Decreasing the angular divergence by placing a smaller aperture at the rear reflector tended to make the image concave up, while increasing the divergence had the opposite effect. With no rear reflector aperture, the image was slightly concave down. The rear reflector aperture was made smaller until

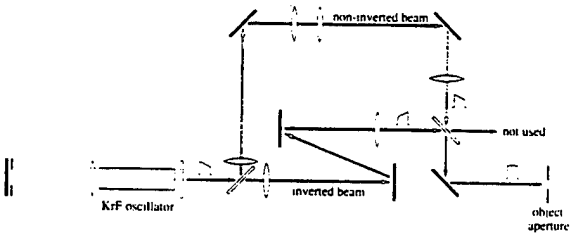


Fig. 3. The optical setup for the oscillator and telescopes.

the concavity was nearly eliminated. For the results in this work, the divergence from the object aperture was limited to ~8 mrad in the vertical direction, and ~14 mrad in the horizontal direction. This divergence is sufficient to overfill the Fourier aperture, which requires only 5 mrad.

The object aperture diameter of 3 mm was limited by the transverse dimension of the laser medium (1 cm by 2 cm), and the Fourier aperture size was limited by the size of the Pockels cells to 0.5 cm × 0.5 cm, which implies a focal length of 1 m. The required angular divergence of the oscillator is $60\lambda/d \sim 5$ mrad, which is less than the measured angular divergence.

Two KD*P Pockels cells in series were used to slice a 4 ns pulse out of the 30 ns oscillator output. The setup is shown in fig. 4. A pair of dielectric polarizers polarized the beam horizontally before the object aperture. The beam then passes through the Fourier aperture, the two Pockels cells, and the remaining sets of dielectric polarizers. The energy contrast ratio, defined as the ratio of the transmitted fluences with and without voltage applied to the Pockels cells, was 3000 to 1.

3. The imager

The imager is a cooled, slow-scan, two-dimensional charge coupled device (CCD) camera. It is capable of measuring the energy profiles to an accuracy of better than 1%, and it has a spatial resolution of 384 by 576. The camera parameters are shown in table 1. The CCD is coated with a phosphor in order to increase its quantum efficiency at 248 nm to ~0.25.

The imager has a vacuum window to prevent formation of frost on the CCD. This window has an antireflection coating on each surface with a power reflectivity of 0.25% at $\lambda = 248$ nm, 0° incidence. This

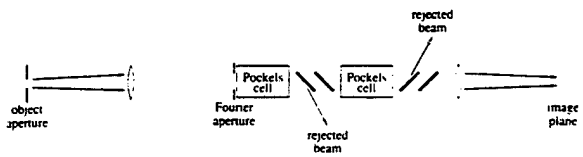


Fig. 4. The pulse slicing system.

Table 1
Parameters of the CCD camera.

Parameter	Value
Format	384 by 576 pixels
Pixel size	23 μm by 23 μm
Readout noise	25 electrons
Charge transfer efficiency	0.99998
Full well capacity	160000 electrons/pixel
Quantum efficiency	0.25 at 248 nm
Dark current	15 electrons/sec/pixel
Exposure time	0.1 s
Digitizer resolution	12 bits
CCD temperature	−45°C
Window thickness	1.0 cm
Window reflectivity	0.25%, each surface

introduced a negligible error in the measurement of the profile because the coherence length of the light (1.9×10^{-2} cm) was much shorter than the window thickness.

For temporally coherent light, the noise in a cooled CCD measurement is due mainly to the statistical nature of the photoelectric process (the electron shot noise) and the preamplifier noise. For the measurements presented the signal was large enough that the preamplifier noise was negligible. Because the number of photoelectrons in one pixel for several measurements of identical light levels has a Poisson distribution, the electron shot noise σ_{ccd} for one measurement is

$$\sigma_{ccd} = \sqrt{N_e} \,. \tag{3}$$

The noise in one pixel is independent of the noise in another.

In addition to the noise in the CCD measurement, there is a random variation σ_i associated with the temporal incoherence of light. If a pixel in the detector is smaller than a spatial coherence zone, then

$$\sigma_i = N_e \, (\tau/T)^{1/2} \,, \tag{4}$$

where T is the laser pulse length, τ is the coherence time, N_e is the average number of photoelectrons in the pixel measured in time T, and σ_i is the shot-to-shot RMS variation of the signal in electrons. In this case, there is correlation in the measurements of adjacent pixels. If on the other hand there are N_c coherence zones in a pixel, then

$$\sigma_i = N_e \, (\tau/N_c T)^{1/2} \,. \tag{5}$$

In this case, the random variation for adjacent pixels is independent. The bandwidth ($\equiv 1/\tau$) of the oscillator has been measured to be 1.6 THz, which gives a coherence time of 0.6 ps. Because the random variation associated with the incoherence is not correlated with the CCD measurement noise, the total random variation σ_y in electrons of the number of electronics in a pixel is given by

$$\sigma_y^2 = \sigma_i^2 + \sigma_{ccd}^2 . \quad (6)$$

Because a typical CCD exhibits a nonuniform response to light, the camera was calibrated by illuminating it with a uniform light source (fig. 5). The oscillator was used as a light source for a commercial integrating sphere. This has the advantage of calibrating the camera for the same wavelength and approximately the same pulse length used in the measurements. The camera was placed ~1 m from the output aperture of the sphere. At that distance, the variation in light across the CCD surface from a lambertian surface is theoretically 1.5×10^{-4}. In an actual setup, the variation will probably be greater; the camera and integrating sphere might not be perfectly aligned, there might be spurious reflections, and the light output from the integrating sphere might not be perfectly uniform.

In order to test the calibration technique, the camera was calibrated at different positions. The distance between the camera and the integrating sphere was varied from 0.4 m to 1.6 m, the camera was moved ±3 mm perpendicular to a line between it and the sphere, and the camera was rotated about that line. The greatest change in the calibration was a 0.3% spatial tilt across the 8.83 mm width of the sensor. This corresponds to a systematic error of less than 0.1% in the calculation of the tilt of a 2 mm diameter profile.

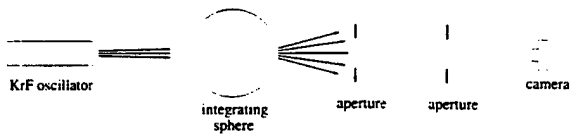


Fig. 5. Calibration of the camera. The oscillator illuminates an integrating sphere. A negative lens (not shown) is used to spread out the input light to avoid damaging the integrating sphere. A tube (not shown) is used in order to reduce stray light. Apertures inside the tube are necessary in order to reduce stray reflections.

During the calibration, there will be the random variation in the measurements given by eq. (6). There were several coherence zones in each pixel, so eq. (5) was used to calculate the random variation due to the incoherence of light. The size δ of a coherence zone at the CCD is approximately

$$\delta = \lambda l / D_l , \quad (7)$$

where λ is the wavelength of light and l is the distance between the CCD and the integrating sphere output aperture, which has a diameter D_l. For $\lambda = 248$ nm, $l = 1$ m, and $D_l = 5$ cm, δ is 4.96 μm. Using the pixel size from the table, the value of N_c (to be used in eq. (5)) is $(23\ \mu m/4.96\ \mu m)^2 = 21.5$.

In order to measure the noise, the output from the integrating sphere was measured 64 times. After correcting for variation in the total energy falling on the CCD, the noise-to-signal ratio (σ_y/N_e) for a given pixel was calculated. At a signal level of 6.6×10^4 electrons, the result ranged from 2.5×10^{-3} to 3.0×10^{-3}, compared to an expected value (from eq. (6)) of 3.8×10^{-3} ($T = 30$ ns, $\tau = 0.6$ ps, and $N_c = 21.5$). This indicates that the camera is capable of measuring the profiles with low noise.

The calibration was determined by summing 64 measurements (with the background subtracted) of the integrating sphere output (flat fields). The average signal level was 1.5×10^5 electrons, or 3825 counts. The calibration factor C_i for pixel i is given by

$$C_i = \frac{1}{N_p f_i} \sum_{j=1}^{N_p} f_j , \quad (8)$$

where N_p is the number of pixels on the CCD, and f_i is the summed flat field signal at pixel i. Because f_i is a sum of 64 measurements, the relative error of C_i calculated by eq. (6) due to the error in f_i is reduced by a factor of 8 from 2.8×10^{-3} to 3.5×10^{-4}.

The relative energy e_i falling on pixel i is given by

$$e_i = C_i(s_i - b_i) , \quad (9)$$

where s_i and b_i are respectively the signal and background at pixel i. The signal levels of the profiles presented in the next section are 6.0×10^4 electrons, or ~1500 counts, while the background levels (which are due to the camera) are ~50 counts. This implies that the RMS measurement error due to the electron shot noise and the calibration error is 0.4% for one

pixel. When observing longer scale length focal profile nonuniformities that cover many pixels, one can reduce this error by averaging over more than one pixel.

4. The profiles

The object aperture is imaged onto the camera through a lens, the Fourier aperture, the pulse slicing system, a demagnifying telescope, and focusing lens. A schematic of the experimental arrangement without the demagnifying telescope is shown in fig. 4. The telescope is located before the second lens, which focuses the beam onto the camera after attenuation by two reflections by uncoated surfaces and transmission through a 95% flat reflector. The telescope demagnifies the image of the object aperture from 3 mm to 2 mm diameter.

The measured profiles were analyzed with the following algorithm:

(i) Determine the centroid (x_c, y_c) of the profile.

(ii) Find the edges of the flat region of the horizontal and vertical cross-sections through the centroid.

(iii) For the flat region of each cross-section: (a) Perform a linear least-squares fit, and calculate the variation of the fit from a flat-top (tilt). (b) Perform a quadratic least-squares fit, and calculate the mean-to-peak variation of the fit from a flat-top. (c) Calculate the RMS deviation of the measured cross-section from the fits.

The centroid is determined by

$$x_c = \sum_{i=1}^{N_p} x_i e_i \Big/ \sum_{i=1}^{N_p} e_i , \qquad (10)$$

where x_i is the x position of pixel i. The calculation of y_c is analogous. The edges of the flat region of a cross-section is determined by the following algorithm:

(i) Find the pixels l and r where the measured value of the cross-section is just greater than a specified fraction f_p of the maximum value.

(ii) The edges of the flat region at t pixels from l and r towards the center of the cross-section. For example, if l is the pixel on the left edge of the profile, and pixel numbers increase from left to right, the the flat region is between pixels $l+t$ and $r-t$, inclusive.

For the results presented, f_p was 0.25, and t was 5 (the cross-section of the profile is 85 pixels across). The results do not depend strongly on f_p or t.

Three aspects of the profiles were evaluated. First, the 4 ns pulse was taken at different times during the oscillator output to check whether the object aperture illumination uniformity changed in time. It was found that the tilts of the profiles do not vary appreciably during the oscillator output. Second, a series of profiles were taken at the same time during the oscillator output to determine the shot-to-shot variation of the profiles. It was found that the observed shot-to-shot RMS variation of the tilts was consistent with the spatial and temporal incoherence of the light. It was also found that the variation of the linear fit from a flat-top profile (a measure of tilt) averaged over the series was $\pm 0.5\%$, and of the quadratic fit (a measure of peaking or concavity) was $\pm 1.4\%$. Finally, the Fourier aperture was partially blocked to simulate the effects of a nonuniform gain in an amplifier. The profiles were found to be insensitive to this partial blocking; blocking half the Fourier aperture changed tilts of the profile by only a few percent.

Figure 6 shows a typical profile and its cross-sections. The RMS deviation of the measurement from the linear least-squares fit is 1.0% along the vertical direction and 0.8% along the horizontal. If the profile were flat, the deviation would be the same as the total random variation σ_y (given by eq. (6)) in each datum. Because a pixel is smaller than a coherence zone, eq. (4) can be used to calculate the variation due to the incoherence of the light. With a coherence time of 0.6 ps, a pulse length of 4 ns, and a signal of 6×10^4 electrons, the expected RMS relative variation of each datum is 1.3%. This indicates that the observed deviation is probably due to the CCD shot noise error and the variation of the light energy due to the laser incoherence and the finite averaging time.

Figure 7 shows the tilts of the cross-sections as a function of time during the oscillator output. Eight images were recorded for seven times spaced 5 ns apart during the oscillator output, for a total of 56 images. Each point on the graphs is the tilt of the horizontal or vertical cross-section of one image. Note that the time variation is negligible. Therefore, the

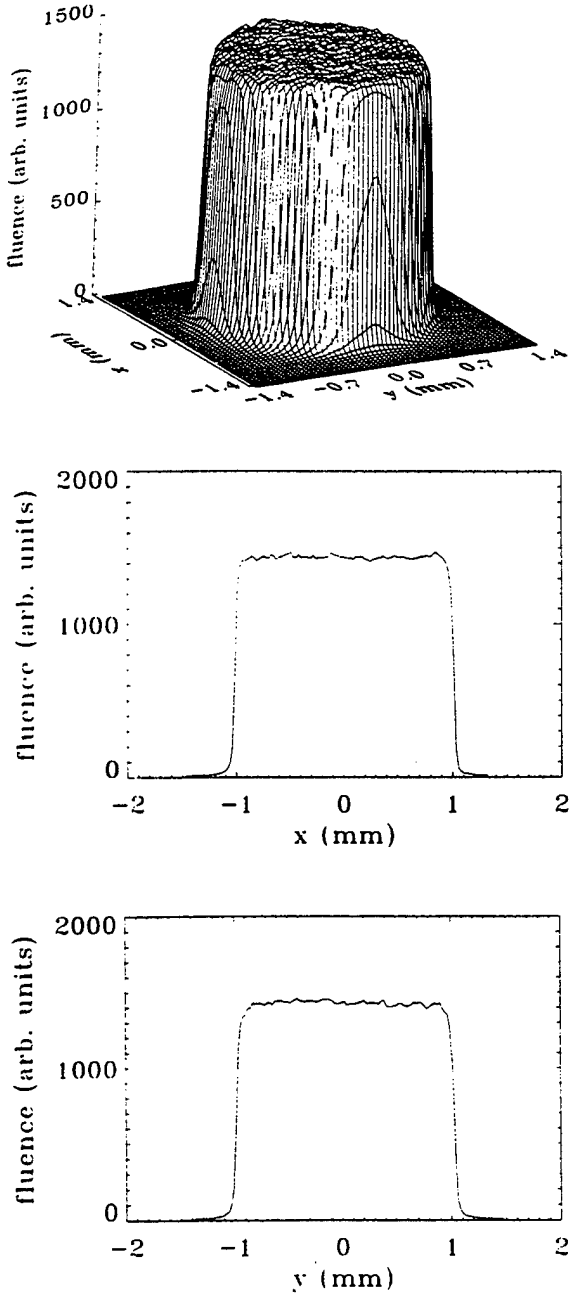


Fig. 6. A typical profile and its cross-sections: (a) three-dimensional view, (b) horizontal cross-section through the centroid, and (c) vertical cross-section through the centroid.

output of the oscillator does not vary appreciably when averaged over 4 ns.

Figure 8 is a histogram of the tilts for 64 images.

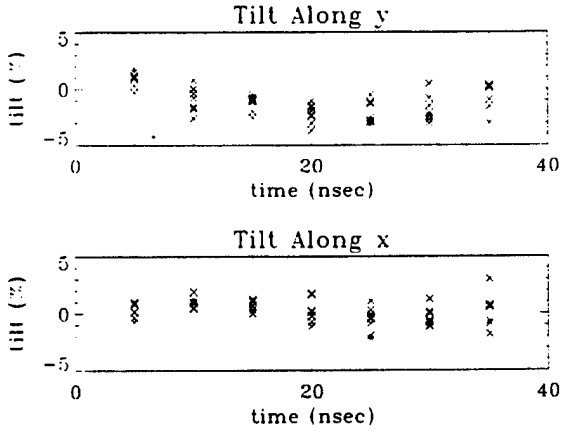


Fig. 7. Tilts of the cross-sections as a function of time during the oscillator output.

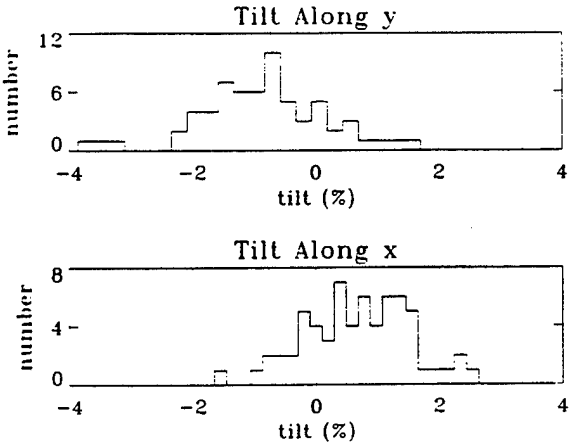


Fig. 8. A histogram of the tilts of the linear fits to the profiles with the Fourier aperture completely unblocked.

Note that the average tilt is very nearly zero; it is −0.75% along the vertical direction, and 0.75% along the horizontal. The RMS variation of the tilt is 1%, which is close to the expected value of 0.7% (this is calculated in the appendix). The RMS variation of the tilt of a single beam of fig. 3 is ~5% along the vertical direction and ~1% along the horizontal. The larger variation along the vertical direction for a single beam indicates that the telescopes shown in fig. 3 are effective in reducing the shot-to-shot variation of the tilt. Furthermore, the average tilt of a single beam is 5 to 10%, which indicates that the telescopes

are also effective in reducing the average tilt.

Figure 9 is a histogram of the mean-to-peak variation of the quadratic fit from a flat-top profile for the 64 images. The mean averaged over the images is $\pm 1.4\%$ along the vertical direction, and $\pm 0.6\%$ along the horizontal. The variation is larger along the vertical direction because there is more shot-to-shot variation of the concavity of the profile, which is not reduced by the telescopes.

Figure 10 shows the tilts of the cross-sections with different parts of the Fourier aperture blocked in order to simulate nonuniform amplifier gain. Sixteen images were recorded with (i) the Fourier aperture completely unblocked, (ii) the top half of the Fourier aperture blocked, (iii) the right half blocked, (iv) the bottom half blocked, and (v) the left half blocked, for a total of 80 images. Note that the average tilt changed by a few percent. Note also that when the top or bottom of the Fourier aperture was blocked, the shot-to-shot variation of the tilt of the vertical cross-section increased, because different parts of the inverted and noninverted beams were blocked. The two beams were not exact inverses of each other, and the the tilt of one did not cancel the tilt of the other. This larger shot-to-shot variation does not occur in the horizontal direction because the angular divergence of the oscillator output in the horizontal direction is larger than the divergence in the vertical direction.

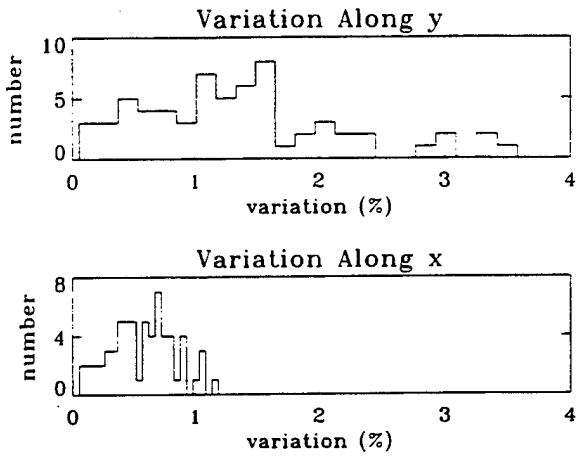


Fig. 9. A histogram of the mean-to-peak variation of the quadratic fits from a flat-top profile with the Fourier aperture completely unblocked.

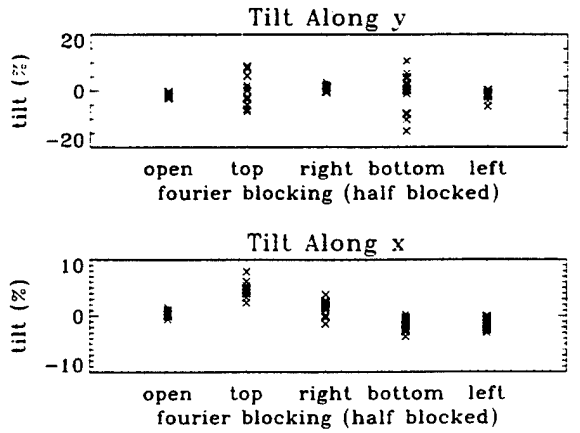


Fig. 10. Variation of the tilts of the cross-section when different parts of the Fourier aperture are blocked.

5. Summary

A laser oscillator system has been developed with a 4 ns pulse output, 3000 to 1 energy contrast ratio, and flat focal profile. Its angular divergence is large enough to produce 60 times diffraction-limited images. The coherence time was 0.6 ps, and the pulse duration was 30 ns. The focal images of 4 ns slices of the beam had the desired flat cross-sections (tilts of $\sim 1\%$). The shape of the focal profile should not depend strongly on the gain profiles of the laser amplifiers if they are placed at the Fourier aperture of the EFISI system. This was tested by partially blocking the Fourier aperture and observing that the shape of the focal profile was not strongly affected (the tilts were typically on the order of a few percent). The top, bottom, left, and right halves of the Fourier aperture were blocked. Partially blocking the Fourier aperture simulates an extremely nonuniform amplifier gain; typical nonuniformities of actual amplifiers used in the NIKE laser are only approximately 20%.

The oscillator system described here comes close to fulfilling the goals for the NIKE system. The existing system is now being used to test the ability of the NIKE system to maintain uniform focal profiles after several stages of amplification. In future work, we will attempt obtain more nearly uniform beams

which are less affected by partial blocking of the Fourier aperture. We will also attempt to produce more than 60 times diffraction limited beams, so that phase aberrations in the laser system will have even less effect on the focal profile.

Acknowledgements

We would like to thank our colleagues for useful discussions and assistance: Mark S. Pronko, Robert H. Lehmberg, Thomas Lehecka, Carl J. Pawley, Warren Webster, Julius Goldhar, and Andrew J. Schmitt. The data was archived on disk using a modified version of National Center for Supercomputer Activities' Hierarchical Data Format. This work was supported by the United States Department of Energy.

Appendix

This appendix presents a calculation of the variance of the slope of the linear least-squares fit. The variance is determined by the variance of the data used to calculate the slope. The random variation of each datum in the measured profile is due to both the CCD measurement noise and the random variation associated with the temporal incoherence of the light. The CCD measurement noise is due mainly to the electron short noise, and is given by eq. (3). The random variation due to the incoherence is given by eq. (4) or (5). The total random variation σ_y (given by eq. (6)) causes the calculated slope m to have a variance σ_m^2.

In order to calculate σ_m, let y_α be a set of N_f random variables, and $y_{\alpha i}$ be the value of sample i of y_α. Let x_α be a set of N_f numbers. The y_α are the measurements at position x_α; N_f is the number of pixels across the profile. Define

$$\bar{y}_\alpha \equiv \frac{1}{N}\sum_{i=1}^{N} y_{\alpha i}\,, \tag{11a}$$

$$\sigma^2_{\alpha\beta} \equiv \frac{1}{N}\sum_{i=1}^{N} (y_{\alpha i}-\bar{y}_\alpha)(y_{\beta i}-\bar{y}_\beta)\,, \tag{11b}$$

$$\beta_1 \equiv \sum_{\alpha=1}^{N_f} x_\alpha\,, \tag{11c}$$

$$\beta_2 \equiv \sum_{\alpha=1}^{N_f} x_\alpha^2\,, \tag{11d}$$

where N is the number of profiles measured. Note that the x_α can be chosen so that β_1 is zero. Note also that in this case, eq. (11d) can be approximated by an integral. Let L be the diameter of the profile, and let $\Delta x = L/N_f$. Then

$$\beta_2 = \frac{N_f}{L}\sum_{i=1}^{N_f} x_\alpha^2 \Delta x\,, \tag{12}$$

$$\approx \frac{N_f}{L}\int_{-L/2}^{L/2} x^2\,\mathrm{d}x \tag{13}$$

$$\approx N_f L^2/12\,. \tag{14}$$

The variance will be calculated by first using the standard formula for the slope of the linear least-squares fit to a set of data. The variance will depend on $\sigma^2_{\alpha\beta}$. However, $\sigma^2_{\alpha\beta}$ will be negligible if the pixels α and β are not close together, because the size of a coherence zone is approximately the same size as a pixel. Using a rough approximation, the ratio σ_m/σ_y is calculated.

The slope m of the linear least-squares fit to the data pairs (x_α, y_α) is a random variable given by (after setting $\beta_1 = 0$):

$$m = \frac{1}{\beta_2}\sum_{\alpha=1}^{N_f} x_\alpha y_\alpha\,. \tag{15}$$

Let σ_m^2 be the variance of m. Because x_α is a number (rather than a random variable),

$$\sigma_m^2 = \frac{1}{\beta_2^2}\sum_{\alpha=1}^{N_f}\sum_{\beta=1}^{N_f} x_\alpha x_\beta \sigma^2_{\alpha\beta}\,. \tag{16}$$

For the results presented, there are 60 coherence zones and 85 pixels across a profile; each coherence zone spans approximately one pixel. In addition, $\sigma_i \sim 750$ electrons (from eq. (4), with $T = 4$ ns, $\tau = 0.6$ ps, and $N_e = 6.0\times10^4$ electrons), while $\sigma_{ccd} \sim 250$ electrons (from eq. (3)). The total variation at one pixel is therefore (eq. (6)) 790 electrons, which is a relative variation of 1.3×10^{-2}. Because the variation due to the measurement (σ_{ccd}) is independent for each pixel, while the variation due to the incoherence of light (σ_i) is somewhat correlated, the rather arbitrary assumption can be made:

$$\sigma^2_{\alpha\beta}=\sigma^2_{\alpha\alpha}\,, \quad \text{if } \alpha=\beta\,,$$
$$=k\sigma^2_{\alpha\alpha}\,, \quad \text{if } |\alpha-\beta|=1\,,$$
$$=0\,, \quad \text{otherwise}\,. \tag{17}$$

Equation (16) then becomes

$$\sigma^2_{\mathrm{m}}=\frac{1}{\beta_2^2}\sum_{\alpha=1}^{N_\mathrm{f}}\left(x_\alpha^2+kx_{\alpha-1}x_\alpha+kx_{\alpha+1}x_\alpha\right)\sigma^2_{\alpha\alpha}\,, \tag{18}$$

where $x_\alpha\equiv 0$ for $\alpha=0$ and $\alpha=N_\mathrm{f}+1$ so that there is no contribution from the nonexistent points 0 and $N_\mathrm{f}+1$.

Because the signal level is nearly the same for all points on the profile, the variances of each of the y_α will all be approximately equal:

$$\sigma^2_{11}=\sigma^2_{22}=\ldots=\sigma^2_{\mathrm{y}}\,. \tag{19}$$

Then

$$\frac{\sigma_\mathrm{m}}{\sigma_\mathrm{y}}=\frac{1}{\beta_2}\left[\sum_{\alpha=1}^{N_\mathrm{f}}\left(x_\alpha^2+kx_{\alpha-1}x_\alpha+kx_{\alpha+1}x_\alpha\right)\right]^{1/2}\,. \tag{20}$$

With $x_\alpha x_{\alpha-1}\approx x_\alpha^2$,

$$\frac{\sigma_\mathrm{m}}{\sigma_\mathrm{y}}\approx\left(\frac{1+2k}{\beta_2}\right)^{1/2}\,. \tag{21}$$

Approximating β_2 by $N_\mathrm{f}L^2/12$, this becomes

$$\frac{\sigma_\mathrm{m}}{\sigma_\mathrm{y}}\approx\frac{2\sqrt{3}}{L}\left(\frac{1+2k}{N_\mathrm{f}}\right)^{1/2}\,. \tag{22}$$

For a 2 mm diameter image with $N_\mathrm{f}=85$ and $k=0.5$, $\sigma_\mathrm{m}/\sigma_\mathrm{y}=0.27/\mathrm{mm}$. With $\sigma_\mathrm{y}/N_\mathrm{e}=1.3\times10^{-2}$, $\sigma_\mathrm{m}/N_\mathrm{e}=3.5\times10^{-3}/\mathrm{mm}$, or a 0.7% shot-to-shot RMS variation of the tilt of the profile. With $k=0.75$, $\sigma_\mathrm{m}/\sigma_\mathrm{y}=0.30/\mathrm{mm}$, and with $k=0.25$, $\sigma_\mathrm{m}/\sigma_\mathrm{y}=0.23/\mathrm{mm}$. This indicates that the value of k does not strongly affect the RMS variation of the tilts.

References

[1] S.E. Bodner, J. Fusion Energy 1 (1981) 221.
[2] R.H. Lehmberg and S.P. Obenschain, Optics Comm. 46 (1983) 27.
[3] R.H. Lehmberg and J. Goldhar, Fusion Tech. 11 (1987) 532.
[4] Ximing Deng, Xiangchun Liang, Zezun Chen, Wenyan Yu and Renyong Ma, Appl. Optics 25 (1986) 377.
[5] A.J. Schmitt and J.H. Gardner, J. Appl. Phys. 60 (1986) 6.
[6] S. Skupsky, R.W. Short, T. Kessler, R.S. Craxton, S. Letzring and J.M. Soures, J. Appl. Phys. 66 (1989) 3456.
[7] D.G. Colombant and A.J. Schmitt, J. Appl. Phys. 67 (1990) 2303.
[8] Y. Kato, K. Mima, N. Miyanaga, S. Arinaga, Y. Kitagawa, M. Nakatsuka and C. Yamanaka, Phys. Rev. Lett. 53 (1984) 1057.
[9] D. Veron, G. Thiell and C. Gouedard, Optics Comm. 97 (1993) 259.
[10] D.M. Penington, M.A. Henesian, H.T. Powell, C.E. Thompson and T.L. Weiland, Conf. on Lasers and electro-optics, 1993, Technical Digest Series, Vol. 11 (1993).
[11] H. Nakano, K. Tsubakimoto, N. Miyanaga, M. Nakatsuka, T. Kanabe, H. Azechi, T. Jitsuno and S. Nakei, J. Appl. Phys. 73 (1993) 2122.
[12] S.P. Obenschain, J. Grun, M.J. Herbst, K.J. Kearney, C.K. Manka, E.A. McLean, A.N. Mostovych, J.A. Stamper, R.R. Whitlock, S.E. Bodner, J.H. Gardner and R.H. Lehmberg, Phys. Rev. Lett. 56 (1986) 2807.
[13] J. Grun, M.H. Emery, C.K. Manka, T.N. Lee, E.A. Mclean, A. Mostovych, J. Stamper, S. Bodner, S.P. Obenschain and B.H. Ripen, Phys. Rev. Lett. 58 (1987) 2672.
[14] A.N. Mostovych, S.P. Obenschain, J.H. Gardner, J. Grun, K.J. Kearney, C.K. Manka, E.A. McLean and C.J. Pawley, Phys. Rev. Lett. 59 (1987) 1193.
[15] S.P. Obenschain, C.J. Pawley, A.N. Mostovych, J.A. Stamper, J.H. Gardner, A.J. Schmitt and S.E. Bodner, Phys. Rev. Lett. 62 (1989) 768.
[16] O. Willi, T. Afshar-rad and S. Coe, Phys. Fluids B 2 (1990) 1318.
[17] T.A. Peyser, C.K. Manka, S.P. Obenschain and K.J. Kearney, Phys. Fluids B 3 (1991) 1479.
[18] D.K. Bradley, J.A. Delettrez and C.P. Verdon, Phys. Rev. Lett. 68 (1992) 2774.
[19] C.J. Pawley et al., Proc. Intern. Conf. on Lasers '90, eds. D.G. Harris and J. Herbelin (1991) p. 491.
[20] M.S. Pronko et al., Proc. Intern. Conf. on Lasers '91, eds. F.J. Duartes and D.G. Harris (1992) p. 691.
[21] S.E. Bodner et al., Fourteenth Intern. Conf. on Plasma physics and controlled nuclear fusion research, IAEA-CN-56/B-2-1(C), (1992).

Appendix B

Nike KrF Laser Project-Status

8:30am Invited

CTuC2 Nike KrF laser project—status

T. Lehecka, A. V. Deniz, J. Hardgrove, S. E. Bodner, R. H. Lehmberg, S. P. Obenschain, C. J. Pawley, M. S. Pronko, J. D. Sethian, J. A Stamper

Science Applications International Corp., MS 2-3-2, 1710 Goodridge Road, McLean, Virgina 22102

The ablative acceleration of spheres and foils for direct drive inertial confinement fusion research requires extremely uniform spatial pressure profiles, which in turn demand extremely uniform laser intensity profiles. Less than two percent peak to valley ablation pressure nonuniformity may be essential to restrict the growth of the Rayleigh-Taylor hydrodynamic instability and to ensure adequate spherical implosion symmetry. During the past ten years, a large part of the direct drive laser fusion effort has been directed toward achieving this laser uniformity and demonstrating control of the Rayleigh-Taylor modes.

Nike is a KrF laser system designed to produce controlled, uniform spatial profiles on flat targets. Echelon-free induced spatial incoherence (ISI) beam smoothing[1] is incorporated into an angularly multiplexed KrF amplifier system that is expected to deliver more than 2 kJ of energy in a 4 ns pulse on the target. Flat target experiments were chosen because they provide better access for diagnostics, and allow direct scaling to high gain targets. Complementary experiments with Nd:Glass lasers and spherical targets are being performed at the University of Rochester's Omega laser[2] and Osaka's Gekko laser.[3]

Echelon-free ISI is basically an image amplification technique. An aperture at the front end of the laser is uniformly illuminated with low intensity radiation from an oscillator. The optical Fourier transform of this object aperture is produced with a simple lens. The Fourier transform is then image relayed through a series of amplifiers. If every point at the object illuminates the Fourier aperture in the same way, the image of the object aperture formed on target is relatively insensitive to gain nonuniformities in the amplifiers. If image distortion effects such as optical imperfections, atmospheric turbulence, and nonlinear optical processes are small, then the image on target is an accurate reproduction of the object aperture. This is why echelon-free ISI does not lend itself well to glass laser systems; the nonlinear B integral effects in the laser glass can be excessive. For Nike, the total B integral, including air propagation and optics, is calculated to be less than 0.4.[4]

When completed, Nike will be a 56 beam angularly-multiplexed system. The final amplifier is a 60 cm aperture, 240 ns pulse length electron beam-pumped device with a total output that should exceed 5 kJ. Forty-four 4 ns target illumination beams and twelve 5.3 ns beams for x-ray backlighter illumination are planned. Conservative estimates for losses at optical surfaces and in air propagation paths result in the design goal of more than 2 kJ on target in the 44 main beams. The 60 cm amplifier is partially constructed and the results of pulse power tests are given elsewhere.[5]

Approximately one half of the Nike system is in place and operational. This includes the oscillator, three discharge amplifier stages, and a 28 beam angularly-multiplexed 20 cm electron beam-pumped amplifier. Details of the 20 cm amplifier are presented elsewhere.[5] Currently, the laser is being operated with twenty-eight 4 ns beams to characterize system performance.

The oscillator has produced intensity profiles that are flat to within ±1% over 65% of the diameter integrated over 4 ns; work is continuing to improve the profile further. Past experiments have shown that this beam can be amplified in a 4 × 4 cm aperture discharge amplifier with negligible distortion of the focal profile. We are currently measuring the image quality after both the discharge amplifiers (output energy of 2 joules from four modules) and from the 20 cm amplifier (output energy greater than 120 Joules). The results of these measurements will be presented. In conjunction with these measurements, experiments are planned to measure the nonlinear index of refraction of various gases under Nike operating conditions. A high pressure gas cell will be used to simulate long propagation paths. The results, which will be used to update the calculations of image distortion due to B integral effects, will also be discussed. A PC-based remote alignment system capable of controlling every optical component in Nike has been developed. CCD cameras and commercial software are used to monitor beam positions, and this information is used in custom software to control stepper motor drivers developed at NRL. The system is low cost, is capable of aligning an entire array of 32 mirrors in approximately 30 seconds, and is currently in use on the 20 cm amplifier system. This alignment scheme, with slight modifications, will also be used for target alignment. Details of the system will be presented.

The Nike optical design is complete. We expect to finish fabrication and installation of the laser in late 1993 and to complete the target facility in 1994.

Code 6730, Plasma Physics Division, U.S. Naval Research Laboratory, Washington, DC 20375

1. R. H. Lehmberg, J. Goldhar, Fusion Technol. 11, 532 (1987).
2. R. L. McCrory, *et al.*, in *Fourteenth International Conference on Plasma Physics and Controlled Nuclear Fusion Research*, (International Atomic Energy Agency, Wurzburg, Germany, 30 Sept.-7 Oct., 1992). Paper B-1-4-4(C).
3. T. Yamanaka, *et al.*, in *Fourteenth International Conference on Plasma Physics and Controlled Nuclear Fusion Research*, (International Atomic Energy Agency, Wurzburg, Germany, 30 Sept.-7 Oct., 1992). Paper B-3-4.
4. R. H. Lehmberg, T. Lehecka, in *1992 Conference on Lasers and Electro-Optics*, Vol. 11, OSA Technical Digest Series (Optical Society of America, Washington, DC, 1992), Paper CFA2.
5. C. J. Pawley, J. D. Sethian, S. P. Obenschain, R. H. Lehmberg, M. S. Pronko, A. V. Deniz, T. Lehecka, M. McGeoch, paper CTuN45, this volume.

Appendix C

Production of High Energy, Uniform Focal Profiles with the Nike Laser

ELSEVIER

15 June 1995

Optics Communications 117 (1995) 485–491

OPTICS COMMUNICATIONS

Production of high energy, uniform focal profiles with the Nike laser

T. Lehecka [a], R.H. Lehmberg [b], A.V. Deniz [a], K.A. Gerber [b], S.P. Obenschain [b], C.J. Pawley [b], M.S. Pronko [b], C.A. Sullivan [b]

[a] *Science Applications International Corporation, MS 2-3-1, 1710 Goodridge Dr., McLean, VA 22102, USA*

[b] *Plasma Physics Division, U.S. Naval Research Laboratory, Washington, DC 20375, USA*

Received 4 January 1995

Abstract

Nike, a KrF laser facility at the Naval Research Laboratory, is designed to produce high intensity, ultra-uniform focal profiles for experiments relating to direct drive inertial confinement fusion. We present measurements of focal profiles through the next-to-last amplifier, a 20×20 cm^2 aperture electron beam pumped amplifier capable of producing more than 120 J of output in a 120 ns pulse. Using echelon free induced spatial incoherence beam smoothing this system has produced focal profiles with less than 2% tilt and curvature and less than 2% rms variation from a flat top distribution.

1. Introduction

Uniform illumination of the spherical pellet is critical to the success of inertial confinement fusion. Intensity nonuniformities can produce ablation pressure variations resulting in an asymmetric implosion and possible failure of the target. For direct drive ICF this problem is reduced by illuminating the pellet in a 'bath' of X-rays produced when high intensity laser beams strike a high-Z material in a hohlraum configuration [1]. One disadvantage of this technique is the relatively low efficiency for conversion of laser light to X-rays absorbed on the pellet, resulting in low overall target gain.

For direct drive ICF the pellet is directly irradiated with laser light and one seeks to obtain symmetry by the combination of uniform laser illumination and thermal smoothing. Laser nonuniformities on the order of 1% tend to degrade the performance of a high gain pellet because the laser nonuniformity seeds the Rayleigh–Taylor hydrodynamic instability. This instability can be reduced by preheating the fuel, but the preheat then reduces the efficiency of the thermonuclear burn and thus the target gain.

The Nike laser facility is a 5 kJ, 56 beam angularly multiplexed KrF laser system located at the Naval Research Laboratory. Nike is designed and built with the goal of producing focal profiles with the least possible nonuniformity to investigate hydrodynamic instabilities and evaluate the uniformity requirements for direct drive ICF. Initial design goals for the system were to produce a focal profile with less than 2% rms intensity nonuniformities. Results presented in this paper show that we have exceeded this requirement in a single beam and should do significantly better when multiple beams are overlapped on the target.

In this paper we describe the operation and evaluation of performance through approximately two-thirds of the Nike laser system. This portion includes an oscillator and Pockels cell capable of producing the desired laser uniformity in a 4 ns pulse, four stages of ampli-

SSDI 0030-4018(95)00179-4

fication in discharge pumped amplifiers and angular multiplexing of 28 beams through a 20 cm aperture electron beam pumped amplifier. It has produced more than 120 J of output distributed in 28 4 ns long pulses. Focal profile measurements of one of these beams have demonstrated that this amplifier can produce the desired uniformity, i.e. $<2\%$ tilt and first order curvature and $<2\%$ rms variation. This focal uniformity is significantly better than that obtained with any other large laser facility. In addition, overlap of multiple beams on target should reduce this to less than 1% rms deviation in a 4 ns pulse. The output of the 20 cm amplifier will be used as the input to a 60 cm aperture final amplifier, which should produce 5 kJ of output.

In the next section we will describe echelon free induced spatial incoherence (ISI), the beam smoothing technique employed on Nike. The experimental setup used to produce and measure the focal profile uniformity is presented in Section 3, and experimental results are described in Section 4.

2. Beam smoothing theory

To produce a uniform beam, Nike is implementing a modified version [2] of the induced spatial incoherence (ISI) technique [3]. An object aperture is uniformly illuminated by broad bandwidth ($\Delta\nu \sim 1$ THz), spatially incoherent light. This aperture is imaged onto the target by relaying its optical Fourier transform plane (pupil plane) through the laser system. The time-averaged focal profile at the target is therefore controlled directly by adjusting the object aperture; for example, it can be shaped to a nonflat profile by using a soft aperture. Light from each point in the object aperture illuminates the Fourier plane with a different ray angle but the same time-averaged intensity distribution. If each amplifier stage is located at or near an image of the Fourier plane, the image of the object formed at the target will be insensitive to the amplifier gain nonuniformities[2].

The time-averaged focal profile will remain relatively insensitive to large scale phase nonuniformities imposed by the laser system if its angular width is many times diffraction-limited (XDL) and large compared to the angular perturbations introduced by the phase aberration. The Nike optical design allows nearly complete compensation of all systematic phase aberrations, such as astigmatism, coma, and spherical aberration. Distortion due to random phase aberration, such as those arising from optical surface imperfections, will be discussed in Section 4. For the 40–100 XDL flat top profiles to be used in Nike planar target experiments, the main effect of random phase aberration is to round off the edges of the profile. Distortion of nonflat profiles can be precompensated at the object aperture as long as the angular perturbations remain relatively small and reproducible.

The instantaneous focal spot intensity ($I(x, t)$ produced by incoherent beam smoothing schemes such as ISI is highly nonuniform speckle; the illumination profile approaches a smooth profile only when this intensity is averaged over time intervals t_{av} much larger than the laser coherence time τ_c. The coherence time is formally defined by

$$\tau_c \equiv \int_{-\infty}^{+\infty} |\gamma_T(t)|^2 \, dt, \quad \gamma(0) \equiv 1, \tag{1}$$

where $\gamma(t)$ is the temporal autocorrelation function of the laser. Using the Wiener–Khinchin theorem [4], one can relate τ_c to the *normalized* optical power spectrum $S(\nu)$:

$$\tau_c = \int_{-\infty}^{+\infty} S^2(\nu) \, d\nu = \frac{\eta}{\Delta\nu}, \quad \int_{-\infty}^{+\infty} S(\nu) \, d\nu \equiv 1, \tag{2}$$

where $\Delta\nu$ is the fwhm bandwidth and η is a numerical factor that depends on the shape of the spectrum (e.g., $\eta \cong 0.664$ in the case of a gaussian spectrum).

In the limit of large spatial incoherence, the rms residual nonuniformity due to speckle can be calculated from the definition

$$\sigma \equiv \sqrt{\langle J(x)^2 \rangle_S - \langle J(x) \rangle_S^2} \, / \langle J(x) \rangle_S, \tag{3}$$

where $J(x)$ denotes the time-averaged intensity and $\langle\ \rangle_S$ denotes a spatial average over many coherence zones. Assuming gaussian statistics (chaotic light) and negligible tilt and curvature across the averaging region, one finds in the limit $t_{av} \gg \tau_c$ [3,4],

$$\sigma \cong \sqrt{\tau_c / t_{av}}. \tag{4}$$

Eq. (4) is the expected result for a superposition of t_{av}/ρ_c statistically independent speckle patterns.

Strictly speaking, t_{av} is the actual averaging time only if the pulse shape remains constant during the interval.

In the results reported here, where the averaging process consists of measuring the time-integrated profiles, the effective averaging time is related to the normalized pulse envelope function $P(t)$ by the integral

$$\frac{1}{t_{av}} \equiv \int_{-\infty}^{+\infty} P^2(t)\,dt\,, \quad \int_{-\infty}^{+\infty} P(t)\,dt \equiv 1\,. \tag{5}$$

3. Experimental arrangement

As discussed above we have been operating the Nike laser system through the 20 cm amplifier to verify that beam uniformity is maintained through the amplifier chain. A block diagram of the experiments described below is shown in Fig. 1. A spatially and temporally incoherent commercial KrF oscillator in the front end illuminates a lambertian reflector. A small portion of the light from this reflector is focussed onto a circular aperture. The Fourier transform of this aperture is formed in the center of the first commercial discharge amplifier by a single lens. After amplification the Fourier image is relayed to a Pockels cell for temporal shaping of the pulse, then relayed again to a second commercial discharge amplifier.

Although the Fourier plane is relayed into the center of each amplifier, the combination of small aperture ($\sim$5 mm) and finite length ($\sim$50 cm) allows some imprinting of gain nonuniformities onto the laser beam at the focal plane. We have observed that these commercial discharge amplifiers have a tendency to impose a small ($\sim$5%) linear tilt on the focussed beam. To overcome this difficulty, the beam is split into two parts after the second amplifier. One of these is focussed directly onto an aperture while the second is inverted in both the horizontal and vertical directions and then focussed onto the aperture. This aperture is the object plane for the rest of the system and therefore will be imaged onto the target. This superposition of the two beams reduces the tilts below the desired 2% level.

A single lens again forms the Fourier transform of the object illumination onto a square aperture which clips the beam to the correct dimensions. The size of the object and Fourier aperture, as well as the focal length of the lens between them, determine the beam divergence or number of times over the diffraction limit. For Nike experiments we plan to use focal spots in the range of 30–100 times diffraction-limited (XDL), with a baseline of 60 XDL. The experiments described below were performed with 60 XDL beams.

The beam is expanded and the Fourier image relayed to a 4×4 cm^2 discharge amplifier [5]. After this amplifier the beam is split into four separate beams to begin the time multiplexing required for the electron beam pumped amplifiers [6]. Two of the beams go directly

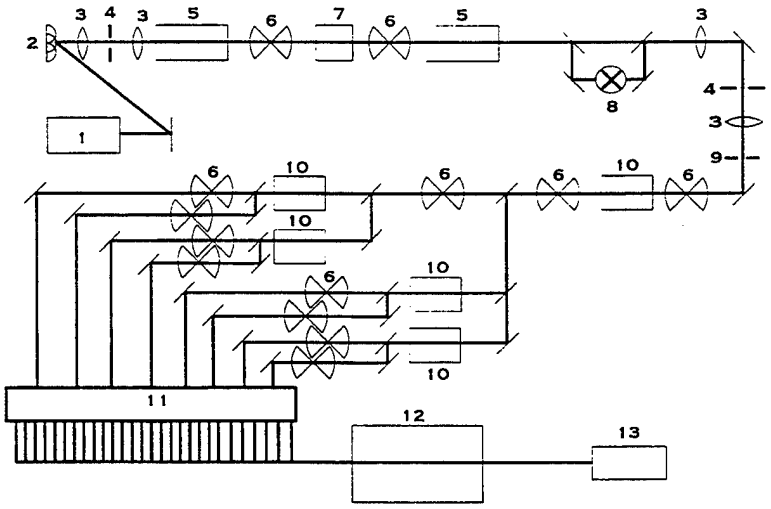


Fig. 1. Block diagram of the Nike laser system through the 20 cm amplifier. 1: Commercial oscillator, 2: diffuse reflector, 3: lens, 4: object aperture, 5: commercial discharge amplifier, 6: image relaying telescope, 7: temporal pulse shaping, 8: beam inverter, 9: Fourier aperture, 10: 4×4 cm^2 discharge amplifier, 11: temporal multiplexing array, 12: 20 cm amplifier, 13: diagnostics (cameras, photodiodes, calorimeters).

to two 4×4 cm^2 discharge amplifiers while the other two are delayed by 64 ns (via propagation delay) before going to another set of two 4×4 cm^2 discharge amplifiers. Telescopes in the beam paths relay the Fourier image from the preamplifier to this set of four discharge amplifiers. After this set of amplifiers the beams are split twofold and each of the resulting eight beams passes through a separate image relaying telescope. These telescopes, in conjunction with the optics that direct the laser beams through the electron beam pumped amplifier stages, were designed to relay the Fourier image from the array of discharge amplifiers to the 60 cm amplifier. Immediately after the telescopes the eight beams are divided by beamsplitters into 28 beams. The temporal spacing for the first six beams (intended for backlighter use) is 5.33 ns, with 4 ns spacing for the remaining 22 beams. This gives an effective pulse length of 120 ns to drive the 20 cm amplifier.

It is important to note here that the 20 cm amplifier is *not* located at a Fourier image plane. It was felt that image relaying the beams from this amplifier to the 60 cm amplifier would be too complicated and costly. Because of the large beam sizes (and correspondingly small divergence angles) involved, the optical distance from the Fourier plane to the 20 cm amplifier is not large. For example, the maximum deviation of extreme rays is calculated to be 8 mm at the 20 cm amplifier for the 60 XDL baseline operation. If gain nonuniformities in this amplifier are less than 20% we would then expect less than a 1% (20%×0.8 cm/20 cm) tilt to be imposed on the beam. Because of this separation, however, amplification in the 20 cm amplifier was thought to be one of the higher risk items in the Nike design, so we were particularly interested in obtaining focal profile data from this amplifier at the earliest possible date. The 60 cm amplifier lies closer to the Fourier plane, so we expect that any gain nonuniformities in that laser will have less of an effect on the focal spot uniformity. The 28 beams emerging from the 20 cm amplifier will be split into 56 beams and then propagated to the 60 cm amplifier. For the measurements presented herein we are directing these 28 beams to calorimeters, photodiodes and focal profile diagnostics.

A previous version of our oscillator has exceeded the beam uniformity requirements discussed above for a 60 XDL beam divergence [7]. The present lambertian reflector arrangement allows us to go to higher divergence while maintaining this beam uniformity. In the next section we will present results from amplification of this beam with the system described above. These data will show that we have successfully amplified the laser beam with insignificant distortion in the focal profile uniformity.

4. Experimental results

The main goal for Nike is the production of uniform high intensity focal profiles for the acceleration of flat targets. For this application we require a flat topped profile. The edges of this profile are rounded by accumulated optical distortions [6]. At the level it occurs on Nike this rounding is desirable for flat target accelerations because it helps reduce edge effects. In our definition of beam uniformity we will therefore restrict our attention to the central region of the focal profile, as discussed below.

For this region of interest we assign three numbers to characterize the beam uniformity: peak-to-valley linear tilt, peak-to-valley first order curvature and rms deviation from a linear fit to the data. The latter arises primarily from residual speckle. These numbers are obtained by performing a two-dimensional linear and second order least squares fit to the (usually, as it was in this case) circular region of interest. This yields a plane of the form $P_1=a_0+a_1x+a_2y$ and a paraboloid of the form $P_2=a_0+a_1x+a_2y+a_3xy+a_4x^2+a_5y^2$. The tilt is defined as $[\max(P_1)-\min(P_1)]/\mathrm{Avg}[J(x,y)]$ and the curvature is defined as $[\max(P_2-P_1)-\min(P_2-P_1)]/\mathrm{Avg}[J(x,y)]$ where $J(x, y)$ is the measured time-integrated focal profile. The rms error is simply the rms deviation of the data from the linear fit divided by the average value of $J(x, y)$. The goals for Nike experiments are a tilt of less than 2% peak to valley, a first order curvature less than 2% peak to valley, and a less than 2% rms deviation for each of the beams in the system. As discussed earlier, overlapping multiple beams on target should reduce this rms deviation even further.

The focal profile data ($J(x, y)$) described below were obtained in the following way. After amplification in the 20 cm amplifier, as described in the previous section, one of the 28 beams is propagated through an aperture located at a Fourier image plane. This aperture limited the beam to 60 times diffraction-limited for

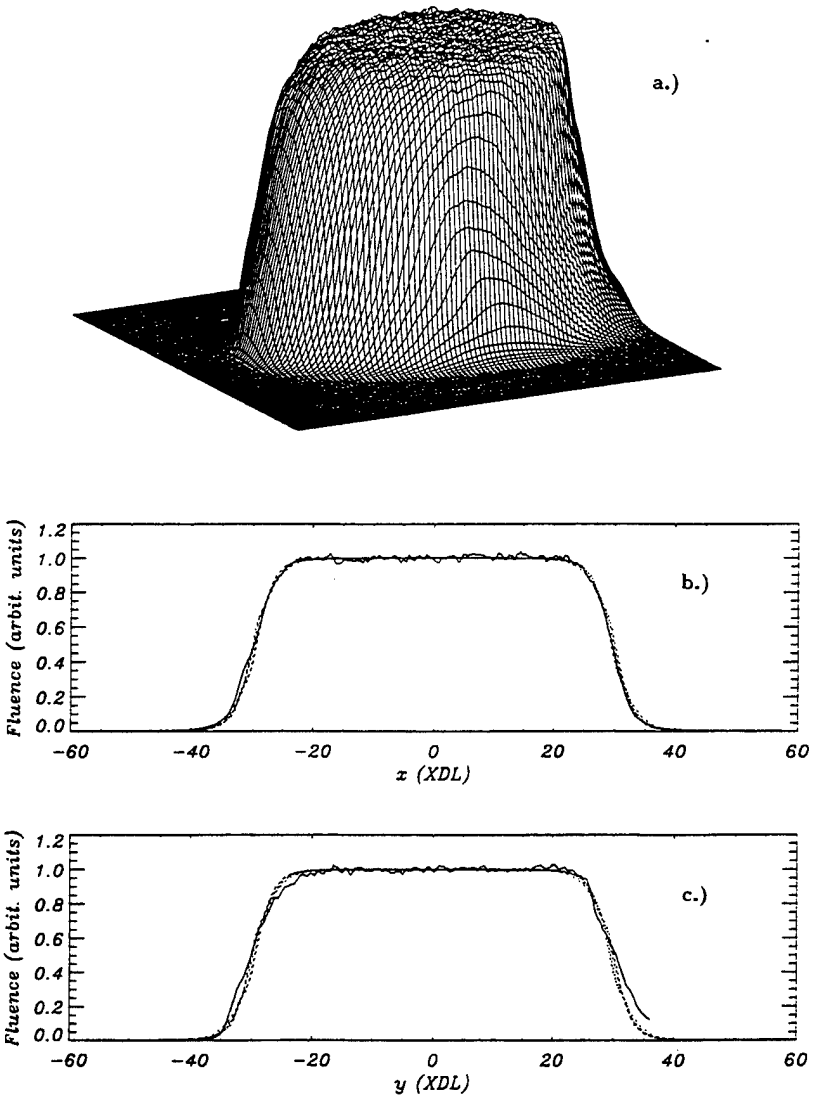


Fig. 2. (a) Surface plot of the measured focal profile. (b), (c) Horizontal and vertical cross sections of the measured focal profile compared with calculated profiles perturbed by random phase distortion. The jagged solid lines are the measured profiles, while the dotted lines (which blend into straight solid lines in the center portion) are the four quadrant profiles of the calculated fluence distribution.

these measurements. (Note that in the final system the 60 cm amplifier will act as the final Fourier image aperture.) This single beam is collimated to a 2.7 cm square size, then focussed onto a CCD camera with an 11 m focal length lens. This camera is a cooled, slow scan device capable of measuring rms errors as low as 0.3% and has been described previously [7]. The $F\#$ of the focussing lens is 400, resulting in an overall focal spot size of 6 mm for these 60 XDL beams. The shortest wavelength speckle is therefore 6 mm/60 XDL, or 100 μm, which is well resolved by the 23 μm pixels of the CCD camera. The beam was attenuated by reflection from three uncoated flats (4% reflection per surface) and propagation through a 95% and a 99% reflector. These attenuation methods have been routinely used in developing the Nike front end.

A surface plot of the focal profile is shown in Fig. 2a, with horizontal and vertical cross sections through the centroid of this profile in Figs. 2b and c. Also shown in Figs. 2b and c are the calculated focal profiles

accounting for optical imperfections. They were calculated by taking the interferograms available for the optics in the system and modelling the net effect of all the optics in the system on the point spread function (PSF). This PSF is then convolved with the initial aperture function. As is evident in the figure the model and data are in good agreement. Both the model and the data indicate that the optical system, described in the previous section, is producing a seven times diffraction-limit PSF at 248 nm.

The seven XDL PSF combined with the 60 XDL beam divergence yields a focal profile which should be uniform over the central 75% of the diameter. The data analysis region is defined by a circle centered on the image centroid. The diameter of this circle is determined by first identifying the 50% contour and then multiplying this diameter by 0.75.

For the focal profile shown in Fig. 2 the time-averaged rms nonuniformity for these 4 ns fwhm pulses is measured to be 1.4% with a maximum linear tilt of 1.8%. The first order curvature is only 1.0%. These meet and exceed the Nike goals of 2% rms deviation and 2% peak-to-valley tilt and curvature discussed previously. For this shot the measured energy was between 4 and 5 J per beam with a total of 125 J in the 28 output beams from the 20 cm amplifier. After propagation through the final aperture the energy is reduced to 2.7 J in a single 4 ns beam.

It should be noted that these results are the best obtained with the Nike system to date, and this level of uniformity was obtained on approximately 10% of the system shots. They do, however, provide a proof of principle for the echelon free ISI concept on Nike. Typically we are obtaining linear tilts of about 4% from the 20 cm amplifier. We are currently in the process of identifying the cause of this nonuniformity. Possible causes include variations in alignment, shot-to-shot irreproducibility in the 20 cm amplifier and thermal variations in the propagation paths.

The measured rms beam nonuniformity was compared to speckle theory (Eq. (4)) for five shots, using the measured spectral and temporal pulse shapes to evaluate Eqs. (2) and (5). For completeness the measured temporal and spectral shapes are shown in Figs. 3a and b. Because the laser pulse tends to have a trailing edge, the calculated t_{av} values (8.5–9.6 ns) were significantly larger than the fwhm pulsewidth t_p ($\sim$4.2 ns); hence the calculated nonuniformities were somewhat lower than one would expect for a 4.2 ns pulse. To reduce the contributions due to residual envelope curvature, we confined the data analysis to a circle of 60% of the fwhm beam diameter. The theory yields $\sigma = 0.75$–0.85%. Fig. 4 shows the comparison of theory and experiment, along with recent nonuniformity measurements reported on large glass lasers [8–10], plus the result (dashed line) that shows the $\sigma \propto 1/\sqrt{\Delta\nu\, t_p}$ scaling. The abscissa values used for the Nike measurements were taken from the measured fwhm values of $\Delta\nu = 0.98$–1.05 THz and $t_p = 4.1$–4.4 ns. The small discrepancy between the measured and calculated rms nonuniformities appears to stem primarily from the residual envelope curvature, which over the five shots varied from 0.3% to 2% within the 60% analysis circle. To the authors' knowledge, our

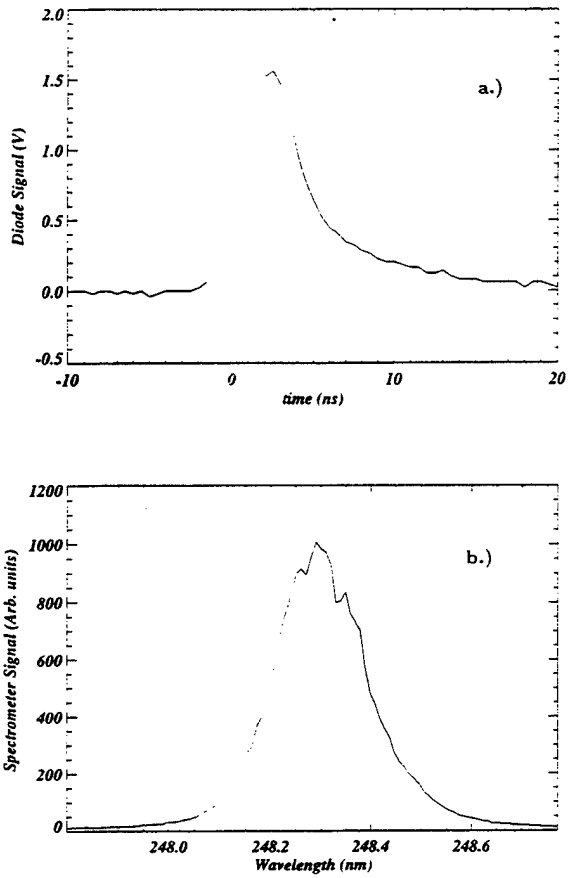


Fig. 3. (a) Temporal pulse shape after amplification in the 20 cm amplifier. (b) Spectral pulse shape after amplification in the 20 cm amplifier.

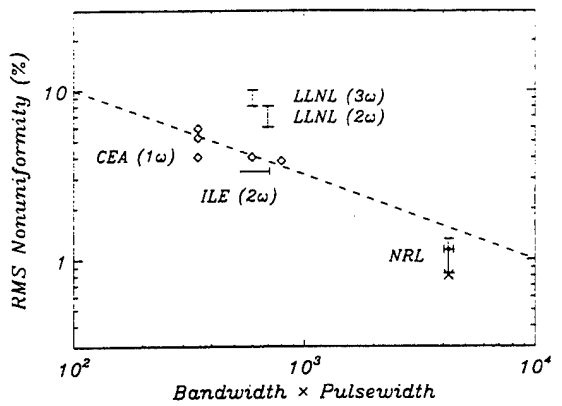


Fig. 4. Comparison of the rms nonuniformities measured after the 20 cm amplifiers with speckle theory (× point) and with measurements on the Nova laser at LLNL, USA [8], the Phebus laser at CEA, France [9], and the Gekko XXII laser at ILE, Japan [10]. The bandwidth and pulsewidth refer to fwhm values. The dashed line, shown for reference, would be strictly applicable to a rectangular spectrum and pulse shape, and shows the $\sigma_f \propto \sqrt{\tau_c/t_p} \propto 1/\sqrt{\Delta\nu t_p}$ scaling.

measurements represent the highest degree of laser uniformity obtained at high energy to date.

5. Conclusion

We have demonstrated the production high energy, ultra-uniform intensity focal profiles with the Nike laser. A flat top beam profile with 1.8% linear tilt, 1.0% first order curvature and 1.4% rms deviation from the linear fit was achieved in a single 2.7 J, 4 ns beam. Overlap of multiple beams on target should reduce the rms variation below 1%. This level of uniformity will allow us to perform detailed hydrodynamic stability experiments with conditions relevant to direct drive inertial confinement fusion.

References

[1] J.H. Nuckolls, Phys. Today 35 (1982) 24.
[2] R.H. Lehmberg and J.Goldhar, Fusion Tech. 11 (1987) 532.
[3] R.H. Lehmberg and S.P. Obenschain, Optics Comm. 46 (1983) 27.
[4] J.W. Goodman, Statistical optics (Wiley, New York, 1985) Sections 5.1, 6.1.
[5] M. Pronko, IEEE J. Quantum Electron. QE-30 (1994) 2147.
[6] R.H. Lehmberg, S.P. Obenschain, C.J. Pawley, M.S. Pronko, A.V. Deniz and T. Lehecka, Proc. SPIE 1870 (1993) 163.
[7] A.V. Deniz and S.P. Obenschain, Optics Comm. 106 (1994) 113.
[8] H.T. Powell, private communication (August, 1993).
[9] D. Veron, G. Thiell and C. Gouedard, Optics Comm. 97 (1993) 259.
[10] H. Nakano, K. Tsubakimoto, N. Miyanaga, M. Nakatsuka, T. Kanabe, H. Azechi, T. Jitsuno and S. Nakei, J. Appl. Phys. 73 (1993) 2122.

Appendix D

Optical Beam Smoothing on the Nike KrF Laser

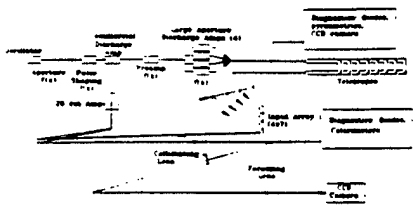
Optical beam smoothing on the NIKE KrF laser

T. Lehecka, A. V. Deniz, J. Hardgrove, S. E. Bodner,* K. A. Gerber,* R. H. Lehmberg,* E. A. McLean,* S. P. Obenschain,* C. J. Pawley,* M. S. Pronko,* J. D. Sethian,* J. A. Stamper,* *Science Applications International Corporation, McLean, Virginia 22102*

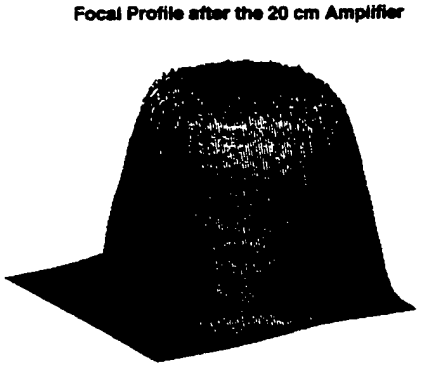
NIKE is an angularly-multiplexed KrF laser that is currently under development at the Naval Research Laboratory to address technological and physics issues of direct-drive laser fusion.[1] It will deliver 2–3 kJ of 248-nm light in a uniform flat-top profile at intensities $>2 \times 10^{14}$ W/cm^2, which will be used to study ablative acceleration of thin planar foil targets under conditions close to the operating regime envisioned for a high gain ICF pellet. The NIKE system consists of a commercial oscillator/amplifier front end, an array of discharge preamplifiers, two e-beam pumped amplifier stages (with apertures of 20×20 cm^2 and 60×60 cm^2), and the optics required to relay, multiplex, and demultiplex the optical beams. Approximately two-thirds of the system, including the 20-cm amplifier, is operational and currently undergoing tests. The output of the 20-cm amplifier is sufficiently uniform, and its energy exceeds the 120 J required to drive the 60-cm final stage.

Direct-drive ICF requires a highly smooth and controllable focal spot irradiance because the required UV light allows only modest smoothing of beam nonuniformities by lateral energy flow in the underdense plasma surrounding the pellet. To produce such a beam, the NIKE optical system is designed to incorporate spatial incoherence.[2] In this technique, the desired time-averaged beam profile is created by a broadband spatially incoherent oscillator, then imaged through the amplifier system and final focusing lens onto the target. Because the amplifier stages lie at or near the optical Fourier-transform plane of the image, their gain nonuniformities will have little effect on the envelope (long time-average) of the focal profile. Focal spot nonuniformities can still arise from either residual speckle (mainly at short transverse scalelengths), or envelope distortion (mainly at longer scalelengths) due primarily to random phase nonuniformities in the optical system. Here, we report on focal distribution measurements of a single NIKE beam at the output of the 20-cm amplifier.

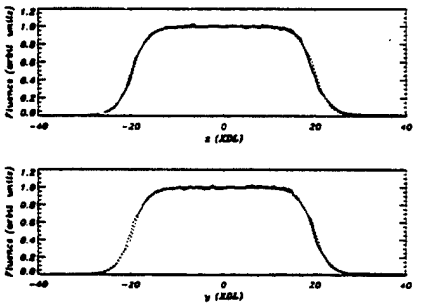
The experimental setup is shown in Fig. 1. A "top-hat" envelope F(x) of 40 x = diffraction-limited width (40 XDL) is created by a circular hard aperture at the output of the broadband incoherent oscillator. The optical Fourier transform is relayed through the pulse shaping optics and preamplifiers onto locations just beyond the 20-cm amplifier in a 4×7 array of sequential 4–5 J, 4-ns beams. One of these beams is then collimated and focused onto a CCD camera. With a 23-μm pixel spacing, the camera fully resolved the shortest speckle wavelength of 70 μm.



CTuQ2 Fig. 1. Experimental setup used to measure focal energy distributions at the output of the NIKE 20-cm amplifier.



CTuQ2 Fig. 2. Isometric view of the measured focal distribution after the 20-cm amplifier.



CTuQ2 Fig. 3. Comparison of the measured horizontal and vertical focal spot profiles with the calculated envelope profiles perturbed by random phase distortion. The heavy solid lines are the experimental data, and the dots (which blend into thin solid lines in the flat regions) are the calculated profiles as viewed from the four quadrant directions.

Figure 2 shows one of the measured focal distributions. A least-squares analysis performed within the central 65% of the diameter gave a planar fit with a tilt of 1.4%. The rms deviation from this planar fit was 1.3%, in good agreement with a theoretical speckle nonuniformity $\sigma_S = 0.815/\sqrt{\Delta\nu \cdot t_{av}} = 1.4\%$ for the measured FWHM bandwidth $\Delta\nu = 0.79$ THz and averaging time $t_{av} = 4$ ns. In the planar foil experiments, which will use 44 independent overlapping beams and an operating bandwidth ≥2 THz, NIKE should thus attain a speckle-level $\sigma_S < 1\%$ rms in $t_{av} = 100$ ps.

Figure 3 compares the measured horizontal and vertical profiles through the centroid with the corresponding envelope profiles calculated for a 40-XDL top-hat beam perturbed by random phase distortion within the laser system.[3] The statistical realization of the phase distortion was constructed by using measured or specified rms phase variations and phase gradients on each of the optical surfaces within the completed portion of the system; this gave a point-spread function containing half of its energy within a width of 7 XDL.

In the completed NIKE system, the point-spread function is expected to broaden to ~10 XDL, and additional envelope distortion may arise from nonlinear refraction in the high intensity air paths beyond the 60-cm amplifier. These effects, along with rotational SRS, two-photon absorption, and group velocity dispersion, are modeled in the NIKEBEAM code.[3] For the 60-XDL focal spots to be used in the baseline target experiments, NIKEBEAM simulations predict envelope distortion <2% over the central 2/3 of the focal spot diameter if the high intensity beams propagate in argon or helium.

* *Plasma Physics Division, U.S. Naval Research Laboratory, Washington, DC 20375*

1. S. E. Bodner, *et al.* in *Fourteenth International Conference on Plasma Physics and Controlled Nuclear Fusion Research* (IAEA-CN-56, Wurzburg, Germany, 1992).
2. R. H. Lehmberg, J. Goldhar, Fusion Technology **11**, 532 (1987).
3. R. H. Lehmberg, S. P. Obenschain, C. J. Pawley, M. S. Pronko, A. V. Deniz, T. Lehecka, Laser Coherence Control: Technology and Applications, H. T. Powell, T. J. Kessler, eds., Proc. SPIE, Vol. 1870, 163 (1993).

Appendix E

Two-Photon Resonantly-Enhanced Negative Nonlinear Refractive Index in Xenon at 248 nm

ELSEVIER

15 November 1995

OPTICS COMMUNICATIONS

Optics Communications 121 (1995) 78–88

Full length article

Two-photon resonantly-enhanced negative nonlinear refractive index in xenon at 248 nm

R.H. Lehmberg [a], C.J. Pawley [a], A.V. Deniz [b], M. Klapisch [c], Y. Leng [d]

[a] *Plasma Physics Division, U.S. Naval Research Laboratory, Washington, DC 20375, USA*

[b] *Science Applications International Corporation, McLean, VA 22102, USA*

[c] *ARTEP Inc., 6432 Elffolk Terrace, Columbia, MD 21045, USA*

[d] *Institute for Plasma Research, University of Maryland, College Park, MD 20742, USA*

Received 4 May 1995; revised version received 7 July 1995

Abstract

We report measurements and ab initio calculations of a large *negative* nonlinear refractive index in xenon for linearly-polarized KrF laser light, and describe a simple experiment that demonstrates xenon compensation of both self-focusing and self-phase modulation. This effect arises because KrF laser frequencies lie just above a two-photon resonance with the Xe $6p[1/2]_0$ state at 249.6 nm. The negative nonlinear index resulted in self-defocusing of 10 ps linearly-polarized 248.4 nm pulses in a 4.6 m cell at Xe pressures up to 1000 Torr. Numerical simulations of this defocusing at pressures up to 266 Torr yield a nonlinear index $n_2 = -(2.06 \pm 0.14) \times 10^{-14}$ esu for 760 Torr and 300 K, but the simulations fail to adequately reproduce the shapes of the measured profiles at pressures significantly above 266 Torr. In the compensation experiment, the cell was first filled up to 1260 Torr with pure carbon dioxide, whose large positive n_2 resulted in self-focusing of the beam. A small amount of Xe was then added to the cell and allowed to mix with the CO_2. Less than 130 Torr of Xe completely compensated the 1260 Torr of CO_2, allowing both the spectrum and beam profile to be recovered. Numerical simulations of the self-focusing measurements yield a nonlinear index of $(1.60 \pm 0.07) \times 10^{-15}$ esu for CO_2, and the compensation experiment gives $n_2(\mathrm{Xe}) = -(1.82 \pm 0.08) \times 10^{-14}$ esu at 760 Torr, 300 K, and 248.4 nm. The combined defocusing and compensation measurements give $n_2(\mathrm{Xe}) = -(1.97 \pm 0.23) \times 10^{-14}$ esu, in good agreement with our ab initio calculation of $-(1.54 \pm 0.30) \times 10^{-14}$ esu.

1. Introduction

One of the limitations of intense KrF laser beams in either short pulse or laser fusion systems is the self-focusing and self-phase modulation that can occur in the optical components or air propagation paths [1–3]. These effects arise from nonlinear refractive indices n_2, which are normally positive in most gases and optical materials. If a material of *negative* n_2 were available, it would allow one to compensate these effects under suitable conditions, and thereby increase the useful KrF power density. In this paper, we report calculations and measurements of a large negative n_2 in xenon at KrF laser wavelengths. These measurements include not only self-defocusing and self-phase modulation in Xe, but also a simple experiment that demonstrates Xe compensation of whole-beam self-focusing and self-phase modulation in CO_2.

The large negative n_2 arises because KrF light around 248 nm lies just *above* a two-photon resonance with the Xe $6p[1/2]_0$ state at 249.6 nm, as shown in Fig. 1. This resonance, which has been studied by several authors [4–7], has also been used to enhance a variety of other four-wave parametric mixing processes driven

SSDI 0030-4018(95)00433-5

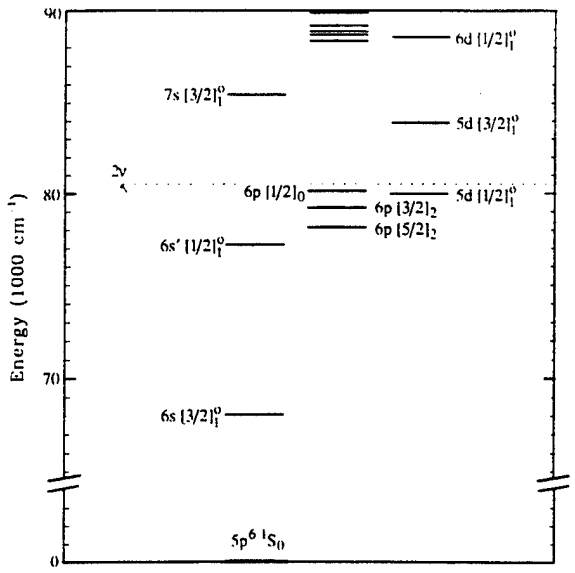


Fig. 1. Lower-lying energy levels of atomic xenon, showing the odd parity states that couple most strongly with the two-photon near-resonant level $6p[1/2]_0$ at 80119 cm^{-1} to produce the negative n_2.

by very high intensity KrF light [6,7]. A similar two-photon resonant enhancement causes a large negative n_2 in cesium at 1.06 μm [8,9]. The phenomena described here arise only from the third order atomic susceptibility; they are distinct from the self-defocusing and blue-shifting caused by photoionization electrons [10], which would occur at much higher intensities. At the intensities used in our experiments (~ 2 GW/cm^2) the refractive index perturbations due to photoionization are negligible.

2. Theory

Applying the usual definition $\delta n \equiv \frac{1}{2} n_2 |\mathscr{E}|^2$, where δn is the nonlinear perturbation of the refractive index and $\mathscr{E}$ is the slowly-varying complex laser field amplitude, one can calculate the general expression for n_2 from the third order atomic susceptibility. For linear polarization, this expression reduces to [8]

$$n_2(\nu) = 2\pi N \sum_{\beta \neq 0} \left[\frac{(P_\beta^-)^2}{E_{\beta 0} - 2h\nu} + \frac{(P_\beta^+)^2}{E_{\beta 0} + 2h\nu} + \frac{(P_\beta^- + P_\beta^+)^2}{E_{\beta 0}} \right] - 2\pi N \left[(P_0^- + P_0^+)R + \frac{(P_0^-)^2 - (P_0^+)^2}{2h\nu} \right], \quad (1)$$

where

$$P_\beta^\pm \equiv \sum_\alpha \left[\frac{\mu_{\beta\alpha}\mu_{\alpha 0}}{E_{\alpha 0} \pm h\nu} \right], \quad (2a)$$

$$R \equiv \sum_\alpha \left[\frac{\mu_{0\alpha}\mu_{\alpha 0}}{(E_{\alpha 0} - h\nu)^2} + \frac{\mu_{0\alpha}\mu_{\alpha 0}}{(E_{\alpha 0} + h\nu)^2} \right], \quad (2b)$$

N is the atomic number density, ν is the laser frequency, $E_{\gamma 0}$ is the energy difference between an excited state $|\gamma\rangle$ and the ground state $|0\rangle \equiv 5p^6\ {}^1S_0$, and $\mu_{\alpha 0} \equiv \langle \alpha | ex | 0 \rangle$ and $\mu_{\beta\alpha} \equiv \langle \beta | ex | \alpha \rangle$ are the allowed dipole matrix elements between the odd ($|\alpha\rangle$) and even ($|0\rangle$, $|\beta\rangle$) parity states. The first term inside the summation in Eq. (1) is enhanced near a two-photon resonance, while the $P_0^- R$ term on the third line would be enhanced near a single-photon resonance. At KrF laser wavelengths, the two-photon near-resonance with the xenon $6p[1/2]_0 \equiv |2\rangle$ level allows one to approximate Eq. (1) by

$$n_2(\nu) \cong \frac{2\pi N (P_2^-)^2}{E_{20} - 2h\nu}, \quad (3)$$

where $E_{20} = 80119\ cm^{-1}$. (For example, $E_{20} - 2h\nu$ is only $-396\ cm^{-1}$ at 248.4 nm.)

The Xe dipole matrix elements in expressions (2a,b) were calculated by the relativistic parametric potential method [11–13] using the "spectroscopic" mode. In this mode, the atomic potential parameters are varied until a minimum is found for the rms deviation between the calculated energy levels and experimentally documented ones. For each parameter set, the code solves the Dirac equation for the wave functions and uses these

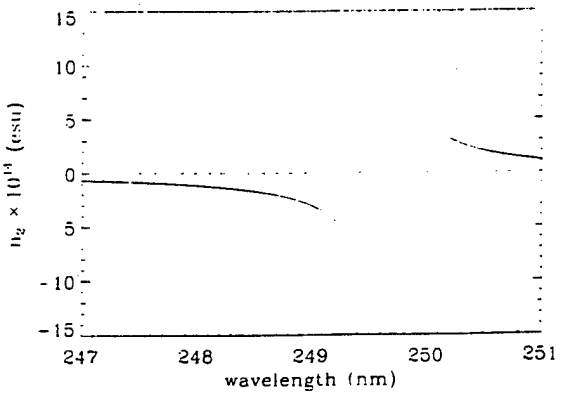


Fig. 2. Calculated nonlinear refractive index n_2 versus wavelength in 760 Torr of Xe at 300 K.

to compute the Hamiltonian matrix elements of the Coulomb interaction, plus the Breit term and QED corrections. In accord with first order perturbation theory, the mean value of the effective potential is subtracted from the energies. The eigenvalues of the matrices are then compared to the experimental energy levels [14], and the parameters are varied until a minimum is found for the rms deviation. Thus the parametric potential method generates consistent wave functions that give the best fit to experimental energies to first order. The resulting wave functions are then used to calculate the dipole matrix elements in Eqs. (2a,b). This procedure has already been used successfully to evaluate transition probabilities [15]. An rms deviation of 700 cm^{-1} was achieved with five parameters for a total of 36 levels ranging in energy from 0 (ground state) to 90000 cm^{-1}.

The summations in Eqs. (1) and (2) included all $J_\alpha = 1$ odd parity intermediate states up to the 9s and 9d levels, and all two-photon accessible ($J_\beta = 0, 2$) even parity states up to the 9p and 5f levels. No attempt was made to include any continuum state contributions. The largest intermediate state contributions came from the lower levels, such as the 6s $[3/2]^o_1$ and 5d $[3/2]^o_1$ levels shown in Fig. 1. At 248.4 nm, expressions (1) and (3) gave nearly the same results, $n_2 = -1.536 \times 10^{-14}$ esu and -1.539×10^{-14} esu, respectively, at 760 Torr and 300 K. From an estimated $\pm 7\%$ uncertainty in the magnitudes of the calculated dipole matrix elements, we obtain $n_2 = -(1.54 \pm 0.30) \times 10^{-14}$ esu. Fig. 2 shows the wavelength dependence of n_2, as calculated from Eq. (1).

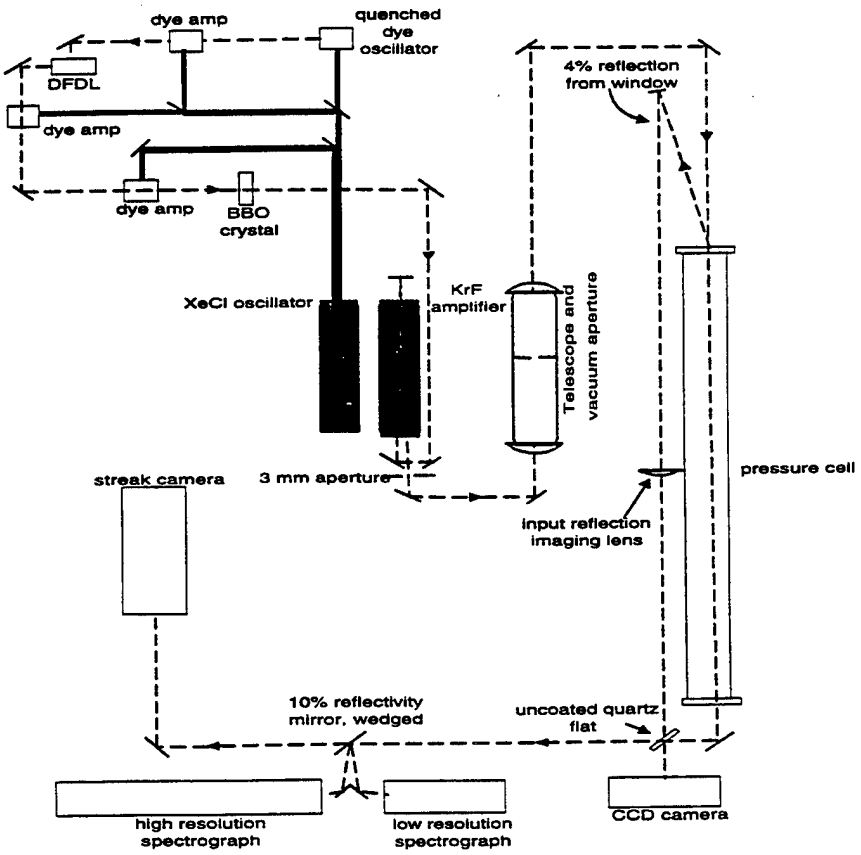


Fig. 3. Experimental arrangement. The short pulse dye laser front-end in the upper left hand corner is pumped by a XeCl laser at 308 nm. The beam is amplified by the KrF laser in the center and then goes to the gas cell and incident beam-imaging optics. The relative positions of the CCD camera, spectrometers, and streak camera at the output are shown.

3. Experiment

To observe the negative n_2 of Xe, we set up a nonlinear propagation experiment using a short pulse laser system [16] at the University of Maryland (Fig. 3). A quenched mode dye oscillator creates a 200 ps pulse, which is amplified and used to pump a distributed feedback dye laser (DFDL). The DFDL is made with a holographic transmission phase grating [17], and generates a single output pulse in each direction normal to the pump pulse. The pulsewidth is dependent on the length of the pumped region, and will vary from shot to shot. After two stages of amplification the pulse is frequency-doubled by a BBO crystal to the KrF gain band. The 248 nm light double-passes a KrF discharge laser with ASE-suppression apertures between passes, producing an output beam of 10 to 14 mJ in a 3.2 cm^2 area. The beam then traverses a 3 mm aperture, followed by a 2× reducing telescope with a 400 μm pinhole at the focal plane (in vacuum), which blocks almost all of the residual ASE. At the output lens, the fwhm beam diameter is approximately 1.5 mm.

Beyond the telescope, the beam propagates through 5.3 meters of air to enter a gas cell. At this point, diffraction has caused the beam to acquire an approximately Gaussian distribution with a 0.85 mm fwhm. The cell entrance window is a 5 mm thick calcium fluoride (low n_2) wedge, whose inner surface is AR coated with a 2 cm clear aperture. The cell is a 4.6 meter long, 8 cm ID steel pipe, whose inner walls were treated with a black anti-rust finish. The small beam size plus the inner wall treatment ensured that no light scattered from the inside cell wall. After exiting the cell through a second calcium fluoride window (uncoated), the beam is measured by a cooled CCD camera, two spectrometers, and a streak camera.

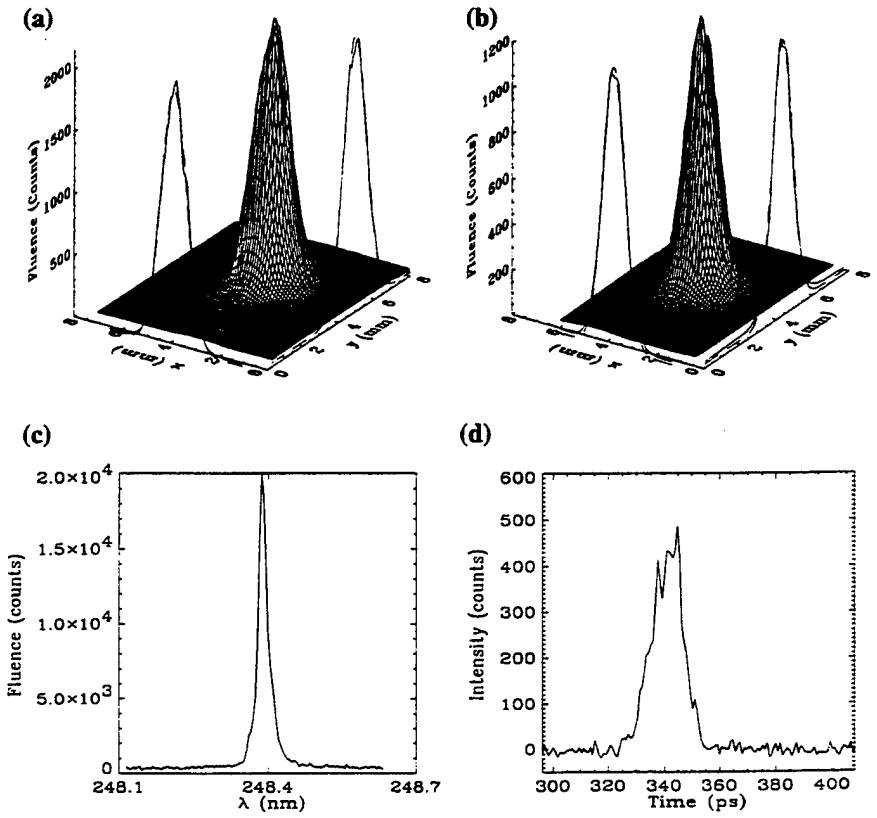


Fig. 4. Magnified incident profile (a) and output profile (b) of the 10 ps KrF beam for the cell under vacuum. The peak intensity of the incident profile is near 1.9 GW/cm^2. Least-squares 2D fits using azimuthally-symmetric Gaussian functions are shown in the dashed lines. The spectrum (c) and streak camera pulse shape (d) are also shown.

The CCD camera acquires the spatial distribution of the output beam, plus a magnified relayed image of the incident beam taken from the 4% reflection off of the uncoated first surface of the cell entrance window. The focus of the lens used to transfer the incident image to the camera was checked by placing a 210 mm wire tied in a 1 mm diameter loop on the input window and comparing it with the image to verify the calculated 1.86 magnification. To ensure that the output image intensity was comparable to the relayed incident image, the output beam was directed onto the CCD camera by reflection off the front face of the first beamsplitter, which was an uncoated fused silica wedge. Additional attenuation was provided by calibrated ND filters that

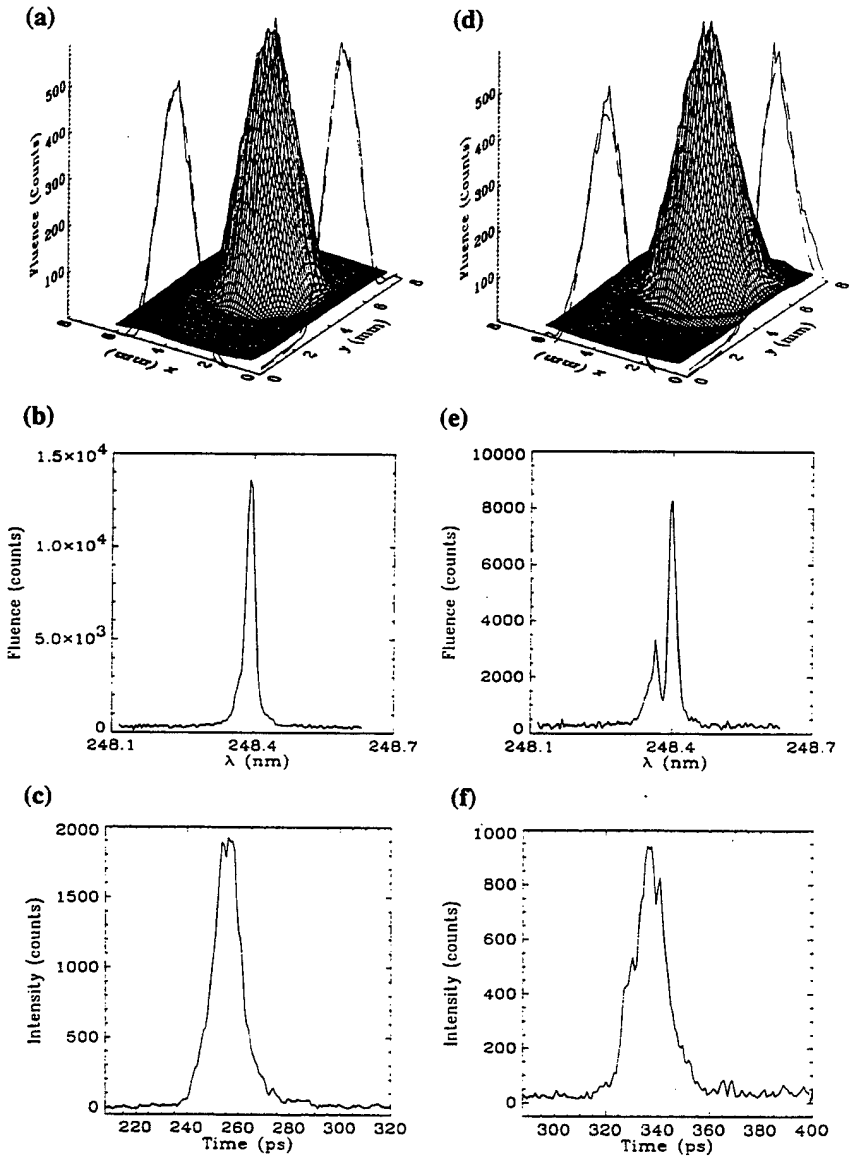


Fig. 5. Output profiles and Gaussian fits, along with the spectra and pulse shapes for 267 Torr (a–c) and 1000 Torr (d–f) of Xe. In the 1000 Torr case, the double-humped spectral structure is consistent with a self-phase modulation of peak value $\sim\pi$ [17]. The input profiles were essentially unchanged from those of Fig. 4.

were shared by the output beam and relayed image. Test shots in which the BBO crystal in the laser was blocked showed that the ASE from the unloaded KrF amplifier was below detectability of the CCD camera to better than 2000:1. The CCD camera was checked for linear response to the short pulse light by comparing the total integrated signal on the camera with the energy on a pyrometer. No evidence of saturation of the CCD response was found.

The portion of the output beam that is transmitted through the first beamsplitter (uncoated wedge) is sent to the spectrometers and the streak camera. One spectrometer, which has high resolution (0.0043 nm), measured the wavelength and the spectral broadening and structure due to self-phase modulation. The second, with lower resolution (0.045 nm) but larger spectral range (242–267 nm), was used to look for light shifted by four-wave parametric optical mixing processes [6,7]. At the intensities (≤ 2 GW/cm^2) and Xe pressures (< 1100 Torr) used in the defocusing and compensation experiments, there was no evidence of any such processes, nor any significant attenuation of the beam within the gas. The streak camera measured the pulse shape after the pressure tube. The pulsewidths exhibited shot-to-shot variation, but generally remained within 9 to 14 ps fwhm. Part of this variation was due to the streak camera itself, whose resolution was ~7 ps with a linear array that was subject to electrical noise from the discharge amplifiers.

Initial experiments were carried out to confirm that the beam was well characterized, and that the incident and output images were not appreciably perturbed by the windows under varying pressures. We first propagated the beam in air with the tube completely removed to verify that the beam did not go through a focus. When the camera was moved in one meter steps from just behind the entrance window of the cell (supported on a mount) to the final position where the output beam was normally measured, the image spread slowly from a diameter of 0.85 mm fwhm at the entrance window to 1.3 mm fwhm at the final position. We then measured the beam with the cell under vacuum and with an argon fill (low n_2 [2]) of 1100 Torr. Fig. 4 shows the incident and output beam profiles, along with the spectrum and pulse shape, for the vacuum case. These results were unaffected by filling the cell with argon.

We measured the broadening of the output beam profile as a function of Xe pressure at various incident intensities for pulses centered around 248.2 nm and 248.4 nm. Fig. 5 shows a sample of this data for 248.4 nm, along with spectral data from the high resolution spectrometer. In the 1000 Torr case, the large negative n_2 has not only significantly broadened the spatial profile (Fig. 5d), but also created the well-known spectral structure (Fig. 5e) due to the self-phase modulation that builds up as the pulse travels through the Xe [18]. The blue shift of the lower spectral peak and (weak) red shift of the main peak are consistent with a negative n_2 and a pulse whose leading edge is steeper than its

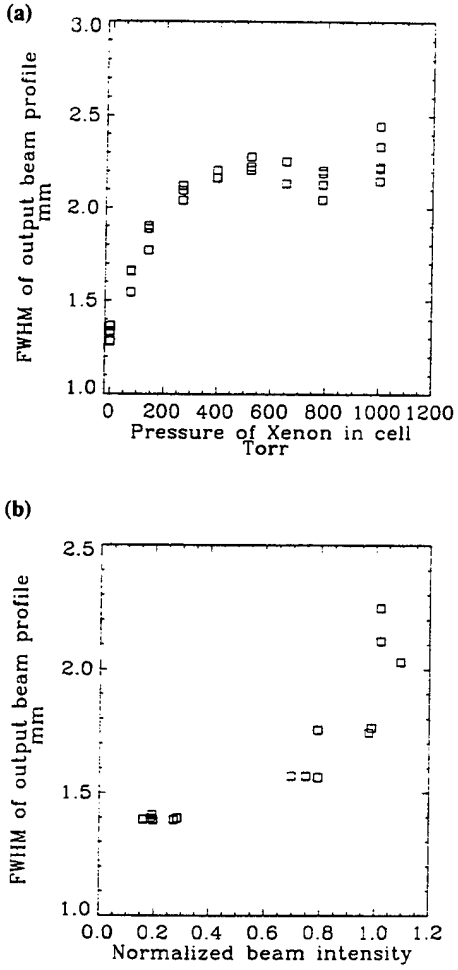


Fig. 6. Fwhm of the Gaussian fits to the output profiles as a function of Xe pressure (a) and normalized incident beam intensity (b). In (a), the Gaussian fits become a poorer approximation to the actual profile, especially in the low intensity skirts, for Xe pressures greater than ~400 Torr. In (b), the fwhm becomes indistinguishable from the vacuum case of 1.4 mm at an intensity of 0.2 GW/cm^2.

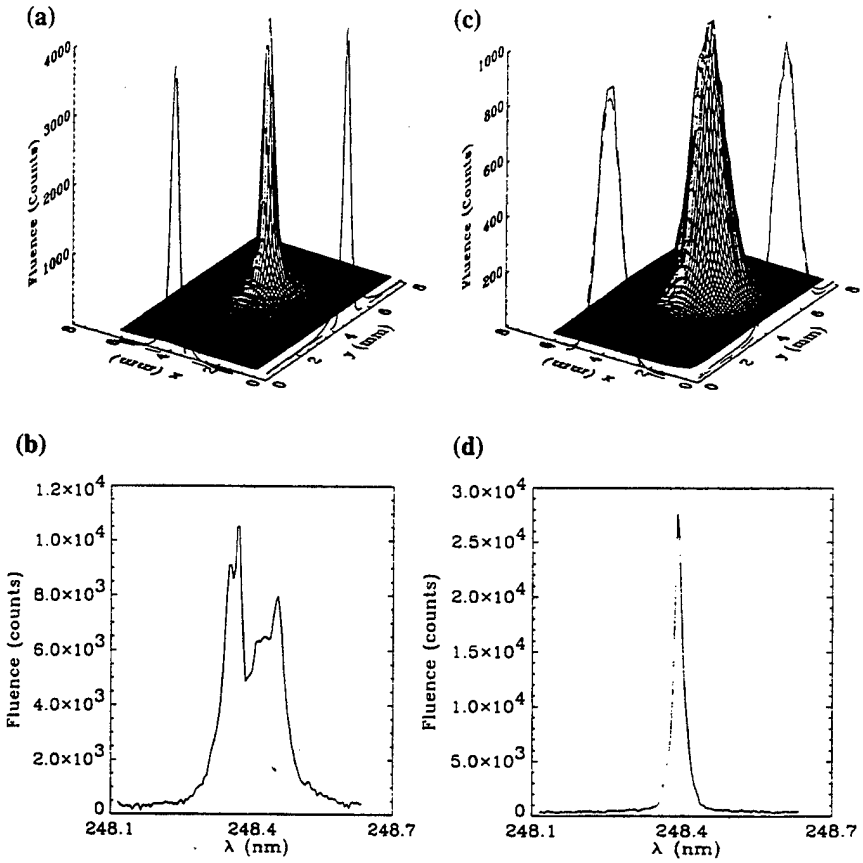


Fig. 7. Compensation of self-focusing in CO_2 by adding Xe. Output profiles with Gaussian fits and spectra are shown for 1260 Torr of CO_2 (a,b), and the CO_2 plus 129 Torr of Xe (c,d).

trailing edge [18]. The spectral shape differs markedly from those created by photoionization [10], indicating that such processes are negligible in our experiments.

As an initial analysis, we plotted the measured fwhm of least squares Gaussian fits as a function of Xe pressure at constant incident intensity (Fig. 6a), and as a function of incident intensity at a constant pressure of 775 Torr (Fig. 6b). These plots show that the defocusing does indeed tend to increase with the incident intensity and Xe concentration, as expected, but the behavior at the highest intensities and pressures is not well understood. Part of the problem stems from the fact that although a Gaussian fit provides an unambiguous number for the fwhm, it becomes a poorer approximation to the actual profile as the nonlinear refraction increases.

The most convincing demonstration of a negative n_2 is the cancellation experiment summarized in Fig. 7. The cell was initially filled in six steps to 1260 Torr with pure carbon dioxide, whose large positive n_2 resulted in whole-beam self-focusing. (Because of the difficulty of accurately modeling molecular dipole matrix elements, no ab initio calculation of n_2 was carried out for CO_2.) As the CO_2 was added, the beam partially focused (Fig. 7a and Fig. 8) and the spectrum (Fig. 7b) developed the double-hump structure due to self-phase modulation [18]. (Note that the positive n_2 has reversed the shifts of the two spectral peaks from those in Fig. 5e.) A small amount of Xe was then added to the cell and allowed to mix with the CO_2. Less than 130 Torr of Xe completely compensated the 1260 Torr of CO_2 (Fig. 8). As shown in Figs. 7c,d, both the beam profile and spectrum were recovered with negligible loss. The Xe pressure at which exact compensation occurs was estimated by using a third order polynomial

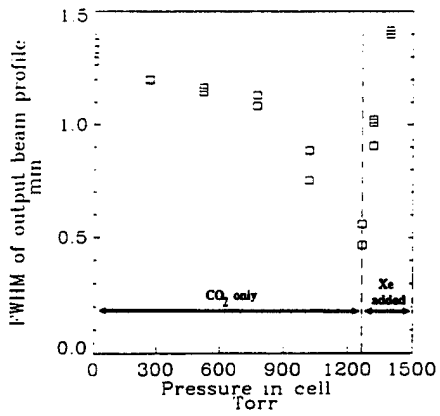


Fig. 8. Fwhm of the Gaussian-fitted output profile versus *total* pressure, which includes Xe in addition to 1260 Torr of CO_2 for all points to the right of the dashed line. The 129 Torr of added Xe slightly overcompensates the beam.

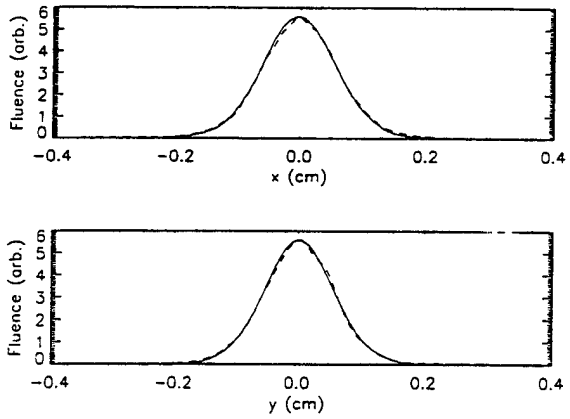


Fig. 9. Comparison of simulated (—) and measured (---) output beam profiles for the vacuum case, showing agreement to ~1.6% RMS of the peak fluence. For this data shot, the optimum curvature of the incident beam was 75 cm^{-2}.

least-squares fit to the data points ⩾ 1260 Torr in Fig. 8. This fit intercepted the average fwhm of the three vacuum shots at 111 Torr, thus giving $n_2(Xe)/n_2(CO_2) = -11.4$ at 248.4 nm.

4. Analysis

To evaluate n_2 from the experimental data, we developed a propagation code that simulates the output profiles from the measured incident beam profiles. It calculates the time-integrated intensity by solving the nonlinear Schrödinger equation for the complex field amplitude $\mathscr{E} = \mathscr{E}(r_\perp, z, t - z/c)$:

$$\left(\frac{\partial}{\partial z} - \frac{i}{2k}\nabla_\perp^2\right)\mathscr{E} = ik\tfrac{1}{2}n_2|\mathscr{E}|^2\mathscr{E}, \qquad (4)$$

where $r_\perp \equiv (x, y)$ and z are the transverse and axial coordinates, respectively, $t - z/c$ is the time measured in the pulse frame, and $k = 2\pi/\lambda$, assuming $n_0 \cong 1$ for the gas. The numerical integration is performed by a split-operator technique [19], which effectively lumps the nonlinear refraction into a stack of evenly-spaced thin windows with the beam propagating freely between them; the free propagation is treated by stan-

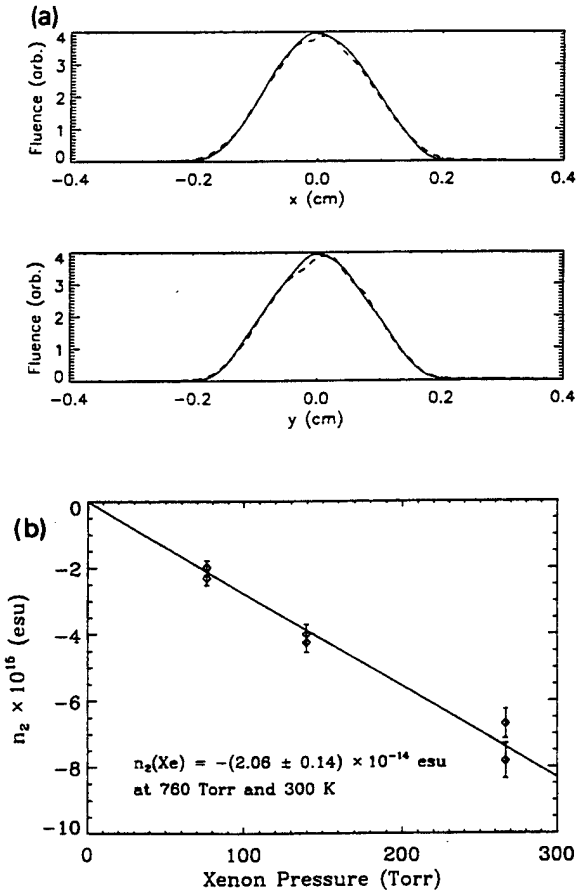


Fig. 10. (a) Comparison of simulated (—) and measured (---) output beam profiles for 140 Torr of Xe, showing self-defocusing. For this shot, the optimum fit (within ~2.1% RMS) was obtained for $n_2(Xe) = -(4.0 \pm 0.3) \times 10^{-15}$ esu. (b) Pressure dependence of the Xe nonlinear refractive indices obtained from the self-defocusing simulations for pressures up to 266 Torr.

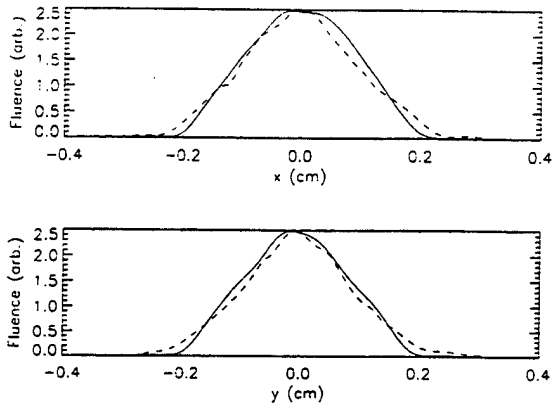


Fig. 11. Comparison of simulated (—) and measured (---) output beam profiles for 785 Torr of Xe, showing self-defocusing. The anomalous behavior, with the simulation being too broad near the axis and too narrow in the skirts, was found in all of the high pressure shots.

dard fast Fourier transform techniques. Most of the calculations used a 128×128 or 256×256 transverse grid, 11–41 z-planes, and 20–30 time steps. The code also models the weak self-focusing in the CaF windows and the free propagation in the 65 cm air path between the cell exit window and CCD camera. To benchmark the code, we modeled the axial growth (for $n_2>0$) of a small ripple $\delta I=\epsilon I_i\cos(Kx)$ superimposed upon a flat-top incident beam profile of intensity I_i, and compared the results with perturbation theory [20]. Agreement between the code calculations and analytic growth rate $K\sqrt{4\pi n_2 I_i/c-K^2/4k^2}$ was excellent (within ~ 1% maximum deviation) for all cases where the ripple amplitude at the output remained small compared to I_i.

The incident complex amplitudes used in the simulations were modeled from the experimental data. The intensity distributions were calculated from the measured pulse shapes and the incident fluence profiles that were imaged onto the CCD camera on each shot. To estimate the effect of the streak camera resolution time $t_{sc}\approx7$ ps, we scaled up the intensities by the deconvolution factor $t_p/\sqrt{t_p^2-t_{sc}^2}$, where t_p is the measured 1/e pulsewidth. This factor typically ranged from 1.1 to 1.3. To remove small-scale noise structure from the fluence profiles without significantly affecting their overall shape, we performed a 2D boxcar average over nearby points. The incident phase distribution $\phi_i(r_\perp)$ was determined by fitting the calculated and measured output profiles for the three shots where the cell was under vacuum. We found that it was well modeled by a combination of linear gradient and diverging quadratic curvature terms; i.e., by the function $\phi_i\cong C_x x+C_y y+C|r_\perp|^2$. The coefficients varied over the three shots, but the optimum fits were found for $|C_x|,|C_y|<2$ cm^{-1} and $C=70$–90 cm^{-2} (i.e., curvature radii of 14–18 m). An example of such a fit is shown in Fig. 9. In these and in subsequent shots, the best fit was determined by minimizing the total rms difference between the calculated and measured fluences for all points where the measured fluence was larger than 10% of its peak value.

We first evaluated n_2 directly from simulations of the Xe self-defocusing data at 248.4 nm and several

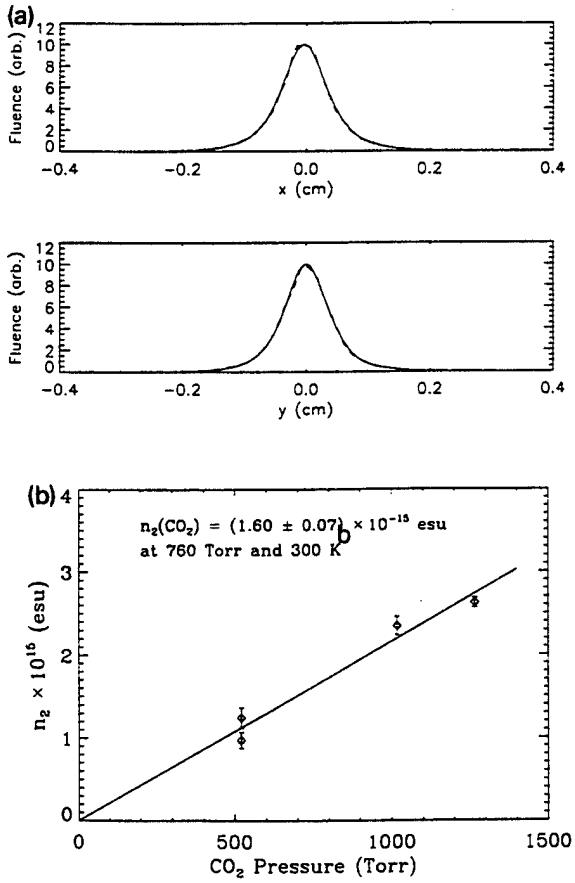


Fig. 12. (a) Comparison of simulated (—) and measured (---) output beam profiles for 1020 Torr of CO_2, showing whole-beam self-focusing. For this shot, the optimum fit (within ~ 1.5% RMS) was obtained for $n_2(Xe)=(2.35\pm0.12)\times10^{-15}$ esu. (b) Pressure dependence of the CO_2 nonlinear refractive indices obtained from the self-focusing simulations.

different pressures. (The ambient temperature was ~20°C.) For each shot, we chose a range of n_2 values that gave the best fit to the measured output fluence profiles, assuming the same 70–90 cm^{-2} range of quadratic phase curvatures that characterized the vacuum shots. At pressures up to 266 Torr, the agreement between measured and calculated profiles was good, with RMS differences around 1.9–3.3% of the peak fluence. Fig. 10a shows an example of this fit, and Fig. 10b shows the pressure dependence of the measured n_2. The linear fit in Fig. 10b scales to $n_2 = -(2.06 \pm 0.14) \times 10^{-14}$ esu at 760 Torr and 300 K.

At pressures much above 266 Torr, the simulations failed to adequately reproduce the shapes of the measured profiles; they were generally too flat near the axis and dropped off too abruptly around the skirts (Fig. 11). The onset of this disagreement coincides roughly with the saturation of the beam widths shown in Fig. 6a. The anomalous behavior could arise from higher order terms in the radial phase distribution, which would become more important in the broader profiles. It could also arise from Xe_2 dimers, whose concentration increases quadratically with Xe pressure. Gornik et al. [4] have observed weak two-photon excitation bands between 247 and 249 nm, which they attributed to Xe_2. (The excitation of the near-resonant atomic $6p[1/2]_0$ state was negligible for $\lambda < 249$ nm.) In our experiments, the fractional Xe_2 concentrations remained small (~0.6% at 1 atm of Xe [21]), and the measured absorption was negligible. The anomalous profiles were apparently not caused by two-photon absorption, or by photoionization from excited Xe_2 states. (Upper limits on the magnitudes of those effects are estimated in the Appendix.) However the close proximity of the excitation bands may resonantly enhance the Xe_2 nonlinear susceptibility enough to have affected our profiles at higher pressures.

Because of the uncertainties encountered in the higher pressure Xe simulations, we also calculated $n_2(Xe)$ indirectly by first evaluating $n_2(CO_2)$ from the CO_2 self-focusing data, then multiplying by the ratio $n_2(Xe)/n_2(CO_2)$ derived from the cancellation experiment. That ratio is unaffected by either profile broadening or high Xe pressure effects. In addition, the self-focused profiles are less sensitive to the uncertain phase variation in the skirts of the incident beam, but are inherently more sensitive to the values chosen for n_2. (The defocusing is a self-limiting process.) The simulations gave excellent fits to the self-focused beams, typically with RMS differences of 1.5–2% of the peak fluence. Fig. 12a shows an example of such a fit, and Fig. 12b shows the pressure dependence of the calculated $n_2(CO_2)$ values. The resulting indices at 760 Torr, 300 K are then $n_2(CO_2) = (1.60 \pm 0.07) \times 10^{-15}$ esu and $n_2(Xe) = -(1.82 \pm 0.08) \times 10^{-14}$ esu.

Table 1
Calculated and measured nonlinear refractive indices (for 760 Torr and 300 K) of xenon at 248.4 nm, compared to that of CO_2 at 248.4 nm, air at 248 nm, and fused silica at 249 nm

Medium (source)	n_2 (10^{-14} esu)
Xe (Eqs. (1) or (3))	-1.54 ± 0.30
Xe (self-defocusing exp.)	-2.06 ± 0.14
Xe (CO_2 compensation exp.)	-1.82 ± 0.08
CO_2 (self-focusing exp.)	$+0.160 \pm 0.007$
Air (Ref. [2])	$+0.029 \pm 0.010$
Fused SiO_2 (Ref. [1])	15–20

5. Summary and conclusion

Table 1 summarizes our Xe and CO_2 results, and compares them with the measured n_2 of air at 248 nm [2] and fused silica at 249 nm [1]. At 248.4 nm, 760 Torr, and 300 K, our combined measurements yield $n_2(Xe) = -(1.97 \pm 0.23) \times 10^{-14}$ esu, in good agreement with the ab initio calculation of $-(1.54 \pm 0.30) \times 10^{-14}$ esu. The magnitudes of the measured xenon indices are 60–75 times larger than the positive n_2 in air, and within a factor of ten of that measured in fused silica. This suggests that under suitable conditions, Xe cells could be used to compensate self-phase modulation and self-focusing in air paths and KrF window materials, thereby increasing the useful KrF power density. The only caveat is the question of what causes the anomalous behavior at higher Xe pressures. If it is caused by Xe_2 dimers, then the problem might be solved by operating at lower pressures, heating the gas, or using longer more energetic pulses to dissociate the dimers. These questions can be addressed in a later paper.

Acknowledgements

The authors wish to thank J. Goldhar, B. Welch, and R. Elton at the University of Maryland and T. Lehecka

at Science Applications International Corporation for their contributions to this project, and to J. Bone and R. Morse for engineering support. This work was supported by the U.S. Department of Energy.

Appendix A

In this appendix, we estimate upper limits on dimer two-photon absorption, and on refraction due to photoionization of the excited dimer states. From the two-photon Xe_2 spectra in Ref. [4], plus those authors' earlier measurement of the weak atomic line at 247 nm, we estimate an upper limit on the two-photon excitation rate $W_{Xe2}\,(s^{-1}) < 3.7\times10^5 I^2\,(GW/cm^2)$ at 248.4 nm. The two-photon absorption coefficient $\kappa_{Xe2} = 2h\nu N_{Xe2} W_{Xe2}/I$ for $N_{Xe2} \propto N_{Xe}^2$ density [21] is then

$$\kappa_{Xe2}(cm^{-1}) < 8\times10^{-5}[p(Torr)/760]^2 I(GW/cm^2).$$

This result predicts less than 1% peak on-axis absorption in our 4.6 m propagation path under all conditions where n_2 was measured ($I \le 2$ GW/cm^2 and $p \le 266$ Torr). Although this absorption is larger at higher pressures, it is not large enough to have been unambiguously measured, or to have created the anomalous beam profiles that we observed. Simulations in which the two-photon absorption term was added to Eq. (4) revealed small losses and negligible profile distortion at all pressures that were used, even when the coefficient was chosen several times larger than the 8×10^{-5} shown above.

The electron density produced by photoionization of the two-photon-excited dimer levels is $N_e \cong \frac{1}{2}(N_{Xe2} W_{Xe2} t_p)(\sigma_{PI} I/h\nu) t_p$, where t_p is the pulsewidth and $\sigma_{PI} \approx 4\times10^{-18}$ cm^{-1} [22] is the photoionization cross section of those levels. The resulting refractive index perturbation is $\delta n_{PI} = -N_e/2N_c$, where $N_c \equiv m\omega^2/4\pi e^2$ is the critical density. To compare this with our result $\delta n = \frac{1}{2} n_2 |\mathscr{E}|^2$ (assuming $n_2 \cong -2\times10^{-14}\, p(Torr)/760$ esu) we evaluate the ratio $\delta n_{PI}/\delta n$:

$$\frac{\delta n_{PI}}{\delta n} < 4\times10^{-8}[p(Torr)/760] I^2(GW/cm^2) t_p^2(ps),$$

which is negligible under all conditions of our defocusing experiments.

We have also estimated the absorption and refraction due to *three*-photon ionization of the *atomic* xenon [5]. Although this process is resonantly-enhanced by the $6p[1/2]_0$ state, its absorption and refraction remain negligible under our experimental conditions. However, its excitation rate scales as I^3, and it is likely to limit operating intensities to a few times 10 GW/cm^2, even if dimer effects can be eliminated.

References

[1] I.N. Ross, W.T. Toner, C.J. Hooker, J.R.M. Barr and I. Coffey, J. Mod. Optics 37 (1990) 555.
[2] M.J. Shaw, C.J. Hooker and D.C. Wilson, Optics Comm. 103 (1993) 153 and references therein.
[3] R.H. Lehmberg, S.P. Obenschain, C.J. Pawley, M.S. Pronko, A.V. Deniz and T. Lehecka, eds. H.T. Powell and T.J. Kessler, Proc. SPIE 1870 (1993) 163.
[4] W. Gornik, S. Kindt, E. Matthias and D. Schmidt, J. Chem. Phys. 75 (1981) 68.
[5] S. Kröll and W. Bischel, Phys. Rev. A 41 (1990) 1340.
[6] A. Tünnermann, K. Mossavi and B. Wellegehausen, Phys. Rev. A 46 (1992) 2707 and references therein.
[7] A. Tünnermann, C. Momma, K. Mossavi, C. Windolph and B. Wellegehausen, IEEE J. Quantum Electon. 29 (1993) 1233.
[8] R.H. Lehmberg, J. Reintjes and R.C. Eckardt, Phys. Rev. A 13 (1976) 1095.
[9] R.H. Lehmberg, J. Reintjes and R.C. Eckardt, Appl. Phys. Lett. 30 (1977) 487.
[10] S.P. Le Blanc, Z. Qi and R. Sauerbrey, Optics Lett. 20 (1995) 312 and references therein.
[11] E. Koenig, Physica (Utrecht) 62 (1972) 393.
[12] M. Klapisch, J.L. Schwob, B.S. Fraenkel and J. Oreg, J. Opt. Soc. Am. 67 (1977) 148.
[13] M. Klapisch, Comput. Phys. Comm. 2 (1971) 239.
[14] Charlotte E. Moore, Atomic Energy Levels, U.S. Nat. Bur. Standards (U.S. GPO, Wash., DC, 1971) Vol. III.
[15] M. Aymar, M. Crance and M. Klapisch, J. Physique 31 (1970) 141.
[16] Y. Leng, V.E. Yun, J. Goldhar and H.R. Griem, Rev. Sci. Instrum., submitted.
[17] D. Hatten, Y. Cui, W.T. Hill III, T. Mikes and J. Goldhar, Appl. Optics 31 (1992) 7042.
[18] J.F. Reintjes, Nonlinear Optical Parametric Processes in Liquids and Gases (Academic Press, Orlando, 1984) Ch. 5, Sec. 4.
[19] J.A. Fleck Jr., J.R. Morris and M.D. Feit, Appl. Physics 10 (1976) 129; 14 (1977) 99.
[20] V.I. Bespalov and V.I. Talanov, JETP Lett. 3 (1966) 307.
[21] Marie-Claude Castex, J. Chem. Phys. 74 (1981) 759.
[22] M.M. McCusker, Excimer Lasers, ed. C.K. Rhodes (Springer, Berlin, 1979).

Appendix F

The Nike KrF Laser Facility: Performance and Initial Target Experiments

The Nike KrF laser facility: Performance and initial target experiments*

S. P. Obenschain,† S. E. Bodner, D. Colombant, K. Gerber, R. H. Lehmberg, E. A. McLean, A. N. Mostovych, M. S. Pronko, C. J. Pawley, A. J. Schmitt, J. D. Sethian, V. Serlin, J. A. Stamper, and C. A. Sullivan
Code 6730 Plasma Physics Division, U. S. Naval Research Laboratory, Washington, DC 20375

J. P. Dahlburg and J. H. Gardner
Laboratory for Computational Physics, Naval Research Laboratory, Washington, DC 20375

Y. Chan, A. V. Deniz, J. Hardgrove, and T. Lehecka
Science Applications International Corporation, McLean, Virginia 22102

M. Klapisch
Artep, Inc., Columbia, Maryland 21045

(Received 10 November 1995; accepted 6 February 1996)

Krypton-fluoride (KrF) lasers are of interest to laser fusion because they have both the large bandwidth capability (>THz) desired for rapid beam smoothing and the short laser wavelength ($\frac{1}{4}$ μm) needed for good laser–target coupling. Nike is a recently completed 56-beam KrF laser and target facility at the Naval Research Laboratory. Because of its bandwidth of 1 THz FWHM (full width at half-maximum), Nike produces more uniform focal distributions than any other high-energy ultraviolet laser. Nike was designed to study the hydrodynamic instability of ablatively accelerated planar targets. First results show that Nike has spatially uniform ablation pressures ($\Delta p/p<2\%$). Targets have been accelerated for distances sufficient to study hydrodynamic instability while maintaining good planarity. In this review we present the performance of the Nike laser in producing uniform illumination, and its performance in correspondingly uniform acceleration of targets. © *1996 American Institute of Physics.* [S1070-664X(96)94705-X]

I. INTRODUCTION

The inventions and refinement of laser beam smoothing[1,2-3] have enhanced the prospects for direct-drive fusion. The combination of a short laser wavelength with laser beam smoothing should control laser plasma instabilities,[4-9] particularly when using moderate laser intensities ($\leq 10^{15}$ W/cm^2). The combination of beam smoothing with precise beam balance should easily provide sufficient low-mode illumination uniformity around a pellet (mode numbers <10).

The most serious remaining challenge to direct-drive fusion[10-13] is the control of the shorter-wavelength asymmetry and hydrodynamic instability growth that can disrupt symmetric pellet implosions. Possible control measures include the following.

(1) Improved laser uniformity, particularly during the early, laser imprinting phase.
(2) Improved target fabrication, particularly at the shorter wavelengths that are most dangerous to Rayleigh–Taylor (RT) growth.
(3) Exotic targets that may reduce the imprint (such as gold-coated foam).[14]
(4) Thicker targets via controlled preheat (shocks, x rays).[15,16]
(5) Lower aspect ratio pellets (initial thickness to diameter), with the concomitant higher laser intensities ($>10^{15}$ W/cm^2).

There is a penalty to pay in the pellet yield when utilizing measures #4 and #5, and possibly #3. Use of a higher intensity (#5) may also reintroduce laser–plasma instabilities. The highest performance route, and also the most conservative in regard to the target physics, will be to optimize items 1 and 2, and then appeal to the other methods.

Nike is a high-energy krypton-fluoride (KrF) laser that was primarily developed to study the physics and means for control of hydrodynamic instability. The KrF laser has unique characteristics for such experiments, particularly in regard to beam smoothing technology. Nike produces the most uniform target illumination of all available high-energy uv lasers (at $\frac{1}{3}$ or $\frac{1}{4}$ μm). In this paper we describe the Nike laser, its performance as a laser, and its initial use in uniform acceleration of planar targets.

II. THE NIKE LASER FACILITY

KrF is an excimer laser that combines two highly desired characteristics for the fusion application.[17,18] First, it has a short (248 nm) wavelength, needed for efficient target coupling and for suppression of laser–plasma instability. Second, it has a broad laser bandwidth capability for rapid beam smoothing. Bandwidths exceeding 3 THz have been demonstrated in large KrF laser systems.[19] Design studies indicate that KrF scales to the megajoule energies that will be required for high-gain laser fusion. KrF also has potential as a driver for the energy application[20] (see the Appendix). KrF poses technical challenges as well as opportunities in developing large systems since the architecture is quite different from that of the more established glass lasers.

*Paper 3IB1, Bull. Am. Phys. Soc. 40, 1693 (1995).
†Invited speaker.

 1070-664X/96/3(5)/2098/10/$10.00

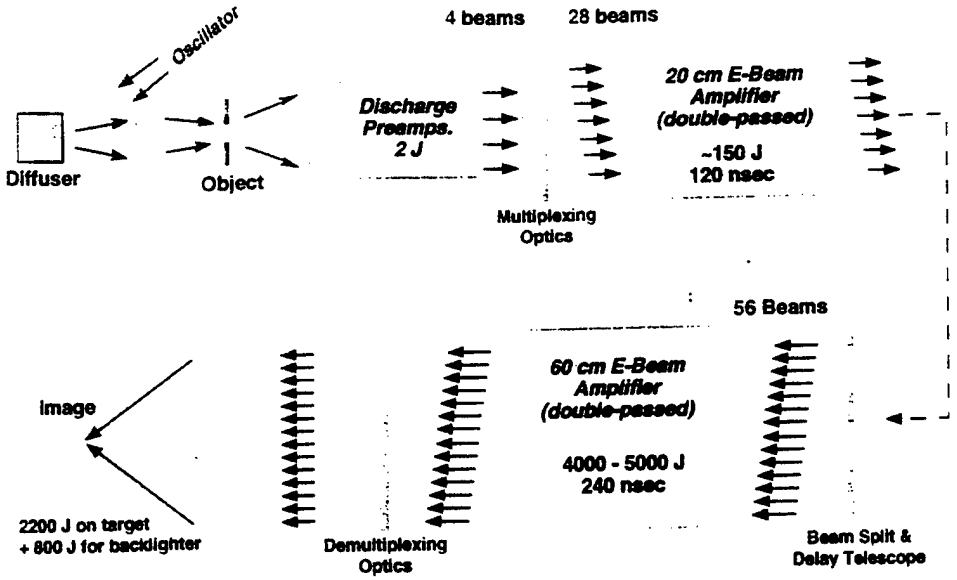


FIG. 1. Nike optical configuration and amplifier staging. A uniformly illuminated aperture in the front end is imaged onto target through the amplifier system (echelon-free ISI).

The Nike laser utilizes *echelon-free ISI (Induced Spatial Incoherence)* beam smoothing: an aperture at the front end of the laser is uniformly illuminated by spatially and temporally incoherent light. This object is then imaged through the amplifiers onto the target.[21] The amplifiers are placed at or near the Fourier transform planes of that aperture, so that the target illumination is not sensitive to amplifier gain nonuniformities. To obtain good beam fidelity from the aperture to the target, the added beam divergence due to cumulative optical phase errors must be small compared to the angular divergence of the smoothed beams. In addition, the nonlinear phase shift (B integral) must be small. The KrF gaseous medium, low saturation fluence in the amplifiers, and optical architecture leads to small B integrals.[22]

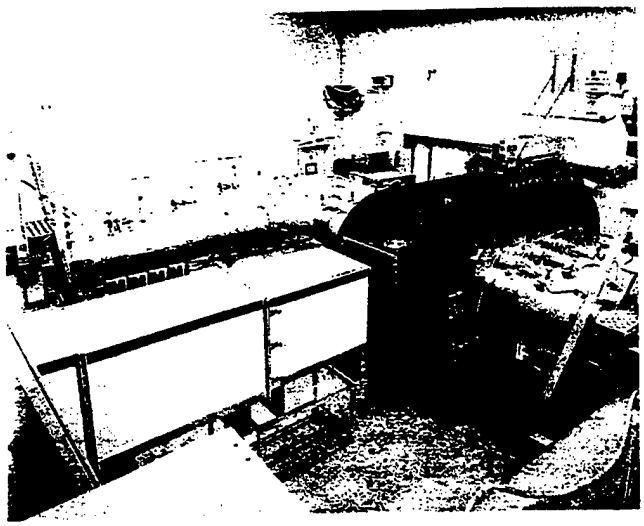


FIG. 2. Photograph of the 60 cm aperture electron-beam pumped final amplifier. Four parallel water lines located on either side of the amplifier provide the pulse power. The laser beams double pass the amplifier. An insulated tunnel protects the laser beams from air turbulence.

Figure 1 depicts the echelon-free ISI concept in the Nike optical system.[23] A broadband multispatial mode, discharge-pumped KrF oscillator illuminates a reflecting diffuser, which in turn illuminates the initial ISI aperture. Because of the diffuser, the energy illuminating the aperture is small (~20 nJ). However the diffuser assures that the aperture is illuminated uniformly (to the statistical limit of such broadband incoherent illumination sources). The beam energy is then amplified by a series of three discharge-pumped amplifiers, followed by an array of four large discharge-pumped amplifiers operated in parallel.[24] The Fourier plane of the ISI aperture is relayed to the center of each of the discharge-pumped amplifiers via a series of relaying telescopes. The gain duration of these discharge pumped amplifiers is about 25 ns full width at half-maximum (FWHM) for the oscillator and first two preamplifiers, and 15 ns for the final amplifiers. Pulse shaping is achieved via Pockels cell shutters located after the first amplifier. The rejected beam from the Pockels cell copropagates with the target beam through the discharge-pumped amplifiers, and is then separated by a combination of polarizers and spatial filtering. The copropa-

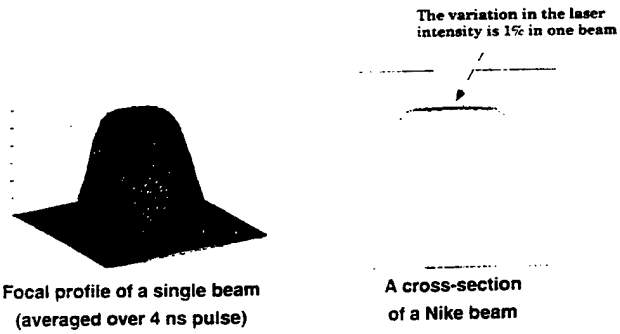


FIG. 3. Nike focal distribution of a single attenuated beam after amplification by the 60 cm amplifier. The FWHM of the focal profile is 750 μm.

FIG. 4. Photograph of optics near the target chamber including: the input array to the 60 cm amplifier (foreground left); the final turning arrays directing beams into the target chamber; and the 44 beam lens array for the target-acceleration beams.

gating pulse preserves the pulse shape and contrast of the laser pulse by maintaining a constant loading of the partially saturated amplifiers. The discharge-pumped amplifier system produces 1.5–2 J in a 4 ns FWHM pulse.

Nike then utilizes two angularly multiplexed[25] electron-beam-pumped amplifiers. The angular multiplexing system uses only mirrors and high $F\#$ lenses to preserve the broad bandwidth capability of the laser. Beams from the front end system are split and delayed by various amounts to obtain 28 sequential 4 ns FWHM beams. Twenty-two beams intended for target experiments are delayed 4 ns center to center, while six beams intended for backlighters have 5 ns delays center to center. The beams sequentially extract energy from the 140 ns gain duration 20 cm aperture amplifier. A large beam splitter and delay telescope after the 20 cm amplifier doubles the number of beams. The 56 beams then sequentially extract their final energy from the 250 ns gain duration

FIG. 5. The Nike target chamber. The laser beams converge on target through the cone-shaped portion of the vacuum chamber seen in the background. The 12-beam focusing array for the backlighter beams is to the left of the spherical target-interaction chamber.

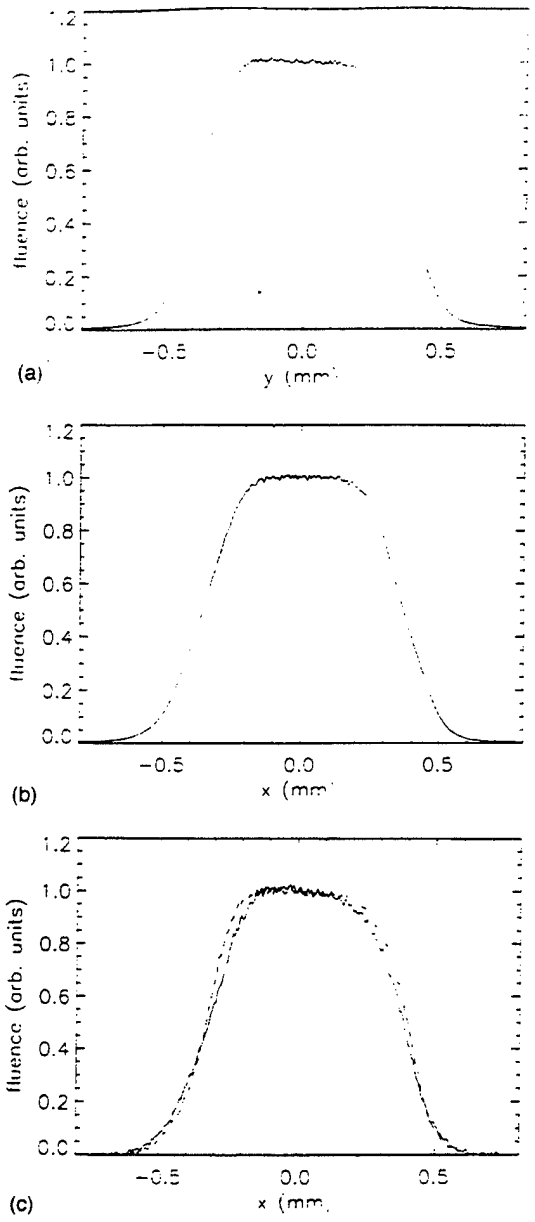


FIG. 6. Focal distributions of 37 overlapped beams in the Nike target chamber with only the front end amplifiers firing. (a) and (b) show perpendicular profiles at the target location while (c) shows a profile 300 μm closer to the lens array (dashed line) and a profile 300 μm farther from the lens array (solid line). Here 77% of the laser energy sampled by the diagnostic (3.5 $\times$2.4 mm^2 field of view at the target plane) is contained within the focal FWHM. The peak-to-valley curvature determined by a best second-order fit across the center 50% of the FWHM of (a) is 2.2% and 2.6% for (b). (c) indicates that the beam overlap is quite good over axial distances longer than the planned target accelerations (200–300 μm).

60 cm aperture amplifier. The KrF lasing media has a short excited state lifetime (a few ns), so one must nearly continuously extract laser energy during the e-beam pumping in order to obtain good efficiency and avoid pulse distortion. The 44 target interaction beams are amplified with little temporal distortion because their 4 ns FWHM pulse lengths match the interbeam delays. The pulse shape of the initially 4 ns FWHM backlighter beams tend to lengthen to match the 5 ns pulse interbeam delays. The front end design allows in-

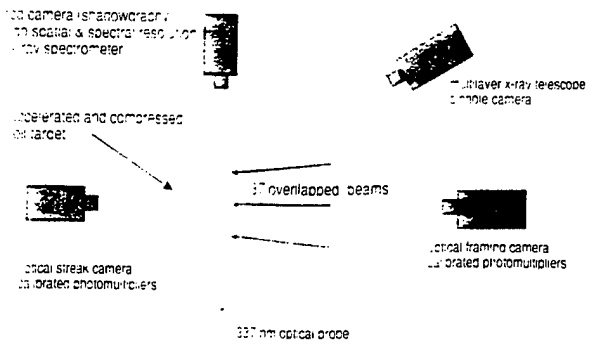


FIG. 7. Diagram of an accelerated foil target and currently fielded diagnostics. The width (~2 mm) of the target foil is chosen wide compared to the laser focal diameter to isolate the relatively cool rear surface from the laser irradiated surface.

stallation of an independent pulse shaping for the backlighter beams if needed.

A photograph of the final 60 cm aperture amplifier is shown in Fig. 2. Both of the large Nike amplifiers employ large, monolithic cold-cathode diodes to electron-beam pump the laser gas from opposite sides of the gas cell. External magnetic fields guide the e beam into the laser gas.[26] Up to 5 kJ laser output has been obtained from the 60 cm aperture amplifier.[27] The amplifier is routinely operated with 3–4 kJ output. Approximately 75% of the laser light in each beam at the amplifier is transported to the target through the remaining optical system.

Figure 3 shows the time-integrated focal distribution of a single 4 ns beam after amplification by the 60 cm amplifier. The total energy in the 56 beams was 3.9 kJ for this laser shot (measured in front of the first mirror array following the 60 cm amplifier). The aperture in the front end was adjusted to provide an angular width of 75× diffraction limit. The beam was propagated a distance typical of that to the target chamber, attenuated, and focused using a 6 m focal-length lens (the same as the target chamber). The image after the 6 m lens was then magnified 4× and placed onto a cooled charge-coupled-device (CCD) camera. The magnification is sufficient with the CCD camera's 23 μm pixel size to detect the smallest interference structures in the focused laser beam. To remove light-sensitivity variation, each camera pixel was calibrated at the laser wavelength.[28] We observed smooth, flat-topped focal distributions with the center 50% of the

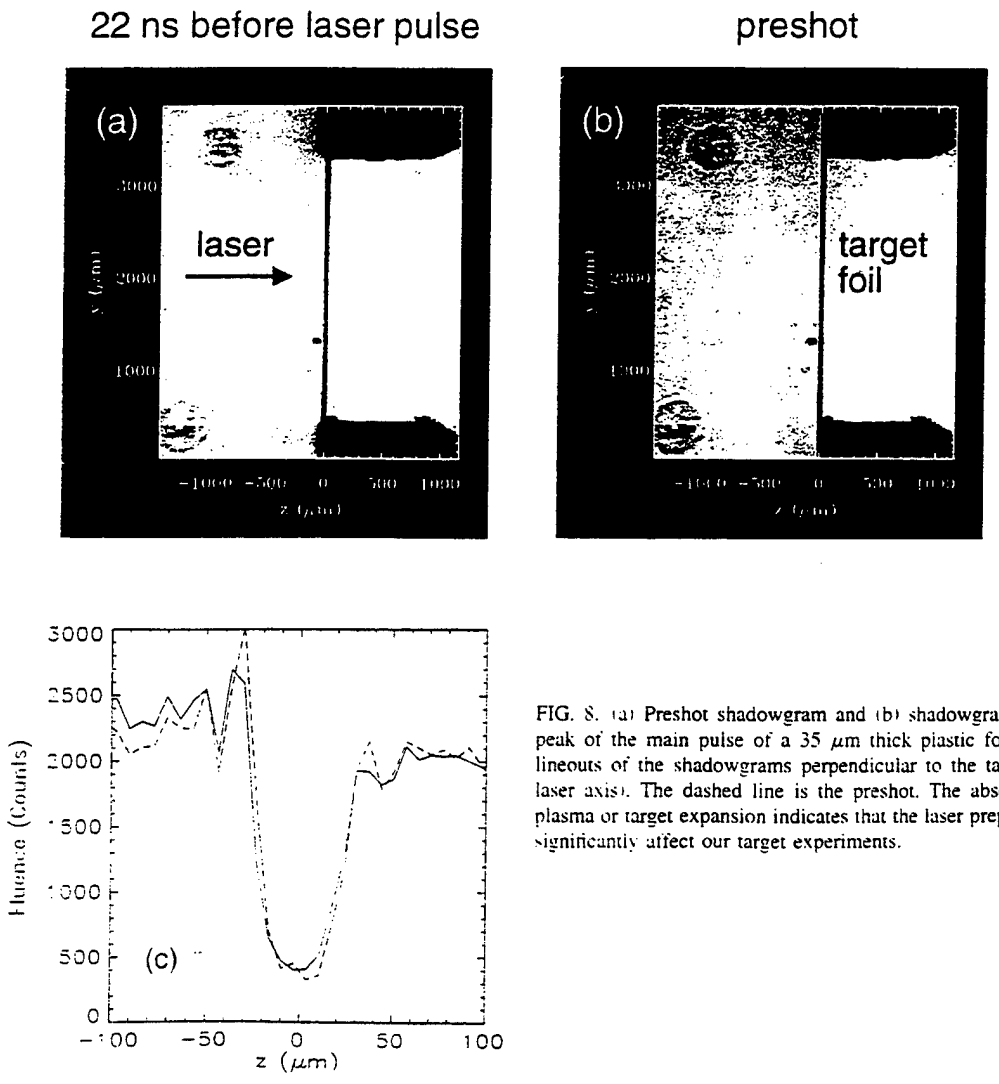


FIG. 8. (a) Preshot shadowgram and (b) shadowgram 22 ns before the peak of the main pulse of a 35 μm thick plastic foil target. (c) shows lineouts of the shadowgrams perpendicular to the target foil (along the laser axis). The dashed line is the preshot. The absence of measurable plasma or target expansion indicates that the laser prepulse is too small to significantly affect our target experiments.

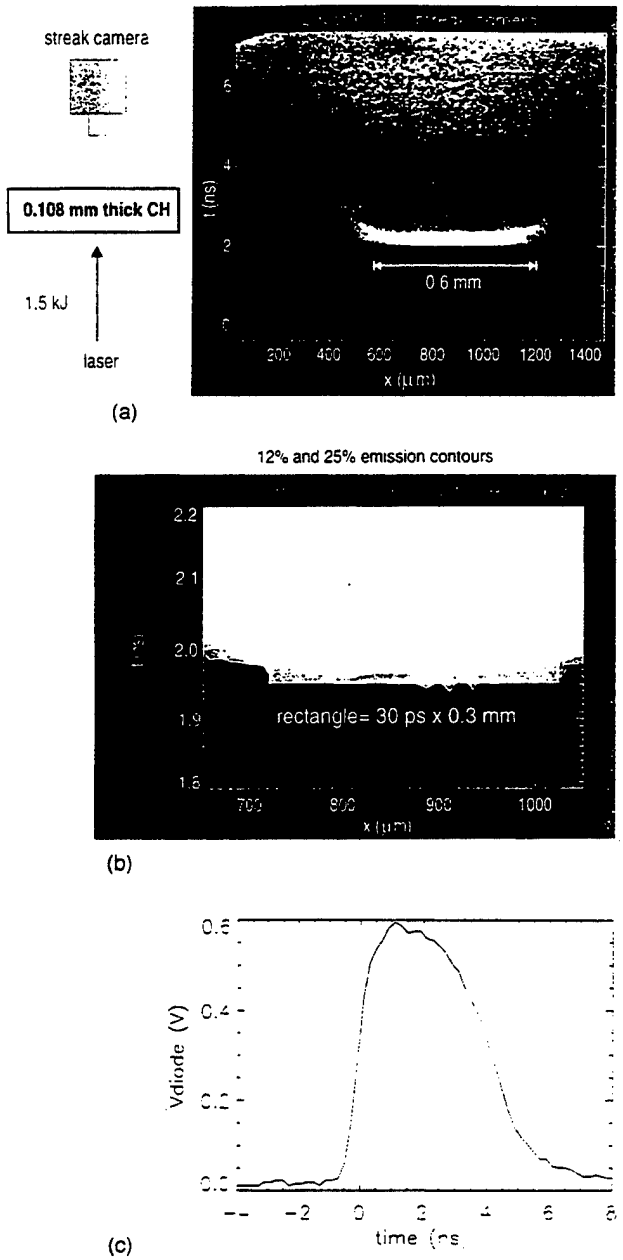


FIG. 9. (a) Streak camera image of the uniform shock breakout obtained from the rear surface of a 108 μm thick Nike illuminated target. The peak intensity is 0.8×10^{14} W/cm^2. (b) 12% and 25% of peak emission contour plots on an expanded time scale. The streak camera system has 25 μm spatial and 30 ps temporal resolution. (c) Diode measurement of the pulse shape of a Nike target beam after the 60 cm amplifier during this target shot. The diode measurement has a 0.5 ns temporal resolution.

FWHM deviating from a flat fit by 1% rms. The center rms nonuniformity is close to the ideal result calculated from $\langle \Delta I/I\rangle \cong (\tau_c/\tau_a)^{1/2} \cong 1.3\%$ using τ_c=laser-coherence time$\cong$ 0.6 ps (bandwidth is 1 THz FWHM) and τ_a=averaging time $\cong$4 ns. The curvature near the edges of the focal distribution is the result of the cumulative phase errors in the optical train. (This phase aberration, which corresponds to an approximately 15× diffraction limit point spread distribution at 248 nm, could be reduced by using deformable mirror techniques similar to those in the Beamlet laser at Lawrence Livermore National Laboratory.[29]) We observe no appreciably enhanced distortion in the laser focal profile when firing the 60 cm amplifier, provided one waits a sufficient time between shots for the amplifier gas to cool: approximately 30 min.

After the final amplifier, the time delay between the beams is removed by the demultiplexing optical system, and the beams are then directed to the target chamber. Figure 4 is a photograph of the optical system near the target chamber, showing several mirror arrays and the structure that supports the target chamber lens array. The mirrors are remotely aligned by a computerized control and alignment system. The optics between the front end and the target chamber are contained in a precisely temperature-controlled propagation bay (45 m×10 m floor area). Once aligned onto the target, the beams tend to stay aligned because of the temperature control ($\Delta T < \pm 1$ F).

Diagnostics have been installed in the Nike system to evaluate and maintain the amplifier system reliability, and to monitor the pulse shape, energy, and focal quality of the laser beams. These diagnostics are all digitized and fed into a UNIX based workstation system via fiber optic links. The energy and pulse shape are monitored for the four beams from the discharge-pumped amplifiers. The focal distribution of one of these beams is monitored via a cooled CCD camera. The laser bandwidth is monitored via a spectrograph. The demultiplexing mirrors that remove the optical delays after the 60 cm amplifier all have fused silica substrates that allow transmission of 248 nm light. This provides measurements of pulse shape and energy using partially transmitted beams. Six of the demultiplexing mirrors have optical grade substrates that allow measurement of the focal distributions of high-energy beams during target shots.

III. NIKE TARGET FACILITY

Forty-four of the 56 laser beams are directed to an array of 6 m focal length lenses for target acceleration experiments. The remaining 12 beams illuminate targets for x-ray backlighting, via an array of 3 m focal length lenses. Figure 5 is a photograph of the Nike target chamber.

For the experiments reported here, 7 of the 44 target acceleration beams were blocked by calorimeters located at the target chamber focusing lens. Here 1.1–1.6 kJ was incident on the laser-accelerated targets. Higher laser energy, (~3 kJ in 44 beams), has become available for such experiments in the recent operation of Nike.

Figures 6(a) and 6(b) show perpendicular cross sections of the focal distribution obtained when overlapping 37 beams onto a target with only the front end firing. (Earlier measurements have shown that the profile of single beams is essentially unchanged when firing the e-beam pumped amplifiers.) The measurement was obtained by placing a uv fluorescent crystal at the target plane and imaging the resulting fluorescent light onto a cooled CCD camera using a multielement lens. The spatial resolution of the measurement is 12 μm. The flatness, overall curvature, and FWHM of the focal profiles are similar to those of a single beam, indicating that good overlap is achieved by the Nike optical and alignment systems. The effective F#(2.6) of the lens array and the focal diameters were chosen large enough that the over-

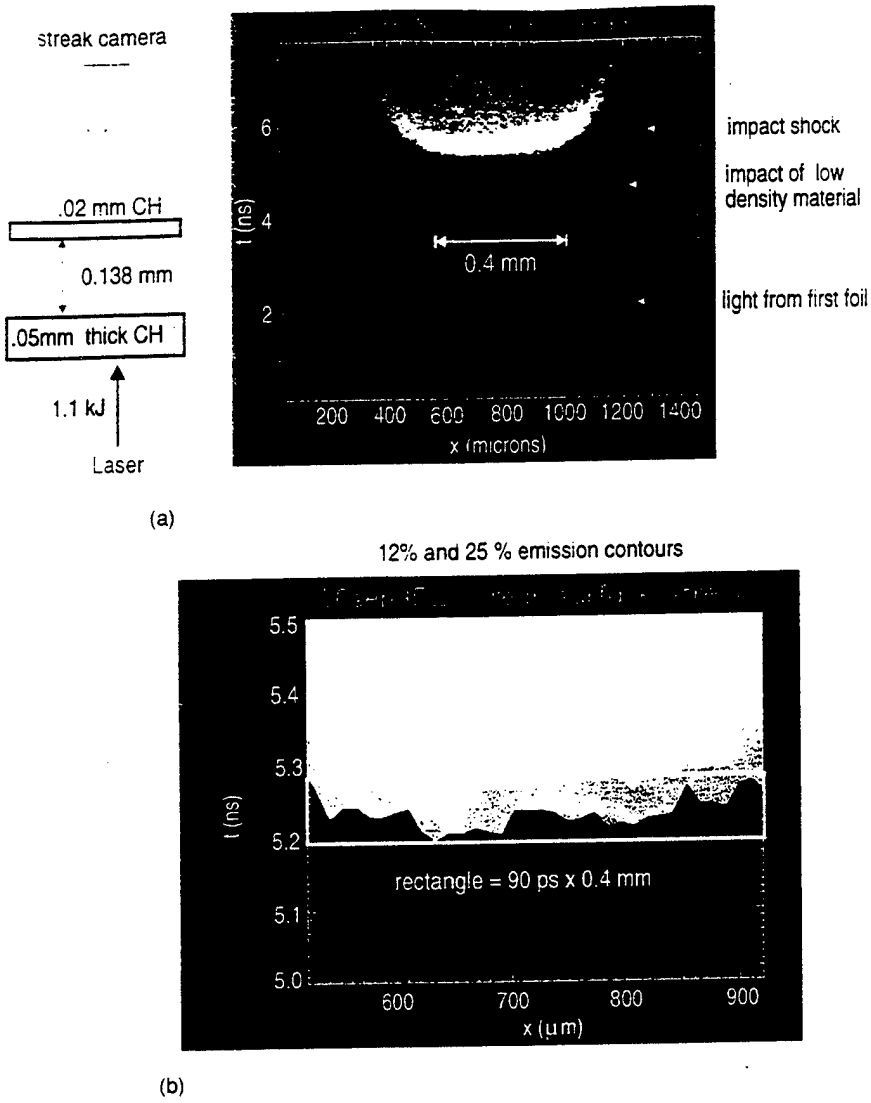


FIG. 10. (a) Streak camera image and (b) 12% and 25% emission contours for shock breakout from a double foil with 138 μm spacing. The peak laser intensity is 0.6×10^{14} W/cm^2. The laser pulse shape is similar to that shown in Fig. 9. Prior to the collision, the CH impact foil is transparent to visible light. The first flash of light in the streak image is light originating from the rear of the accelerating foil that is transmitted through the impact foil. The second flash is due to low-density (precursor) material from the laser-accelerated foil impacting the surface of the second foil. These early signals are not observed when using opaque pyrolytic carbon targets as impact foils. The light emission from the shock breakout after the accelerated target impacts the second foil is much more intense than these early emissions.

lapped beams can maintain uniform illumination of targets accelerated several hundred μm along the laser axis. Figure 6(c) shows that composite focal profiles maintain good illumination uniformity at distances ±300 μm along the laser axis from the position shown in Fig. 6(a).

Overlapping many beams improves the illumination uniformity for the spatial-wavelength regime that is most dangerous for laser imprinting on the target, and the subsequent Rayleigh–Taylor growth. Overlapping the multiple smoothed beams shifts most of the fine-scale interference structure to shorter wavelengths (1–7 μm). We do not expect the laser accelerated targets used in our experiments to respond to this very fine structure, because those wavelengths are shorter than the in-flight (compressed) thickness of the targets. The illumination nonuniformity at the intermediate wavelengths (~10–~100 μm), most important to Rayleigh Taylor growth, should be reduced by a factor of $N^{1/2}$, where N is the number of beams overlapped onto the target. The effective illumination nonuniformity at wavelengths important for imprinting may therefore be as small as 0.4% rms for Nike with averaging time of 1 ns and the current 1 THz FWHM bandwidth. This uniformity could be improved further by increasing the laser bandwidth. The Nike optical system was designed to accommodate bandwidths up to the ~5 THz theoretical limit available with KrF.

IV. NIKE TARGET EXPERIMENTS

Nike was designed to study laser imprinting and hydrodynamic instability.[30] Figure 7 illustrates the Nike experimental conditions, along with a diagram of the presently fielded diagnostics. We try to obtain planarity in the target acceleration experiments by limiting the target acceleration distance to less than half of the focal diameter. The goals of

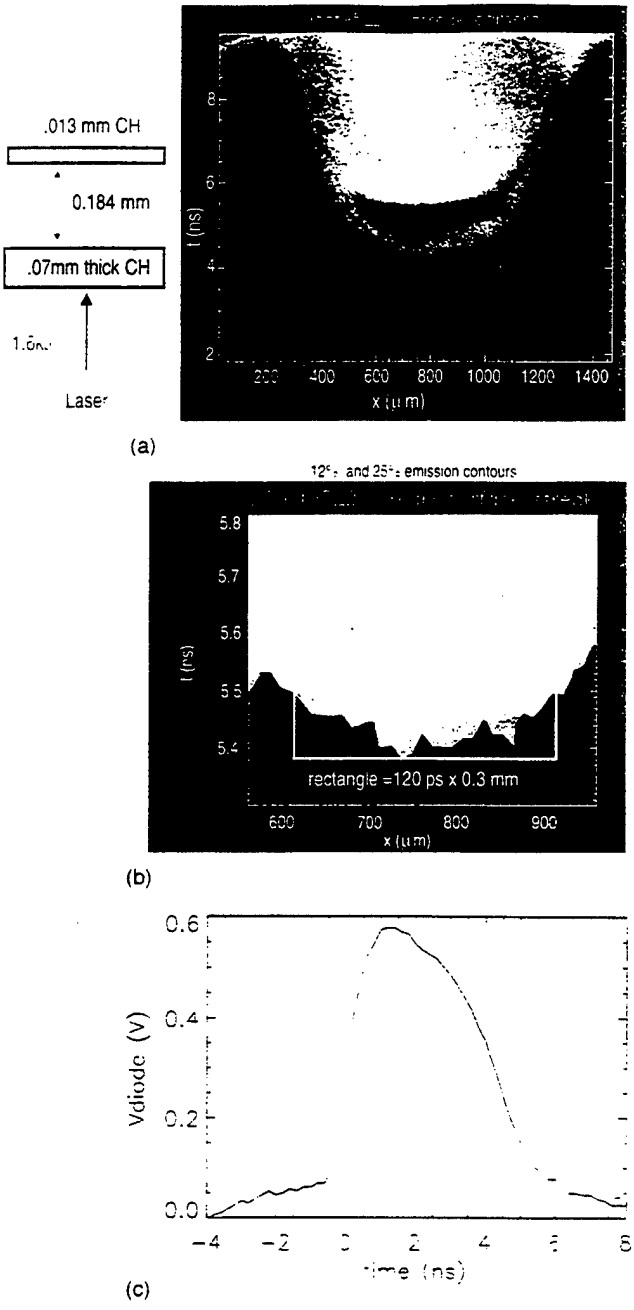


FIG. 11. Similar double-foil measurement to Fig. 10, but with a thicker first foil (69 μm), increased laser intensity (0.8×10^{14} W/cm^2), and increased foil spacing (184 μm). (c) Diode measurement of a Nike beam for this target shot. A 10% amplitude prepulse was introduced for this shot to precompress the laser-accelerated first foil.

the initial target experiments include (i) determining if uniform laser illumination actually produces uniform ablation pressures; and (ii) determining whether planar targets can be accelerated for several hundred microns, with good lateral uniformity, and without excessive edge effects. As shown below the results were quite positive.

Figures 8(a) and 8(b) are a shadowgram of a 35 μm thick by 2 mm wide plastic (CH) foil taken as a preshot, and also 22 ns before arrival of the main laser pulse. The shadowgraphy system employs a 337 nm laser. F-12 collection optics and has a 6 μm spatial resolution. At the time shown the target has been exposed to most of the prepulse that can occur because of amplified spontaneous emission and beam cross-talk from the two large electron-beam pumped amplifiers. To the 6 μm resolution of the shadowgraphy diagnostic, the target foil shows no measurable preheat-induced expansion or acceleration of the target surfaces. We estimate that the diagnostic could detect 10 μm axial scale length plasmas with densities as small as 10^{19} electrons/cm^3 corresponding to less than 0.5% of critical density for 248 nm light. In earlier measurements shadowgraphy did detect plasma on the laser illuminated side of the target (but not the rear surface) due to prepulse. [Cold CH targets have an absorption length at 248 nm that is short (2 μm) compared to the target thickness.] Measurements in the target chamber with the electron-beam amplifiers off, revealed a 5×10^{-7} prepulse (fraction of peak intensity) at 25 ns before the main pulse. The prepulse fell to less than 5×10^{-8} at 150 ns before the main pulse. The duration of the prepulse indicated that beam-to-beam scattering from optics in the 20 cm amplifier, rather than the 60 cm amplifier, were the major source of cross-talk-induced prepulse. Two well-used (>1 yr), partially fogged windows in the 20 cm amplifier were replaced and the prepulse-induced plasma detected by shadowgraphy disappeared. Shadowgraphy is now routinely used to monitor that a particular target is not distributed by prepulse. Photomultipliers also monitor the front and rear surface emission of our targets with the capability of detecting preplasmas with brightness temperatures as low as 0.3 eV. The above will soon be supplemented by beam diagnostics that measure the prepulse during target power shots down to the $<10^{-8}$ level.

Figure 9(a) shows a streak camera image of 500 nm light emission from the rear surface of a laser-irradiated 108 μm thick plastic (CH) target. The breakout time of the shock is very uniform over the center region of the image. Figure 9(b) shows a contour plot with an expanded time scale. Over the center 300 μm, the shock breakout time is uniform to the resolution of the camera (30 ps). The time of flight of the most intense shock through the foil is approximately 2 ns (predicted by our hydrocode simulations). The 30 ps temporal uniformity in shock emission indicates the shock velocity is uniform to within 1.5%. Assuming shock velocity is proportional to the square root of the ablation pressure, the long-wavelength pressure nonuniformity in the central 300 μm must be less than 3% (peak to valley). The rather large (108 μm) initial thickness of the target probably contributes to the rounding near the edges that can be observed in Fig. 9.

Figure 10(a) shows a measurement of the uniformity of target acceleration via the double-foil diagnostic.[31,32] The lateral uniformity of the target acceleration can be deduced from the variation in impact time with a closely spaced second foil, as seen by streak camera measurements of the shock breakout. Figure 10(b) shows a contour plot on an expanded time scale. The two foils are separated by 138 μm, the accelerated target is 49 μm thick plastic (CH) while the impact foil is 20 μm thick CH. The shock breakout time from the second foil is uniform to within 90 ps across the center 400 μm. The time of flight to impact is approximately 5 ns. The ratio between shock breakout-time nonuniformity and time of flight indicates that the target velocity across the

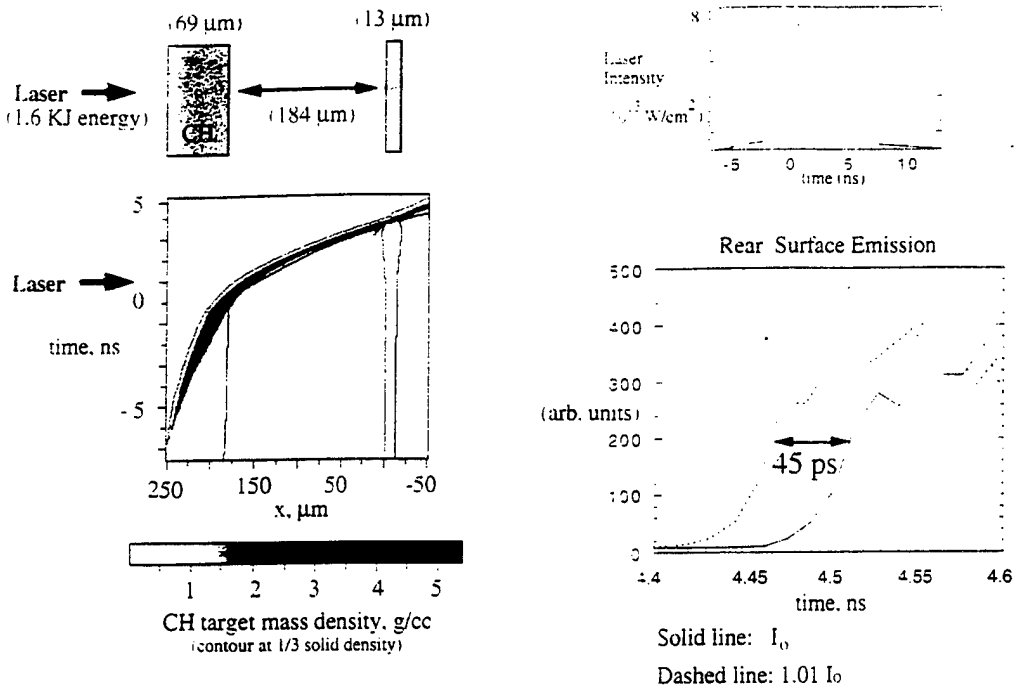


FIG. 12. One-dimensional simulation of the acceleration and impact of a double foil under the conditions similar to the experiment shown in Fig. 11.

center 400 μm is uniform to within 2% (peak to valley).

Figures 11(a) and 11(b) show a double-foil measurement with larger foil separation (184 μm), thicker first foil (69 μm), and higher laser energy than Fig. 10. The shock breakout time is still highly uniform in the center 300 μm (120 ps total variation), but with noticeably more curvature than Fig. 10, near the edges of the accelerated target. The distance from target front surface to impact foil is quite long here (253 μm) and finite focal-spot-diameter effects should become more important. The Rayleigh–Taylor experiments will concentrate on perturbation wavelengths below 100 μm: these experiments should not be affected by the small levels of long-scale-length curvature observed here. For the experiment shown in Fig. 11 a prepulse was used to precompress the laser-accelerated foil target. In future experiments we will be able to improve on the planarity of the target acceleration by employing a larger focal diameter during the low-intensity, prepulse portion of such laser pulses. This is

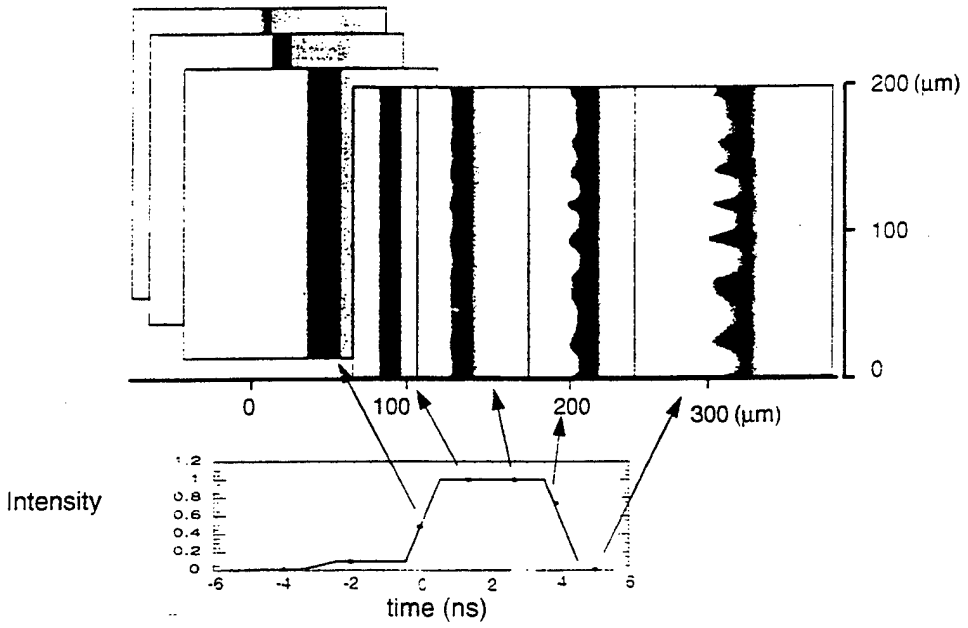


FIG. 13. Two-dimensional simulation of the acceleration of a single CH foil target under conditions similar to Fig. 11. The foil target is assumed to be perfectly smooth. The perturbations arise from the residual nonuniformity in the overlapped Nike beams. The simulation predicts significant growth of Rayleigh–Taylor instability seeded by the laser imprint.

straightforward to implement with the beam smoothing scheme used by Nike, where the focal diameter on target is determined by apertures in the front end.

Figure 12 shows a one-dimensional hydrocode simulation of the acceleration and impact of a double foil target under the laser and target conditions of Fig. 11. The code includes super transition array (STA) opacities. The predicted x-ray preheat of both foils in this particular calculation used local thermodynamic equilibrium (LTE) atomic physics and a variable Eddington transport, with the a laser pulse shape similar to that used in the experiment. The shock preheat of the laser-accelerated foil is low and the target compresses to about 4× solid density before impact. Control of shock-induced preheat will be important in future experiments that study hydrodynamic instability. Figure 12 includes a prediction of the difference in impact times (45 ps) that one would observe when the laser changes 1% in intensity (equivalent to 0.8% $\delta p/p$). The 120 ps nonuniformity in the impact time observed experimentally in Fig. 11 across the center 300 μm thus corresponds to a 2.1% (peak-to-valley) nonuniformity in ablation pressure.

The curvature in the central region of the double-foil impacts shown in Figs. 10 and 11 is consistent with the 2%–3% peak-to-valley curvature in the illumination observed across the center 375 μm (50% FWHM) of our focal profiles (see Fig. 6).

Two calibrated photomultipliers supplemented the streak camera measurements of optical emission from the rear surface of foil targets.[33] The photomultipliers have 1 ns time resolution and employ interference filters centered at 480 and 700 nm. These measurements indicate that the preheat level is low for our target acceleration experiments. Typically, the rear surface brightness temperature remained below 2 eV for accelerated single foil targets under conditions similar to those employed for the double-foil experiments.

V. CONCLUSIONS AND PLANS FOR FUTURE EXPERIMENTS

A series of experiments using single and double foils have demonstrated that targets can be accelerated at least 200 μm with good lateral uniformity and low preheat. Simulations indicate that this acceleration distance is sufficient for studies of imprinting and Rayleigh–Taylor instability. Figure 13 shows a multimode two-dimensional (2-D) imprinting calculation for conditions similar to those of Fig. 11, but with a single-foil target. The code models the small imprinting level by the overlapped Nike beams and the subsequent growth of RT. Even with the very smooth ISI beams, the simulation predicts that imprint grows to appreciable amplitudes after accelerating 200 μm. Note, however, that the rear surface of the target remains relatively flat and uniform, and the double-foil technique probably cannot properly resolve this structure. In the near future we will begin x-ray backlighting and sidelighting diagnostics to measure the net effect of imprinting and Rayleigh–Taylor growth.

Our initial experiments indicate that Nike will be a good test bed for studying the physics and the means for controlling imprinting and Rayleigh–Taylor instability under conditions relevant to high-gain laser fusion (long pulse durations compared to the laser's coherence time, and thick targets). The laser intensity available for the hydrodynamic experiments is limited to $\sim 2\times10^{14}$ W/cm^2 by the need for large focal diameters. This intensity is comparable with those of earlier RT experiments with glass lasers, but is somewhat less than the peak intensity ($3-7\times10^{14}$ W/cm^2) envisioned for high-gain targets. However, it is generally agreed that laser intensity is not a very sensitive parameter for such experiments. We plan a close synergism between Nike experiments and pellet simulation codes to ensure that the regimes explored are relevant to high-gain fusion. Our experience with Nike has also shown that a KrF laser can be a reliable and valuable tool for fusion research.

ACKNOWLEDGMENTS

The authors thank J. Bone, D. Hardesty, W. Webster, D. Gibson, S. Swann, N. Nocerino, L. Granger, D. Williams, and K. Goins for their technical support. We thank J. F. Seely, Y. Aglitsky, G. Holland, and C. Brown for fielding x-ray diagnostics for Nike experiments. Los Alamos National Laboratory contributed to the Nike optical and target chamber designs. Commonwealth Technology, Inc. and East Coast Engineering were responsible for most of the Nike mechanical engineering. I. Smith, M. McGeoch, J. Shipman, and J. Goldhar contributed extensively to the design of the Nike amplifier systems. We thank J. Jacob, K. Ueda, and M. Shaw for useful discussions on KrF technology.

This work was supported by the U.S. Department of Energy.

APPENDIX: ENERGY APPLICATIONS

KrF lasers may be practical for the energy application. High repetition rates are possible because the gaseous lasing media can be cooled by a continuous flow. We have observed intrinsic efficiencies (ratio of laser energy out to electron energy deposited in the laser gas) of 8% with the final Nike amplifier. Experimental and theoretical studies of KrF indicate that intrinsic efficiencies of 12%–14% can be attained when one uses stage gains of ~20× (rather than the ~80× with Nike) and thereby extract a larger fraction of the amplifier energy.[34,35] The wallplug efficiency of existing single-shot high-energy KrF amplifiers is more modest, typically 1.5% (Nike laser) to 2% (ASHURA laser)[18] due to high-gain staging and various losses in the pulse-power systems. Much higher wallplug efficiencies seem attainable for KrF systems designed for the energy application. In a recent conceptual design study (SOMBRERO),[20] a 7% KrF laser efficiency and pellet gain of 118 was found sufficient for a commercial energy reactor.

Y. Kato, K. Mima, N. Miyanaga, S. Arinaga, Y. Kitagawa, M. Nakatsuka, and C. Yamanaka, Phys. Rev. Lett. 53, 1057 (1984).
R. H. Lehmberg and S. P. Obenschain, Opt. Commun. 46, 27 (1983).
S. Skupsky, R. W. Short, T. Kessler, R. S. Craxton, S. Letzring, and J. M. Soures, J. Appl. Phys. 66, 3456 (1989).
S. P. Obenschain, J. Grun, M. J. Herbst, K. J. Kearney, C. K. Manka, E. A. McLean, A. N. Mostovych, J. A. Stamper, R. R. Whitlock, S. E. Bodner, J. H. Gardner, and R. H. Lehmberg, Phys. Rev. Lett. 56, 2807 (1986).
A. N. Mostovych, S. P. Obenschain, J. H. Gardner, J. Grun, K. J. Kearney,

C. K. Manka, E. A. McLean, and C. J. Pawley, Phys. Rev. Lett. **59**, 1193 (1987).
[6]S. P. Obenschain, C. J. Pawley, A. N. Mostovych, J. A. Stamper, J. H. Gardner, A. J. Schmitt, and S. E. Bodner, Phys. Rev. Lett. **62**, 768 (1989).
[7]O. Willi, T. Ashar-rad, S. Coe, and A. Giulietti, Phys. Fluids B **2**, 1318 (1990).
[8]A. J. Schmitt, Phys. Fluids **31**, 3079 (1988).
[9]W. Seka, R. E. Bahr, R. W. Short, A. Simon, R. S. Craxton, D. S. Montgomery, and A. E. Rubenchik, Phys. Fluids B **4**, 2232 (1992).
[10]M. H. Emery, J. H. Gardner, R. H. Lehmberg and S. P. Obenschain, Phys. Fluids B **3**, 2640 (1991).
[11]J. P. Dahlburg, J. H. Gardner, and R. J. Taylor, Bull. Am. Phys. Soc. **40**, 1750 (1995).
[12]J. H. Gardner, J. P. Dahlburg, A. J. Schmitt, and S. E. Bodner, Bull. Am. Phys. Soc. **40**, 1750 (1995).
[13]J. D. Kilkenny, S. G. Glendinning, S. W. Haan, B. A. Hammel, J. D. Lindl, D. Munro, B. A. Remington, S. V. Weber, J. P. Knauer, and C. P. Verdon, Phys. Plasmas **1**, 1379 (1994).
[14]T. Afshar-rad, M. Desselberger, M. Dunne, J. Edwards, J. M. Foster, D. Hoarty, M. W. Jones, S. J. Rose, P. A. Rosen, R. Taylor, and O. Willi, Phys. Rev. Lett. **73**, 74 (1994). M. Dunne, M. Borghesi, A. Iwase, M. W. Jones, R. Taylor, O. Willi, R. Gibson, S. R. Goldman, J. Mack, and R. G. Watt, Phys. Rev. Lett. **75**, 3858 (1995).
[15]C. P. Verdon, Bull. Am. Phys. Soc. **38**, 2010 (1993).
[16]J. H. Gardner, S. E. Bodner, and J. P. Dahlburg, Phys. Fluids B **3**, 1070 (1991).
[17]M. J. Shaw, Laser Part. Beams **9**, 309 (1991).
[18]Y. Owadano, I. Okuda, Y. Matsumoto, M. Tanimoto, T. Tomie, K. Koyama, and M. Yano, Lasers Part. Beams **7**, 383 (1989).
[19]R. C. Sze, S. J. Thomas, N. A. Kurnit, and C. W. Patterson, *1990 Conference on Lasers and Electro-Optics*, OSA Technical Digest Series (Optical Society of America, Washington, DC, 1990), Vol. 7, Paper CWF46.
[20]I. N. Sviatoslavsky, M. E. Sawan, R. R. Peterson, G. L. Kulcinski, J. J. MacFarlane, L. J. Wittenberg, H. Y. Khater, E. A. Mogahed, and S. C. Rutledge, Fusion Technol. **21**, 1470 (1992).
[21]R. H. Lehmberg and J. Goldhar, Fusion Technol. **11**, 532 (1987).
[22]R. H. Lehmberg, S. P. Obenschain, C. J. Pawley, M. S. Pronko, A. V. Deniz, and T. Lehecka, Proc. SPIE **1870**, 163 (1993).
[23]T. Lehecka, R. H. Lehmberg, A. V. Deniz, K. A. Gerber, S. P. Obenschain, C. J. Pawley, M. S. Pronko, and C. A. Sullivan, Opt. Commun. **46**, 485 (1995).
[24]M. Pronko, IEEE J. Quantum Electron. **QE-10**, 2147 (1994).
[25]J. A. Hanlon and J. McLeod, Fusion Technol. **11**, 634 (1987).
[26]L. A. Rosocha and K. B. Riepe, Fusion Technol. **11**, 576 (1986).
[27]J. D. Sethian, C. J. Pawley, S. P. Obenschain, V. Serlin, T. Lehecka, W. Webster, I. D. Smith, P. A. Corcoran, and R. A. Altes, "The Nike electron-beam-pumped amplifiers," *Digest of Technical Papers*, 10th IEEE International Pulsed Power Conference, Albuquerque, NM, 1995. (Institute of Electronics and Electrical Engineering, New York, in press).
[28]A. V. Deniz and S. P. Obenschain, Opt. Commun. **106**, 114 (1994).
[29]B. M. Van Wonterghem, J. H. Murray, D. R. Speck, and J. H. Campbell, Fusion Technol. **26**, 702 (1994). See National Technical Information Service Document No. UCRL-LR-105821-95-1 (B. M. Van Wonterghem, J. T. Salmon, and R. W. Wilcox, Inertial Confinement Fusion Quarterly Report 5, 42, 1994). Copies can be ordered from the National Technical Information Service, Springfield, VA 22161.
[30]S. G. Glendinning, S. V. Weber, P. Bell, L. B. DaSilva, S. N. Dixit, M. A. Henesian, D. R. Kania, J. D. Killkenny, H. T. Powell, R. J. Wallace, P. J. Wegner, J. P. Knauer, and C. P. Verdon, Phys. Rev. Lett. **69**, 1201 (1992). J. Grun, M. E. Emery, C. K. Manka, T. N. Lee, E. A. McLean, A. Mostovych, J. Stamper, S. E. Bodner, S. P. Obenschain, and B. H. Ripin, Phys. Rev. Lett. **58**, 2672 (1987).
[31]S. P. Obenschain, J. Grun, B. H. Ripin, and E. A. McLean, Phys. Rev. Lett. **46**, 1402 (1981).
[32]A. V. Bolotin, V. V. Gavrilov, S. M. Gol'berg, A. Yu Gol'tsov, V. N. Kondrashov, N. G. Kovalsky, A. L. Velikovich, and S. V. Zavyalets, Phys. Fluids B **4**, 2596 (1992).
[33]E. A. McLean, S. H. Gold, J. A. Stamper, R. R. Whitlock, H. R. Griem, S. P. Obenschain, B. H. Ripin, S. E. Bodner, M. J. Herbst, S. J. Gitomer, and M. K. Matzen, Phys. Rev. Lett. **45**, 1246 (1980).
[34]A. E. Mandl, D. E. Klimek, and E. T. Salesky, Fusion Technol. **11**, 542 (1987).
[35]A. Suda, F. Kannari, and M. Obara, J. Appl. Phys. **65**, 4532 (1989).

Appendix G

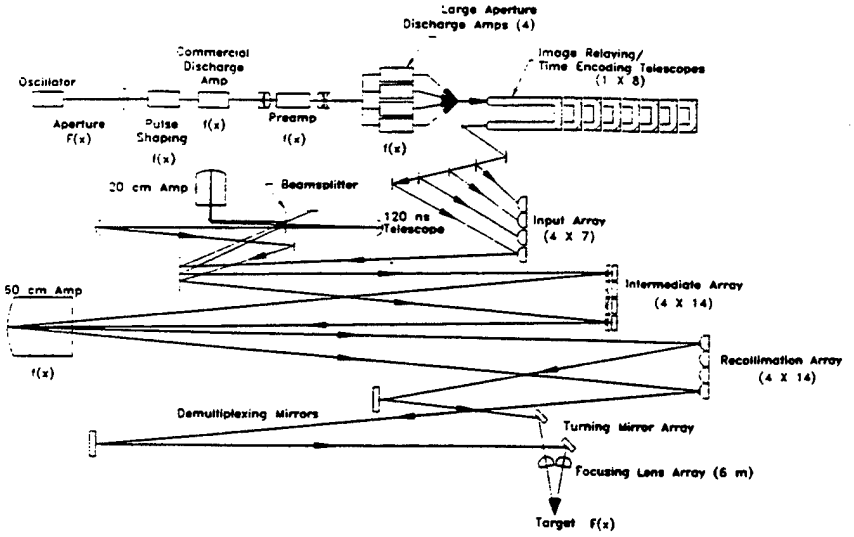
Initial Results from the Nike KrF Laser Facility

Initial results from the Nike KrF laser facility

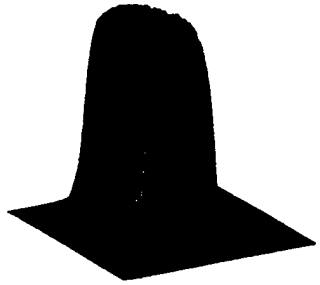
T. Lehecka, A. V. Deniz, J. J. Hardgrove, S. E. Bodner, K. A. Gerber, R. H. Lehmberg, E. A. McLean, S. P. Obenschain, C. J. Pawley, M. S. Pronko, J. D. Sethian, J. A. Stamper, C. A. Sullivan, *Applied Physics Operation, Science Applications International Corp., M.S. 2-3-1, McLean, Virginia 22102*

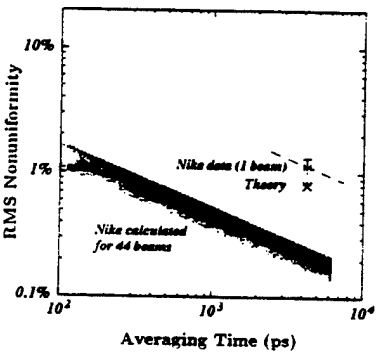
Nike is a 5-kJ KrF laser facility at the U.S. Naval Research Laboratory. The primary goal for the system is the production of ultrauniform, high-intensity ($\geq 2 \times 10^{14}$-W/cm^2) focal distributions that will be used to accelerate planar targets under conditions similar to those expected for direct-drive inertial-confinement fusion (ICF). Nike achieves intensity uniformity through the use of echelon-free induced spatial incoherence.[1] This technique is essentially an image-amplification scheme which the desired focal distribution is produced by diffusely illuminating an object aperture with the output from a broadband, spatially incoherent oscillator. This distribution is then image relayed through multiple amplification stages, resulting in a high-intensity image of the initial object aperture on target. Because the amplifiers are located at the Fourier-transform plane of the object aperture, the focal profile is relatively insensitive to gain nonuniformities in the amplifiers.

A simplified block diagram of Nike is shown in Fig. 1. The system consists of a commercial oscillator/amplifier, an array of discharge amplifiers, and two electron-beam (e-beam)-pumped amplifier stages (with apertures 20×20 cm^2 and 60×60 cm^2). Angular multiplexing is used to produce a train of 4-ns pulses through the e-beam amplifiers. The



CFJ3 Fig. 1. Simplified block diagram of the Nike laser system.

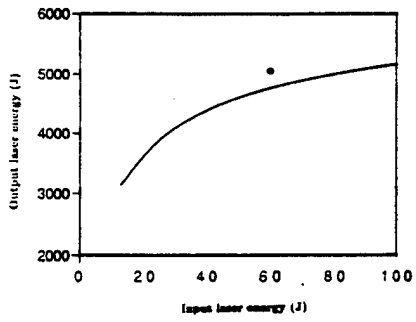




CFJ3 Fig. 2. Single beam focal profile after the Nike 20 cm amplifier and predicted focal spot uniformity for a single beam and 44 overlapped beams on target.

entire amplifier chain is complete, along with optics to deliver 30 of the 56 beams to target. (Delivery of 26 final turning mirrors is pending.)

Previously, we reported the production of uniform focal profiles with the 20-cm amplifier with 40-times-diffraction-limited (XDL) beams.[2] We have extended this result to a 60-XDL beam, the desired spatial incoherence for initial target experiments. An example of the focal profile of a single beam obtained with the 20-cm amplifier is shown in Fig. 2, along with the additional beam smoothing expected on Nike when 44 beams are overlapped on target. This shows that we expect to achieve an rms nonuniformity of less than 1% for averaging times of more than 400 ps.



CFJ3 Fig. 3. Kinetics code prediction (solid line) and measured value of energy out of the Nike 60 cm amplifier.

For the profile shown, the measured uniformity was 1% peak-to-valley tilt, 1.0% peak-to-valley second-order curvature, and 1.4% rms deviation for the linear fit. This performance is better than the design goals of 2% nonuniformity.

The output from the 20-cm amplifier is used to drive the 60-cm amplifier. With approximately 60 J of laser input energy and 70 kJ of e-beam energy in the laser gas we have met the design goal of 5 kJ out of the 60-cm amplifier.[3] This result is in good agreement with kinetics-code predictions, as shown in Fig. 3.

We are currently aligning the output beams from the final amplifier onto target. Diagnostics to measure pulse shape and energy of all 56 beams and the focal-profile uniformity of several beams are being installed. We discuss the results of these measurements along with beam-pointing stability and image quality on target. Results from initial target acceleration experiments are also presented.

Plasma Physics Division, U.S. Naval Research Laboratory, Washington, D.C. 20375

1. R. H. Lehmberg, J. Goldhar, Fusion Tech. **11**, 532 (1987).
2. T. Lehecka *et al.*, in *Conference on Lasers and Electro-Optics*, 1994 Technical Digest Series, Vol. 8 (Optical Society of America, Washington, D.C., 1994), paper CTuQ2.
3. J. D. Sethian *et al.*, presented at this conference.

Appendix H

X-Ray Emission from Plasmas Created by Smoothed KrF Laser Irradiation

X-RAY EMISSION FROM PLASMAS CREATED BY SMOOTHED KrF LASER IRRADIATION

Y. Aglitskiy, T. Lehecka, A. Deniz, and J. Hardgrove
Science Applications International Corporation, McLean VA 22102

J. Seely, C. Brown, U. Feldman
Space Science Division, Naval Research Laboratory, Washington DC 20375

C. Pawley, K. Gerber, S. Bodner, S. Obenschain, R. Lehmberg, E. McLean, M. Pronko, J. Sethian, J. Stamper, A. Schmitt, and C. Sullivan
Plasma Physics Division, Naval Research Laboratory, Washington DC 20375

G. Holland and M. Laming
SFA Inc., 1401 McCormick Drive, Landover MD 20785

Abstract

The x-ray emission from plasmas created by the NRL Nike KrF laser was characterized using imaging and spectroscopic instruments. The laser wavelength was 1/4 μm, and the beams were smoothed by induced spatial incoherence (ISI). The targets were thin foils of CH, aluminum, titanium, and cobalt and were irradiated by laser energies in the range 100 J - 1500 J. A multilayer mirror microscope operating at an energy of 95 eV recorded images of the plasma with a spatial resolution of 2 μm. The variation of the 95 eV emission across the 800 μm focal spot was 1.3% rms. Using a curved crystal imager operating in the 1-2 keV x-ray region, the density, temperature, and opacity of aluminum plasmas were determined with a spatial resolution of 10 μm perpendicular to the target surface. The spectral line ratios indicated that the aluminum plasmas were relatively dense, cool, and optically thick near the target surface. The absolute radiation flux was determined at 95 eV and in x-ray bandpasses covering the 1-8 keV region. The electron temperature inferred from the slope of the x-ray flux versus energy data in the 5-8 keV region was 900 eV for an incident laser energy of 200 J and an intensity of $\approx 10^{13}$ W/cm^2.

26 April 1996. *Physics of Plasmas.* PACS 52.50.Jm, 52.70.La

I. INTRODUCTION

The Nike KrF laser was designed to produce uniform irradiation and acceleration of planar foils for direct-drive fusion studies. The Nike laser produces an energy up to 3 kJ of energy on target in a 4 nsec pulse at a wavelength of 0.248 μm. The excellent beam uniformity is achieved by the induced spatial incoherence (ISI) optical smoothing technique and by overlapping up to 44 beams in the focal spot.[1,2,3]

Some of the initial experiments carried out using the Nike laser have focused on characterizing the x-ray emission from the plasmas. These studies are necessary to determine the uniformity and brightness of the x-ray emission from plastic and metal foil targets. The acceleration of plastic targets will be studied in detail using the smooth Nike laser irradiation, and metal targets will be used for backlighter diagnostics.

The initial experiments that are described here determined the uniformity of the x-ray emission, the absolute x-ray flux, and the spatially-resolved plasma properties such as density, temperature, and opacity. In addition to being useful for the design of future foil acceleration and backlighter experiments, the characterization of the x-ray emission is also necessary to test the validity of computer codes that model the hydrodynamics and radiation physics of the plasmas.

The layout of the Nike target chamber and the x-ray diagnostics are shown schematically in Fig. 1. The inner diameter of the target chamber is 1.85 m. The target foils were mounted at the center of the target chamber and perpendicular to the laser axis. The incident laser energy was adjusted to control the plasma temperature with the lowest energy shots below 100 J and a maximum of 1500 J. The laser pulse duration was 4 nsec

full width at half maximum. The focal distribution was 750 μm FWHM with a flat top over 400 μm wide for all shots, and this resulted in the intensity ranging from $5x10^{12}$ to $8x10^{13}$ W/cm^2. These laser conditions are relevant to a number of direct-drive and backlighter applications.

A multilayer mirror microscope, operating at an energy of 95 eV, viewed the front side of the target foil at an angle of 40° to the laser axis. An imaging x-ray spectrometer and pinhole x-ray cameras viewed at an angle of 45° to the laser axis and from the side of the target foil at an angle of 90° to the laser axis. A survey x-ray spectrometer, with wide energy coverage, viewed the target foil at an angle of 45° to the laser axis. X-ray filter arrays were positioned to measure the x-ray flux at various locations in the target chamber. Diagnostic instruments could not be positioned near the laser axis owing to the cone angles of the incoming laser beams.

II. HIGH-RESOLUTION IMAGING AT 95 eV

Images with high spatial resolution were recorded at an energy of 95 eV (wavelength of 130 Å) by a multilayer mirror microscope. The microscope consisted of a concave primary mirror (P in Fig. 1) and a convex secondary mirror (S) in a Cassegrain type optical configuration. The primary and secondary mirrors had multilayer coatings consisting of alternating layers of molybdenum and silicon. The peak reflectance was 50% at an energy of 95 eV, and the width of the reflectance bandpass was 3 eV.[4] The mirrors were positioned in an evacuated tube that was external to the target chamber. The primary mirror was 1.65 m from the target, and this large distance prevented significant damage to the primary mirror by target debris. The primary-secondary mirror separation was 0.51 m,

and the film holder (D in Fig. 1) was 0.96 m behind the primary mirror. This resulted in a magnification factor of 2.33.

The magnified images were recorded on Kodak 101 or 104 film. The effective grain size of 101 film is approximately 10 μm, and the resolution in the target plane was 4 μm. The 104 film has an effective grain size less than 4 μm, and the resolution in the target plane was 2 μm.[4]

It is necessary to cover the entrance aperture of the film holder with a filter (F1 in Fig. 1) that is opaque to visible light. Since this filter is near the focal plane of the microscope, this filter must be protected from damage by laser light that is backscattered from the target plasma and focused by the optical system. This was accomplished by placing a second filter (F2) behind the primary mirror which covered the central hole in the primary mirror. The two filters were also designed to attenuate the extreme ultraviolet radiation with energy lower than 60 eV that was reflected by the top layers of the multilayer coating and focused by the mirrors.

The film holder was covered by a filter consisting of 2000 Å of titanium and 2000 Å of Lexan. This filter was opaque to visible light and had a transmittance of 1.5% at 95 eV. A 1 μm beryllium filter covered the hole in the primary mirror and protected the Ti/Lexan filter from damage by focused light. This filter was opaque to the laser light and had a transmittance of 20% at 95 eV.

The 95 eV images of CH, aluminum, titanium, and cobalt targets were recorded. The 95 eV emissions from the titanium and cobalt targets were quite intense, and the film exposure was farther reduced by adding a second 1 μm beryllium filter between the

primary mirror and the film holder. The image, recorded on 104 film, of a 10 μm thick cobalt foil is shown in Fig. 2. The laser energy incident on the target foil was 1500 J.

The microscope viewed the target from a location that was 22.8° from the laser axis in the horizontal plane and 33.8° below the equator of the target chamber. The resulting angle with respect to the laser axis was 40.0°. The shape of a circle on the target foil of 1 mm diameter is shown in Fig. 2 as it would appear from the microscope location. The laser beams were incident from the top left direction in Fig. 2.

The 104 film exposure was converted to flux density in units of photons/μm^2 using a film calibration function.[5] Shown in Fig. 2 is a lineout through the center of the focal spot at the location indicated by the fiducials at the edge of the image. The most intense emission originated from a region on the target foil with 800 μm diameter, and this was consistent with the expected focal spot diameter of the overlapping laser beams. A quadratic curve was fitted to the data points in the central 800 μm region using the least squares technique, and the rms deviation of the data points from the fitted curve was 0.128 photons/μm^2. A second quadratic curve was fitted to the data points at the beginning of the lineout (outside the exposed region of the image), and the rms deviation of the data points from the fitted curve was 0.038 photons/μm^2. This latter value represents the noise level of the 104 film, and the inferred rms smoothness of the data points in the focal spot region is 1.3%.

The flux incident on the primary mirror was calculated from the flux incident on the film by accounting for the reflectance of the two multilayer mirrors, the transmittance of the filters, and the illuminated area (37.7 cm^2) of the primary mirror. For the image shown

in Fig. 2, the flux incident on the primary mirror, within the 3 eV wide bandpass centered at 95 eV that was defined by the multilayer coating, was 6.9×10^{10} photons/cm^2. Assuming that the emission was approximately isotropic into the hemisphere facing the microscope, the total energy radiated in the 95 eV bandpass was 0.18 J. For comparison, the flux in the 95 eV bandpass from a 100 μm thick CH foil that was irradiated by 1400 J was 0.15 J.

III. PINHOLE IMAGING IN THE 1-5 keV ENERGY REGION

Images of CH and aluminum foils were recorded by two pinhole cameras with bandpasses in the 1-5 keV range. The images were recorded on Kodak DEF film. As shown in Fig. 1, the pinhole cameras viewed the target foils at an angle of 45° to the laser axis and from the side of the foils at an angle of 90° to the laser axis. The pinholes were fabricated by laser drilling tantalum foils that were 15 μm thick. The 15 μm tantalum foil stopped x-rays with energies below 15 keV. The pinholes had diameters of 5 μm or 10 μm, and most of the images were recorded through 10 μm pinholes. The pinholes were positioned 4.6 cm from the focal spot.

Thin K-edge filters covered the pinhole and established the energy bandpass. Well-exposed images of 40 μm CH foil targets were obtained using a filter consisting of 11 μm of aluminum with a average transmittance of 15% in a 1.2-1.5 keV bandpass. Well-exposed images of 25 μm aluminum foil targets were obtained using each of the following two filter sets (the energy bandpass and average transmittance are in parenthesis): 12 μm saran and 11 μm aluminum (2% transmittance in the bandpass 1.3-1.5 keV), and 25 μm titanium and 13 μm beryllium (20% transmittance in the bandpass 4.0-5.0 keV). Additional exposures from aluminum targets through each of the following two filters sets

were quite weak: 25 μm nickel and 13 μm beryllium (30% transmittance in the bandpass 7.3-8.3 keV), and 25 μm molybdenum and 13 μm beryllium (60% transmittance in the bandpass 12-20 keV).

Shown in Fig. 3(a) is the side view of a 25 μm aluminum foil in the 1.3-1.5 keV bandpass. The laser beams were incident from the left in Fig. 3(a), and the energy incident on the target was 200 J. The emission of the aluminum plasma backlit a nickel mesh with 35 μm wire diameter and 320 μm wire period. The image of the aluminum plasma was magnified by a factor of 7.9. The nickel mesh was closer to the pinhole and was magnified by a factor of 15. The effective wire period projected to the center of the focal spot was 610 μm. The brightest emission from the aluminum plasma originated in a region approximately 800 μm in size. The emission region near the foil surface has a concave shape that may result from the translation of the 25 μm aluminum foil during the 4 nsec laser pulse. Preliminary calculations using a hydrodynamics code predict that a 25 μm aluminum foil moves during the laser pulse, while aluminum foils with substantially higher thickness are stationary during the laser pulse.

An image in the 1.2-1.5 keV bandpass of a 40 μm CH foil is shown in Fig. 3(b). The pinhole camera was positioned in the equatorial plane of the target chamber and was 45° from the laser axis. The laser beams were incident from the left in Fig. 3(b), and the energy incident on the target was 200 J. The magnification factor was 7.9. A circle on the target foil 1 mm in diameter as viewed at an angle of 45° is superimposed in Fig. 3(b). The brightest emission originates in a region of diameter 800μm.

We conclude from the x-ray images recorded in the 1-5 keV energy range that the brightest emission originated in a plasma region of diameter 800 μm. The x-ray emission was quite smooth and featureless. These results are consistent with the 95 eV images that were recorded by the multilayer mirror microscope and discussed in Section II.

IV. SURVEY ALUMINUM X-RAY SPECTRA

The x-ray spectra from aluminum foils were recorded by a RAP crystal that was bent to a convex, cylindrical shape. The purpose of this instrument was to provide survey spectra with wide wavelength coverage of the hydrogen-like and helium-like aluminum spectral lines and the recombination continuum radiation. As shown in Fig. 1, the spectrometer viewed the targets at an angle of 45° to the laser axis. The spectra were recorded on DEF film and were time and spatially integrated. The Bragg angle at the middle of the crystal was set to 14.5° to record the first-order spectra in the 5.0-8.5 Å wavelength region which included the lines of hydrogen-like (Al^{12+}) and helium-like (Al^{11+}) ions. A typical spectrum is shown in Fig. 4(a), and some of the line identifications and wavelengths[6] are listed in Table 1. The laser energy incident on the 25 μm aluminum foil was 200 J.

The film exposure was converted to flux incident on the film using a DEF film calibration[7] and accounting for the transmittance of the filters (35 μm saran and 25 μm beryllium). The electron temperature was determined from the slope of the recombination continuum in the 5.1-5.6 Å region, and the inferred value was 360 eV. The electron density was determined from the Inglis-Teller expression for the merging of the Al^{11+} Rydberg lines with the continuum, and the inferred value was $1x10^{21}$ cm^{-3}. The

temperatures and densities derived from the survey spectra were space and time integrated values.

Shown in Fig. 4(b) is the aluminum spectrum calculated using the FLY computer program[8] for an electron temperature of 360 eV and an electron density of $1x10^{21}$ cm^{-3}. The overall peak line intensities of the experimental and calculated spectra are in good agreement. The calculated lines are narrower than observed because the instrumental broadening was not included in the calculation.

V. SPATIALLY-RESOLVED ALUMINUM X-RAY SPECTRA

X-ray emissions from aluminum targets were recorded by an imaging spectrometer. The purpose of this instrument was to provide spectra with excellent spatial resolution. The spectrometer had a mica crystal that was bent to a concave spherical shape with a radius of curvature of 10 cm. A flat piece of DEF film was positioned tangent to the Rowland circle, and the plasma x-ray source was outside the Rowland circle. As discussed in Ref. 9, images were recorded with spatial resolution in the direction perpendicular to the dispersion direction. The Bragg angle at the center of the crystal was set to 49° to record images of the emission in the 6.9-8.3 Å (1.5-1.8 keV) region using the second diffraction order of the mica crystal. The magnification factor was determined by the Bragg angle and was 0.2 in this case. The mica crystal also reflected in other diffraction orders, and these orders were partially attenuated by a 12 μm saran filter. In addition, a 6.4 μm beryllium filter covered the film. For future experiments, this instrument will be modified to record two dimensional monochromatic images using the techniques discussed in Ref. 9.

Shown in Fig. 5(b) is a portion of the image in the 7.6-7.9 Å region that includes the Al^{11+} resonance line w, the intercombination line y, and the dielectronic satellite line j (see Table 1 and Fig. 4). The instrument viewed the aluminum foil from the side as shown by the target configuration (not to scale) in Fig. 5(a). The thickness of the aluminum foil was 50 μm, and hydrodynamic modeling indicated that foils of this thickness moved less than 30 μm during the 4 nsec laser pulse. The spectral image in Fig. 5(b) is spatially resolved in the horizontal direction which corresponds to the direction perpendicular to the foil. The spatial scale shown in Fig. 5(b) applies to the spectral image and not to the target configuration shown in Fig. 5(a), which is greatly enlarged.

As shown in Fig. 5(a), a nylon fiber with diameter 150 μm was positioned on the side of the foil nearest the spectrometer and perpendicular to the line of sight of the spectrometer. The purpose of the fiber was to provide a fiducial for the surface of the target foil. The fiber was backlit by the aluminum plasma. X-ray radiation passed through the thinnest section of the fiber nearest the foil surface and provided on the film a fiducial for the surface. Shown in Fig. 6 are the continuum intensities measured in the direction perpendicular to the surface on shots (a) with the fiber and (b) without the fiber. For curve (a), the data points are 10 μm apart in the target plane, and the shadow of the 150 μm fiber is clearly visible. After accounting for the change in the thickness of the fiber with distance from the foil surface and for the 2 μm DEF film grain size, the change in the intensity of the x-ray flux through the fiber as a function of distance from the surface indicates a spatial resolution of 10 μm in the target plane. The surface position for curve (b), recorded without the fiber, was established by aligning curve (b) with curve (a) in the

spatial region above the fiber (at distances from the surface greater than 200 μm). Then lineouts of the spectrum recorded without the fiber were taken at the spatial positions indicated by the data points on curve (b) in Fig. 6. Spectral scans made at distances of 50 μm, 200 μm, 600 μm, and 1000 μm above the surface are shown in Fig. 7.

Description of the Spectra

In the spectra of Fig. 7, the changes in the relative line intensities with distance above the target surface indicate that the plasma properties, such as temperature, density, and opacity, are rapidly changing with distance. The Al^{11+} intercombination line y is quite intense and actually exceeds the intensity of the Al^{11+} resonance line w in the region near the surface. The dielectronic line J is more intense than the Al^{12+} resonance lines Ly1 and Ly2 near the surface. As a function of distance from the surface, the Al^{12+} lines Ly1 and Ly2 increase in intensity relative to the Al^{11+} lines w and y near the surface and decrease relative to the Al^{11+} lines far from the surface. Near the surface and far from the surface, the intensity of Ly1 is lower than Ly2 which is the opposite of what is usually observed.

In Fig. 7, the only spectral line that is intense over the entire range of distances above the target surface is the Al^{11+} intercombination line y. This line is isolated from other strong lines and has lower opacity than the resonance lines owing to its small oscillator strength. The spectral width of line y was measured as a function of distance above the surface. The electron density was inferred from the line widths that were calculated using the FLY computer program. After removing the instrumental width that was determined from the line widths observed far from the surface, the electron density was found to have a peak value of $1x10^{22}$ cm^{-3} at 100 μm above the surface. This electron density may be

compared to the critical electron density of the 0.248 μm KrF laser radiation (1.6×10^{22} cm^{-3}).

The absence of the Al^{12+} Ly1 and Ly2 lines near the target surface implies a relatively low temperature near the surface. The line ratios Ly1/w and Ly2/w have peak values in the 700-800 μm region and may indicate a maximal temperature, although this needs to be confirmed by atomic and hydrodynamic code calculations.

Some of the spectral features in the 7.3-7.7 Å region that appear within 200 μm of the target surface (see Figs. 7(a) and (b)) may be inner-shell transitions in Al^{10+} and lower charge states. This is under investigation using atomic code calculations.

Interpretation of the Spectra

The low w/y line ratio within 100 μm of the target surface resulted from the opacity of the resonance line w in the dense, relatively cool plasma near the surface. The optical depth was estimated using the escape factor approximation[10] and the electron density 1×10^{22} cm^{-3}. Assuming that half of the ions were in the Al^{11+} ground state, the optical depth across the 400 μm half-width of the focal spot was estimated to be of order 100 for the resonance line w. Additional cool plasma surrounding the focal spot, resulting from secondary heating by radiation from the focal spot, would increase the opacity. In any case, for the viewing direction parallel to the foil surface, the optical depth of the resonance line w was much greater than for the intercombination line y.

Based on the statistical weights of the $2p_{1/2}$ and $2p_{3/2}$ levels, the Ly2/Ly1 intensity ratio is expected to be 0.5. In the spectra of Fig. 7, the Ly2/Ly1 intensity ratio exceeds unity when first observed at 200 μm from the surface and falls to unity in the 300-500 μm

region. Above 600 μm, the Ly2/L1 ratio increases to 2.2 at a distance of 1600 μm from the surface. The unexpectedly large Ly2/Ly1 ratio near the target may result from the contribution of dielectronic satellites with n>2 which would add to the Ly2 intensity to a greater degree than to the Ly1 intensity.[11] The large Ly2/Ly1 ratio far from the surface results from recombination in the expanding plasma. An Ly2/Ly1 ratio up to 1.2 may be explained by radiative recombination, with the $2s_{1/2}$ population being transferred to the $2p_{1/2}$ level by collisions with other ions in the plasma. The estimated rate coefficient is 10^{-9}-10^{-8} cm^3s^{-1}. The $2s_{1/2}$ population may also be transferred to the $2p_{3/2}$ level, but the rate coefficient for this process is much lower. Ratios up to about 0.8 may also be produced by electron collisional excitation, with the same population redistribution by ion impacts as above.

The Ly2/Ly1 ratios greater than 1.3 may result from polarized excitation of the atomic levels. If directed electrons were present in the plasma, the Ly1 line could be polarized (Ly2 is unpolarized in the absence of spin polarized electrons). At a Bragg angle of 49°, $\cos^2(2\theta)=0.02$ of the radiation that is incident on the crystal with its electric field vector in the plane of dispersion will be reflected. Hence the anomalous Ly2/Ly1 intensity ratio could be due to some of the Ly1 radiation being blocked by the crystal. For this to occur, the Ly1 radiation must be polarized with its electric field vector in the plane of the target foil, that is, at 90° to the laser axis. Such polarization has previously been observed by Kieffer et al.[12] in the Al^{11+} w line excited by a 1 psec pulse of 1.053 μm laser radiation. Scaling the polarization they observed to our case of an H-like ion, we would expect a 15% reduction in the intensity of Ly1 due to polarization. Our experimental conditions are

considerably different from those of Ref. 12, and a more detailed investigation of such polarization effects is warranted.

Rydberg Spectra

The Al^{11+} Rydberg lines from the levels n=5 through n=11 were recorded while viewing the edge of the aluminum foil in the third diffraction order of the mica crystal at a Bragg angle of 66°. For this Bragg angle, the magnification factor was 0.65. A spatially-resolved image of the emission in the 5.95-6.20 Å region is shown in Fig. 8. The continuum between the lines was quite intense near the target. The maximum intensity of the lines above the continuum occurred at a distance of 400 μm from the target surface. The lines were broader near the surface than far above the surface. The electron density was derived from the width of the line from the n=8 level as described above for the line y. The electron density was $1x10^{20}$ cm^{-3} at the surface, increased to a peak value of $2x10^{20}$ cm^{-3} at a distance of 200 μm from the surface, and decreased to less than $5x10^{18}$ cm^{-3} at 800 μm. Electron densities in this range are consistent with the observed merging of the Rydberg lines into a continuum at the n=9 level near the target surface and at the n=11 level 600 μm from the surface. Similar techniques were used by Dyakin et al.[13]

These electron densities are lower than those derived from the Al^{11+} line y, and this may indicate that the emission in the Rydberg lines is characteristic of the lower-density recombining plasma at a later time. By comparing Figs. 8 and 5(b), the Rydberg lines extend farther from the target surface and are therefore more intense in the blowoff plasmas. These conclusions are preliminary, and the time dependence of the emission

recorded by the imaging spectrometer will be studied in the future using a time-resolving framing camera.

Spectra Recorded at 45 Degrees to the Laser Axis

Spectral images were recorded while viewing the aluminum foil at an angle of 45° to the laser axis. These images were spatially resolved in the direction perpendicular to the line of sight of the instrument. As shown in Figs. 5-7, the most intense continuum emission comes from the plasma within 200 μm of the foil surface, while the images discussed in Sections II and III indicate that the width of the emission in the plane of the foil is 800 μm. Thus for the 45° view, the transverse width of the strip of continuum on the film essentially corresponds to the 800 μm width of the focal spot viewed at an angle or 45°. The emission recorded on the target side of the continuum, analogous to the left side of the continuum in Fig. 5, is from the plasma near the surface of the foil at the edge of the focal spot closest to the instrument. The emission recorded on the laser side of the continuum, analogous to the right side of the continuum in Fig. 5, is from the blowoff plasma.

Shown in Fig. 9 are the spectra from (a) the edge of the focal spot closest to the instrument, (b) the center of the focal spot, and (c) the blowoff plasma. The laser energy incident on the aluminum foil was 100 J. These spectra may be compared to the spatially-integrated spectrum of Fig. 4(a) recorded at 45° to the foil surface and with the spatially-resolved spectra of Fig. 7 recorded parallel to the foil surface. In contrast to the side-view spectra near the target (Fig. 7(a)), the resonance lines w and y in Fig. 9 are more intense than the nearby satellite and intercombination lines and are in qualitative agreement with

the spatially-integrated spectrum of Fig. 4(a). This is additional evidence that the resonance lines w and Ly are optically thick in the side-view spectra near the target (Fig. 7).

The preliminary analysis of the experimental data presented here leads to the following conclusions. The aluminum plasma is relatively dense, cool, and optically thick (for side viewing) within 200 μm of the target surface. In the region 100-300 μm above the surface, the density is decreasing and the temperature is increasing. The relative emission from the highest charge state (Al^{+12}) peaks in the 700-800 μm region, perhaps implying a maximal temperature, and decreases at greater distances. Beyond 1000 μm from the surface, the plasma is strongly recombining.

A complete interpretation of the aluminum spectra will require detailed modeling of the hydrodynamics, atomic kinetics, and opacity of the plasmas. A simple 1-D code has been used to provide an initial estimation of the electron density and temperature for comparison with the experimental results. This code used cold opacities for the radiation transport and local thermodynamic equilibrium (LTE). These approximations are probably the least valid in the low density blow-off region of the plasma and give more accurate results within 500 μm of the target surface. Figure 10 shows the calculated results at 3 ns after the rise of the laser pulse, just before the pulse begins to fall. At this time, the temperature and density scale lengths are at their maximum. These results are in good agreement with the observations previously discussed. The initial front surface of the foil is at 0 on the plot. The first 100 μm of plasma is cold and dense with an electron temperature of about 50 eV and an electron density of order 10^{22} cm^{-3}. For greater

distance from the target, the temperature rapidly rises and has a plateau near 400 μm. Beyond this distance from the target, the 1-D approximation is probably no longer valid owing to the more spherical plasma expansion, and in addition the LTE approximation is not expected to give a reliable temperature. However, it is expected that the temperature is maximal in the plasma region approximately 400 μm from the target surface. Beyond 400 μm, the temperature that is inferred from line ratios can be in error owing to "freezing-in" effects in the rapidly expanding plasma.

VI. SPATIALLY-RESOLVED TITANIUM X-RAY SPECTRA

Spatially-resolved spectra from a titanium foil, irradiated by 1500 J of laser energy, were recorded using the imaging spectrometer with the spherically bent mica crystal. The Bragg angle was set to 66° to record the 2.40-2.65 Å (4.7-5.2 keV) emission in the seventh diffraction order. This wavelength region included the helium-like Ti^{20+} resonance line w, the intercombination line y, and the nearby dielectronic satellites. The wavelengths of these titanium lines are listed in Table 1.

The filter set consisted of the following materials: 47 μm saran, 16 μm polypropylene, and 13 μm beryllium. The filter set had an average transmittance of 10% in the 2.40-2.65 Å wavelength region but also had comparable transmittance in the 4.5-6.0 Å region above the Cl L absorption edge of the saran filter. The DEF film had emulsion on both sides of its polyester base, and the emulsion on the side facing the crystal was saturated by the intense 4.5-6.0 Å (2.1-2.8 keV) continuum. The emulsion facing away from the crystal showed intense lines and a weak continuum. Thus the polyester base, with a thickness of 185 μm, served as an additional filter that stopped the 4.5-6.0 Å continuum

which saturated the front emulsion. The average transmittance of the filter set and the polyester base was 3% in the 2.40-2.65 Å wavelength region. The front emulsion was removed from the polyester base leaving only the rear-side image.

The imaging spectrometer viewed from the side of the titanium foil at an angle of 90° to the laser axis. The spectra observed at distances of 10 μm, 100 μm, 150 and 250 μm from the surface of the foil are shown in Fig. 11. The lines are identified in Fig. 11(c) and in Table 1.

As a function of distance from the target surface, the intensity of the lines increased to peak values at a distance of 150 μm and then decreased for greater distances from the surface. The titanium spectra were quite weak beyond 300 μm from the surface. By comparison, the aluminum spectra extended to 1600 μm from the surface. The smaller extent of the titanium spectra may have resulted from the higher radiative cooling rate in the high-Z titanium plasma. In addition, the titanium resonance line w is more intense than the intercombination line y throughout the titanium plasma. This indicates that the titanium plasma was not as optically thick as the aluminum plasmas, and radiation could more readily escape from the titanium plasma. Using the escape factor approximation, the estimated optical depth of the Ti^{20+} line w was of order unity.

VII. HARD X-RAY EMISSION

The x-ray fluxes from aluminum targets were determined using an array of filters with bandpasses in the 5.0-8.3 keV energy range. As shown in Fig. 1, the filter array was positioned at an angle of 45° to the laser axis and 4.6 cm from the target. The K

absorption edges of the filter materials defined the bandpasses. Listed in Table 2 are the filter materials, the K edge energies, and the transmittances at the K edge energies.

For each energy bandpass, the flux incident on the filter was determined from the DEF film calibration curve and the filter transmittance. Listed in Table 2 is the flux incident on the filters from a 25 μm aluminum foil irradiated by 200 J of laser energy ($I \approx 10^{13}$ W/cm^2). X-rays with energies as high as 8.3 keV were detected. The slope of the flux versus energy curve indicates a temperature of 900 eV. The electron temperature derived from the hard x-ray flux is expected to peak during the laser pulse and is higher than the temperature (360 eV) that was derived from the recombination continuum in Section IV. Listed in the last column in Table 2 is the energy emitted from the target within the filter bandpass based on the assumption that the emission was approximately isotropic into the hemisphere facing the filter array. The emitted energy within the x-ray filter bandpasses decreased with increasing photon energy and was much smaller than the energy within the 95 eV bandpass discussed in Section II.

VIII. CONCLUSIONS

The primary purpose of the work reported here was to characterize the smoothness and extent of the x-ray emission, determine the time-integrated x-ray flux, and infer the plasma properties such as density, temperature, and opacity from the spectral line ratios and widths. Much of the work was directed toward developing suitable backlighter conditions using a portion of the available Nike laser energy. Future work will implement time-resolving framing and streak cameras and multi-channel fast photodiode detectors.

The images from the multilayer mirror microscope operating at an energy of 95 eV and from the pinhole cameras operating in the 1-5 keV range indicated focal spot diameters of 800μm, comparable to the 750 μm FWHM of the laser focal profile. The 95 eV images indicated that the rms value of the smoothness in the focal spot region was 1.3%. The absolute radiation flux in the 95 eV bandpass was 0.18 mJ from a cobalt target irradiated by 1500 J of laser energy and 0.15 J from a CH target irradiated by 1400 J.

The electron density in aluminum plasmas was determined from x-ray spectra with spatial resolution of 10 μm perpendicular to the target surface. The peak electron density measured from Stark broadening was $1x10^{22}$ cm^{-3} at a distance of 100 μm from the target surface. The spectral line ratios indicated that the aluminum plasma near the surface was relatively dense, cool, and optically thick (for side viewing). Compared to the aluminum plasmas, the titanium plasma had a smaller extent perpendicular to the surface, lower opacity, and higher radiative cooling rate.

The absolute fluxes in bandpasses in the x-ray energy range 5-8 keV were determined. For aluminum targets irradiated by 200 J of laser energy, the electron temperature inferred from the slope of the x-ray flux versus energy data was 900 eV.

In the spatially-resolved spectra recorded by viewing parallel to the target surface, the spectral line ratios indicated that the optical depths of the aluminum resonance lines were quite large. This resulted from the relatively high density and low temperature near the surface, the rather large diameter (800 μm) of the focal spot, and the excellent uniformity of the laser irradiation. The large, uniform plasmas produced by the Nike laser have unique properties that will be quite useful for future opacity studies.

ACKNOWLEDGMENTS

This work was supported by the U. S. Department of Energy. The construction and implementation of the Nike laser were carried out by a large team of talented engineers and technicians. The spherically bent mica crystal that was used to record the spatially-resolved spectra was provided by A. Ya. Faenov.

REFERENCES

[1]R. H. Lehmberg and S. P. Obenschain, Opt. Commun. **46**, 27 (1983) and R. H. Lehmberg and J. Goldhar, Fusion Technology **11**, 532 (1987).

[2]T. Lehecka, R. H. Lehmberg, A. V. Deniz, K. A. Gerber, S. P. Obenschain, C. J. Pawley, M. S. Pronko, and C. A. Sullivan, Optics Commun. **46**, 485 (1995).

[3]S. P. Obenschain, S. E. Bodner, D. Colombant, K. Gerber, R. H. Lehmberg, E. A. McLean, A. N. Mostovych, M. S. Pronko, C. J. Pawley, A. J. Schmitt, J. D. Sethain, V. Serlin, J. A. Stamper, C. A. Sullivan, J. P. Dahlburg, J. H. Gardner, Y. Chan, A. V. Deniz, J. Hardgrove, T. Lehecka, and M. Klapisch, Phys. Plasmas **3** (1996), in press.

[4]J. F. Seely, G. E. Holland, and J. V. Giasson, Appl. Opt. **32**, 6294 (1993).

[5]L. Dwivedi, G. J. Tallents, I. C. E. Turcu, and P. Trenda, in Central Laser Facility Annual Report, Rutherford Laboratory, p. 231 (1994).

[6]E. V. Aglitskii, V. A. Boiko, S. M. Zakharov, S. A. Pikuz, and A. Ya. Faenov, Sov. J. Quant. Electron. **4**, 500 (1974).

[7]B. L. Henke, F. G. Fujiwara, M. A. Tester, C. H. Dittmore, and M. A. Palmer, J. Opt. Soc. Am. B **1**, 828 (1984).

[8]R. W. Lee, *private communication.*

[9]A. Ya. Faenov, S. A. Pikuz, and A. I. Erko, Physica Scripta **50**, 333 (1994), and Dyakin *et al.*, Physica Scripta **52**, 201 (1995).

[10]H. R. Griem, *Spectral Line Broadening by Plasmas*, Academic Press, New York (1974).

[11]J. F. Seely, At. Data Nucl. Data Tables **26**, 137 (1981).

[12]J. C. Kieffer et al., Phys. Rev. E **48**, 4648 (1993).

[13]V. M. Dyakin, A. Ya. Faenov, A. I. Magunov, J. Makowski, P. Parys, T. A. Pikuz, T. Pisarczyk, I. Yu. Skobelev, J. Wolowski, and E. Woryna, Phys. Scripta **52**, 201 (1995).

Table 1. Transitions and wavelengths.

Designation	Transition	Wavelength (Å)	
		Al	Ti
Ly1	$1s\ ^2S_{1/2}$ - $2p\ ^2P_{3/2}$	7.1703	
Ly2	$1s\ ^2S_{1/2}$ - $2p\ ^2P_{1/2}$	7.1759	
J	$1s2p\ ^1P_1$ - $2p^2\ ^1D_2$	7.2745	
w	$1s^2\ ^1S_0$ - $1s2p\ ^1P_1$	7.7565	2.6101
d	$1s^23p\ ^2P_{3/2}$ - $1s2p3p\ ^2D_{5/2}$	7.7735	2.6141
y	$1s^2\ ^1S_0$ - $1s2p\ ^3P_1$	7.805	2.6229
k	$1s^22p\ ^2P_{1/2}$ - $1s2p^2\ ^2D_{3/2}$	7.8709	2.6319
j	$1s^22p\ ^2P_{3/2}$ - $1s2p^2\ ^2D_{5/2}$	7.8752	

Table 2. Bandpass filters, K edge energy, transmittance, and x-ray flux.

Number	Filter	Energy (keV)	T (%)	Flux (ph/μm^2)	Energy (mJ)
1	50 μm Ti + 13 μm Al	5.0	6.2	27.8	4.8
2	50 μm V + 13 μm Al	5.5	5.6	5.8	1.2
3	25 μm Co + 13 μm Al	7.7	27.0	1.0	0.06
4	25 μm Ni + 13 μm Al	8.3	29.0	0.52	0.03

FIGURE CAPTIONS

Fig. 1. Schematic of the Nike target chamber and the x-ray imaging and spectroscopic diagnostic instruments.

Fig. 2. (Top) The image of a cobalt foil recorded by the multilayer mirror microscope at an energy of 95 eV. Superimposed on the image is a circle on the target foil of 1 mm diameter as it would appear from the microscope viewing position. (Bottom) The lineout through the image at the location of the fiducials shown at the edge of the image.

Fig. 3. (a) The x-ray image of an aluminum foil viewed from the side by a pinhole camera. The plasma backlit a nickel mesh with a wire spacing of 610 μm when projected to the plane of the plasma. (b) The x-ray image of a CH foil recorded by a pinhole camera. Superimposed on the image is a circle on the target foil of 1 mm diameter as it would appear from the camera viewing position at an angle of 45° to the laser axis. In both images, the laser beams were incident from the left.

Fig. 4 (a) The spatially-integrated spectrum from an aluminum foil recorded by the cylindrically bent RAP crystal that viewed at an angle of 45° to the laser axis. (b) The calculated spectrum for an electron temperature of 360 eV and an electron density of $1x10^{21}$ cm^{-3}.

Fig. 5. (a) The target configuration showing the aluminum foil and the nylon fiber that was backlit by the aluminum plasma. (b) The spatially-resolved spectral image from the aluminum foil that was recorded by the spherically bent mica crystal viewing parallel to the surface of the foil. The spatial scale refers to the image, and the target configuration is greatly enlarged.

Fig. 6. The continuum intensity as a function of the distance from the foil surface for the two cases of (a) with the nylon fiber and (b) without the nylon fiber.

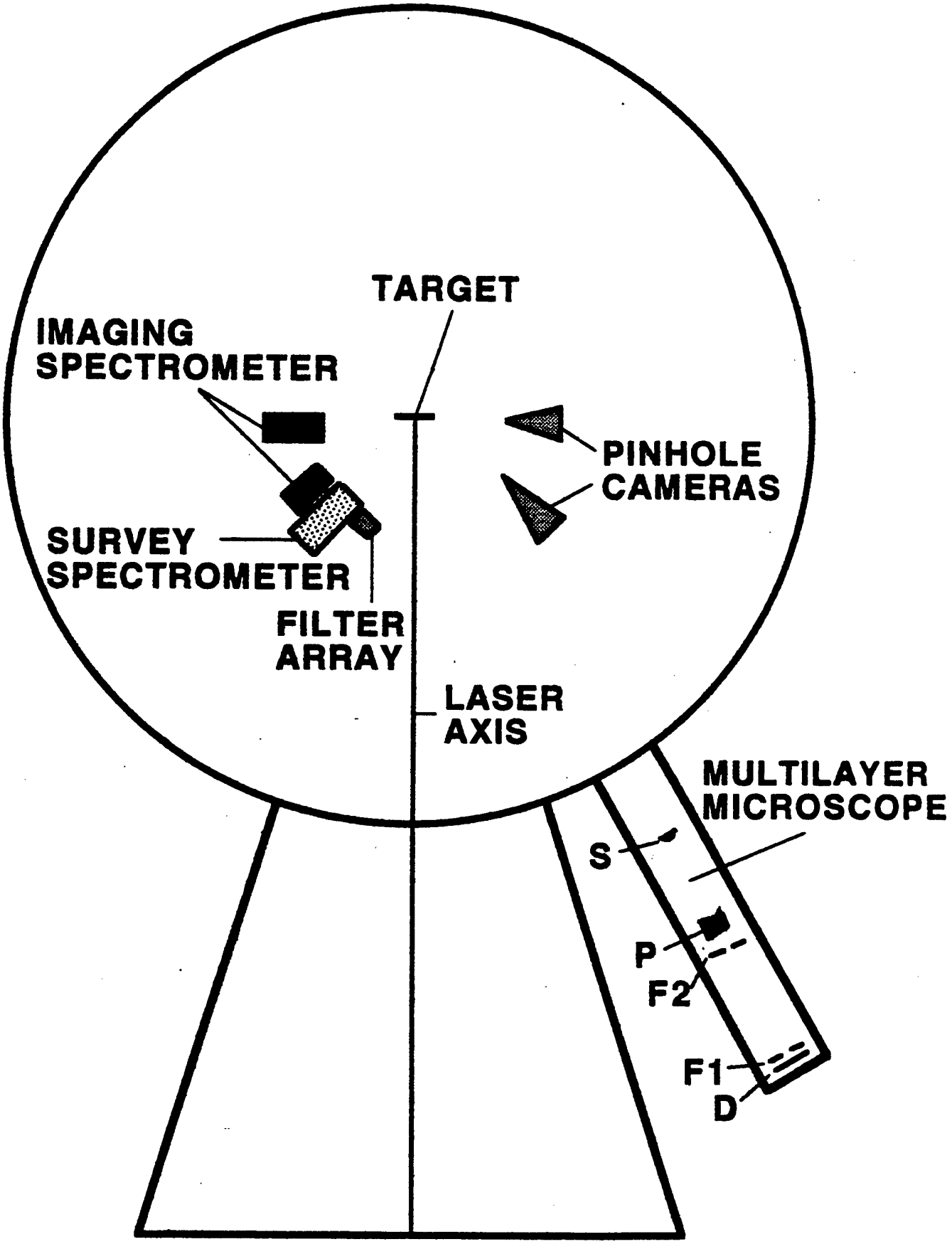
Fig. 7. The spectra from an aluminum foil at distances from the foil surface of (a) 50 μm, (b) 200 μm, (c) 600 μm, and (d) 1000 μm.

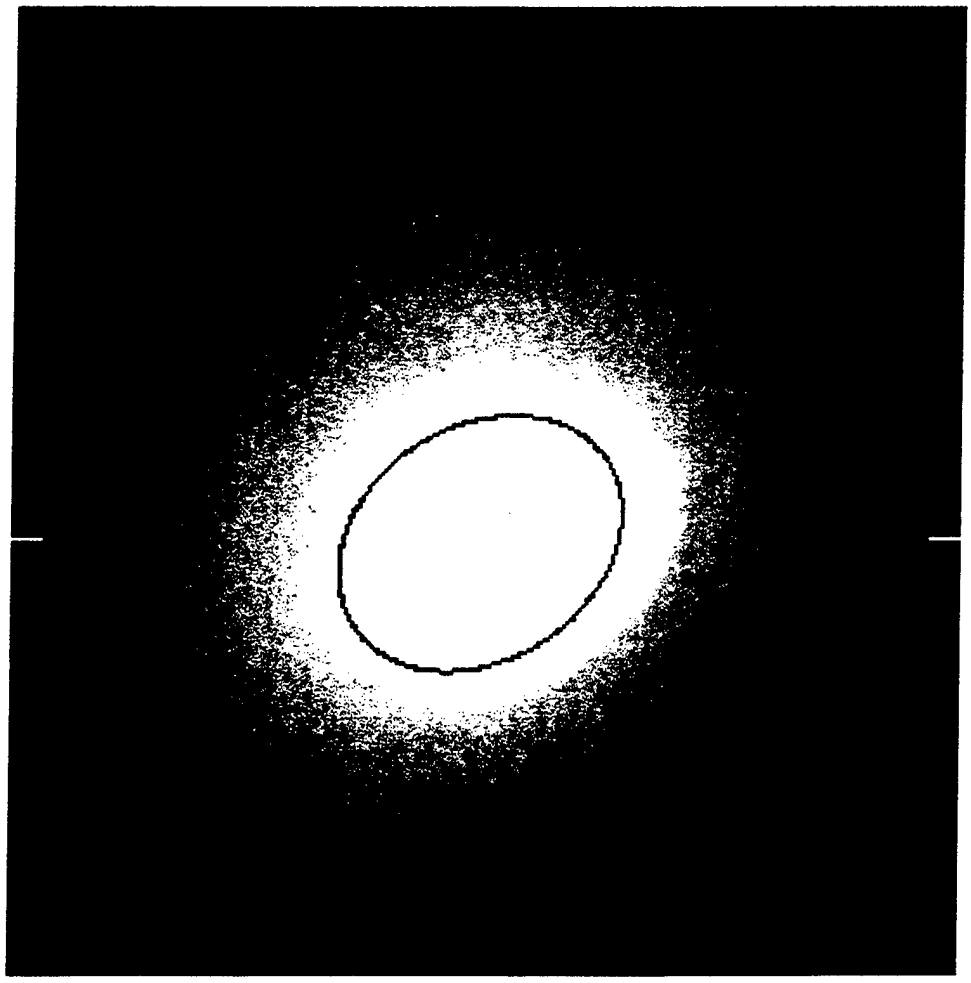
Fig. 8. The Rydberg lines emitted from an aluminum foil observed from the side of the foil by the spherical mica crystal.

Fig. 9. The aluminum spectra recorded by the spherical mica crystal viewing at an angle of 45° to the laser axis. The spectra are from (a) the plasma near the edge of the focal spot nearest the spectrometer, (b) the middle of the focal spot, and (c) the blowoff plasma.

Fig. 10. The 1-D code results at one time step near the end of the flat-top portion of the laser pulse. The electron temperature scale is to the right, and the electron density scale is to the left. Zero on the horizontal axis is the original position of the foil surface.

Fig. 11. The spatially-resolved spectra from a titanium plasma recorded by the spherical mica crystal viewing parallel to the foil surface. The spectra are from distances above the target surface of (a) 10 μm, (b) 100 μm, (c) 150 μm, and (d) 250 μm.


TARGET
IMAGING SPECTROMETER
PINHOLE CAMERAS
SURVEY SPECTROMETER
FILTER ARRAY
LASER AXIS
MULTILAYER MICROSCOPE
S
P
F2
F1
D



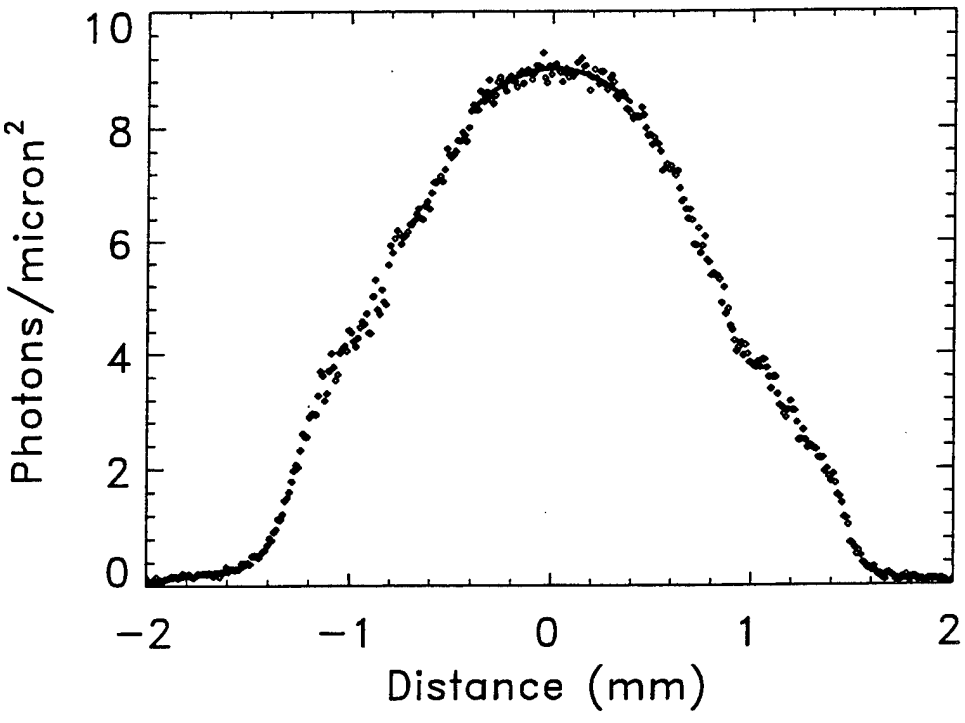
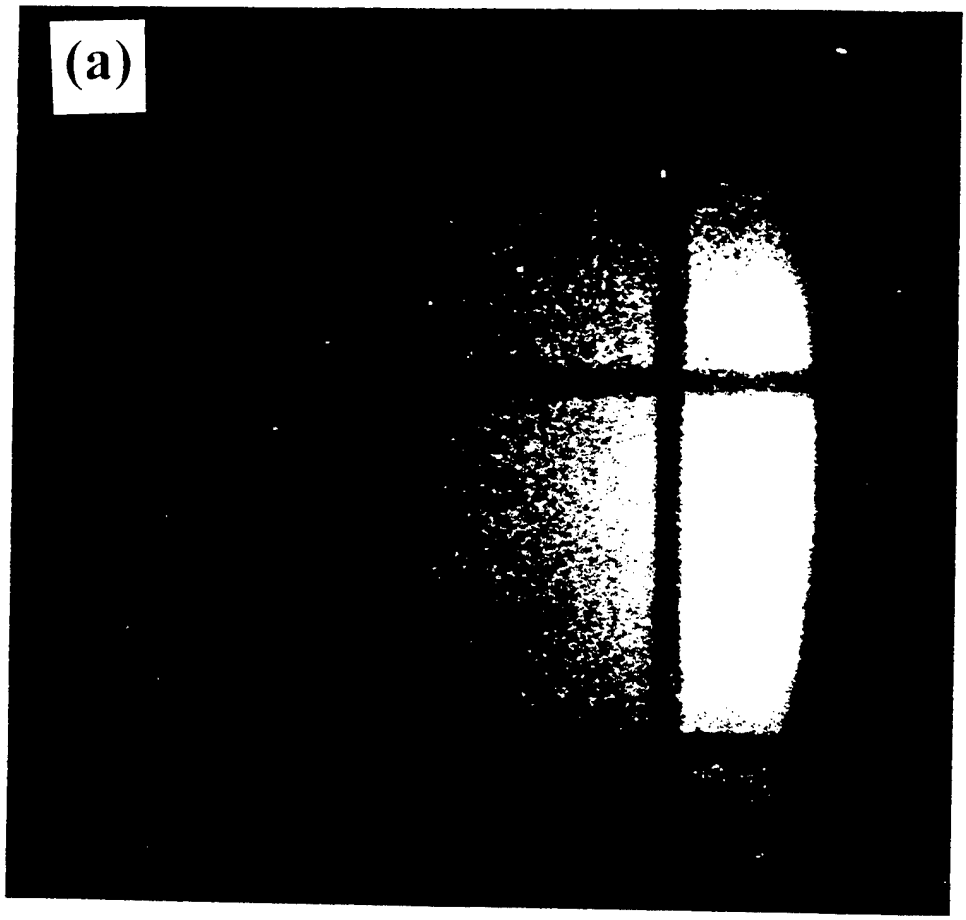


Figure 2



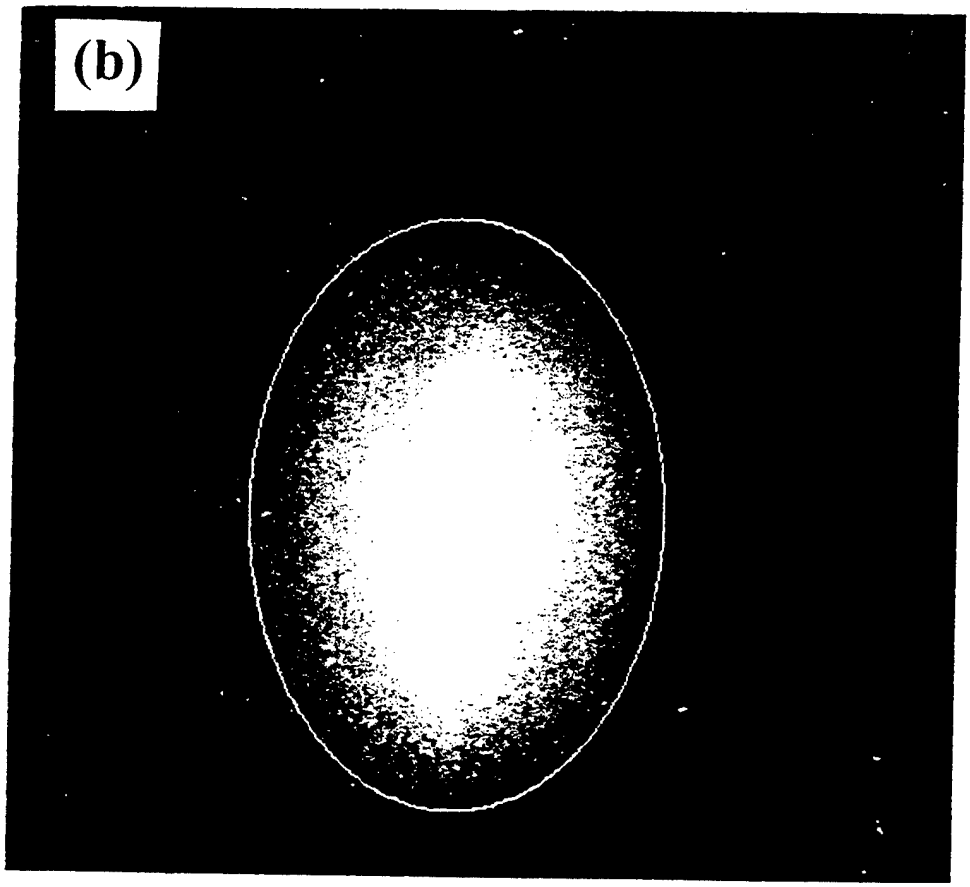


Figure 3

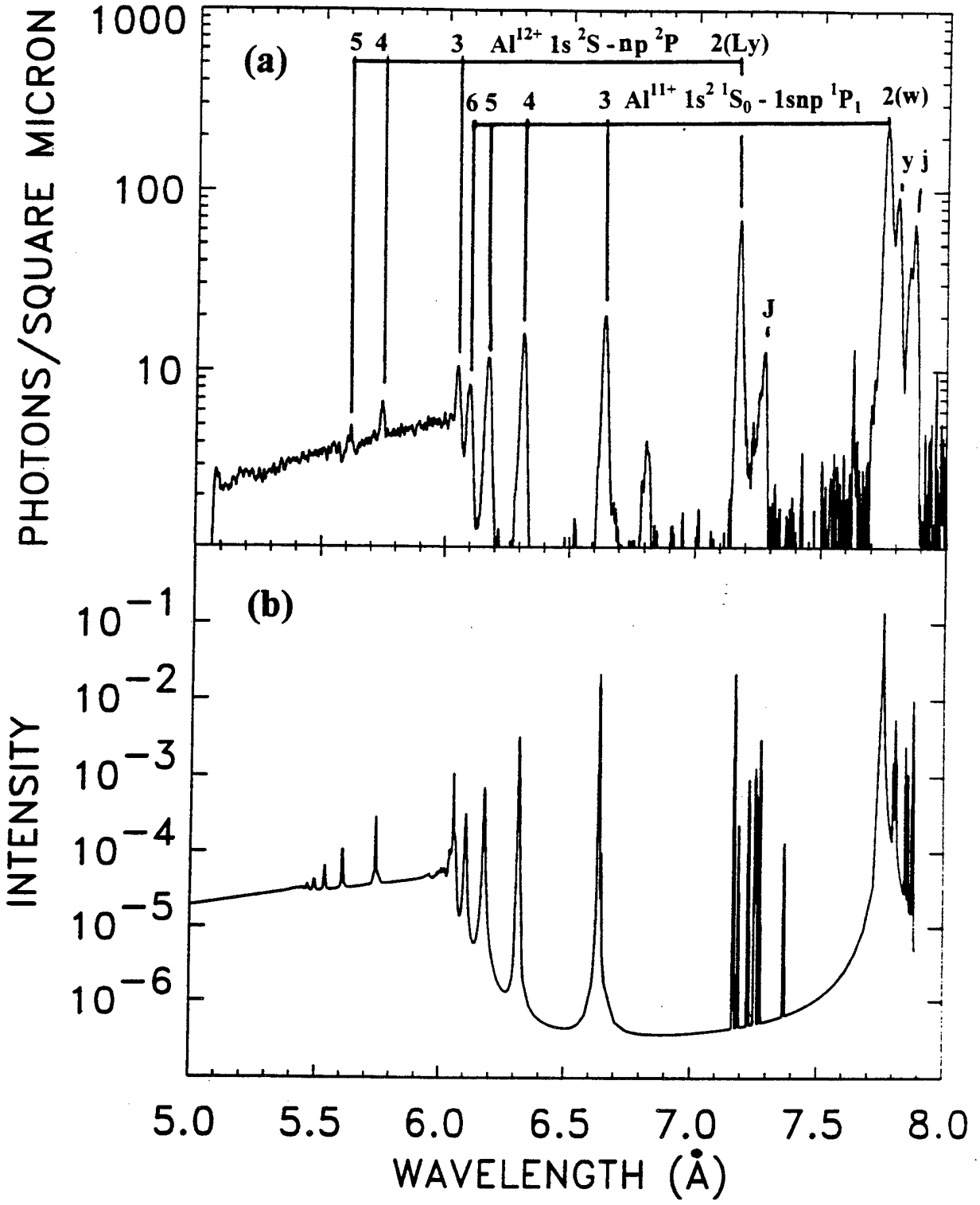

(a)
(b)
5 4 3 Al12+ 1s 2S - np 2P 2(Ly)
6 5 4 3 Al11+ 1s2 1S0 - 1snp 1P1 2(w)
y j
J
PHOTONS/SQUARE MICRON
INTENSITY
WAVELENGTH (Å)
1000
100
10
10−1
10−2
10−3
10−4
10−5
10−6
5.0
5.5
6.0
6.5
7.0
7.5
8.0

Figure 4

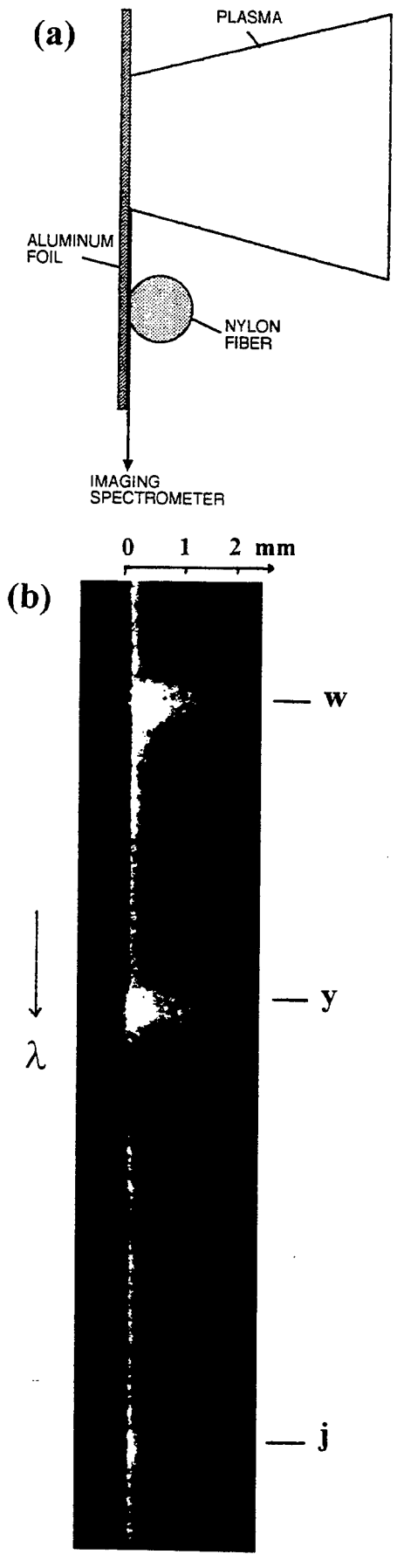

(a)
PLASMA
ALUMINUM FOIL
NYLON FIBER
IMAGING SPECTROMETER
0 1 2 mm
(b)
w
y
λ
j

Figure 5

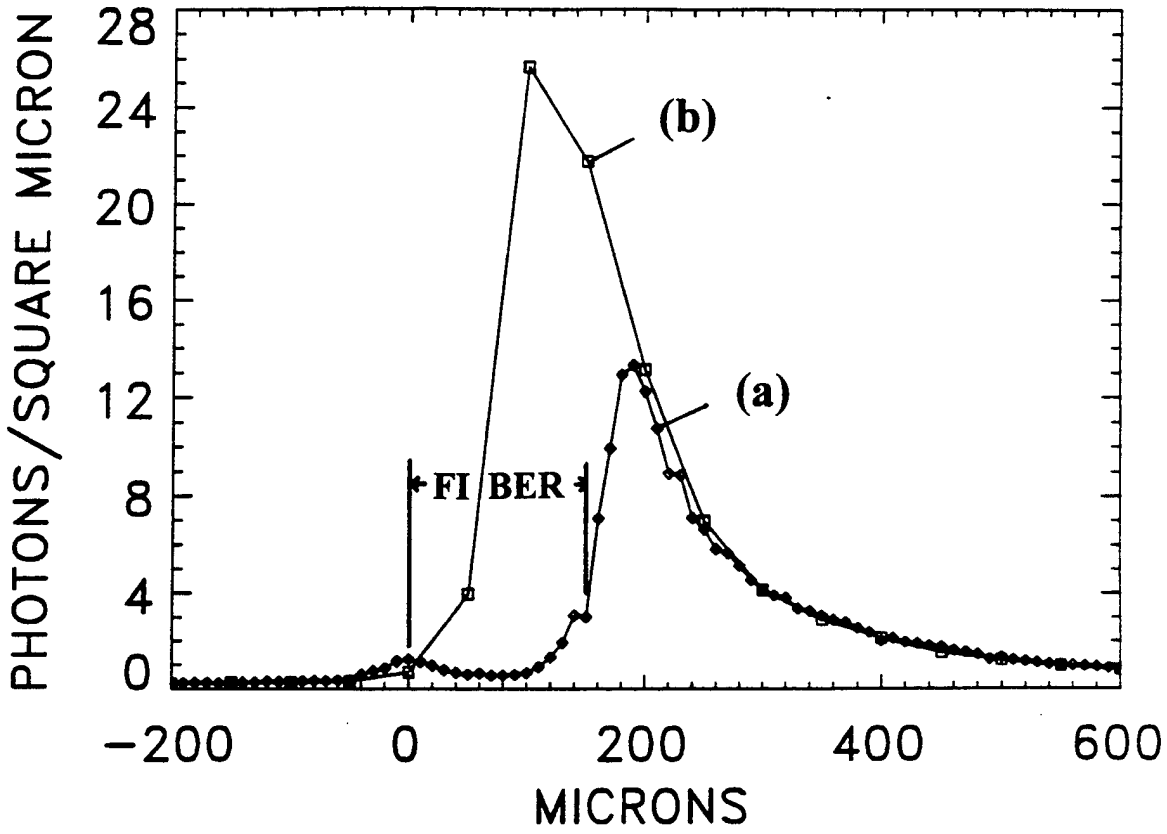

PHOTONS/SQUARE MICRON
28
24
20
16
12
8
4
0
(b)
(a)
FIBER
−200
0
200
400
600
MICRONS

Figure 6

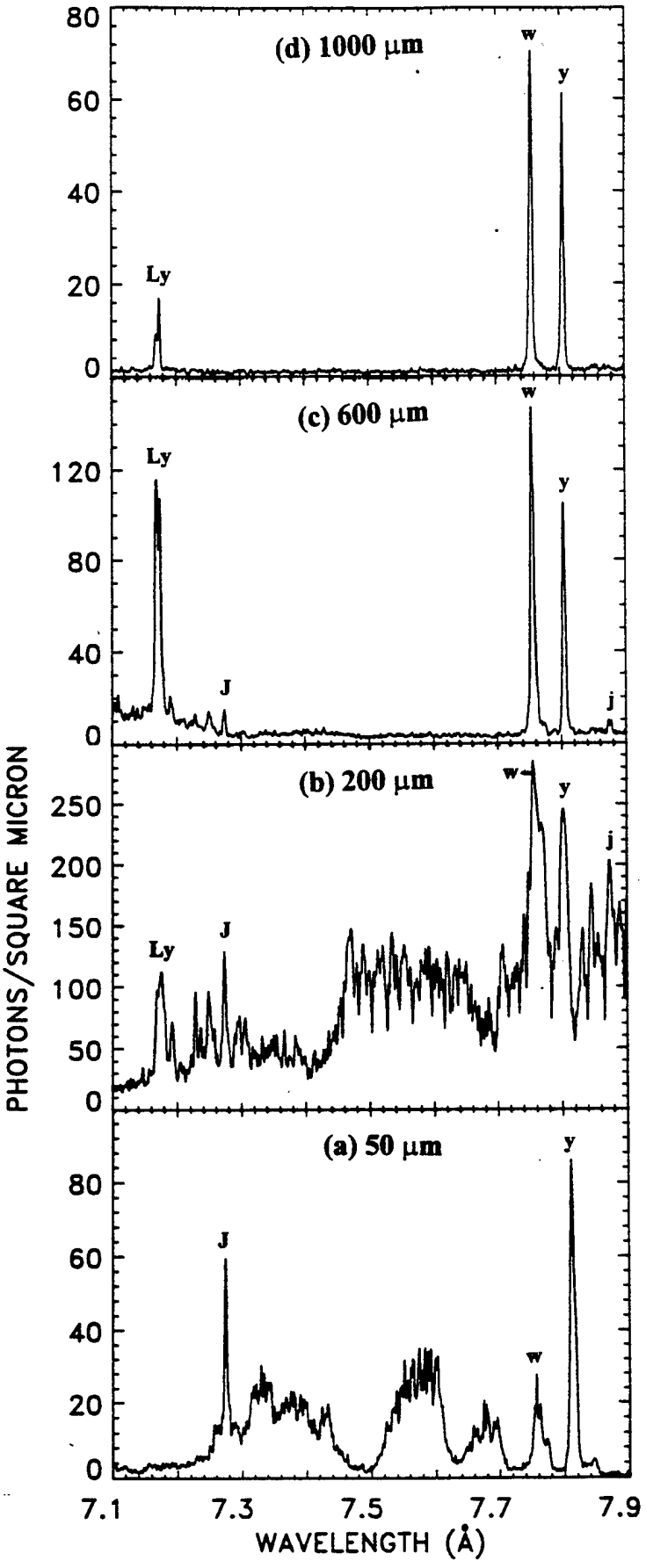

(d) 1000 μm
(c) 600 μm
(b) 200 μm
(a) 50 μm
Ly
J
w
y
j
PHOTONS/SQUARE MICRON
WAVELENGTH (Å)
7.1
7.3
7.5
7.7
7.9

Figure 7

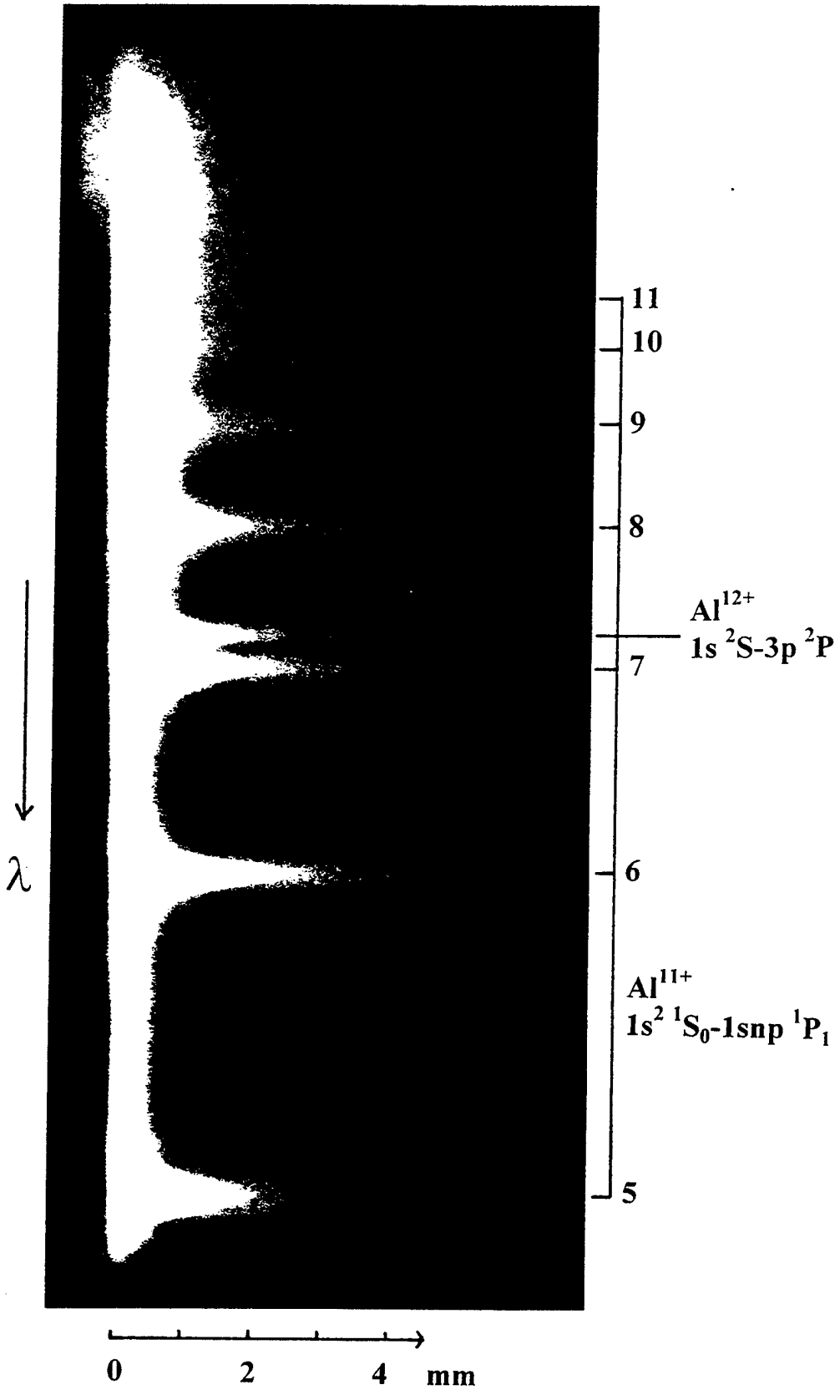

11
10
9
8
Al^{12+}
$1s\ ^2S\text{-}3p\ ^2P$
7
6
Al^{11+}
$1s^2\ ^1S_0\text{-}1snp\ ^1P_1$
5
λ
0
2
4
mm

Figure 8

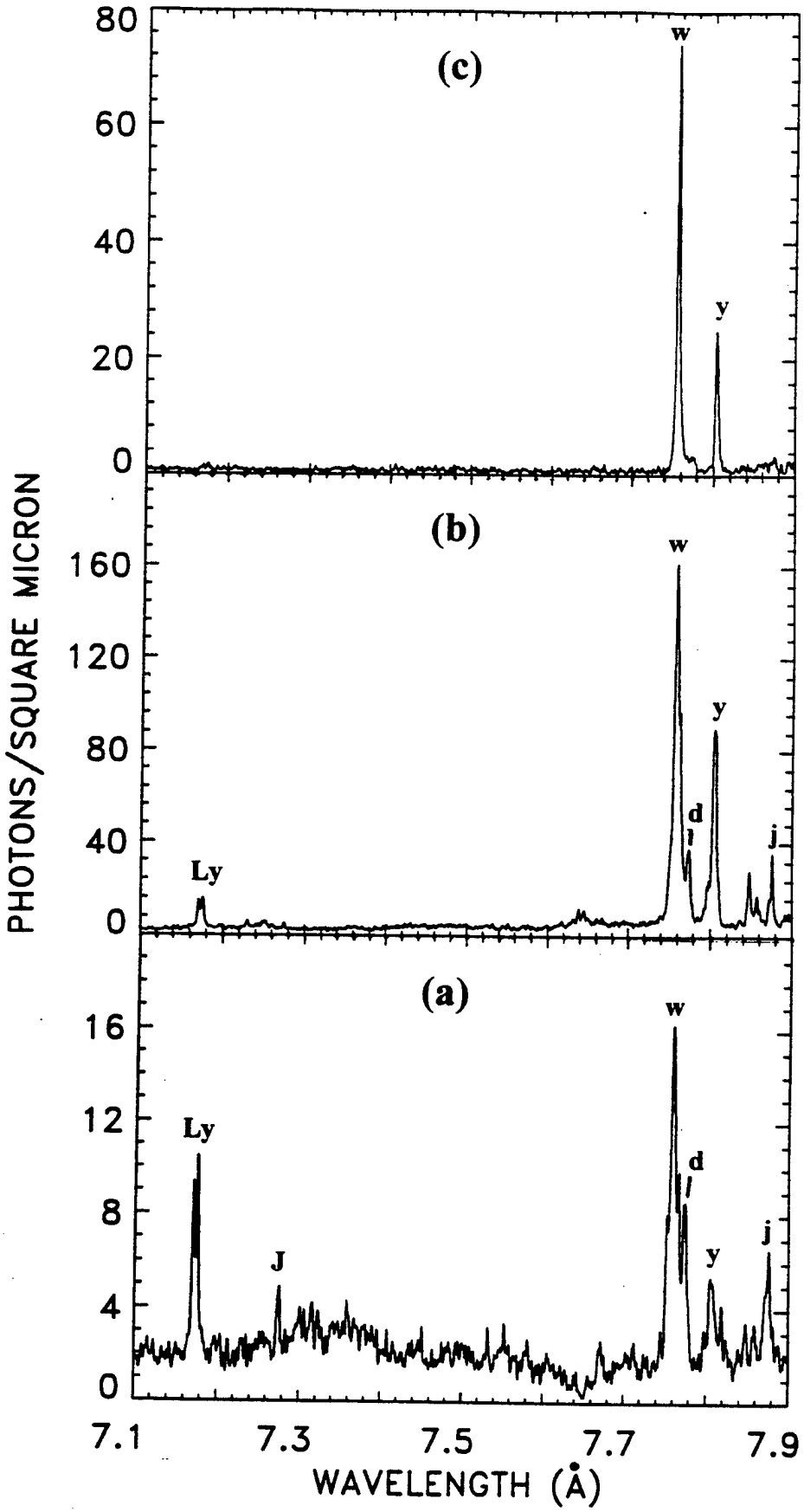

(c)
(b)
(a)
w
y
d
j
Ly
J
80
60
40
20
0
160
120
16
12
8
4
7.1
7.3
7.5
7.7
7.9
PHOTONS/SQUARE MICRON
WAVELENGTH (Å)

Figure 9

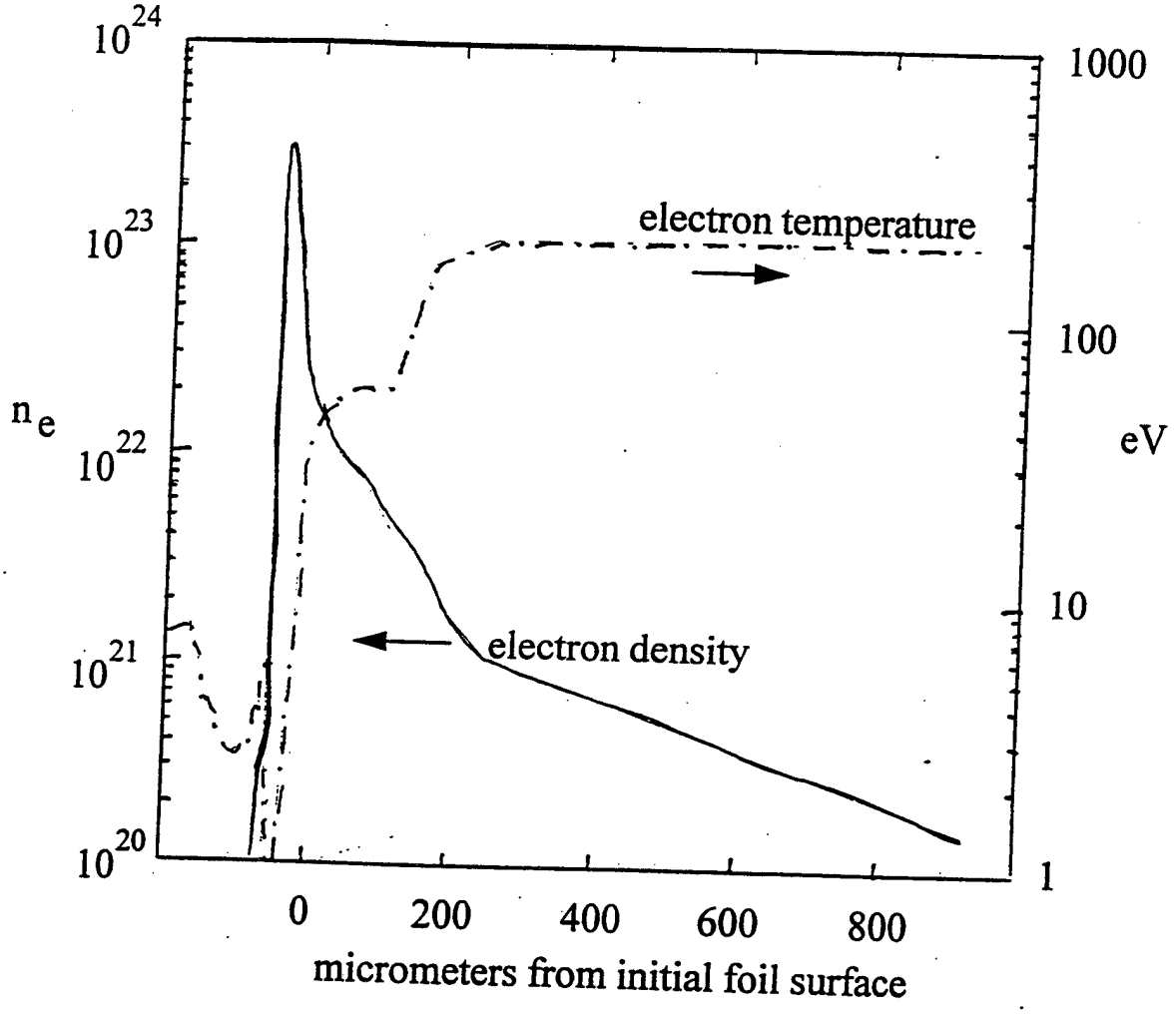

10^{24}
10^{23}
10^{22}
10^{21}
10^{20}
n_e
1000
100
10
1
eV
electron temperature
electron density
0
200
400
600
800
micrometers from initial foil surface

Figure 10

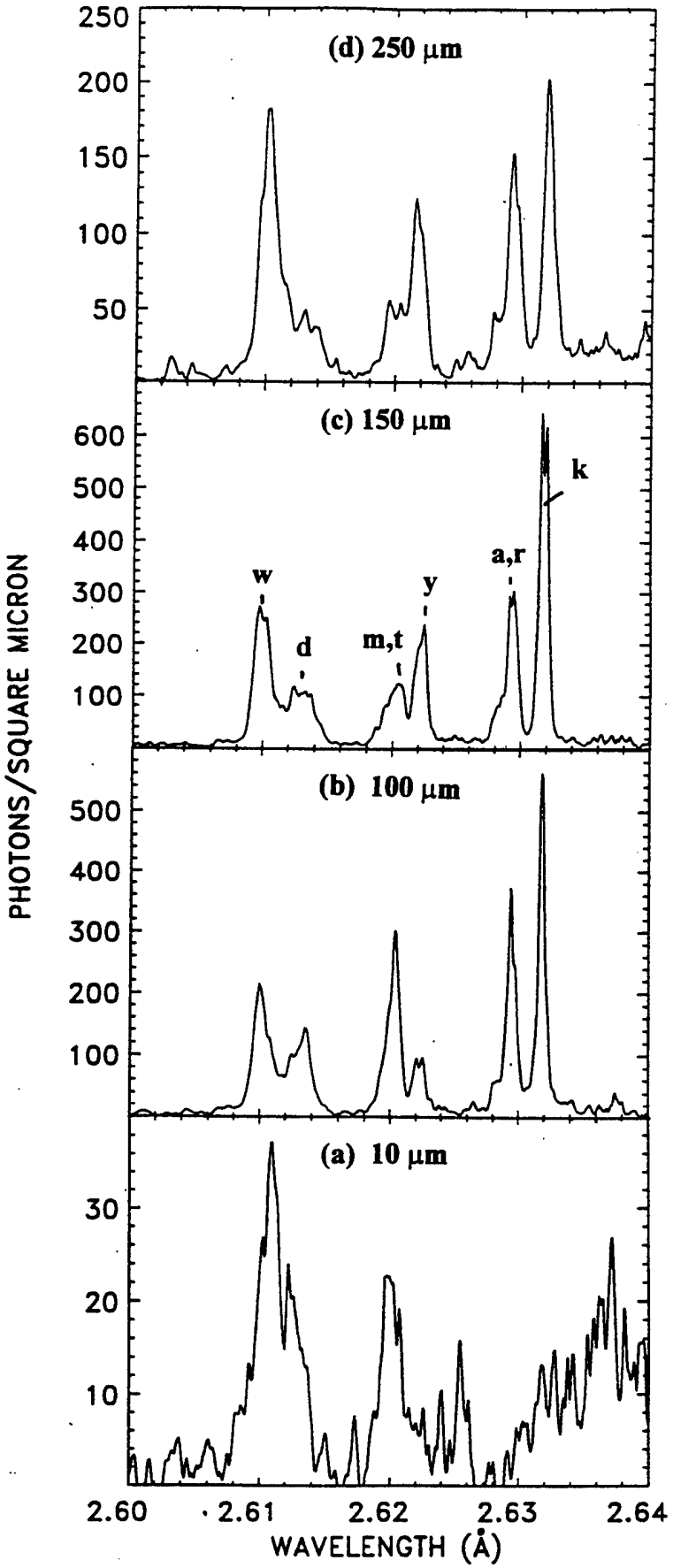

(d) 250 μm
(c) 150 μm
(b) 100 μm
(a) 10 μm
w
d
m,t
y
a,r
k
PHOTONS/SQUARE MICRON
WAVELENGTH (Å)
2.60
2.61
2.62
2.63
2.64

Figure 11

Appendix I

Conceptual Design of a 2 MJ KrF Laser Fusion Facility

CONCEPTUAL DESIGN of a 2 MJ KrF LASER FUSION FACILITY

M. W. McGeoch
PLEX Corporation

P. A. Corcoran, R. G. Altes and I. D. Smith,
Pulse Sciences, Inc.

S. E. Bodner, R. H. Lehmberg, S. P. Obenschain and J. D. Sethian
Plasma Physics Division, Naval Research Laboratory

ABSTRACT

A detailed KrF amplifier model is first benchmarked against Nike laser data, then used to design higher energy modules with segmented pumping. It is found that segmentation with unpumped regions does not carry with it any penalty in efficiency because the distributed geometry has reduced losses from amplified spontaneous emission (ASE) which counteract the fluorine absorption of unpumped regions. A 68 kJ module is designed, incorporating a new water line geometry and a combined switch/bushing. The electrical parameters of the module have been calculated in detail. The effect of multiplexed beamline energy on facility size is discussed and an energy of 100 - 200 kJ is found to be optimal. Two 68 kJ modules are combined in a 136 kJ multiplexed beamline incorporating incoherent spatial imaging that fits within a compact beam tunnel. A total of 16 such beamlines are arranged on four floors to deliver 64 beams to a target, the net energy being 2.0 MJ. Detailed calculations of prepulse ASE energy are given, and the levels are designed to be low enough not to initiate a prepulse plasma. The basic geometrical uniformity of target illumination is shown to be better than 0.3% for a 64 beam illumination geometry which has a high degree of symmetry. A test of the 68 kJ module would be necessary to facilitate more detailed design of this fusion laser.

INTRODUCTION

Laser fusion is an attractive option for a fusion reactor: a small, round pellet located far from the walls of a large reactor chamber, with the high-technology laser located outside the radioactive enclosure. By directly illuminating the pellet, one can obtain a relatively high coupling efficiency. Typically, the pellet can absorb 70-90% of the laser energy (compared to 10-15% for a hohlraum). The energy gains can therefore be in the range of 100-200 using a few-megajoule laser. More reliable predictions of energy gains will have to await further near-term laser-target experiments, but the estimated energy gains are clearly consistent with a reactor.

KrF is the most attractive of the currently available high energy lasers. The short wavelength, 1/4 micron, provides a high overall coupling efficiency to the pellet, and good collisional damping of the deleterious laser-plasma instabilities. KrF as a gas laser is consistent with optical smoothing using echelon-free induced spatial incoherence (ISI) [1], [2]. This optical smoothing technique is simple to implement, flexible, and produces random fluctuations with no residual long-time-average nonuniformities. The broad bandwidth of KrF, 2-3 THz, produces very rapid optical smoothing; this appears to be necessary to minimize the imprinting of laser nonuniformities on the pellet. Experiments with the Nike laser at NRL [3], [4] using 1.3 THz FWHM bandwidth have demonstrated single beam focal nonuniformities of only 1% rms, averaged over the 4 ns pulse. This measured nonuniformity is identical to that expected from the theoretical modeling.

KrF also may meet the various requirements for a power reactor: high efficiency, moderate cost, good rep-rate, reliable. The intrinsic efficiency of KrF can be up to 12%, with overall efficiencies up to 7.5%. Reactor design studies [5] indicate that KrF would be competitive with heavy ion drivers in the overall cost of electricity, assuming that the projections on pellet performance and reactor design prove accurate.

Two directions in the laser fusion program will be necessary to validate this fusion concept: (a) demonstration of a reliable, cost effective and efficient laser system; (b) demonstration of high energy gains from a pellet. The second of these objectives is probably best addressed with a laser facility that operates single pulse. That is the subject of this paper, which describes a 2MJ KrF laser facility design with a new amplifier concept that has an overall cost of approximately $950M, including $500M for the laser itself. Although the size of the laser is much greater

than the 3kJ Nike system, the extrapolation in optics and amplifier technology is modest, and easily evaluated by building one amplifier module.

Our 2MJ laser has an overall efficiency of 4.5% (energy delivered to target divided by energy stored in a capacitor bank). With ISI optical imaging, the nonuniformity is <0.3% rms, excluding high-mode beam-to-beam interference, and the laser has excellent temporal beam shaping capabilities. The laser architecture is that of angular multiplexing, in which time- and angle- coded beams pass through the electron beam pumped KrF laser medium [6], [7].

Each of the largest amplifier modules has a crude energy of 34kJ in a $1m^2$ aperture. These modules are stacked in pairs to form 68kJ, and then ganged again to produce a unit of 136 kJ that feeds one angularly-multiplexed beamline cluster (Figures 1 and 2). There are 16 separate beamlines, producing the total of 2 MJ (16 x 136, reduced to 2MJ by losses in the final optical leg). The amplifiers have been segmented to reduce the inductance of the electron beam generator and thereby limit the pump pulse to 250 ns. 36-fold multiplexing reduces the laser pulse duration to 7 ns at the target.

These amplifier parameters can be compared with the 1991 KrF design proposed by Los Alamos scientists for a laser microfusion facility (LMF) [8]. The los Alamos design used an amplifier module with an energy output of 400 kJ and a pump pulse duration of 1080 ns, as compared to our 68 kJ and 250 ns. Our shorter multiplexing lines have a profound impact on facility size and cost. (The shorter pump pulse duration also reduces the pre-pulse to safe levels). The overall cost of our design compares favorably with the earlier Los Alamos design using a KrF laser, and also with the recent LLNL NIF using a frequency-converted glass laser [9], [10] The costing rules used here were copied and scaled from the costing rules used in those studies.

This report thus outlines a path from the current technology of the NRL Nike laser to a 2MJ laser fusion facility. The next experimental step would appear to be construction and testing of the 68 kJ module.

FACILITY SIZE

The goal is to minimize the cost of the 2.0 MJ facility. One of the major cost elements is the facility building, including its foundations and site preparation. By a trial process we can show that the building size decreases with increasing beamline energy. Also, the number of system components decreases, leading to an

"economy of size". However, as the beamline energy is increased above 200 kJ the building size does not continue to become smaller. This is because:

1. A longer amplifier pump pulse is needed, leading to longer optical paths in the de-multiplex process. This is because the higher voltage of larger diodes leads to more inductance, while at the same time their greater current (larger area) leads to lower impedance. The voltage rise time is lengthened (inductance / impedance) and hence the whole pulse has to be lengthened, to maintain the same usable "flat top" fraction.
2. Sixty or more separate beams must converge synchronously on target. As the number of beamlines drops below 60, due to an increasing module size, a distribution problem sets in, in which increasing mirror area is needed to subdivide each beamline and bring beams separately to target. This effect is measured by the average number of reflections experienced after the demultiplex mirror. In the point design, with a beamline energy of 136 kJ, this number is 2.5. For comparison, when a (hypothetical) beamline energy of 550 kJ is used, we require an average of 4.25 reflections. Not only does this increase the cost and complexity of the optical system, but that system now takes up more space, although space is saved elsewhere by the existence of fewer beamlines and amplifier modules.

In Figure 3 we show the facility floor area (as opposed to ground plan area) as a function of beamline energy. Appropriately smaller modules have been considered for the smaller energy cases. At the low end we considered a novel 7.5 kJ module, grouped in fours to give a 30 kJ beamline. Even although this module was extremely compact, and had a pump duration of only 60 nsec, it could not be packed efficiently in the quantity of nearly 300 that would have been required. The higher energy cases all involved segmented diodes with aperture sharing, generically like the point design. In each case a layout was made to find the least area that would be required.

Apart from its effect on facility size, other factors enter into a decision on module size. These are:

1. What is technically feasible with only a modest extension of present day technology.
2. What is safe from the point of view of energy release during fault modes.
3. More difficult service dis-assembly and re-assembly in large devices.

When the foregoing is all taken into account a beamline energy of approximately 100 kJ seems to be indicated. From the distribution point of view, the net system energy of 2.0 MJ requires 31 kJ per target beam in 64 beams. To

achieve this we could consider 125 kJ in each of 16 beamlines, with additional distribution reflections required to generate the 64 target lines. The latter (136 kJ crude) choice was made for the point design. Each beamline is fed by two 68 kJ, $2m^2$ aperture modules. Each module is comprised of a stacked pair of $1x1m^2$ sub-apertures.

AMPLIFIER MODEL

The size, complexity and cost of the facility can all be greatly reduced if improvements can be made in the design of the final KrF amplifier module. For this reason it is important to have as a design tool the best possible quantitative model of the amplifier. Over several years the Nike program at NRL has developed and validated a comprehensive amplifier model that can now be used in the design of the 2MJ facility. Using this new model we are able to perform detailed calculations on a segmented amplifier which has the electron beam broken into four sections along its optical axis. The smaller individual electron beams have less inductance, translating into faster rise and fall times and consequently an energy delivery pulse as short as 250 nsec. Such relatively fast energy delivery reduces the size of the whole facility because shorter optical delays can be used in the multiplexed beamline. Prior to discussion of the segmented amplifier the basis of the model will be reviewed.

Optical propagation in two directions through the amplifier is coupled locally to the KrF gain medium kinetics. In order to reduce the size of the computation we use a simplified picture of the KrF gas kinetics which nevertheless captures sufficient detail to accurately predict available gain, loss and fluorescence efficiency measurements [11 to 15]. The species considered and their reactions are listed in Table I. A detailed discussion of the selection process has been given in [16]. For many reactions we were able to consider a generalized rare gas atom "G", representing either Kr or Ar. For example, Ar^+ and Kr^+ both yield KrF*, because ArF* rapidly reacts with Kr to form KrF*. Fast reactions that go to completion in < 0.5 nsec at typical densities were not separately considered and the model is therefore limited to pulses longer than about 1nsec. It is valid for mixture densities in the range 0.5 to 2 Amag and pump rates from 0.1 to 1.5 MW cm^{-3}, at krypton fractions up to at least 50%.

Because amplified spontaneous emission (ASE) can be a 10% to 50% energy loss, it was modeled in detail by considering, in addition to the principal beam to

be amplified, the propagation of spontaneous emission in a full range of angles and in 20 spectral bands. ASE in directions close to the amplifier axis is particularly important because of its contribution to pre-pulse on the target. However it represents a minority of all ASE loss. To accurately know the axial ASE at the same time as considering a full range of ASE angles, a geometrically decreasing solid angle set was chosen with a smallest "core mode" at the solid angle subtended by the amplifier aperture at twice the amplifier length. Altogether five or six solid angle "ASE modes" were considered in a typical calculation.

Integration of the coupled kinetic and photon equations was performed with reference to Figure 4, using an algorithm which related the distance h between planes in the amplifier to the kinetics time step Δt by $h = c\Delta t$ where c is the velocity of light. Typically a stable and accurate integration was achieved with time steps less than 0.13 - 0.26 nsec, representing a spatial resolution of 4 cm to 8 cm. A detailed discussion of the amplifier model has been given in [17].

Comparison of model with Nike amplifier data Amplifier performance on both the 20 cm and 60 cm Nike apertures has been modeled to an accuracy of ±10%. The comparison with 20 cm data has been presented previously [17,18]. It represents a good test of the ASE calculation because agreement was obtained on the lowest input shots which had ASE losses as high as 45%. The amplified pulse shape was accurately described on the nsec time scale. The pump rate was up to 580 kW cm^{-3}.

A full description of the 60 cm data comparison is given here, because it represents the strongest validation at high amplifier saturation levels and high gain length.

The time-dependent amplifier model was used to calculate 60cm amplifier output for the measured depositions of shots taken between 2nd Nov 1994 and 19th May 1995. The specific shots analyzed are listed in the Table II. They include all shots with double-sided pumping and fresh fluorine fills.

The 60cm amplifier was operated with 56 x 4nsec beamlets (=224nsec) containing a total energy of 50J injected into the 60cm x 60cm aperture. The input intensity was therefore 62kW cm^{-2}. The following optical losses were applied to the crude laser energy calculation, giving a correction factor of 0.862:

a)	Obscuration due to 3.6 mrad multiplex angles	8.3%
b)	Coherent light removed by 8% reflection off tilted window at mirror end	5.0%
c)	Loss at output window due to imperfect anti-reflection coating	1.0%
	combined factor	0.862

The deposited energy as measured by the pressure jump was corrected to remove energy deposited before and after the 224 nsec extraction span. This correction was made by analysis of the voltage and current traces in the electron guns. The energy extracted in the laser output beam and by amplified spontaneous emission was added to the measured deposition. For these conditions the amplified spontaneous emission is calculated to represent approximately 20% of the laser output. Lastly, the deposition was increased by 5% above the pressure jump measurement to account for radiation loss. The following components are expected to contribute to radiation loss:

a)	Kr_2* radiation	<2%
b)	KrF* + Kr_2F* radiation	1.7%
c)	X-ray loss	≤1%
	combined factor	0.953

No correction was made for atmospheric absorption between the 60cm output window and the 16 calorimeters which sampled beamlets spaced throughout the 56 pulse train. This absorption is thought to have been less than 5%. The calorimeter reading is believed to have ±5% accuracy.

The measured and calculated laser output energies are graphed in Figure 5. One of the measured points (labeled in graph) consists of the average of 5 shots.

Comparing theory and experiment we see that at the lowest depositions there is substantially less output than predicted. This is due to the use of reduced voltage for these shots, leading to non-uniform deposition, with very little in the center of the aperture and much more adjacent to foil windows. At higher depositions the most efficient laser shots lie on or within 5% of prediction, but there is a scattering of shots which lie between 70% and 95% of prediction. Because of the observation of declining output on the second and third shots of a given fluorine fill, we know that there was the potential for fluorine depletion which could reduce the output. It seems likely that the less efficient first fill shots were also caused by fluorine loss. A materials change has since eliminated most of the fluorine losses.

In summary, the strongest data point, which is an average of five shots, is predicted almost exactly by the calculation. If we believe that the most efficient

shots more truly represent the performance of the system when well passivated to fluorine, then it appears that the amplifier code comes very close to accurately predicting the laser output. The deposition measurement on which the code prediction is based could carry a systematic error of as much as 10% whereas the laser energy is probably measured to within 5% accuracy. At present we can say that the code agrees with experiment to within the experimental accuracy of ±10%. This result implies that larger amplifiers can be designed with confidence using the amplifier code.

SEGMENTED AMPLIFIER

In this section we show that it is possible to have an efficient KrF amplifier with a segmented electron gun. The segmentation consists of alternate pumped and unpumped regions as we travel along the optical axis. Previously this scheme has been rejected because without detailed calculation it has appeared that the fluorine absorption in the unpumped regions would seriously reduce the amplifier efficiency. This would be true of an oscillator, particularly one with low output coupling, but application of the time- and space- dependent amplifier model shows that there is only a slight penalty in an amplifier. Firstly, the absorption is only effective on the exit pass (of a double pass amplifier) and even then is only fully effective just before the exit window. Secondly, an unexpected benefit of segmentation is that the amplified spontaneous emission (ASE) loss is geometrically reduced. This happens because more of the spontaneous emission reaches the (absorbing) chamber walls without passing through a high gain region.

Using the amplifier code a square cross section amplifier was considered with a number of pumped segments each 56 cm in the optical direction (Figure 6). (The design shown in more detail below has 48 cm pumped segments). There are two principal ways to apply segmentation constraints. To begin with we considered a fixed pump rate of 700 kW cm^{-3} and a fixed mixture of 0.8 Amag (Ar , 30% Kr, 0.6% F_2), and increased the unpumped segment gaps from zero to 40cm. The decrease in energy output with segmentation was less than expected (Figure 7), amounting to only 4%. A larger decrease in output (9%) with segmentation was seen when ASE was artificially turned off. This lead to the conclusion that segmentation was actually cancelling some of the losses from ASE. In a second calculation we considered a constant amplifier length and constant electron gun power, but varying degrees of electron beam spreading by means of magnetic field

coils or electrostatics (implying an increased pump intensity as the segments are formed). Figure 8 shows that the output actually increased from 36.8 kJ to 37.3 kJ as segments were formed in this example. The improvement was mainly due to reduced ASE loss, as the efficiency was not strongly dependent on the pump rate and the fluorine concentration was near optimal in all cases.

Lastly, assuming 48 cm segments separated by 48 cm gaps, the output was studied as a function of the number of (equally) pumped segments. The results are listed in Table III in which the four-segment result applies to the present point design. Here the aperture was 100 cm x 100 cm and the mixture was 0.8 Amag: {Ar, 30% Kr, 0.6% F_2}, pumped for 250 nsec at 800 kW cm^{-3} and extracted by 350J of input radiation. Each additional segment added 48 + 48 = 96 cm to the amplifier length. The decline in efficiency between 3 and 5 segments is mainly due to excited state absorptions, and occurs both with and without segments.

For the design number of segments the amplifier intrinsic efficiency is 9.5%, which is apparently better than would be expected on the basis of the calculated small signal gain and loss coefficients. The improvement is due to a) having exceeded the saturation intensity for the saturable loss of Kr_2F^* and b) recycling of energy following (non-saturable) absorption out of G*, where G represents the general rare gas atom. These effects only contribute to output when the kinetic behaviour is computed in real time along with the amplified flux, as in the present code. A separate calculation has been performed to verify the effects.

The above results make segmentation look very feasible and eliminate the need to use magnetic material or electrostatics to spread the electron beams. The beams can be guided as usual by a parallel magnetic field. Because there are now two return current planes in each diode segment the self-field is greatly reduced. This allows guide fields of less than 2 kG to be used, which reduces the diode closure velocity. More important than this, the diode inductance is also reduced, allowing a pump duration as short as 250 nsec.

68 kJ MODULE

We have shown in the previous section that a segmented diode can drive an efficient amplifier without requiring the beam spreading that has always been a source of difficulty in the design. This opens up tremendous possibilities for amplifier improvement, particularly in the direction of reduced diode inductance and pulse duration. Compared to unsegmented designs of the same output energy

the pulse duration can be approximately halved. This translates into half the optical delay length for the multiplexed beamlines, and substantially fewer mirrors, making a high energy facility significantly less expensive.

There are two good reasons for designing an amplifier with vertical, as opposed to horizontal segments. Firstly, the aperture is pumped much more uniformly when the segments are perpendicular to the optical axis, and secondly, the vertical segments match the preferred design for the water pulse line, in which the electrodes can be introduced and removed through the top surface of the water and the presence of breakdown-initiating bubbles and particles adjacent to the electrodes is minimized by making most of the sensitive electrode areas vertical.

The scale of this module is determined principally by the largest size of good optical quality fused silica window that is currently available, a square aperture of 100 cm x 100 cm. (an aperture of 60 cm x 60 cm is used on the Nike large amplifier). The water plate size scales in the vertical direction, the only limit really being the energy that we are willing to store per plate and feed to a single diode. At reasonable stored energy in the water plate (less than 100kJ) we find that we can feed a height of 100 cm, matching the window height. For economy of structure and space-saving we place another 100 cm square aperture on top, making an optical aperture that is nominally 100 cm wide by 200 cm high.

The window is not expanded into a 2 x 2 array because the diode voltage would have to increase in order to penetrate the deeper gas volume. Along with this would come decreased specific energy deposition because the stopping power and also the gas pressure would be reduced. The module energy would therefore have to be made up either by an increase in current density or by an increase in pump duration. Either way we would enter a diode regime which has not been explored and the risk would be too great. As it is, keeping the depth at 100 cm allows an 800 kV diode together with 0.8 Amag mixtures that are within the scope of traceable kinetic modeling. Two separate 100 cm x 200 cm modules are combined by aperture sharing (see below) to generate a useful square aperture beamline with a nominal output energy of 4 x 34kJ = 136 kJ.

The design of the 68 kJ module is shown in Figures 9, 10 and 11. A laser gas volume that totals 356 cm long by 100 cm wide by 267 cm high is pumped in four bands (shaded), each of length 48 cm in the optical direction. Between these bands are unpumped regions each of length 48 cm, and at each end there is an unpumped length of 10 cm. The pumping is from two opposing sets of eight vertically oriented diodes, each with nominal cathode dimensions of 48 cm x 100 cm. Each

diode is fed by a water PFL of depth 1.0 m and length 5.0 m (corresponding to a pulse duration of 300 nsec to supply a 250 nsec 'flat top' duration) [19]. Each center plate is charged to -1.64 MV in 1.8 μsec by a Marx bank (not shown in this view) before being switched through two parallel laser-triggered switches to the diode load. The classical "Z-stack" bushing of the Nike 60 cm machine and its separate elaborate switch modules have been greatly simplified by their combination into an integrated switch/bushing, of design shown in Figure 12. Apart from reducing the cost, this simplification leads to a lower inductance which translates into a faster diode risetime.

We have simulated the circuit using a 150 nsec long drive line (300 nsec nominal pulse) with 1.28 ohms impedance, getting best results after making the last 20 nsec a 1 ohm peaking section. This circuit represents a single column of two water plates, four output switches and two diode segments. As exemplified in Figure 13, when the driver is switched out at 1.64 MV into a diode with initial spacing 5.5 cm and 2.3 cm μsec^{-1} closure velocity, the diode voltage and current ramp down and up by about 6% during the pulse, averaging 790 kV and 690 kA, with a peak power of 0.55 TW. Each column of two plates is charged by an 11-stage Marx with ± 90 kV on two 3.0 μF capacitors in series in each stage. Resistive damping is included similar to Nike's.

Considering the voltage waveforms at both ends of the driver line, we estimated an operating field for the water based on a 32-module laser (2.2 MJ output). The formula most often used to predict water breakdown in uniform fields (F, $MVcm^{-1}$) at large areas (A, cm^2) in terms of stress time (t, μs) is

$$Ft^{1/3}A^{-0.058} = 0.23$$

The form of this equation is due to AWE, Aldermaston workers, and the values of the constants were suggested by Eilbert and Lupton at NRL (unpublished, but listed in [20]). Several workers have reported that very large areas (e.g. the Sandia PBFA II) can operate safely at or very near 100% of the "breakdown" level predicted by this formula, because it overestimates the area term for large machines. The 2 MJ KrF laser has a larger area than any previous water dielectric system, and 100% by the formula corresponds to a 17.6 cm spacing. The ability to operate safely at this spacing is supported by another formula ($Ft^{1/2} = 0.09$) that predicts that even if one electrode were replaced by a very sharp point breakdown will not occur with spacings greater than 16 cm.

Table IV shows the risetime and 90% width of the diode voltage pulse as a function of the total inductance between water line and cathode in one column

consisting of two plates and diodes. The driver pulse is nominally 300 nsec long in all cases. The peaker has been used to increase the 90% pulse width. It is seen that even with >100 nH series inductance (for a vertical pair of plates and cathodes in parallel "column") the 250 nsec flat top is still obtained. The preliminary design of Figures 10 to 12 has an inductance of approx. 75 nH.

Many of the module design parameters are listed in Table V. The energy stored at 1650 kV in a single water plate is 83 kJ, which is nominally the energy delivered to one of the diodes. The energy delivered in the 250 nsec flat top portion (> 90% voltage) is only 75 kJ. This is of the same order as the energy presently delivered to a single Nike diode on the 60 cm machine, an energy that has been shown to be tolerable if released anywhere within the machine in a fault mode, provided the water plates are constructed of 0.14" stainless steel. The self magnetic field is 1.7 kG, again comparable to that in the Nike machine.

One difference from the Nike machine is that an energy transfer efficiency of 0.7 is assumed from stored energy to laser gas, in contrast to the 0.4 - 0.5 performance on Nike. There is reason to believe that 0.7 will be achieved in the new design. The span of the foil support structure across one segment is 48 cm, less than the 60 cm free span on Nike. The support members can therefore be lighter. In fact, it is proposed to use a "suspension bridge" design very similar to that demonstrated by Los Alamos on the Mercury amplifier [21]. That demonstration involved a suspension bridge span of 41 cm, not much less than that proposed for the present module. A "geometrical" transmission exceeding 87% was measured by comparison with another less transmissive support design. To this loss must be added the reflection that takes place due to back-scattering from the laser gas itself, which amounts to an 85% effective transmission for the mixes of interest here. The overall figure of 70% assumes other edge losses, switch losses and unspecified electrode surround losses totalling 5% in addition to transmission and reflection losses. Although the suspension bridge is very promising, the full transmission was not measured; the effect of a concave un-extracted region was not investigated; and the system lifetime was not determined.

The choice of electron beam voltage (800 kV, average) is determined by several factors of approximately equal importance. Firstly, the electron beam deposition profile in the 100 cm aperture is approximately uniform for gas mixtures at approximately 0.8 Atmosphere, according to Monte Carlo simulations of deposition at the nearby voltage of 700kV.This relatively low gas pressure eases the window and foil design. A mixture of approximately the correct stopping

power is 0.8 Amag: Ar, 30% Kr. This mixture is used in calculations of the amplifier extraction efficiency. Secondly, the decision to keep a 96 cm diode module width (including unpumped spaces) leads to a space restriction on the rear side of the cathode shell which limits the radius of the bushing, with a corresponding limit on its voltage. Thirdly, there is experimental data on laser triggered switches at up to 1300kV on the Nike 60 cm device and to 1200kV on the RAL laser [22], also the DARHT [23] and AIRIX injectors operate very reliably with 1500 kV laser-triggered switches so the present line charge voltage of 1650 kV is not too great an extrapolation from prior experience.

The electric stress on the cathode surround is 150 kV cm^{-1}, which is consistent with the trend for historical KrF experiments (stress criterion = $70\tau^{-0.5}$ kV cm^{-1} [24], where the pulse duration τ is in μs, giving 140 kV cm^{-1} at 250 ns). At this surface field there will be field emission from the curved part of the cathode shell, but the guide magnetic field will prevent current from flowing to the nearby ground plane. The distance to the back plane is larger, so the main loss will be toward the anode surrounds. The current loss will be limited by space charge to about 5% percent of the total current.

The device is designed for easy access to all internal parts. There is a break plane where the water tanks mate with the vacuum box. The water / oil tanks on a given side move together at all times. When they are moved back there is access to all eight foils on one side. Also by removing the cathode shells there is access to all the switch/bushings. The latter component is designed to be removed from the machine through the ground plane (once the water has been lowered).

MULTIPLEXED ISI BEAMLINE

With the ISI type of beam smoothing, pioneered on Nike [3], a highly uniform, low coherence source is imaged via various stages of amplification onto the target. Because a 7 nsec pulse is needed and delivery of pump energy to the main amplifier takes 250 nsec, the amplifier is traversed by 36 x 7nsec beamlets in an angle- and time-coded sequence. The beamlets are then de-multiplexed before superposition on target, using optical delays of varying length up to 37 m (Figure 14).

The beamline can contain more than one main amplifier module. A beamline energy of more than 100 kJ is desirable in order to minimize the facility size but the largest available amplifier subunit consists of a $1 m^2$ aperture producing 34 kJ

(crude). It is possible to combine four of these $1m^2$ apertures in a compact beamline of 136 kJ. This is best achieved by stacking the 34kJ apertures in pairs, to make a 68 kJ basic module, of which there are two per beamline (Figure 14). The crude energy of 136 kJ is quoted, although optical losses through the final focusing lens reduce the net energy to 90% of this.

Rather than generate several beamlines within the common space of a large hall, as proposed in the Los Alamos LMF design, we prefer to contain each separate 136 kJ beamline in a separate rectangular duct of dimensions 4m wide x 5m high x 55m long. The ducts are helium-filled to avoid atmospheric Raman scattering. This arrangement reduces the helium requirement, gives more convenient access, and gives local containment in the unlikely event of a fluorine leak from an amplifier module. Also it allows complete beamline testing at the level of 136 kJ, removing uncertainty in extrapolation to a larger system. Separate testing of subunits is also advantageous in the commissioning of a 2 MJ facility, where there would be 16 such beamlines.

The amplifier sub-aperture is set at $1m^2$ by the size of available fused silica windows, as well as by the energy requirement. It is noted that 34 kJ in a $1m^2$ area is 3.4 J cm^{-2}, which is below the damage threshold for coated KrF optics. (This is discussed in the LANL LMF document [8]: the LMF was designed to 5 J cm^{-2}, and coatings have been tested to >20 J cm^{-2} in 15 nsec pulses).

There are three factors that influences the size of the beamlet optics. The first and most important is that the final lens choice of fused silica imposes an intensity limit due to two-photon absorption that corresponds to a fluence of 4.0 J cm^{-2}. Because there are strong advantages to having collimated beam transport to the final lens, the lens fluence is equal to the fluence on the de-multiplex mirrors. The size of the individual demultiplex (and recollimation) mirrors is therefore of the order of a square of side 15cm. The second factor relates to the need to image through the amplifier system from a finite source at the front end onto a target of diameter up to 0.4 cm. Ray calculations show that the size of the de-multiplex elements and the final lenses should be increased to 18cm x 18cm in order to avoid edge losses. Thirdly, ISI beam smoothing with incoherent light does not require any additional safety factor to allow for hot spots due to coherent interference.

Beyond the mirror size and optical delay length which have already been discussed, there are two other factors determining the layout of an ISI beamline. The first of these is the angle between multiplexed beamlets. A minimum angle has already been built into the design by the ratio of the demultiplex mirror size to the

total beamline length. For the 15 cm and 55 m figures of the point design this angle is 2.7 mrad. This angle should be sufficiently large to ensure low scattering at optical surfaces from one beam into another [8]. The second factor comes from geometrical optics. The amplifier feed mirror for each beamlet is projected by the main amplifier's spherical reflector onto the recollimation mirror. In principle, because much less energy is handled, and diffraction is not an issue, the feed mirrors could be much smaller than the recollimation mirrors (which are 15 cm by 15 cm). However, when designs are laid out, the most compact design appears to be achieved with the feed array in the same plane as the recollimation array, i.e. at a magnification of unity, so the feed array mirrors are also 15 cm x 15 cm. In the case of aperture-sharing (discussed below) this choice allows symmetrical feeding of the apertures, without the feed array getting in the path of the recollimated beamlets.

Aperture sharing We need >100 kJ beamlines (e.g. 136 kJ), but our largest optical aperture is $1m^2$, yielding only 34 kJ (crude output). The solution is to aperture share. Although it would reduce the number of optical elements to have true aperture sharing, in which four $1m^2$ apertures are treated as a giant 2m x 2m amplifier (by means of plane turning mirrors), in practice it would be difficult to ensure that the four $1m^2$ spherical reflectors would have both precisely the same radius of curvature and the positional flexibility to bring their centers of curvature into alignment. For this reason, although we use a common feed array for the four apertures, we split the re-collimation and de-multiplex optics into four separate sets, related respectively to each of the four $1m^2$ apertures. The size of each recollimation mirror is therefore approximately $1m^2/36$, as noted above. There are 4 x 36 = 144 de-multiplex elements in all. The disposition of feed and recollimation mirrors is shown detail in Figures 15 and 16. In fact two of the four recollimation arrays are set forward by 2.1m in order to avoid interference by some of the elements of the de-multiplex array. This leads to the offset seen in the two 68 kJ amplifiers.

There is a single feed array (3 elements wide by 12 high) projected separately by each of the main amplifier mirrors onto four identically shaped recollimation arrays (Figure 16). The arrays have a vertical aspect because it is possible to transmit greater vertical than horizontal angle excursions through the 34 kJ amplifier subunits without incurring vignetting losses. This is because the electron beam foil windows are an obscuration precisely at the edge of the 100 cm aperture,

whereas it is possible to have a 110 cm aperture in the vertical direction without obscuration. For the present design the vignetting loss is <3%.

Front end. The function of the "front end" is to deliver to each of the 16 driver amplifiers an array of 36 time- and angle-coded shaped-pulse beamlets, carrying (in the whole array) an energy of >20J. The total energy from the front end is 16 x 20 = 320J, increased to 500J to allow for optical losses, including overfilling of the driver amplifier. One possiblility for coding is shown in Figure 17. Coding is done at the 10J level, prior to a preamplifier of gain approximately 50x. Because long paths begin at the encoder (the 245 nsec maximum time delay represents a path difference of 73.5 m) we have expanded the beam up to 15 cm prior to the encoder, ensuring large Fresnel numbers. Entering the encoder is a single pulse which has been shaped just prior to the encoder by one or several Pockel's cells at 15 cm aperture. This pulse has to have an energy of approximately 20J in order to compensate for losses in the encoder beamsplitters and produce the desired 10J encoded energy. Several stages of amplification will be necessary to raise the initial few μJ of incoherent light accepted by the first system lens up to the 20J needed by the encoder. The initial stages might be discharge amplifiers such as the 3cm aperture, 15 nsec design used on Nike. Further amplification at 15cm aperture will probably require one or two short pulse electron-beam pumped stages.

Figure 18 shows how the "front end" encoded pulse array is distributed in the sub-floor to 16 vertical passages leading to the upper floor beamlines. The preamplifier is followed by an array of 36 Pockels cells in order to remove amplified spontaneous emission. After reflection off concave mirrors the 36 beamlet array is split by partial reflectors into four lines. These enter sub-floor beamsplitters to direct an appropriate fraction up to the driver amplifier of each 136 kJ beamline. There is an intermediate focus in the sub-floor region, but it occurs at relatively low intensity (10^8 W cm^{-2}), so it will not be necessary to use a vacuum chamber. Further foci are avoided until after the final lenses.

Diffraction We estimate the effect of diffraction following the final lens aperture by comparing the full width half maximum (FWHM) diameter of the target beam to the radius of the first sinc function minimum, the latter being approximately equal to the FWHM of a diffraction-limited beam at the focal distance. This ratio is referred to as XDL. The target beam diameter is matched to the pellet diameter d_T, thus

$$XDL = \frac{dd_T}{\lambda f}$$

where d is the beam size at the final (square) lens aperture, f is the lens focal length and d_T is the target diameter. For: d = 15 cm ; f = 750 cm and d_T = 0.36 cm we have XDL = 290. Factors affecting the choice of system XDL are discussed in Appendix 1.

Throughout the optical system the Fresnel number ($d^2/\lambda L$) is characteristically very large (L = distance between relaying optical elements). The smallest Fresnel number occurs between the preamplifier and the driver amplifier, and is 820. The system has been simulated using ray optics in the paraxial approximation. The design, which is based on the demonstrated beam-smoothing of Nike [3] and prior calculations [2], has the following features:

1. Uniform illumination is achieved via imaging onto the target of a uniform source located in the amplifier front end.
2. The optical system must accept a large angular spread of light from the incoherent source. A full cone angle of 25 mrad is used.
3. The target size for distortion-free illumination should be up to 0.4cm diameter.
4. Amplifiers are located at or near to the Fourier transform plane of the image or target so that spatial non-uniformity of pumping (particularly in the front end discharge amplifiers) does not create an appreciable fluence non-uniformity on target.
5. The main amplifier is "overfilled" by the beams from the driver amplifier, in order to suppress parasitic oscillation.
6. The driver amplifier is "overfilled" by the beams from the preamplifier for the same reason.
7. The preamplifier is imaged (or very nearly so) onto the main amplifier in order to satisfy the Fourier transform requirement at the relatively large separation between these amplifiers. The driver amplifier is not exactly at a transform plane.
8. An intermediate focus is allowed between the preamplifier and driver amplifier, but not at later stages because of the danger of plasma formation.
9. The beams downstream of the main amplifier must not strike the metalwork at the edges of the recollimation, de-multiplexing and final lens optical elements. Overfilling of optical elements not lying at the image's Fourier transform plane would result in image distortion. It would also be a source of inefficiency as well as a potential problem due to plasma formation. In order to satisfy this it is found that the de-multiplexing mirrors and the final lens aperture must be 18cm x 18cm,

which is larger than the 15cm x 15cm size of the recollimation elements.
10. In order to control aberrations the angles of incidence onto spherical reflectors must be kept as low as possible. If necessary, the convex mirrors of the main feed array can be replaced by planar mirrors with tilted concave lenses in front, to cancel off-axis aberrations, as in the Nike laser [25].
11. Each of the 36 multiplexed beamlets should pass through an approximately equal number of beamsplitter elements in order to share the optical inhomogeneity budget.

Ray calculations showed that the system in Figure 14 combined with the front end of Figure 17 satisfied the above criteria. With a 120 cm focal length lens immediately after the source the complete amplifier system magnifies by 1.5 times. The 0.36 cm diameter target can therefore be illuminated by a 0.24 cm diameter front end source. The distribution of rays on the final lens is illustrated in Figure 19. This is a relative distribution giving the number of rays in 1cm x 1cm areas within only one quadrant of the lens (the lens center is at top left). This result refers to the average beamlet (number 18) and a source of 0.24 cm diameter. The distribution is used below to estimate the two-photon absorption loss in the silica lens material.

Statistics for Optical Elements A large area of optics is inevitable in MJ lasers with relatively low fluence limits. In this system the strongest fluence limit (4 J cm^{-2}) was that caused by two-photon absorption in the final focusing lens (of fused silica). At 2.2 MJ incident on these lenses this implies a minimum lens area of 5.5 x 10^5 cm^2. The actual lens area needed to avoid edge losses is 7.5 x 10^5 cm^2. In the multiplex process there are five reflections in the beamline (including the turning mirrors (Figure 14) and 2.5 further reflections (on average) on the way to the final lenses. Also, the power optical beam transits three windows, including the final lens and two in the main amplifier. The total mirror area is 6.4 x 10^6 cm^2 and the total lens and window area is 3.5 x 10^6 cm^2 (Table VI).

The cumulative mirror area as a function of the characteristic dimension of the elements is plotted in Figure 20. Also shown in that figure is the same compilation for the 3 MJ Los Alamos LMF design. Our design is at lower fluence, 4J cm^{-2} vs 5J cm^{-2} for the LMF, tending to increase the area, but at less total energy.

The major step on these cumulative distributions represents the feed, recollimation and de-multiplex mirrors. Our characteristic dimension is 15 cm or 18 cm for these mirrors vs 40 cm in the LMF design. Our largest mirrors are the

turning mirrors necessary for the aperture share. This is the price paid for a much more compact facility design through use of 136 kJ beamlines rather than the 68 kJ beamlines that would result from not using turning mirrors.

CONTROL OF AMPLIFIED SPONTANEOUS EMISSION

In a time-multiplexed amplifier system there is a flux of amplified spontaneous emission (ASE) reaching the target ahead of the main energy pulse. The duration of this flux varies from approximately zero for the first beamlet of the sequence up to the amplifier pump duration for the last beamlet. The superposition of these fluxes is a "ramp" culminating in N times the ASE intensity of a single beamlet, where N is the number of multiplexed beamlets in a beamline.

In 64-beam illumination of a spherical target the ASE from each target beam is effectively amplified up to 16-fold at any point on the target surface. Depending on its absorption coefficient the target surface will heat, evaporate and ionize during the ASE exposure, unless that exposure is controlled below certain intensity and fluence limits. Detailed estimates show that an aluminum surface will withstand up to 1.5 J cm^{-2} at 249 nm (in a 250 nsec ramp) before plasma formation and that fluences up to 1 J cm^{-2} can be absorbed by plastic surfaces provided that the absorption depth exceeds 4 microns.

It is possible to design the 2 MJ system to have such low levels of ASE on target that plasma formation may be avoided. The recommended design has a fluence of <0.6 J cm^{-2} and a peak ASE intensity of 5 MW cm^{-2} reached at the end of the 250 nsec ramp. If necessary, further reduction is possible.

ASE from the main amplifier contributes less than 10% of the total ASE on the target, as verified by detailed ray calculations. The ASE intensity is instead determined by the product of a) the ASE brightness at the exit of the driver amplifier, and b) the efficiency with which ASE propagates through the main amplifier and down the ISI optical train to the target. It is possible in principle that a geometrical change in the driver amplifier can change both factors at once. These two components of on-target ASE intensity are discussed in sequence below.

ASE Brightness A detailed variation of parameters was performed to dissect out the influence of such factors as driver amplifier pump rate, gain, aperture, length and input intensity. Using the amplifier code, the axial ASE brightness at the exit of a two-pass amplifier was found to *only depend on the amplifier gain* via the

relation

$$B_{ASE} = 2 \times 10^4 \times \text{Gain} \quad W\,cm^{-2}\,Sterad^{-1},$$

regardless of the amplifier length, aperture, output energy or pump rate. This result is true for both very large and very small electron beam pumped KrF amplifiers. One of the parameter variations is illustrated in Figure 21, in which the gain is varied in a trial 3kJ driver amplifier. An analytic derivation of the same result is given in Appendix 2.

ASE transport to the target surface. It was found using ray calculations that variation of the driver amplifier aperture 'W' at constant distance 's' to the main feed array, or variation of its distance at constant aperture, had no effect on the ASE intensity *at the target plane*. This distance could therefore be set by other considerations, such as the angle between driver amplifier multiplexed beamlets, combined with the offset of neighbouring feed array mirrors. The driver amplifier aperture 'W' could also be set by other considerations, such as the need for a certain volume to obtain the design energy at a specific output of approximately 10 J l^{-1} and the requirement of a relatively small aspect ratio for efficient multiplexing. The on-target ASE fluence varies in a simple manner in response to a few of the system parameters, as derived in the following section.

Analytic formula for on-target ASE. For N (=36) beam paths between the driver amplifier of aperture W (=50cm) and the target, the maximum ASE power on target (immediately prior to the energy pulse) is

$$P_N = NG_{Tot}TW^2 B_{ASE}\Delta\Omega_T$$

in which G_{Tot} (=100) is the gain of the final amplifier,T (=0.63) is the optical transmission due to overfilling of the final amplifier, and $\Delta\Omega_T$ is the scaled solid angle subtended by the target as seen from the driver exit plane, given by

$$\Delta\Omega_T = \left(\frac{b}{W}\right)^2 \cdot \frac{A_T}{f^2}$$

Here b (=15 cm) is the beam width at the target lens, f (=750cm) is the target lens focal length and A_T ($\sim\pi(0.18)^2$) is the target cross sectional area. The maximum ASE intensity on target is therefore

$$I_N = \frac{P_N}{A_T} = NG_{Tot}TB_{ASE}\frac{b^2}{f^2}$$

From the previous section, $B_{ASE} = 2\times10^4 G_D$, where G_D is the driver gain. Therefore,

$$I_N = 2\times10^4 G_D G_{Tot} TN \frac{b^2}{f^2}$$

and the fluence in the ASE intensity ramp of duration τ_P (=250 ns) is

$$E_N = \frac{\tau_P I_N}{2} = 10^4 \tau_P G_D G_{Tot} TN \frac{b^2}{f^2}$$

Evaluated for the present parameters (listed in Table VII) E_N = 145 mJ cm^{-2} is the on-target ASE fluence for one of the 16 beamlines. In deriving the figure 580 mJ cm^{-2} onto a spherical surface we take into account that any point on the spherically illuminated target sees 16 times the intensity of one target beam and therefore 4 times the intensity of the 4 target beams associated with one 136 kJ beamline.

Another important factor is the $\tau_p N$ product, which depends on the square of the amplifier pump duration. The 136 kJ beamline (250 nsec) is a little worse in this respect than Nike (176 nsec). The earlier LMF design, at 1080 nsec, would have had a much lower (coherent/ASE) ratio.

Most importantly, as shown above, the on-target ASE is decreased by keeping the driver amplifier gain (G_D) low. In Nike the driver (20 cm amplifier) gain is suitably low, and a factor T = 1/2 is introduced via the beamsplitters between it and the main amplifier. In the 136 kJ case the driver amplifier has the same duration as the main amplifier, but there is substantial overfilling of the main amplifier, and therefore T = 0.63 in this case.

Coherent/ASE ratio vs energy input to driver amplifier It is interesting to look at the (coherent/ASE) fluence ratio as a function of the input energy E_0 to the driver amplifier. The on-target coherent energy from one multiplexed beamline is

$$E_c = E_0 G_D T G_{Tot}$$

For a target of cross sectional area A_T the ASE energy is

$$E_{ASE} = E_N A_T$$

The coherent/ASE ratio R is therefore

$$R = \frac{E_C}{E_{ASE}} = \frac{E_0 G_D T G_{Tot}}{E_N A_T} = \frac{2E_0 f^2}{A_T \tau_p N b^2 K}$$

in which 'b' is the mirror dimension of a beamlet and K = 2×10^4 $Wcm^{-2}sterad^{-1}$. This expression shows that the driver amplifier input energy is the best lever to obtain an increase in R, given the constraints that are associated with each of the other parameters (but see Appendix 1). If there are several stages of amplification, it is the first amplifier and the first feed array dimension that are important. In the present design ASE is removed from each beamlet after the preamplifier by

Pockels cell switches, making the driver amplifier the dominant ASE source.

In the baseline 2 MJ design we have chosen the following parameters:

f (focal length of final lens)	750 cm
A_T (Target cross section, dia. = 0.36 cm)	0.10 cm^2
τ_p (Amplifier pump duration)	250 ns
N (Number of multiplexed beamlets)	36
b (aperture of feed array mirrors)	15 cm

This set of parameters in the above expression leads to $R = 2.9 \times 10^5 E_0$. Consider for example the requirement for an ASE fluence of < 1J cm^{-2} on a target of surface area 0.407 cm^2 in a 2MJ facility. The coherent-to-ASE energy ratio has to equal $2.0 \times 10^6/(1 \times 0.407) = 4.91 \times 10^6$. Equating this to R from above, we determine that the driver amplifier input energy E_0 must be at least 18J to keep the ASE fluence below 1.0 J cm^{-2}. The design assumes E_0 = 30J, implying an ASE fluence of 0.6J cm^{-2}.

As E_0 is increased to achieve a specified ratio R, the capabilities of discharge-pumped preamplifiers (to generate E_0) are exceeded. An electron beam pumped preamplifier is needed for the 2 MJ system, in which E_0 is 30 J for each driver amplifier, and there are 16 required, making a total requirement of 480J, not including losses. Following this preamplifier each of the 36 beamlets has to be 'cleaned up' by Pockel's cell switching in order to remove the ASE, before beamsplitting to the 16 driver amplifiers. This involves 36 cells each of 15 cm aperture, a requirement that seems very feasible in view of the development of plasma-switched Pockels cells [26] and the availability of large electro-optic crystals.

AMPLIFICATION FIDELITY

We investigate amplification fidelity through a 34 kJ (1m^2) sub-aperture amplifier representing 1/2 of the 68 kJ building block module. One pulse shape treated in detail is that for a "baseline" low-isentrope, high gain target using a 3.3 MJ laser, provided by A. Schmitt. It is scaled in time according to $E^{1/3}$, where E is the energy on target. Figure 22 shows the unmodified pulse and three examples of a derived pulse labeled 'mod2', scaled to three different energies. The 'mod2' shape is identical to the baseline shape on the rise, but departs from it on the trailing edge.

The full width at half maximum of the mod2 pulse can be the same as the baseline pulse (e.g. 8.0 nsec, for 3.3 MJ), or it can be set to different values. The reasoning behind the 'mod2' shape is discussed below.

The first set of calculations using the amplifier code involved the unmodified baseline profile, scaled to 1.8 MJ (i.e. reduced in length from 8ns to 6.5 ns by a scaling factor of $(3.3/1.8)^{1/3} = 1.224$). For a train of pulses the amplification fidelity was satisfactory in the region of 6.5 nsec separation (Figure 23), growing worse for lesser or greater separations. This was because the input pulse train had substantial fluctuations and the amplifier was therefore not uniformly "loaded". As a general rule of thumb, a local variation in the pulse shape of less than a factor of two is probably not important for the target performance, as long as the variation does not change the integrated energy delivered to the pellet. However, as we shall show, it is not difficult to tune the rising part of the pulse shape with extreme accuracy, well beyond the requirements of the target designs.

The 'mod2' shape was derived from the baseline shape by the process illustrated in Figure 24. The idea is to modify only the trailing edge of the pulse to make it complementary to the rising edge, so that the sum of the rise of one pulse plus the fall of the previous pulse is a constant. In Figure 24, the portion of the rise below half maximum (labelled (a)) is reflected in a horizontal line HH at half maximum to give segment (a'). This segment is then translated by a time Δ equal to the pulse separation desired, to become the upper half of the trailing edge of the 'mod2' pulse. The same process is applied to the upper half of the rise (segment (b)), which reflects to become (b') and is translated by Δ to become the lower half of the 'mod2' trailing edge. (A small amount of adjustment is required to smooth the resulting profile in cases where the "foot" extends for several pulse durations). This idea is tested in the calculations presented in Figure 25. This time the 6.5 nsec separation together with a mod2 pulse width of 6.5 nsec, gave excellent fidelity.

The 'mod2' profile can be amplified without distortion provided its FWHM is set equal to the pulse separation. However, distortion is apparent when this condition is not satisfied. The most viable route to a flexible facility appears to be to compose pulses with the correct rise for the chosen implosion, but always extend them to have a FWHM equal to that of the pulse separation built in to the facility optics, and then give the pulses a trailing edge according to the formulation of Figure 24. The penalty associated with this approach is a certain amount of inefficiency, related to the unused part of the profile near the trailing edge. As an example, a 1.0 MJ profile (time factor 1/1.489 and FWHM = 5.3 nsec) is extended

to 6.5 nsec FWHM and given a complementary trailing edge. The amplification shows excellent fidelity (Figure 26). The unused energy amounts to about 15% of the total.

The multiplexed facility, once set to a given pulse separation, is therefore limited to a certain peak power that cannot be exceeded. Experimental programs involving the use of accurately tailored pulses must therefore be designed under a peak power constraint.

In the 2 MJ facility there are at least two electron beam-pumped amplifier stages. Each stage has to have the same interpulse separation for propagation through the whole system without distortion. The present calculation was for a single stage, but multiple stage propagation follows subject to the same constraints on pulse shape, as has been verified in calculations for the Nike system.

In conclusion, arbitrary smooth pulses, including the low-isentrope profile, can be amplified with excellent fidelity provided they are built up to the correct FWHM and have a complementary trailing edge. The facility can study lower energy, smaller (or less massive) target implosions at a relatively small cost in overall efficiency.

FACILITY DESIGN: GENERAL

Given the 136 kJ multiplexed beamline as a building block, the following requirements and constraints were taken into account in the design of the 2 MJ facility:

1. Illumination must be spherically uniform from 60 or more directions. As discussed below, symmetrical illumination (0.3% rms deviation) can be achieved with 64 lens arrays, fed by a laser on four floors, at sixteen lens arrays per floor.
2. The pulse shape should be variable but > 90% of energy should be delivered in 7 nsec. Amplification fidelity is discussed above.
3. The target chamber should have an inner radius of 6 m, to diminish the intensity of blast loading on the inside wall.
4. The line from an entry lens through the target should not project to a lens on the opposite side of the chamber.
5. First wall neutron moderation has to be provided.
6. X-ray, gamma ray and neutron shielding is required in the form of a heavy concrete enclosure with indirect beamline penetrations that are as small as

possible.

7. The chamber should be able to be assembled (by welding) in situ from simple components which are easy to fabricate. This is achieved with a cylindrical body capped by truncated cones. The cylinder and cones may be approximated by 32 plane elements.
8. The 4x36 =144 beamlets of a given beamline (total 136 kJ crude) emerge, after de-multiplexing, in an array 6 wide x 24 high, of 18 cm square mirrors. Its total extent, allowing for spaces, is 1.2m wide by 4.8m high. This array is propagated in a helium atmosphere to one of the beam penetrations in the target chamber enclosure.
9. Although kept to a minimum, each beam experiences on average 2.5 more reflections before entering a target chamber lens. These are necessitated by the need to introduce the correct relative delay and to distribute to four different principal lens arrays from each beamline.

Target chamber and illumination geometry Illumination from 60 or more directions immediately brings to mind the "buckyball" structure with its 60 vertices at almost perfectly uniform angular spacing from one another. However, the symmetry of this structure does not mate well with the symmetry of conceivable compact arrangements of multiplexed beamlines. By increasing the number of target beams to 64 it is possible to avoid a complex "optical switchyard" and feed the beamlines to target very directly and symmetrically from a building with twofold symmetry and four floors.

The 64 beam geometry is shown in Figure 2. There are 16 beams per floor and four floors. In terms of polar angle (α) and azimuthal angle (β) the 64 beam directions are:

Polar angle α (degrees)	Azimuthal angleβ (degrees)
24.25	11.25 + 90n
34.25	56.25 + 90n
44.25	11.25 + 90n
54.25	56.25 + 90n
64.75	78.75 + 90n

73.75	33.75 +90n
80.75	56.25 + 90n
86.75	11.25 + 90n
93.25	78.75 + 90n
99.25	33.75 + 90n
106.25	56.25 + 90n
115.25	11.25 + 90n
125.75	33.75 + 90n
135.75	78.75 +90n
145.75	33.75 + 90n
155.75	78.75 + 90n

where n = 0 --> 3.
These divide into four sets of 16 which are arranged on four floors. The two middle floors have identical plans while the upper and lower floors have mirror image plans. All the beams strike the opposing inside wall of the chamber at locations away from lens arrays.

The low-mode uniformity of illumination that can be achieved with this arrangement is 0.3% rms deviation from the mean, ignoring fluctuations within each laser beam. This can be derived by considering a test direction (θ,ϕ) which makes an angle ψ with a beam direction (Figure 27). Then

$$\cos\psi = \sin\theta\cos\phi\sin\alpha\cos\beta + \sin\theta\sin\phi\sin\alpha\sin\beta + \cos\theta\cos\alpha$$

If $\cos\psi$ is positive then the beam contributes to the intensity pattern at an effective radius of $\sin|\psi|$, for a target surface at unit radius.

We consider a hypergaussian (N) intensity distribution in a beam

$$I(\rho) = e^{-\frac{\rho^N}{\gamma^N}}$$

and γ is the 1/e intensity radius of a beam.
For the summed intensity we evaluate for $\cos\psi > 0$

$$\sum_{\alpha,\beta} e^{-\frac{|\sin\psi|^N}{\gamma^N}} \cos\psi$$

For example, with $\gamma = 1$ and $N = 4$ we obtain an average intensity of 11.95 vs the theoretical 16 which would occur for $\gamma << 1$, and an rms deviation of 0.3% from

this average, ignoring fluctuations within each beam profile and ignoring refraction. More detailed studies of uniformity will be presented in later papers.

Final lens arrays and two-photon absorption Each of the 64 target illumination beams corresponds to a crude energy of 34 kJ at an amplifier module. By the time a beam has reached the entry surface of a final focusing lens array, mirror losses have reduced this to 97% of 34 kJ = 33 kJ. The beam energy at the inner surface of one array of focal lenses is 2.0 MJ / 64 =31.2 kJ. There is therefore provision for a final lens transmission of 31.2 / 33 = 94%. Allowing 0.5% loss due to imperfect anti-reflection coatings the balance of the loss (5.5 %) is assigned to two-photon absorption in the fused silica lens material.

The two-photon absorption coefficient at 249 nm of fused silica is 0.06 cm GW^{-1} [28] whereas that for the alternate lens material of calcium fluoride is much lower (0.008 cm GW^{-1} [29]). The aperture of the lens for a single beamlet is 18cm x 18cm and the whole 36 lens array is 120 cm x 120 cm, when the centers of adjacent lenses are spaced by 20 cm (Figure 28). The square lenses are each cut from a disc which has to be 27.5 cm in diameter. Although CaF_2 lenses can be made to this size, there is much more experience with fused silica. The baseline design therefore considers fused silica.

Two-photon absorption, rather than optical damage, is the determinant of the fluence on the final lens. The lenses have to be at least 0.28 cm thick to give sufficient optical power, but the vacuum load demands considerably greater thickness than this. Without detailed mechanical calculation we choose 1.0 cm to be the maximum lens thickness, tapering to 0.7 cm at the edges.

Ray trace calculations have given the intensity distribution on the final lenses (Figure 19). For 33 kJ/36 incident in a baseline pulse shape on a given lens, the central intensity is I_{max} = 0.5 GW cm^{-2}. The two photon absorption is therefore

$$I_{max} \times 0.06 \times 1.0 \times 2 = 0.06, \text{ or } 6\%$$

The factor of two comes from the second moment of the distribution for polarized thermal light $P(I) = (1/I_0)\exp(-I/I_0)$, where I_0 is the average intensity.

The above numbers correspond to a fluence at the lens center of 4 J cm^{-2}.

Structure of target chamber. We have not separately considered neutronics or radiation leakage in this study, but these subjects are discussed more fully in the Los Alamos LMF report [8]. Our target chamber design is modular, intended for in-situ assembly. It has hollow walls of thickness 1.5 m which can in principle be

filled with borated water to moderate fusion neutrons. The walls consist of flat sections with sufficient area to hold the lens arrays. Aluminum 5083 alloy with low iron content (<0.01%) has been proposed [30] for the construction of a low-activation target chamber.

The target chamber is contained within a concrete radiation shield through which there are only indirect penetrations for the optical beams (Figure 2). Higher radiation attenuation can be achieved if necessary by the introduction of an additional turning mirror on each beam, and a further right-angle jog in each of the beam paths. Whether this will be necessary will depend on the outcome of more detailed radiation calculations.

The "first wall" within the target chamber has to be blast and radiation resistant. In front of the lenses it may be necessary to deploy light shrapnel shields (presumably of anti-reflection coated fused silica). The rest of the walls may have to be covered by replaceable tiles of refractory material.

Access to the targets is mainly from above and below via the two polar regions that are free of lenses and mirrors.

The major system parameters for the LANL and NRL designs are compared in Table VIII.

EXPERIMENTAL TEST OF 68 kJ MODULE

Before a KrF microfusion facility can go into detailed design it will be necessary to build and test an amplifier module with a nominal output of 68 kJ.

The assembly of two of these modules into a 136 kJ beamline is not considered to carry much risk, and therefore it can represent one of the programmatic stages in the construction of a 2.0 MJ facility. The information that construction of a beamline will generate relates principally to the accuracy and cost efficiency of the optical engineering design, including such elements as figuring, mounting, automated alignment and the provision of a helium atmosphere. Modifications can then be made before 16 beamlines are built.

The 68 kJ module represents more than a factor of 10 increase in energy from previous experience with the Nike 60cm amplifier, which has generated more than 5 kJ. It has the same stage gain (100) as the 60 cm amplifier. The new elements in the design (relative to Nike) are:

1. Higher pump rate (800 kW cm^{-3} vs 400 kW cm^{-3}). This is a safe extrapolation of the amplifier theory as the model gives a very accurate fit to the Nike 20 cm data

at 580 kW cm^{-3} , and it fits gain and absorption coefficients measured in several other laboratories at pump intensities between 0.2 MW cm^{-3} and 1.5 MW cm^{-3}.

2. Larger aperture, coupled with higher pump rate, implies a moderate transverse gain that could in principle contribute to parasitic oscillations in the transverse direction. The worst case gain along one transverse diagonal ($100\sqrt{2}$ cm) occurs in the segment adjacent to the amplifier end mirror. It is calculated to be 28x, a level that can be controlled by absorbing surfaces. Larger diagonal gain (100x) is present in directions close to the optical axis, but this gain is comparable to that handled in the Nike 60 cm amplifier without parasitic oscillations.

3. Unpumped regions (3 x 48 cm). The use of segmentation is new, but does not involve much risk. The absorption from unpumped regions is predictable, being due to ground state fluorine molecules. The advantage of segmentation that it partially suppresses ASE losses has yet to be demonstrated experimentally. This is an effect at the 5% level.

4. The energy from a single cathode (83kJ vs 70kJ). The energy that can go into a fault mode is comparable to that in Nike.

5. The electron beam voltage (800kV vs 600kV). This is a small increase, not involving new technology.

6. The pump pulse duration (250 nsec vs 176 nsec). This might concern us via its additional effect on diode closure, but the closure velocity is the same as in Nike, because the magnetic guide field is the same, and the initial gap is also the same, so the additional 0.17 cm of closure will not be a significant change.

7. A new design of combined switch/bushing. This is proposed because a total of 1,024 switches and bushings are required in the facility. The new design, if successful, will substantially reduce the cost and complexity of the electron gun. If not successful the proven technologies of separate switch and bushing will have to be used.

8. A higher transmission electron beam window (70% vs 50%). This is made possible through the use of small spans (48 cm) and the "suspension bridge" design of foil support structure. The suspension bridge design has been prototyped in tests at Los Alamos in a span of 40 cm, with an estimated 87% geometrical transmission. The energy transmission would be less than this because of electron reflection from the laser gas.

9. Larger window aperture (100 cm square vs 60 cm square). The largest size of fuzed silica window made to date (with adequate optical homogeneity) is 100 cm square. The manufacturing availability of large quantities at 100 cm has yet to be

proved.

10. The specific lasing energy (17 J l^{-1} vs 8 J l^{-1}). Much higher specific energies have been achieved in prior work. The estimate for the limiting parameter of fluorine consumption comes from the kinetic code which has been accurately benchmarked to data up to 10 J l^{-1}.

In review, none of the technology extensions is particularly risky in its own right, and there is not apparently any synergy of risk between the different technology extensions listed above. The risk of the 68kJ module is only moderate because so many small technology extensions are incorporated simultaneously. Tests of the 68 kJ module stand out as the only major demonstration of feasibility needed for the 2.0 MJ laser fusion facility.

COST ESTIMATE

In order to allow comparison with the cost estimates for prior LMF designs (based on both KrF and glass lasers) we are using the framework previously generated by Bechtel in their report [10]. The Work Breakdown Structure (WBS) in Table IX follows the numbering system and principal categories used by Bechtel. In their report they separated the driver and driver-associated facilities from the overall site costing. We label with asterisks on the WBS all those (1992) costs or cost categories that were taken directly from their report. All costs quoted for years from 1992 up to 1995 are escalated by a 3.5% annual compound increase up to 1996. The contingency reserve of 25% that was applied to the present KrF facility cost is the same as that applied to all the 1992 designs, although it could be argued that KrF laser design has improved since 1992 and a lesser contingency reserve might be appropriate.

The present laser system is much more compact than the (3MJ) LANL KrF design of 1991 [8] and therefore fits into a smaller building. The need for a transition building is eliminated because the beamlines are generated close to the target chamber and brought directly to the chamber in a symmetrical geometry. The target building itself is much smaller, partly because the yield will be closer to 100 MJ than 1000MJ and the final optics can therefore be closer.

Support systems in the WBS have been scaled back from earlier LANL or LLNL estimates roughly in proportion to the laser energy.

Vendor facilitization in our case relates mostly to improving the manufacturing capacity for large dimension fused silica windows. Other categories

include minor efforts on pulsed power components such as the switch/bushing and laser beam control components.

Laser amplifier cost. A "bottom-up" estimate was made based on the drawings of Figures 9 through 12. The assembled and tested hardware for the prototype 68kJ module was estimated to cost $9.8m (net of design, procurement and management overheads) while the quantity of 32 was priced at $7.1m per unit. These overhead costs are included separately in WBS 1.0, as indicated in [10].

Optics cost. Two costing approaches were followed. In the first we used algorithms developed for the production of 900-quantity KrF elements [31]. In the second we started from the quoted actual "Beamlet" (LLNL) optics cost [9], and scaled the cost per unit area for larger or smaller elements in linear proportion to the transverse dimension of the element. We applied a quantity cost reduction factor of 2 to the Beamlet actual costs. In [9] a total cost reduction factor of 2.67 was proposed, resulting from both increased quantity and vendor facilitization.

There was reasonable agreement between the two optics costing approaches. The larger estimate ($132.92m, from the BDM algorithm) is used in order to obtain the most conservative cost estimate. The scaled Beamlet approach led to an estimate of $91.45m for the optics of the 2MJ KrF laser. The difference between these estimates is mostly due to a different estimate for cost of the large amplifier fused silica windows. Very small numbers of 100 cm fused silica windows have been made, so the cost of quantity production is not well known, and neither algorithm can be said to have been verified at this optical dimension. Most of the vendor facilitization (WBS item 4.3) relates to the development of techniques for the quantity production of these windows.

APPENDIX 1

Design choices. In some cases arbitrary design choices were made that could be modified by more detailed considerations. For example, the focal length of the final lenses was chosen as 7.5m, principally to reduce the size and cost of the target chamber building. If, for example, one chose a final lens focal length of 15m, rather than 7.5, the XDL would be reduced by a factor of 2, to a value closer to that used with Nike. The fewer XDL beams would be less sensitive to amplifier gain non-uniformity (although this is already largely eliminated by amplifying at the

Fourier transform plane), and the ASE and beam cross-talk due to scattering would be reduced by a factor of 4, but one would have to specify a higher optical quality for the optical train.

APPENDIX 2

Analytic estimate of ASE brightness. Consider KrF* to have a Gaussian spontaneous emission spectrum $\exp(-Q^2 \ln 2)$ where $Q = (\nu - \nu_O)/\nu_H$ represents the frequency difference from the central frequency normalised by the half width half maximum of the distribution. The associated stimulated emission cross section is

$$\sigma(\nu) = \sigma_0 e^{-Q^2 \ln 2}$$

and the spontaneous emission power per atom, per steradian, per unit frequency interval can be written as $P(\nu) = \beta\sigma(\nu)$

where

$$\beta = \frac{h\nu\sqrt{\ln 2}}{4\pi^{3/2}\sigma_0 \nu_H \tau_R}$$

and τ_R is the radiative lifetime.

At the exit of an amplifier of width W and length 2L (which may be folded) the ASE power per steradian per unit frequency interval can be written as

$$P_{ASE}(\nu) = \int_0^{2L} G(\nu,z)N(z)P(\nu)W^2 dz$$

where N(z) is the KrF* density at position z and $G(\nu,z)$ is the gain at frequency ν from position z to the exit of the amplifier, given by

$$G(\nu,z) = \exp[\gamma(\nu,z)]$$

where

$$\gamma(\nu,z) = \int_0^z [N(z')\sigma(\nu) - \alpha]dz'$$

in which α is the loss coefficient.

The ASE brightness per unit frequency interval at the amplifier exit is

$$\begin{aligned} B_{ASE}(\nu) = \frac{P_{ASE}(\nu)}{W^2} &= \int_0^{2L} G(\nu,z)N(z)P(\nu)dz \\ &= \beta\int_0^{2L} \exp\left\{\int_0^z [N(z')\sigma(\nu) - \alpha]dz'\right\} N(z)\sigma(\nu)dz \\ &= F\beta\int_0^{\gamma(\nu,2L)} \exp(\gamma)d\gamma \end{aligned}$$

where F is a factor slightly greater than unity. At the line center $F = \overline{N}\sigma_0 / (\overline{N}\sigma_0 - \alpha)$

where $\overline{N}$ is the spatially averaged KrF* population density. Thus

$$B_{ASE}(\nu) = F\beta\{\exp[\gamma(\nu,2L)]-1\} \cong F\beta G(\nu,2L)$$

The ASE brightness integrated over frequency is

$$B_{ASE} = \int_0^\infty B_{ASE}(\nu)d\nu = F\beta\int_0^\infty G(\nu,2L)d\nu$$
$$= F\beta\int_0^\infty \exp\left\{\int_0^{2L}[N(z')\sigma(\nu)-\alpha]dz'\right\}d\nu$$

To approximately evaluate the integral we note that the maximum of the integrand is $G_0 = G(\nu_0,2L)$, the line center gain. The frequency half width of the ASE distribution, $Q_A\nu_H$, is found to good accuracy by solving

$$\exp\left[\ln G_0 e^{-Q_A^2 \ln 2}\right] = \frac{G_0}{2}$$

Approximating the integral by $2Q_A\nu_H G0$, we arrive at

$$B_{ASE} \cong \frac{G_0 h\nu}{2\pi\sigma_0\tau_R}\left(\frac{\overline{N}\sigma_0}{\overline{N}\sigma_0-\alpha}\right)\sqrt{\frac{1}{\pi}\ln\left(\frac{\ln G_0}{\ln G_0/2}\right)}$$

In this expression the square root term is a very weak function of G_O, so the ASE brightness depends linearly on the amplifier gain, as found in the detailed numerical simulation discussed above. Using $\sigma_O = 2.6 \times 10^{-16}$ cm^2; τ_R = 6.5 ns and a typical ratio $\overline{N}\sigma_0/\alpha = 7$, the following values of B_{ASE} are calculated in units of (Wcm^{-2}sterad^{-1}):

$G_0 = 25 \Rightarrow B_{ASE} = 2.45\times10^4 G_0$

$G_0 = 50 \Rightarrow B_{ASE} = 2.20\times10^4 G_0$

$G_0 = 100 \Rightarrow B_{ASE} = 2.00\times10^4 G_0$

These compare very closely with the code predictions given in Figure 21 for a typical amplifier on the scale of the facility driver amplifier. In summary, it is a good approximation for all electron beam pumped KrF amplifiers to calculate the exit axial ASE brightness from $B_{ASE} = 2\times10^4 G_O$ (Wcm^{-2}sterad^{-1}).

ACKNOWLEDGEMENTS

This work was supported by the US Department of Energy.

REFERENCES

[1] R. H. Lehmberg and S. P. Obenschain, "Use of Induced Spatial Incoherence for Uniform Illumination of Laser Fusion Targets", Opt. Commun., **46**, 399 (1983).

[2] R. H. Lehmberg and J. Goldhar, "Use of Incoherence to Produce Smooth and Controllable Irradiation Profiles with KrF Fusion Lasers", Fusion Tech. **11**, 532 (1987).

[3] T. Lehecka, R. H. Lehmberg, A. V. Deniz, K. A. Gerber, S. P. Obenschain, C. J. Pawley, M. S. Pronko and C. A. Sullivan, "Production of High Energy, Uniform Focal Profiles with the Nike Laser", Opt. Comm. **46**, 485 (1995).

[4] S. P. Obenschain, S. E. Bodner, D. Colombant, K. Gerber, R. H. Lehmberg, E. A. McClean, A. N. Mostovych, M. S. Pronko, C. J. Pawley, A. J. Schmitt, J. D. Sethian, V. Serlin, J. A. Stamper, C. A. Sullivan, J. P. Dahlburg, J. H. Gardner, Y. Chan, A. V. Deniz, J. Hardgrove, T. Lehecka and M. Klapisch, "The Nike KrF Laser Facility: Performance and Initial Target Experiments", Physics of Plasmas **3**, 2098 (1996).

[5] W. R. Meier, ""OSIRIS and SOMBRERO, Inertial Fusion Power Plant Designs - Summary, Conclusions and Recommendations", Fusion Eng. Des. (Netherlands) **25**, 145 (1994), also W. R. Meier et al., US Department of Energy Report DOE/ER/54100-1 WJSA(1992).

[6] J. J. Ewing, R. A. Haas, J. C. Swingle, E. V. George and W. F. Krupke, "Optical Pulse Compressor Systems for Laser Fusion", IEEE J. Quant. Electron., **QE-15**, 368 (1979).

[7] J. H. Parks, "Conceptual Design for an Angularly Multiplexed Rare Gas Halide Laser Fusion Driver", DoE Report DOE/DP/40113-1, Avco Everett Research Laboratory (1980).

[8] D. B. Harris, J. A. Sullivan, J. H. Pendergrass, N. A. Kurnit, E. A. Rose and J. McLeod, "KrF Laser-Driven Laboratory Microfusion Facility" Report LA-

UR-91-2915, Los Alamos Laboratory, Dec. 1991.

[9] "Manufacturing Readiness Plan for the National Ignition Facility", Lawrence Livermore National Laboratory Report NIF-LLNL-94-204 (April 1994).

[10] "Laboratory Microfusion Capability Cost Study", Bechtel Corporation (US Department of Energy, August 1991).

[11] D. C. Thompson, R. Fedosejevs, A. A. Offenberger, J. P. Santiago and H. R. Manjunath, "Measurements of Gain and Absorption Saturation in an Electron-Beam Pumped KrF Amplifier", IEEE J. Quant. Electron. **QE-25**, 2161 (1989).

[12] A. Mandl, D. Klimek and J. H. Parks, "KrF Laser Kinetics Studies", J. Appl. Phys. **55**, 3940 (1984).

[13] E. Salesky and W. Kimura "Experimental Support for KrF Laser Modelling", Final report from Spectra Technology to LANL under contract 9-X65-W1478-1 (1989).

[14] K. Ueda, A. Sasaki, H. Yoneda, and H. Takuma, "Scalability of High Power KrF Laser Systems", LANL Workshop on KrF Laser Technology, Santa Fe (1989).

[15] C. B. Edwards, F. O'Neill and M. J. Shaw, "Absorption and Gain Measurements in the KrF Laser Medium at High Pump Rate", Appl. Phys. Lett. **38**, 843 (1981).

[16] M. W. McGeoch, "A Simplified KrF Kinetics Model", Proc. 4th Intnl. Workshop on KrF Laser Technology, Annapolis, (U. Maryland / NRL 1994).

[17] M. W. McGeoch, "Time-Dependent Amplifier with Kinetics and ASE", ibid. ref 16.

[18] C. Pawley, "Measurements of Electron Beam Pumping and Gain", ibid. ref 16.

[19] J. D. Sethian, S. P. Obenschain, M. W. McGeoch, I. D. Smith, P. A. Corcoran and R. A. Altes, "Pulsed Power Requirements for a High Power KrF Laser for a High Gain Inertial Confinement Fusion System", to be published in the proceedings of the 22nd Intnl. Power Modulator Symp., Boca Raton, Florida, June 24-27 (1996).

[20] R. J. Adler, "Pulsed Power Formulary", North Star Research Corporation, Albuquerque, NM (1989).

[21] G. York, "An Advanced Electron Beam Hibachi - Design and Performance", ibid. ref 16.

[22] A. Kidd, "Pulsed Power for the 42 cm TITANIA Amplifier", ibid. ref 16.

[23] J. Fockler, B. Bowen, V. Carboni, P. Corcoran, J. Kishi and R. Kuenning, "A 4 MV ± 1% Flat-Top Electron Diode Driver", p 177, Proc. Eighth IEEE Intnl. Pulsed Power Conf., San Diego, CA (1991).

[24] M. W. McGeoch, unpublished.

[25] R. H. Lehmberg, T. Lehecka and A. V. Deniz, "Optical Design of the Nike Laser", Proc. 2nd Workshop of KrF Laser Technology (Banff, Alberta, 1990, sponsored by U. Alberta).

[26] Y. Leng, V. Yun, J. Goldhar, A. Deniz, M. Pronko, and S. Obenschain, "Operation of Small and Fast Pockels Cells with Plasma Electrodes", ibid. ref 16.

[27] M. A. Rhodes, B. Woods, J. J. DeYoreo, D. Roberts and L. J. Atherton, "Performance of Large-Aperture Optical Switches for High-Energy Inertial Confinement Fusion Studies", Appl. Optics **34**, 5312 (1995).

[28] I. N. Ross, W. T. Toner, C. J. Hooker, J. R. M. Barr and I. Coffey, "Nonlinear Properties of Silica and Air for Picosecond Ultraviolet Pulses", J. Mod. Optics **37**, 555 (1990).

[29] A. J. Taylor, R. B. Gibson and J. P. Roberts, "Two-photon Absorption at 248 nm in Ultraviolet Window Materials", Opt. Lett. **13**, 814 (1988).

[30] R. F. Bourque, E. T. Cheng, R. L. Creedon, K. R. Schulz and D. L. Sonn, "Innovative Design Concepts for the LMF Target Chamber and Related Systems", report GA-A19651 (June 1989).

[31] "Cost Reduction Study for the LANL KrF Laser-Driven LMF Design", BDM report BDM/ABQ-89-0747-TR (Oct 1989).

Table I KrF kinetic processes and rates in simplified model (cgs units)

$k_2 = 2\times10^{-6}$	$G^+ + F^- \rightarrow KrF^*$
$k_3 = 2.5\times10^{-31}$	$G^+ + 2G \rightarrow G_2^+ + G$
$k_4 = 5\times10^{-8}$	$G_2^+ + e^- \rightarrow G^* + G$
$k_5 = 7.2\times10^{-10}$	$G^* + F_2 \rightarrow KrF^* + F$
$k_6 = 2\times10^{-9}$	$F_2 + e^- \rightarrow F^- + F$
$k_7 = 1\times10^{-6}$	$G_2^+ + F^- \rightarrow KrF^*$
$k_{7a} = 1\times10^{-6}$	$G_2+ + F^- \rightarrow Kr_2F^*$
$k_8 = 8\times10^{-10}$	$KrF^* + F_2 \rightarrow Kr + 3F$
$k_9 = 3\times10^{-7}$	$KrF^* + e^- \rightarrow Kr + F + e^-$
$k_{10} = 2\times10^{-31}$	$KrF^* + Kr + G \rightarrow Kr_2F^* + G$
$k_{11} = 8\times10^{-32}$	$KrF^* + Ar + G \rightarrow Kr_2F^* + G$
$k_{12} = 2.6\times10^{-16}$	$KrF^* + h\nu_{249} \rightarrow Kr + F + 2h\nu_{249}$
$k_{13} = 1.54\times10^{8}$	$KrF^* \rightarrow Kr + F + h\nu_{249}$
$k_{14} = 1.6\times10^{-18}$	$Kr_2F^* + h\nu_{249} \rightarrow 2Kr + F$
$k_{15} = 6.1\times10^{6}$	$Kr_2F^* \rightarrow 2Kr + F + h\nu'$
$k_{16} = 1\times10^{-10}$	$Kr_2F^* + F_2 \rightarrow 2Kr + 3F$
$k_{18} = 5.6\times10^{-18}$	$F^- + h\nu_{249} \rightarrow F + e^-$
$k_{19} = 3\times10^{-18}$	$G_2^+ + h\nu_{249} \rightarrow G^+ + G$
$k_{20} = 1.5\times10^{-18} \times (Ar + 5*Kr) / (Ar + Kr)$	$G^* + h\nu_{249} \rightarrow G^+ + e^-$
$k_{21} = 1.4\times10^{-20}$	$F_2 + h\nu_{249} \rightarrow 2F$

(significant absorbers: G^*, F^-, F_2, G_2^+, Kr_2F^*)

Table II: Nike 60 cm amplifier data comparison

Shot#	Corrected Deposition (kJ)	Specific Power ($kWcm^{-3}$)	Mix (Total/Kr/F_2) (Amag/fract./%)	Measured Output (kJ)	Calc. Output (kJ)
471	55.93	346	1.08/0.31/0.47%	3.637	4.24
473	68.74	426	"	5.352	5.32
516	52.70	326	1.16/0.35/0.32%	2.732	3.70
527	48.69	302	1.16/0.35/0.35%	2.976	3.40
528	58.81	365	"	4.244	4.17
533	62.82	389	"	3.812	4.48
534	44.16	274	"	2.784	3.04
536	53.29	330	"	3.229	3.74
568	39.22	243	1.16/0.35/0.32%	2.24	2.38
571	30.09	187	1.02/0.27/0.37%	1.80	2.12
573	22.51	139	0.95/0.21/0.39%	0.85	1.54
577	24.84	154	1.02/0.27/0.37%	1.0	1.67
583	51.71	321	1.16/0.35/0.32%	3.7	3.5
584	"	"	"	3.9	3.7
587	"	"	"	3.5	3.3
588	"	"	"	3.6	3.4
589	"	"	"	3.7	3.5
5 shot average				3.68	3.48

Table III: Amplifier output and efficiency vs number of segments

Number of Segments	Deposited Energy (kJ)	Output Energy (kJ)	Intrinsic Efficiency (%)
3	288	29.0	10.1
4	384	36.6	9.5
5	480	42.9	8.9

Table IV: Risetime and flat top as a function of inductance for 150 nsec pulseline with 20 nsec peaker

Inductance (nH)	Risetime to 90% (ns)	Width at 90% (ns)	Peak diode voltage (kV)
55	38	271	810
75	55	265	794
95	62	260	782

Table V: Summary of 68 kJ module parameters

Optical window panel	100 cm x 100 cm
Optical aperture (2 window panels)	100 cm wide by 267 cm high
Pumped segments (dead space)	4 x 48 cm (48 cm)
Total amplifier length	356 cm
Aspect ratio	3.6
Typical laser fill	0.8Amag. Ar , 30% Kr, 0.6% F_2
Cathode of a segment	48 cm wide by 100 cm high
Pump duration (total)	300 nsec
Extraction duration	250 nsec
Input energy into 100 cm x 200 cm	700J
Output energy from 100 cm x 200 cm	68 kJ (73kJ calculated, 7% contingency)
Fluence on window	3.4 J cm^{-2}
Diode voltage	800 kV
Current density at cathode	68 A cm^{-2}
Energy stored in one plate	83 kJ
Transfer efficiency into gas	0.7
Specific pump rate	800 kW cm^{-3}
Rise time	55 nsec
Line impedance (to one cathode)	2.56 Ω
Self magnetic field	1.7 kG
Initial / final anode-cathode gap	5.5 cm / 4.9 cm

Table VI Summary of Optics for 16 x 136 kJ Beamlines with Relay to Target

Element	Optic Type	Quantity	Shape	Width x Height (cm) x (cm)	Thickness (cm)	Extended Area (cm^2)
Blast Shields	Flat transmitter	256	square	60 x 60	0.5	9.2×10^5
Final Lens	Spherical bi-convex	2304	square	20 x 20	1.5	9.2×10^5
Final Mirrors	Flat mirror	4608	rectangle	20 x 28	-	2.58×10^6
Demultiplex mirrors	Flat mirror	2304	square	18 x 18	-	7.4×10^5
Recollimation mirrors	Spherical convex R = 19.3m	2304	square	15 x 15	-	5.2×10^5
Aperture share turning mirrors	Flat mirror	64	rectangle	110 x 150	-	1.06×10^6
Main amplifier windows	Flat transmitter	128	rectangle	110 x 120	5	1.68×10^6
Main amplifier reflectors	Spherical concave mirror R = 61.9m	64	square	110 x 110	-	7.7×10^5
Main amplifier feed array mirrors	Spherical convex mirror R= 6m	576	square	15 x 15	-	1.3×10^5
Driver amplifier windows	Flat transmitter	32	square	55 x 55	2	0.96×10^5
Driver amplifier reflectors	Spherical concave mirror R = 64m	16	square	55 x 55	-	0.48×10^5
Transport mirrors	Flat mirror	1728	rectangle	15 x 22	-	5.7×10^5
Beamsplitters	Flat transmission/ partial reflector	108	rectangle	10 x 15	1	0.16×10^5
Driver amplifier feed array mirrors	Spherical concave mirror R = 45m	36	square	15 x 15	-	0.08×10^5
Misc. small mirrors (front end)	-	-	-	-	-	-

Total mirror area = 6.4×10^6 cm^2 Total lens and window area = 3.5×10^6 cm^2

Table VII: Parameters in ASE calculations

	Nike	136 kJ Beamline
Beamline Energy	5 kJ	136 kJ (four beams on target)
Pulse duration	176 nsec	250 nsec
Driver amp gain	56	67
Driver amp Output	80J	2,000J
Beamsplit/Overfill factor T	0.5 (time doubling)	0.63 (overfilling and attenuation)
Beamlet count	44	36
Dimension of main amp feed array mirror	12 cm	15 cm
Main amp gain	60	100
Focal length	6 m	7.5 m
ASE brightness at driver amp exit ($W\ cm^{-2}\ sterad^{-1}$)	1.13×10^6	1.34×10^6
Ray trace ASE fluence	44 mJ cm^{-2}	-
Analytic ASE fluence	53 mJ cm^{-2}	145 mJ cm^{-2} (**580 mJ cm^{-2}** for 64 beams on spherical target)

Table VIII: Comparison of Facilities

Parameter	LANL LMF Dec 1991	NRL June 1996
Primary illumination geometry	Two-ended Conical	Spherical
Energy on target	3.0 MJ	2.0 MJ
Pulse (FWHM)	6.75 nsec	7.00 nsec
Amplifier duration	1080 nsec	250 nsec
Number of multiplexed beamlets	160	36
Pump rate	190 kW cm^{-3}	800 kW cm^{-3}
Module energy	380-412 kJ	68 kJ
Fluence on final lens	5 J cm^{-2}	4 J cm^{-2}
Total area of mirrors	8.2 x 10^6 cm^2	6.4 x 10^6 cm^2
Typical optical path length	900 m	260 m
Helium volume	150,000 m^3	28,000 m^3
Area of building	20,000 m^2	6,500 m^2
Volume of building	482,000 m^3	230,000 m^3
Facility Total Cost Estimate (same basis)	$1,059m (1996)	**$950m** (1996)

Table IX. Summary of Cost Elements for 2 MJ KrF Laser Facility.

WBS	Item	(1992) ($m)	(1996) ($m)
2.0	Site Improvements and Utilities		
2.1	Site improvements	1.95*	2.24
2.2	Utilities	4.21*	4.84
3.0	Support Facilities		
3.1	Driver dependent facilities		
	3.1.1 Driver buildings		52.30
	3.1.2 Target building		22.08
	3.1.3 Support buildings		5.00
3.2	Driver Independent facilities		
	3.2.1 Target fabrication building	2.76*	3.17
	3.2.2 Office	10.33*	11.88
	3.2.3 Laboratory	4.80*	5.52
	3.2.4 Shops	3.25*	3.74
	3.2.5 Warehouse	3.89*	4.47
	3.2.6 Security building	0.30*	0.35
3.3	Support Equipment (oil, DI, gas, shop, etc)		15.00
3.4	Radioactive waste storage and disposal	5.20*	5.98
4.0	Driver		
4.1	Laser systems		
	4.1.1 Main Amplifiers		226.24
	4.1.2 Driver Amplifiers		25.44
	4.1.3 Front End		4.88
	4.1.4 Optics		132.92
	4.1.5 Helium Equipment		5.00
4.2	Driver Diagnostics		14.90
4.3	Vendor Facilitization		10.0

Table IX, continued

WBS	Item		(1992)	(1996)
5.0	Experiment area equipment			
	5.1 Target chamber			8.98
	5.2 Diagnostics			17.25
	5.3 Target assist equipment			5.00
	5.4 Other			15.00
6.0	Support systems (computers/DaQ)			15.00
	Subtotal			617.18
1.6	Construction management*	3.00%		18.51
	Total field cost			635.69
1.0	Project office			
1.1	Management and administration*	3.50%		22.25
1.2	Systems engineering*	6.00%		38.14
1.3	Design engineering*	6.00%		38.14
1.4	Title III engineering*	1.50%		9.54
1.5	Procurement*	1.00%		6.36
1.7	ES&H Management*	0.50%		3.18
1.8	Quality Assurance*	1.00%		6.36
	Total project office			123.97
	Contingency*	25%		189.91
	Total		$825M (1992)	**$950M** (1996)

FIGURE CAPTIONS

Figure 1 Floor plan of 2.0 MJ KrF laser facility

Figure 2 Plan and elevation of one side of 2 MJ KrF laser facility showing major structural elements

Figure 3 Floor area of 2MJ KrF laser facility under various assumptions for the multiplexed beamline energy

Figure 4 Algorithm coupling kinetics and optical propagation

Figure 5 Comparison of experiment and theory for Nike 60cm amplifier

Figure 6 Side views of segmented square cross section amplifier

Figure 7 Output energy as a function of segment gap

Figure 8 Effect of segmentation of constant power electron beams. Double pass amplifier of 100cm x 100cm cross section, mix 0.8 Amag (Ar, 30% Kr, 0.6% F2), duration 250 nsec and input 140 kW cm^{-2} (350 J). The specific pump rate depends on the dimension of the pumped segments.

Figure 9 Plan and elevation of 68 kJ module, showing overall dimensions

Figure 10 Detailed plan of diode region. (Note pumped segments = 48 cm = spaces)

Figure 11 Detailed elevation of diode region

Figure 12 Bushing/switch. (a) mechanical, (b) electrostatics before switching, (c) electrostatics after switching and (d) after switching detail showing grad increments.

Figure 13 Current, voltage and power for 75 nH case. Note, two diodes per column.

Figure 14 Schematic of time and angle multiplex system. Distances and optics for 136 kJ beamline

Figure 15 Layout of 136 kJ beamline

2 x 68 kJ amplifier modules per 136 kJ beamline.
4 beamlines per floor.
4 floors each 6,500 sq. m (70,000 sq. ft).
64 beams to target, access in 4 tiers of 16.
6m internal radius target chamber.
Concrete shielding.

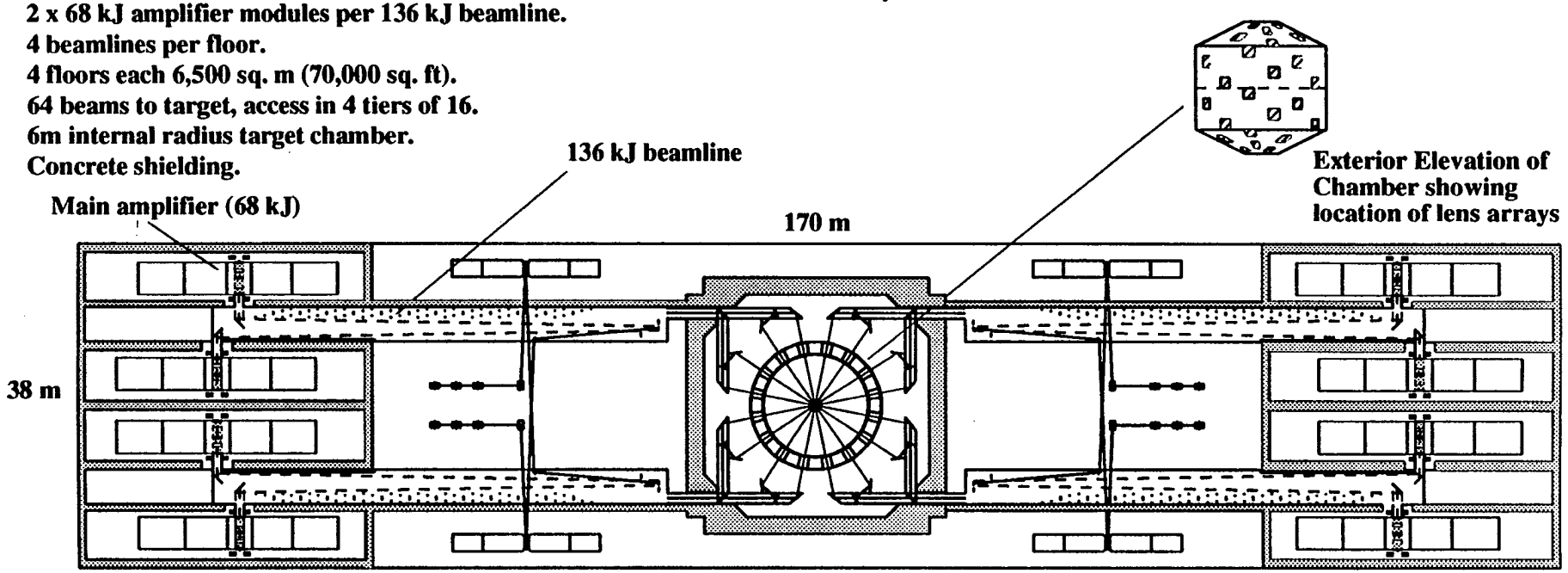


Figure 1 Floor Plan of 2.0 MJ KrF Laser Facility

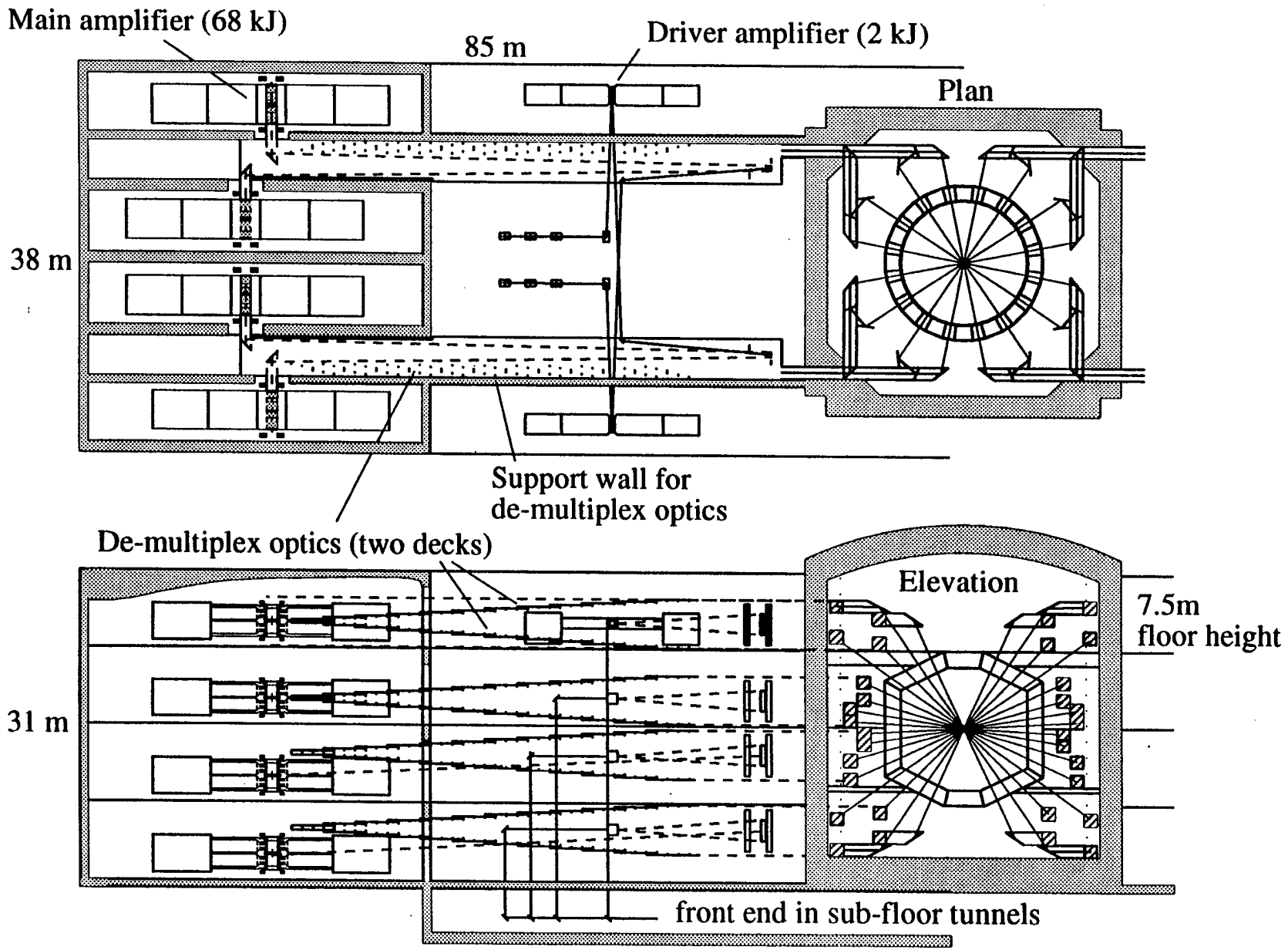


Figure 2 Plan and Elevation of one side of 2 MJ KrF Laser Facility showing major structural elements.

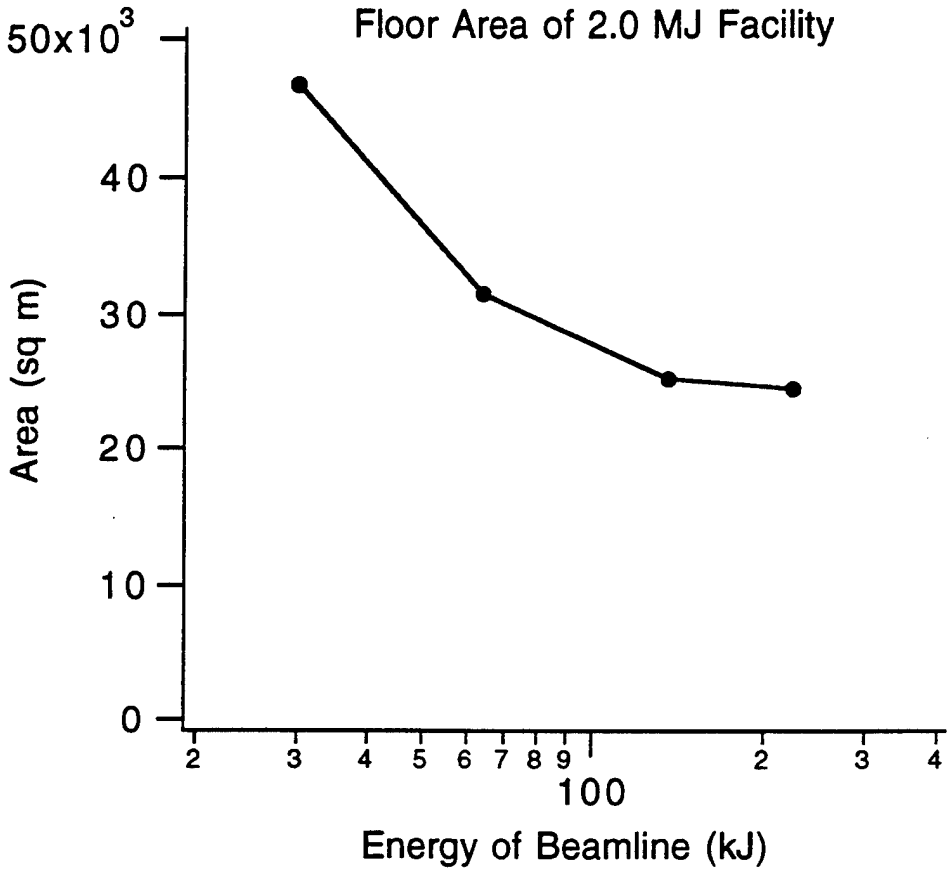


Figure 3 Floor area of 2 MJ KrF laser facility under various assumptions for the multiplexed beamline energy.

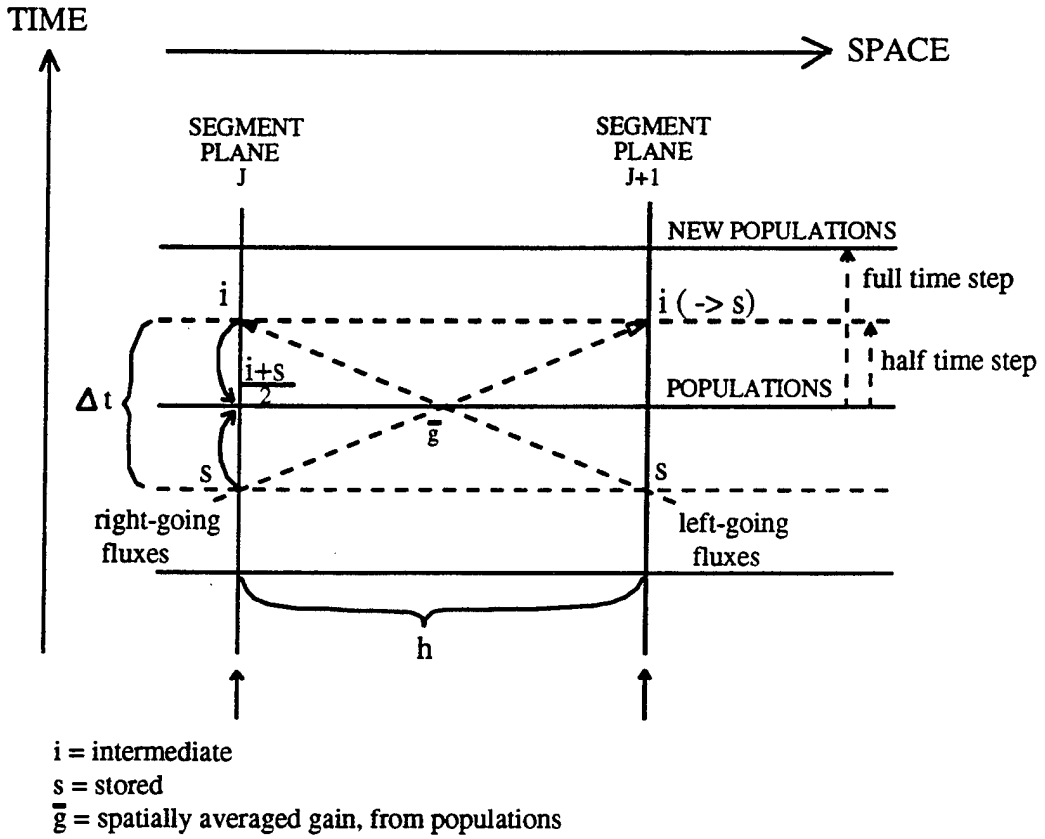


Figure 4 Algorithm coupling kinetics and optical propagation

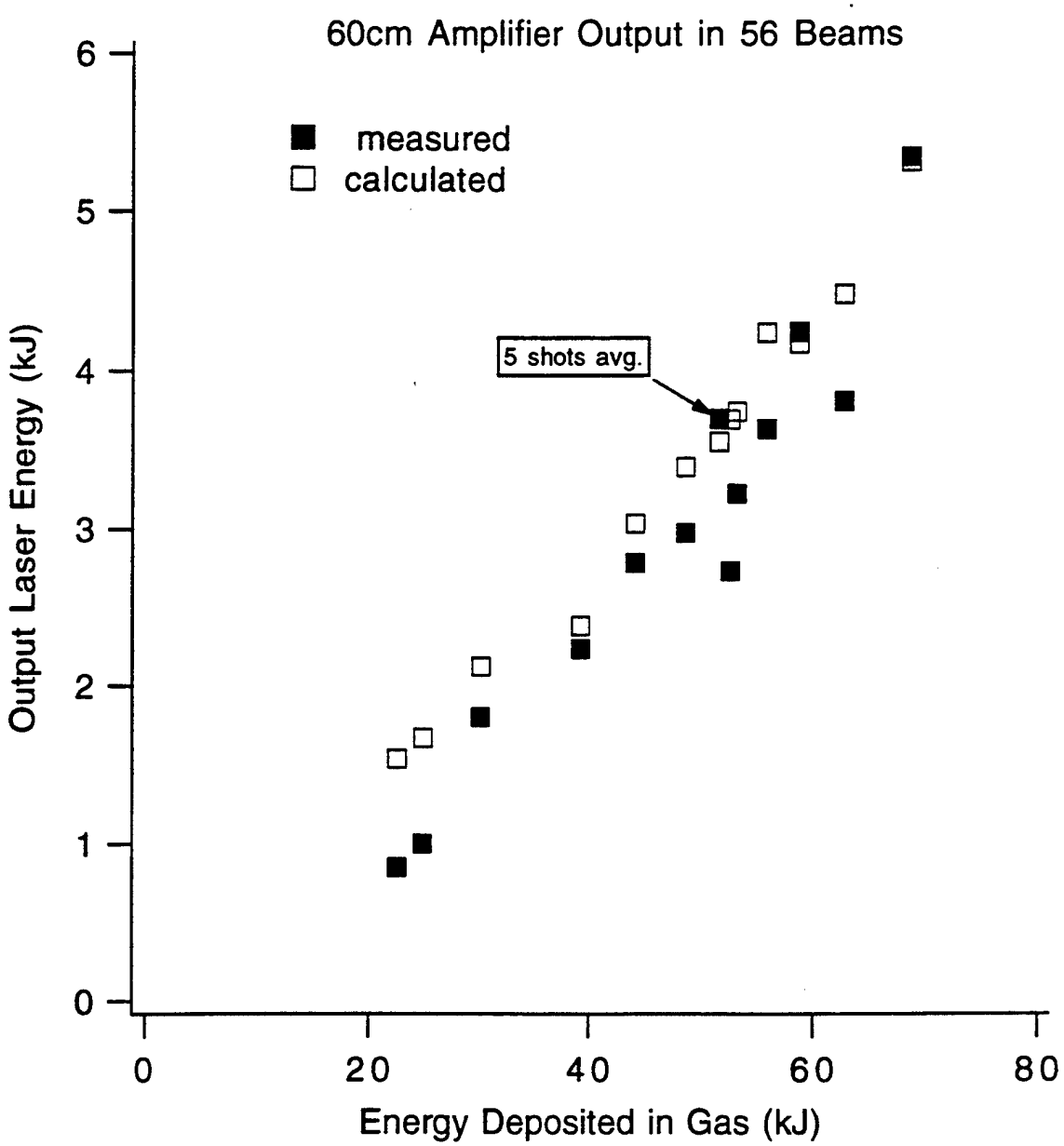


Figure 5 Comparison of experiment and theory for Nike 60cm amplifier

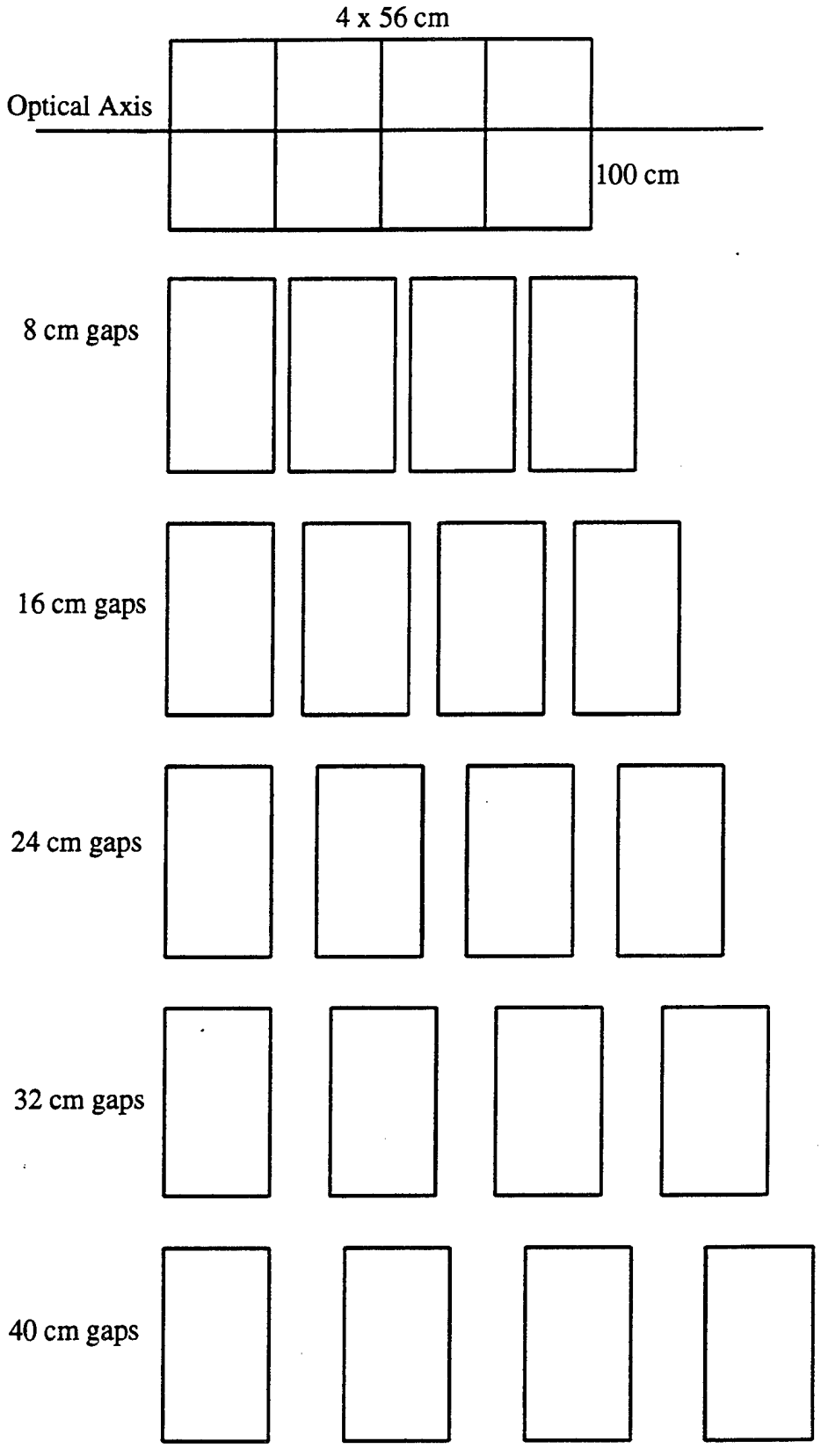


Figure 6 Side views of segmented square cross section amplifier

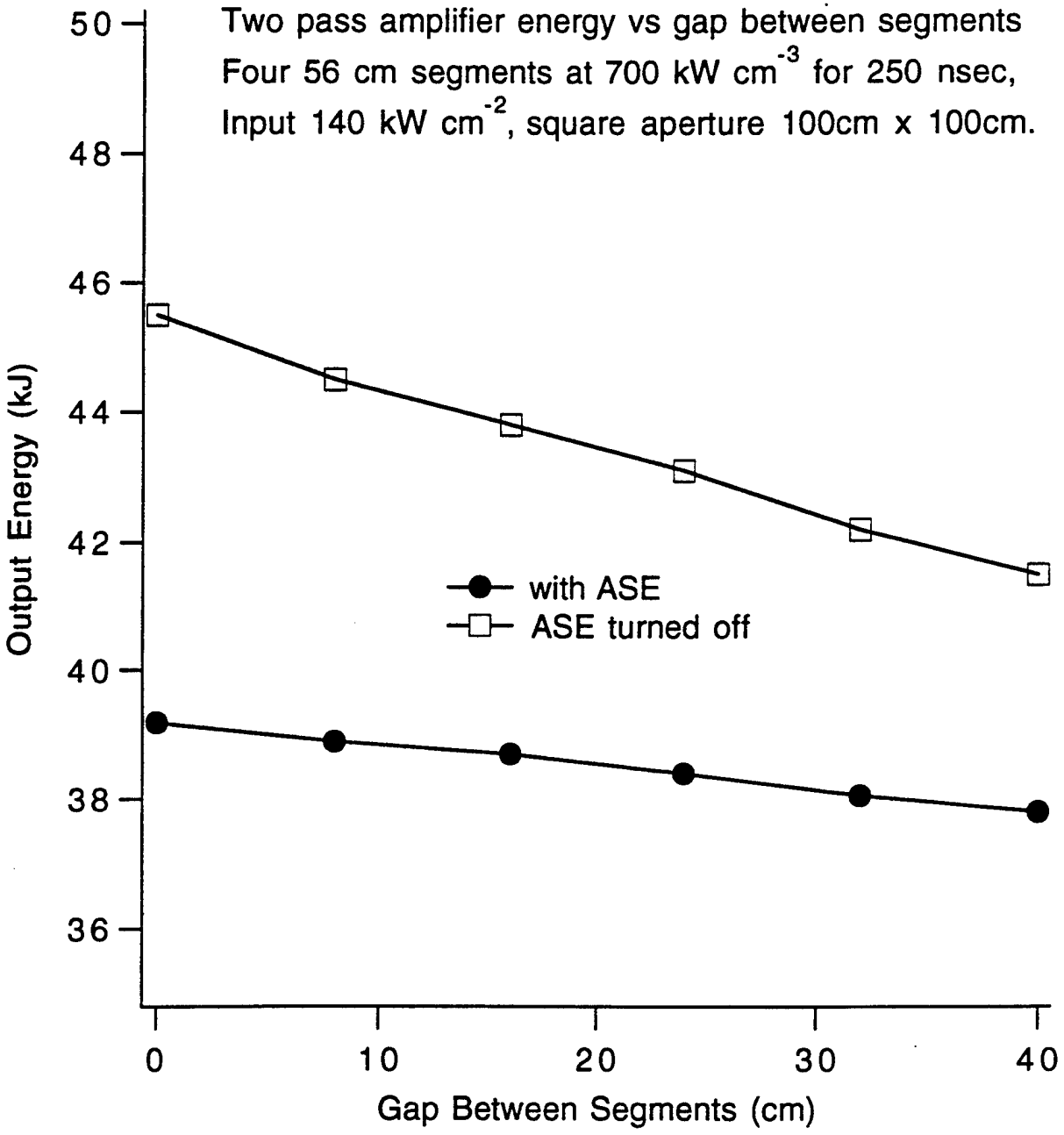
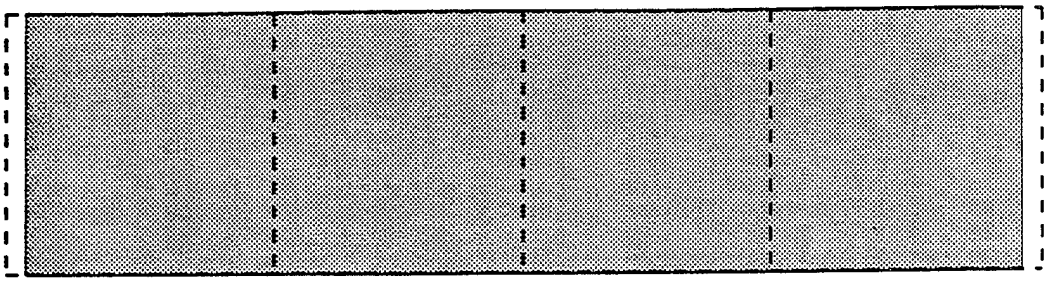
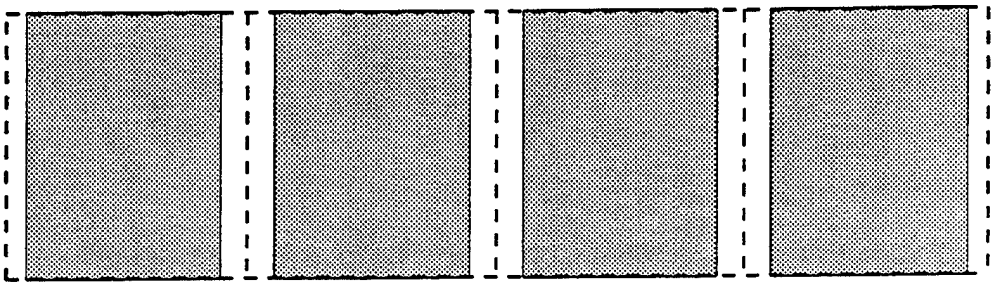


Figure 7 Output energy as a function of segment gap

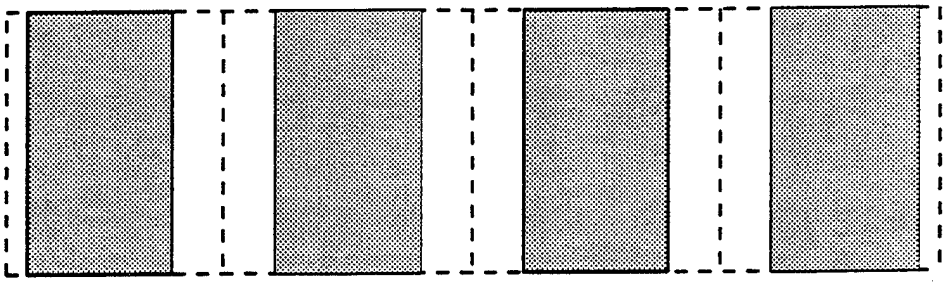
4 x 96 cm segments, no gaps, 0.41 MW cm-3, output 36.8 kJ



4 x 76 cm segments, 20 cm gaps, 0.51 MW cm-3, output 37.5 kJ



4 x 56 cm segments, 40 cm gaps, 0.70 MW cm-3, output 37.8 kJ



4 x 48 cm segments, 48 cm gaps, 0.82 MW cm-3, output 37.3 kJ

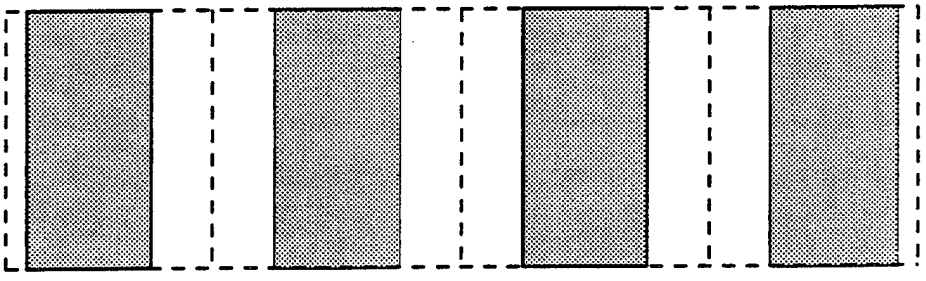


Figure 8 Effect of segmentation of constant power electron beams. Double pass amplifier of 100 cm x 100 cm cross section, mix 0.8 Amag (Ar, 30% Kr, 0.6% F2), Duration 250 nsec, Input 140 kW cm-2 (350 J). The specific pump rate depends on the dimension of the pumped segments.

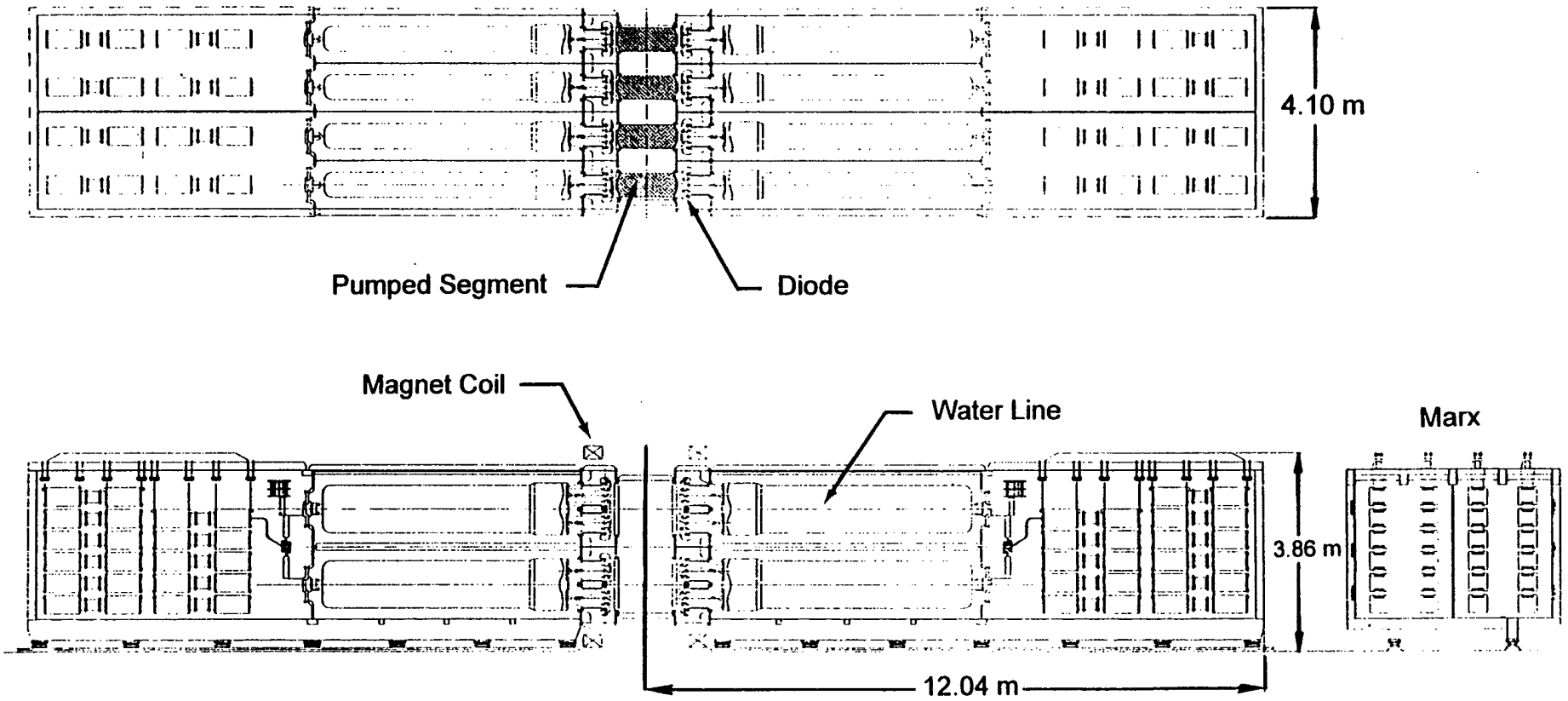
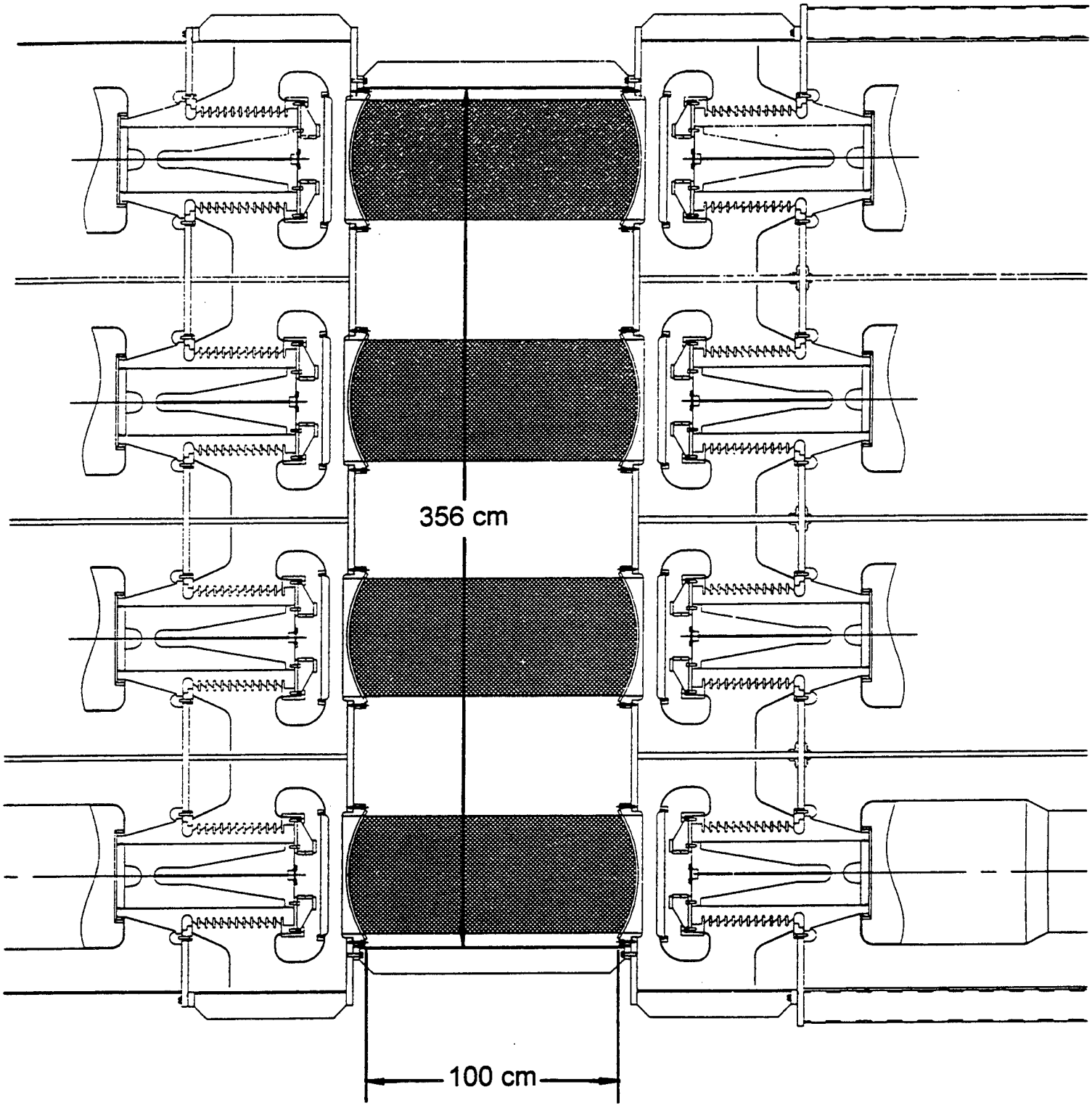


Figure 9. Plan and elevation of 68 kJ module showing overall dimensions.



(Note: pumped segments = 48 cm = spaces).

Figure 10. Detailed plan of diode region.

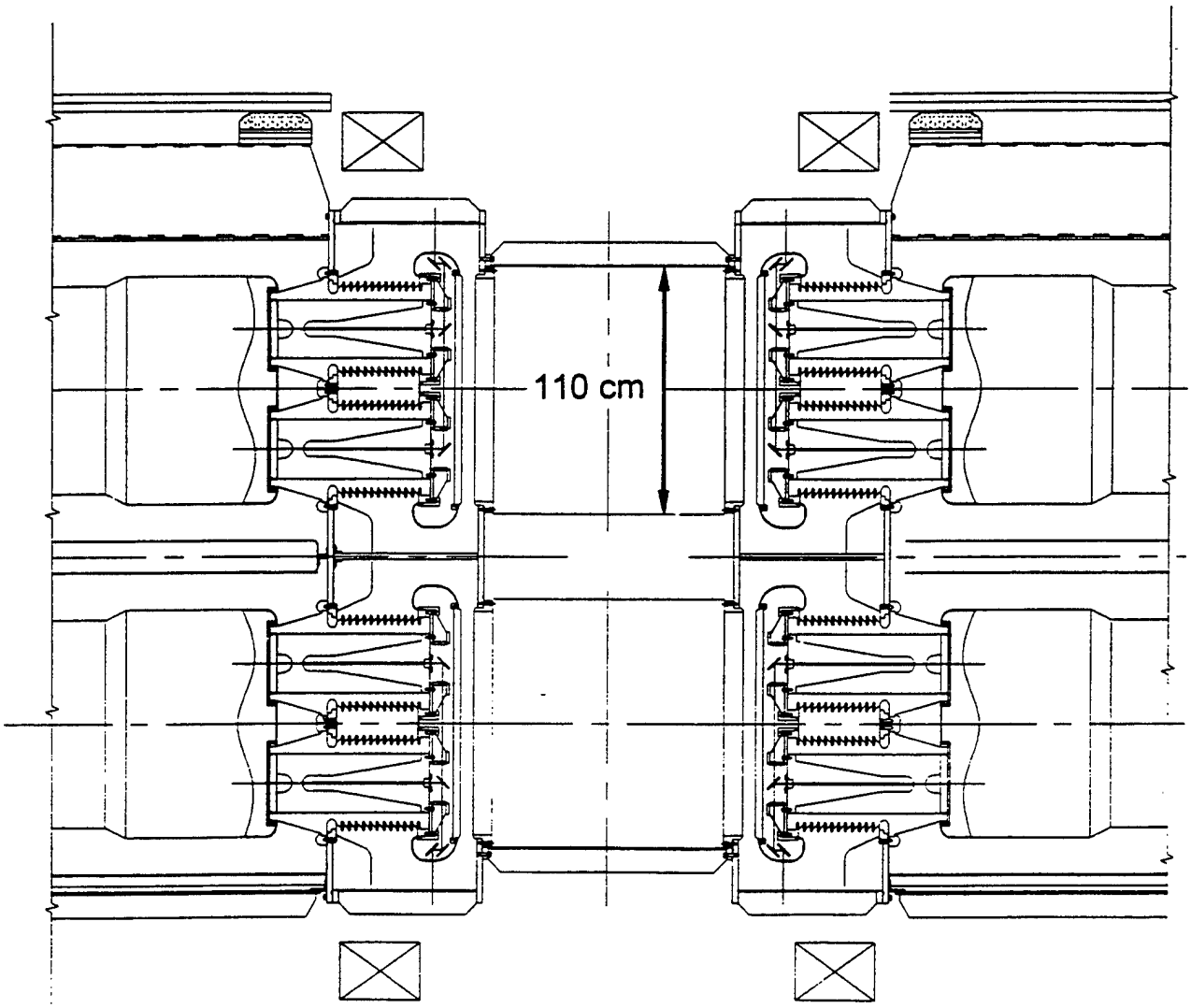
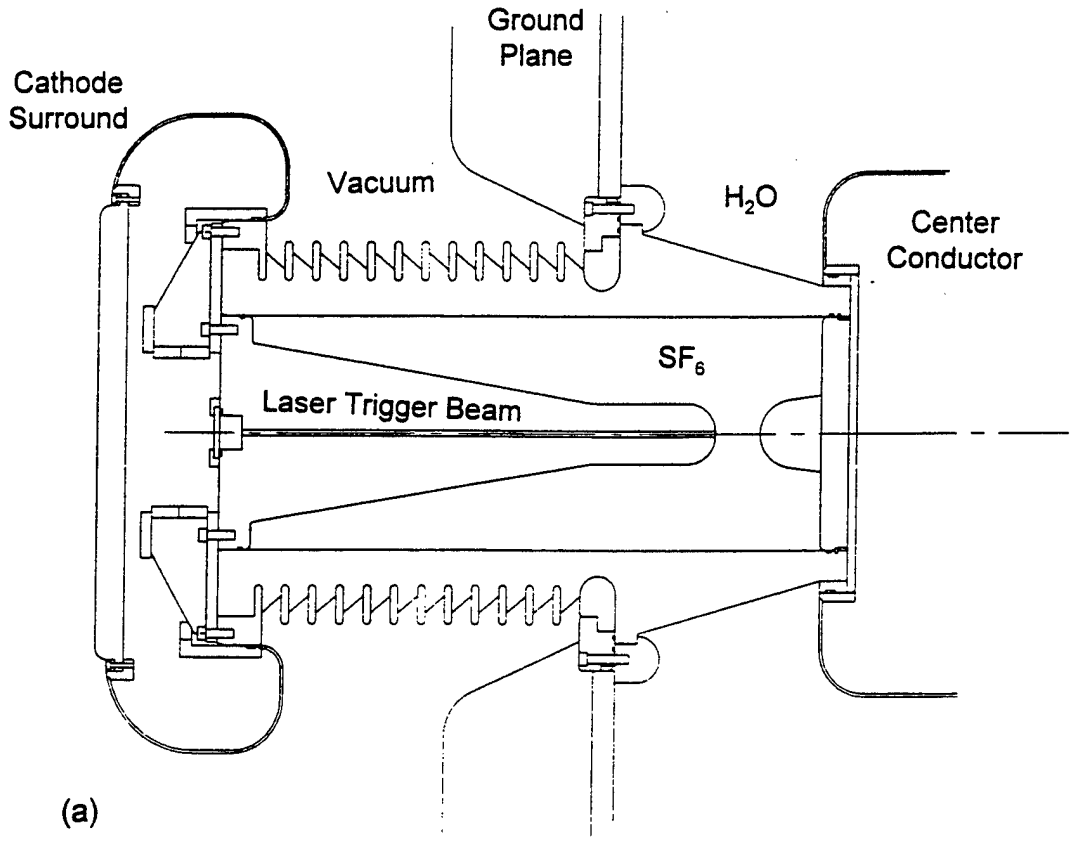
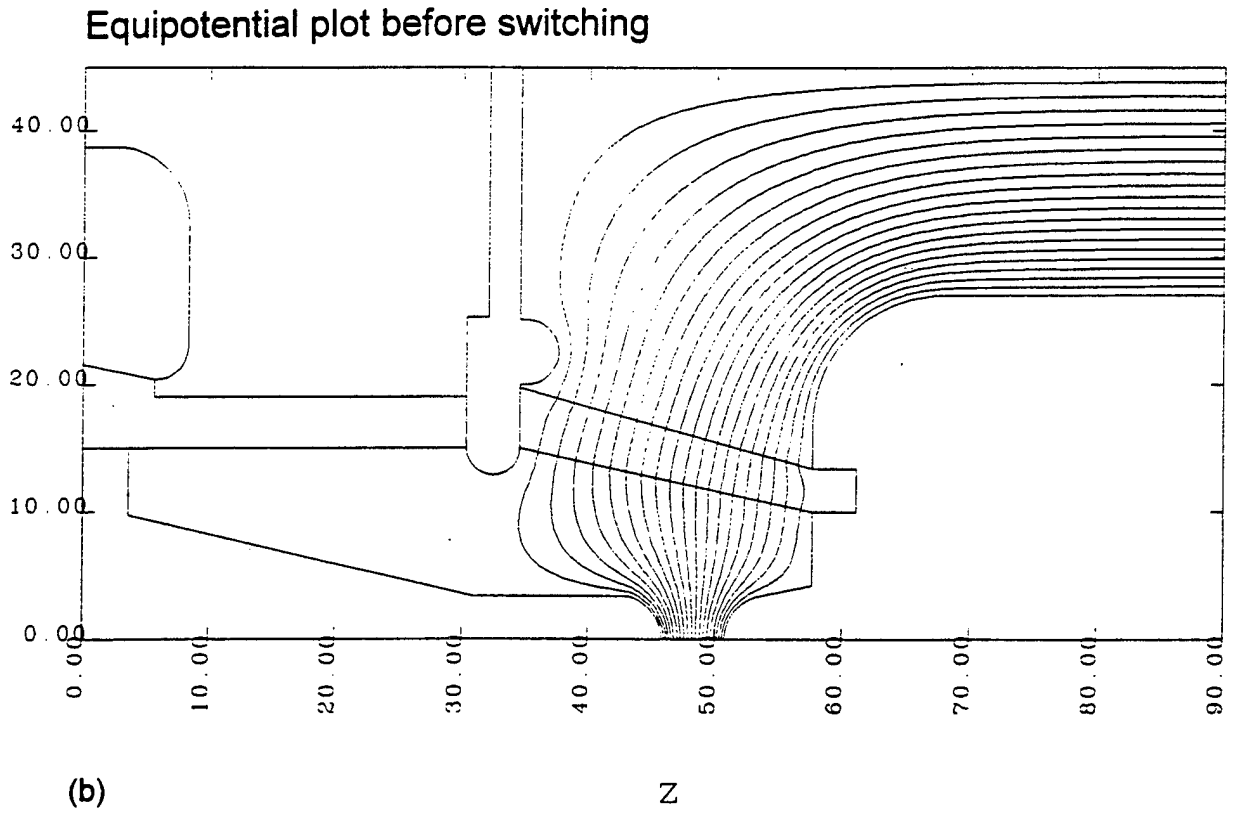


Figure 11. Detailed elevation of diode region.



(a)



(b)

Figure 12. Bushing switch (a) mechanical and (b) electrostatic before switching.

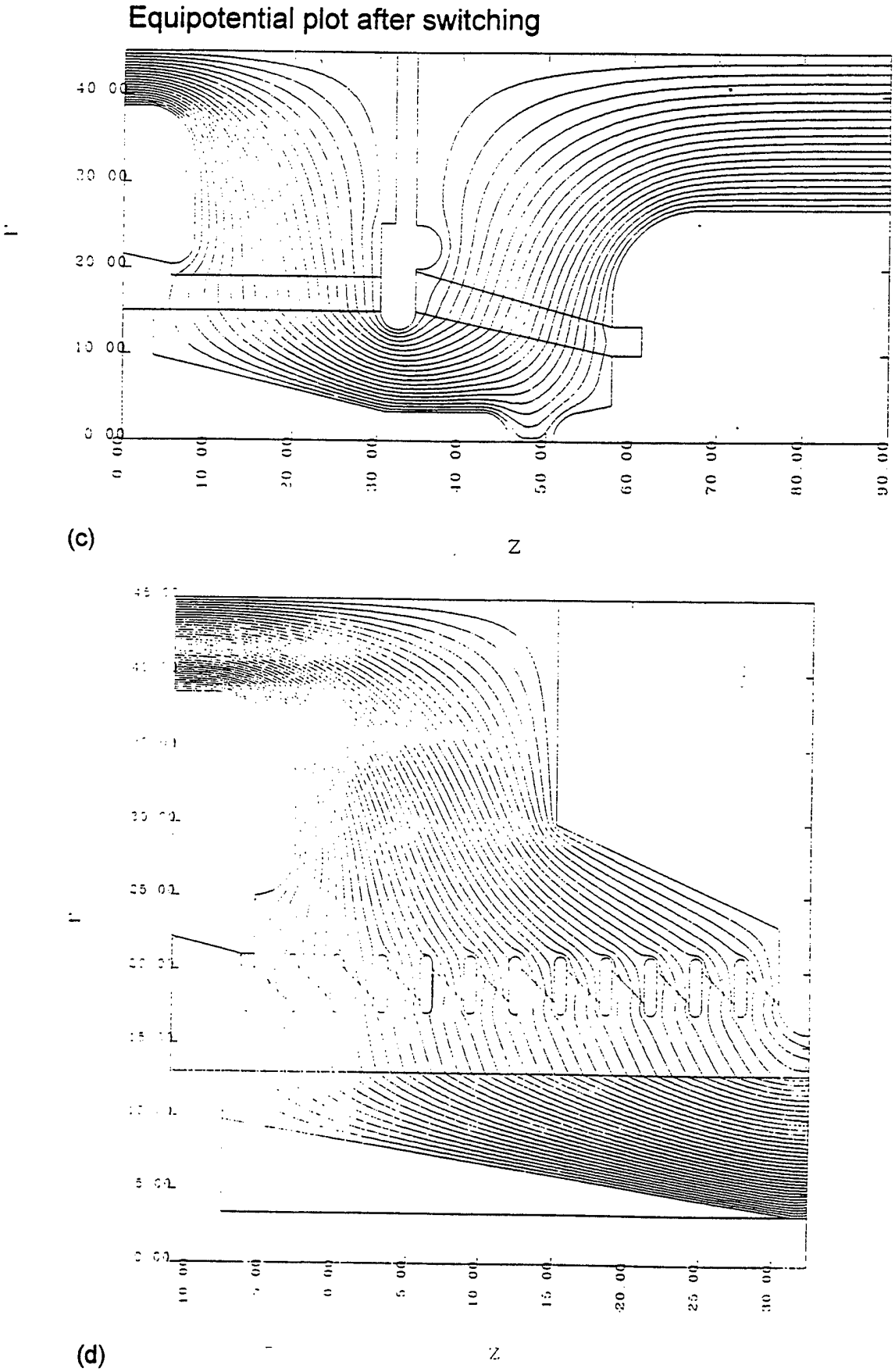


Figure 12. Bushing switch (c) electrostatic after switching and (d) after switch detail of tube showing grad increments.

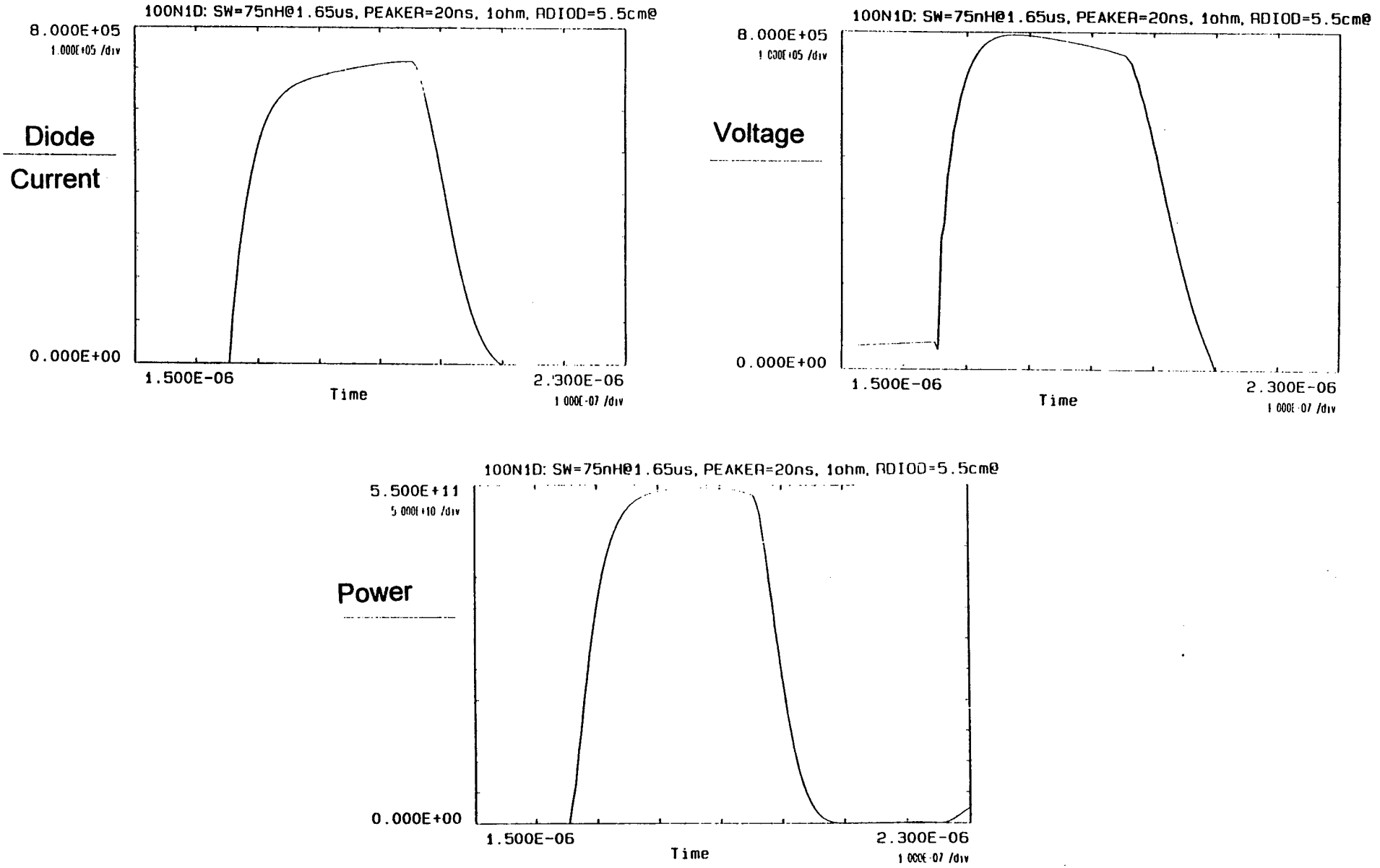


Figure 13. Current, voltage, power for 75 nH case. Note two diodes per column.

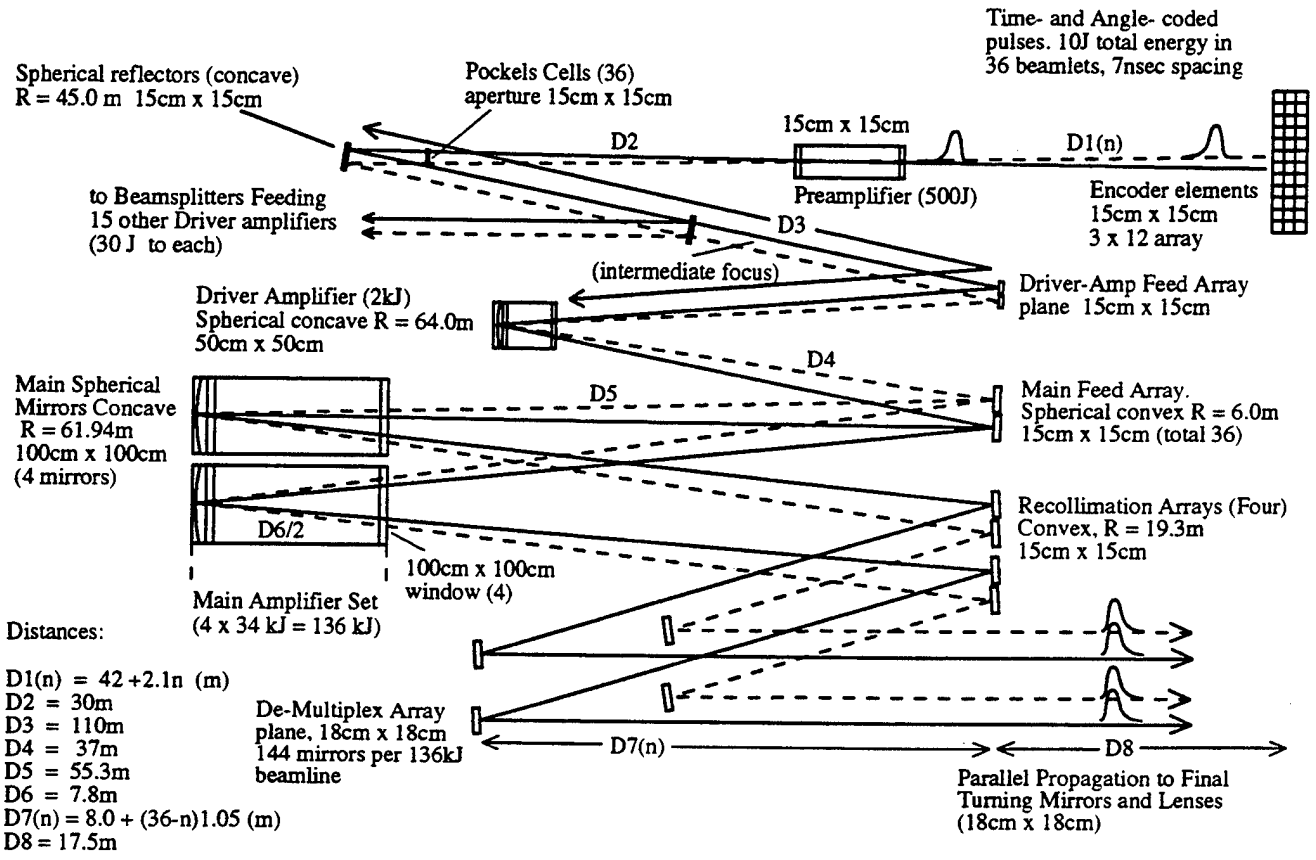


Figure 14 Schematic of time and angle multiplex system. Distances and optics for 136kJ beam line

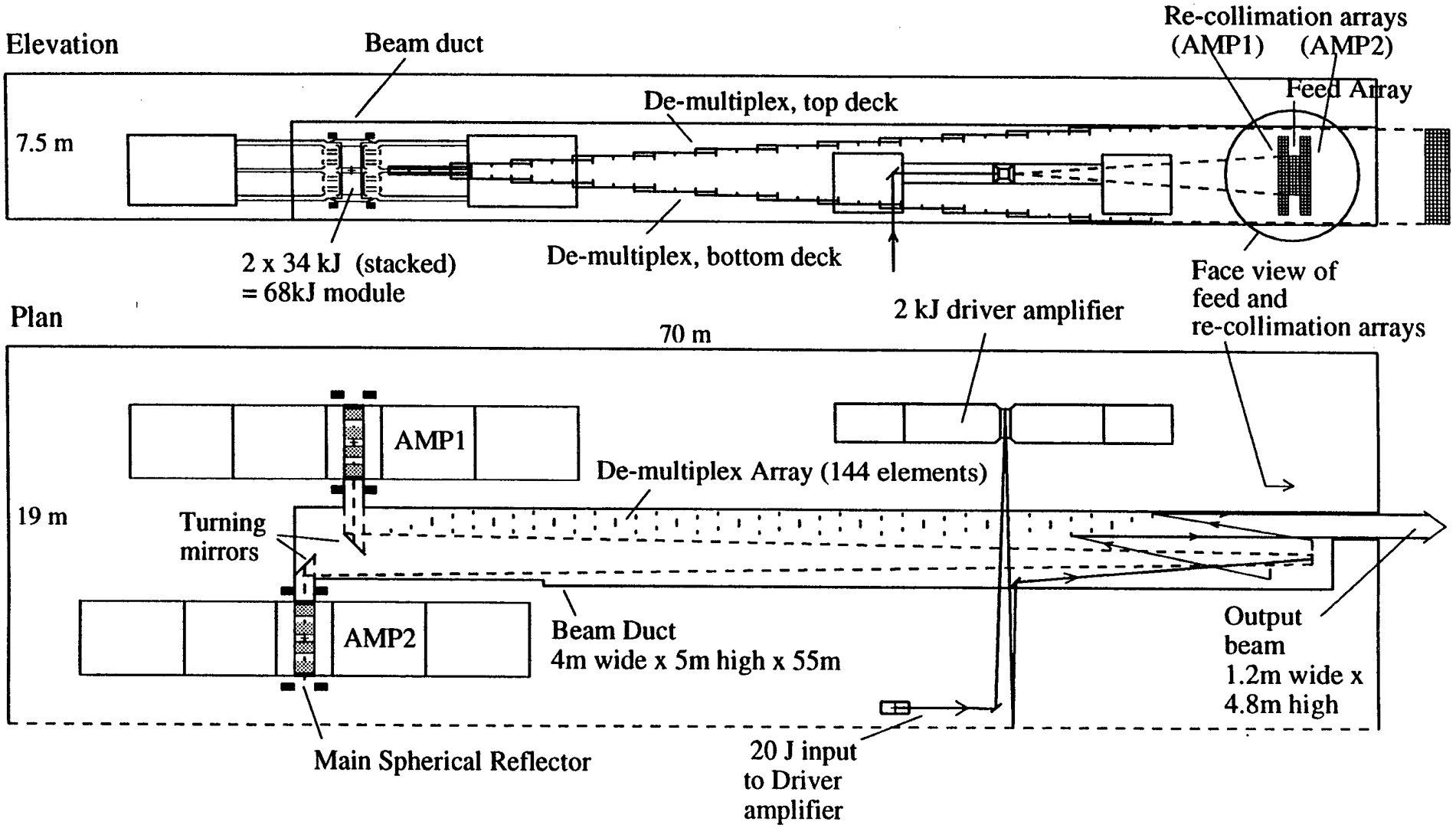


Figure 15 Layout of 136 kJ beamline

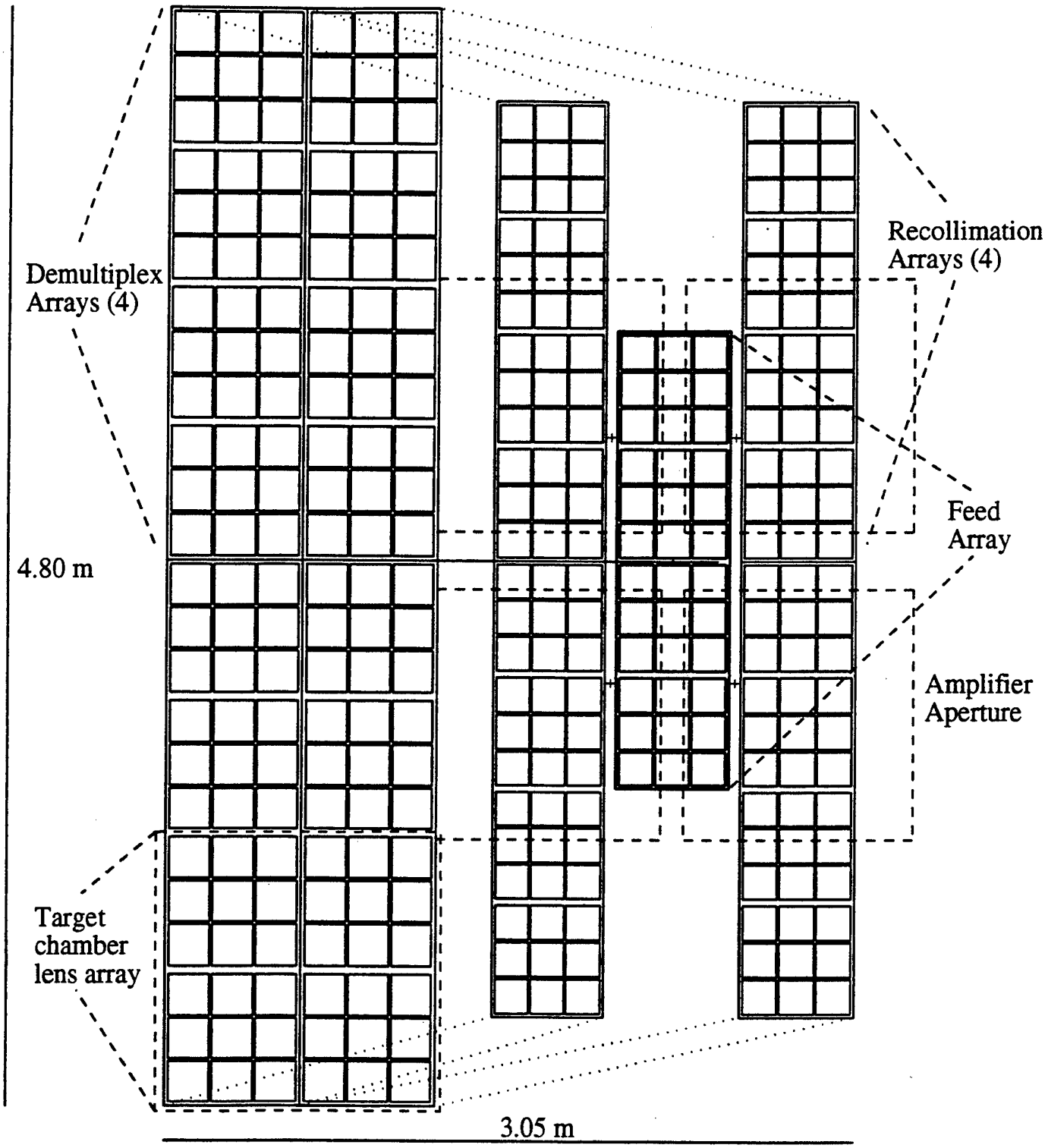


Figure 16 View down 136 kJ beamline from 4 x 1m^2 amplifier apertures toward feed and recollimation arrays.

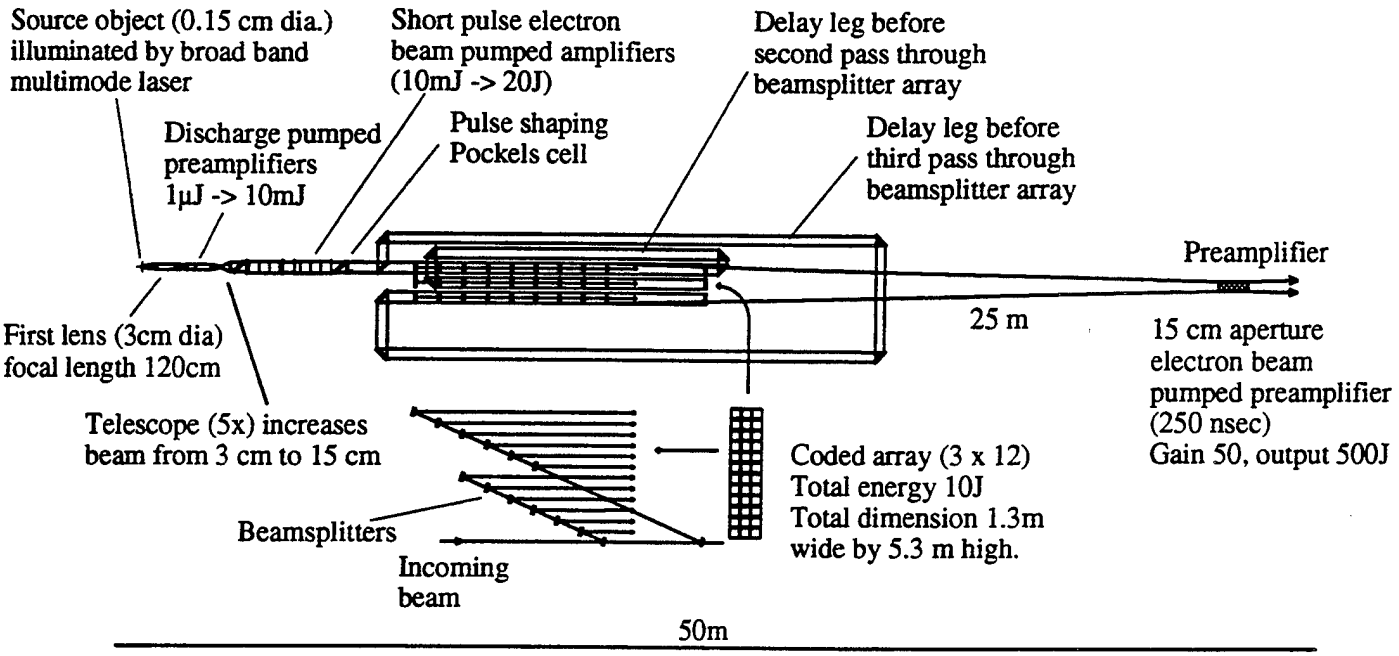


Figure 17 Schematic of amplifier front end. Not consistently to scale. Vertical scale enlarged for clarity.

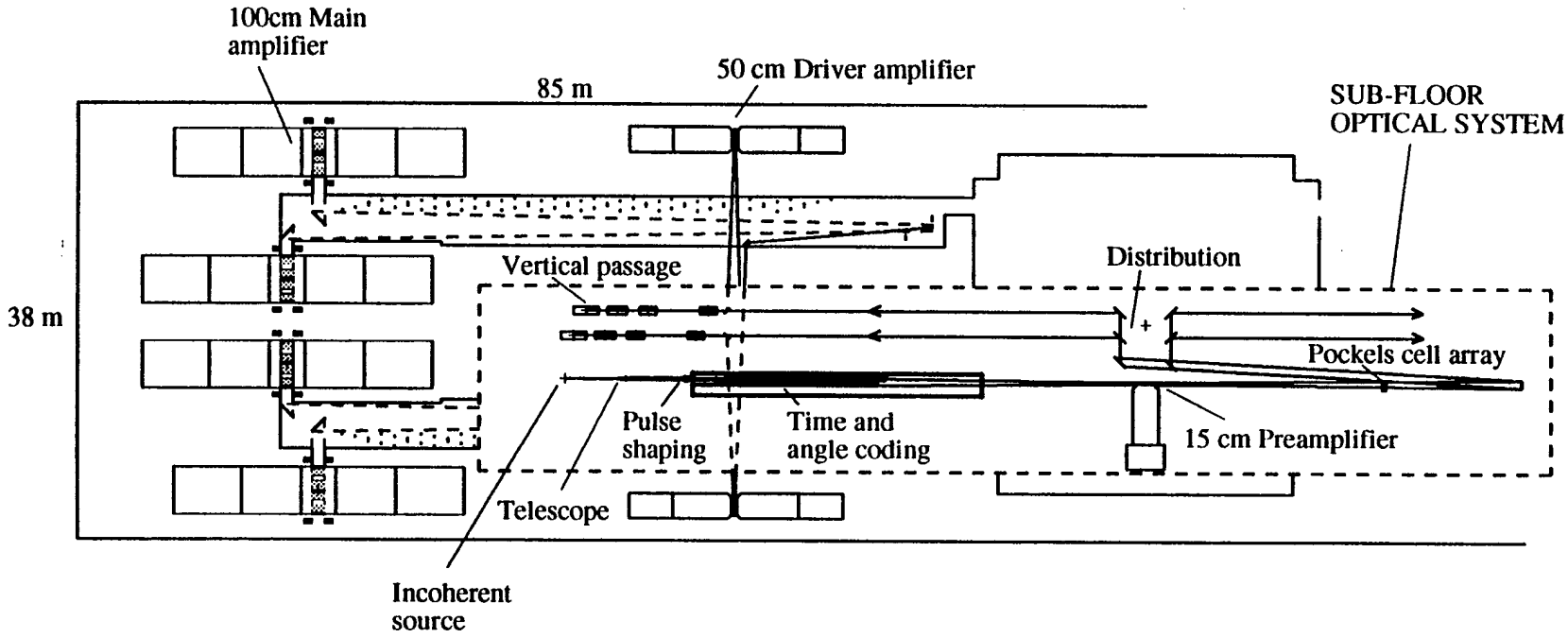


Figure 18 Layout of "Front End" in sub-floor and relationship to driver and main amplifiers

12 cm

12 cm

1.00	0.98	1.00	1.01	0.94	0.92	0.80	0.77	0.33	0.00	0.00	0.00
0.98	0.96	0.97	0.98	0.91	0.89	0.77	0.75	0.31	0.00	0.00	0.00
1.00	0.97	0.99	1.00	0.92	0.90	0.78	0.76	0.31	0.00	0.00	0.00
1.01	0.98	1.00	0.99	0.92	0.89	0.78	0.75	0.30	0.00	0.00	0.00
0.94	0.91	0.92	0.92	0.85	0.82	0.72	0.68	0.27	0.00	0.00	0.00
0.92	0.89	0.90	0.89	0.82	0.79	0.69	0.66	0.25	0.00	0.00	0.00
0.80	0.77	0.78	0.78	0.72	0.69	0.61	0.57	0.22	0.00	0.00	0.00
0.77	0.75	0.76	0.75	0.68	0.66	0.57	0.54	0.20	0.00	0.00	0.00
0.33	0.31	0.31	0.30	0.27	0.25	0.22	0.20	0.06	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Figure 19 Relative distribution of rays from 0.24 cm diameter incoherent front end source on one quadrant of 18 cm x 18 cm final lens. Resolution is 1 cm x 1 cm. Center of lens is at top left.

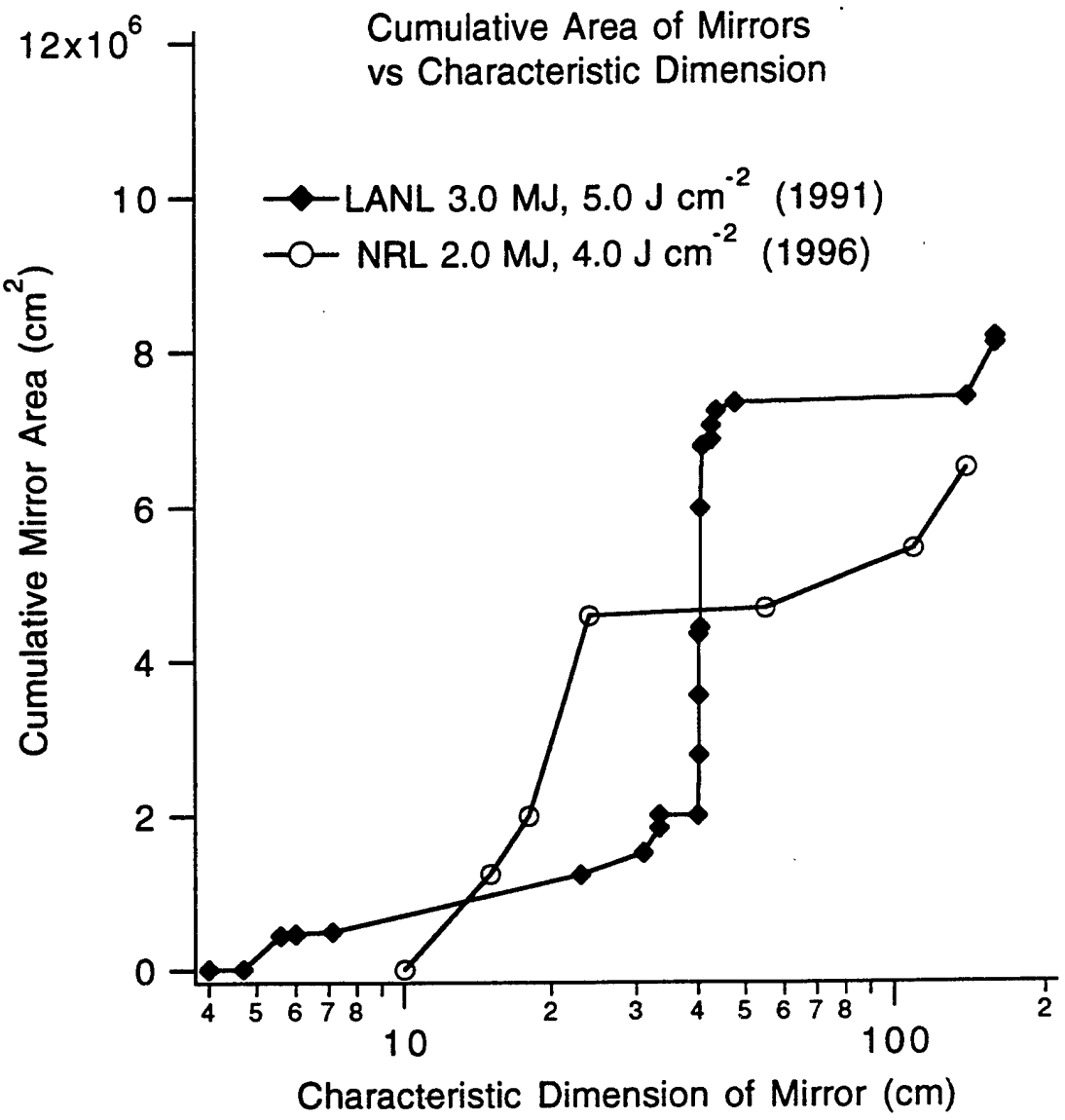


Figure 20 Cumulative mirror area vs characteristic dimension for present 2MJ design and prior 3MJ design from LANL.

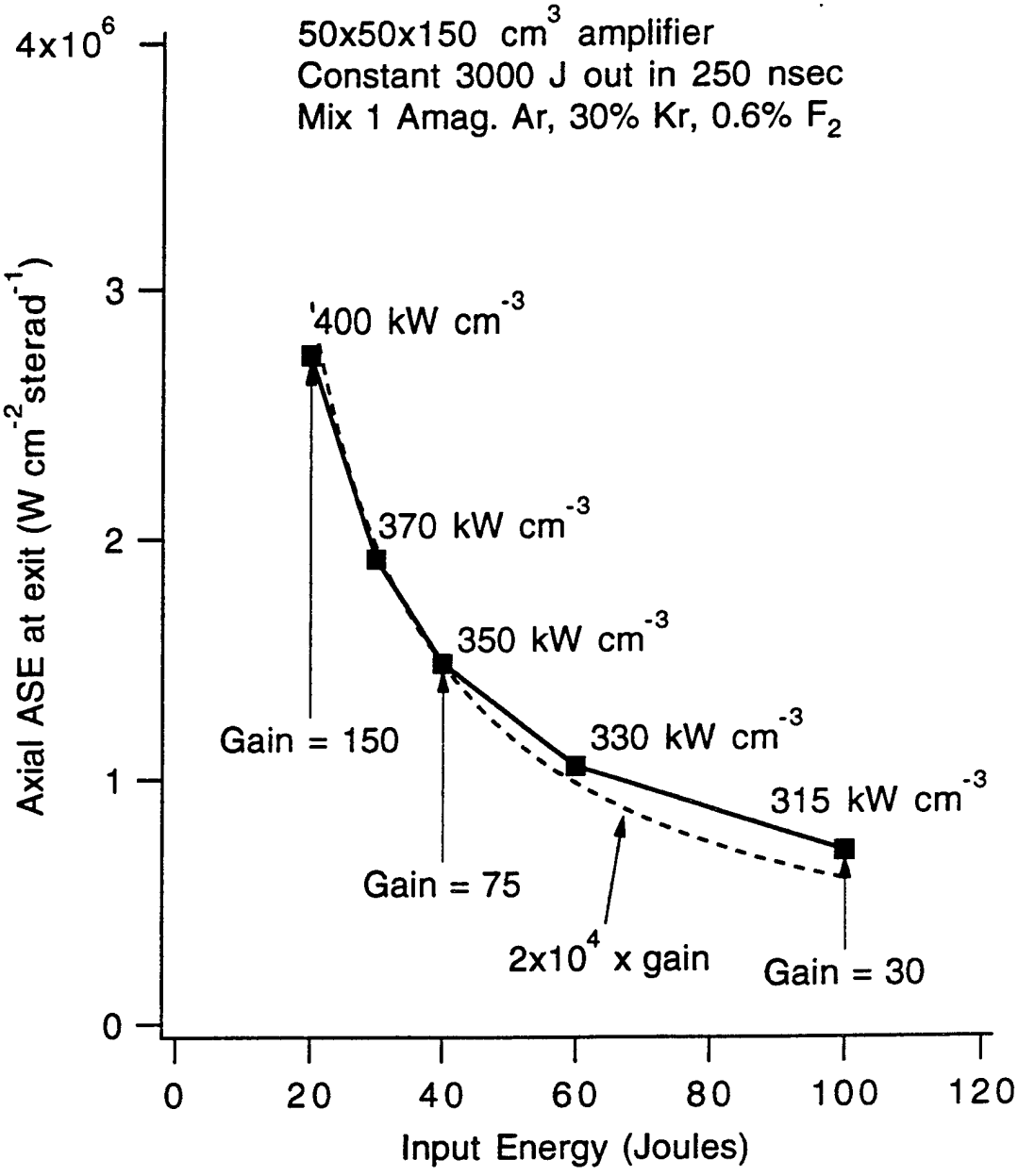


Figure 21 Example of ASE calculations. In this run the output energy was maintained constant as the input was varied, with pump rate adjusted appropriately.

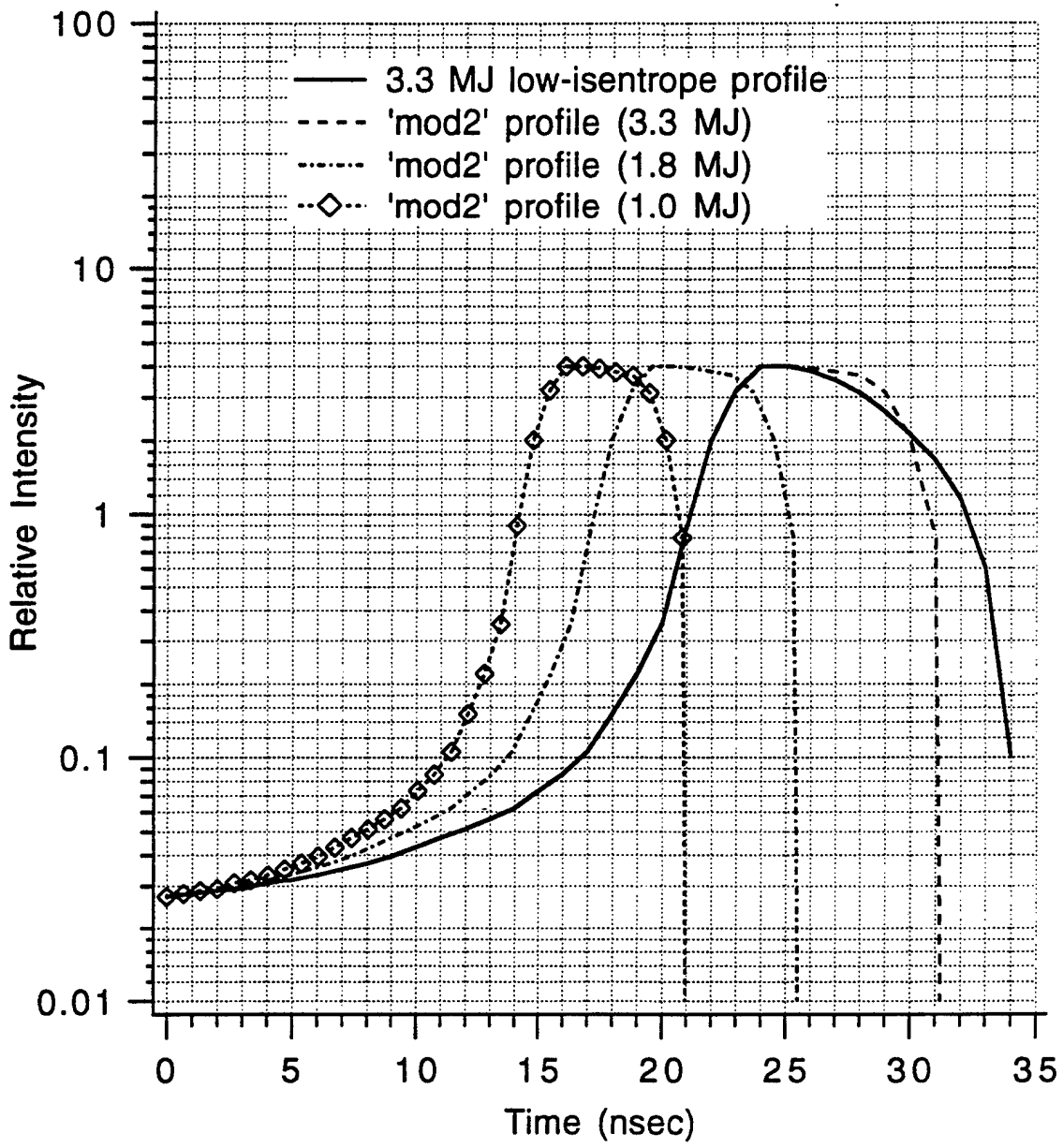


Figure 22 Illustration of the 3.3 MJ low-isentrope profile and three mod2 variants with FWHM equal to that of the energy-scaled pulse.

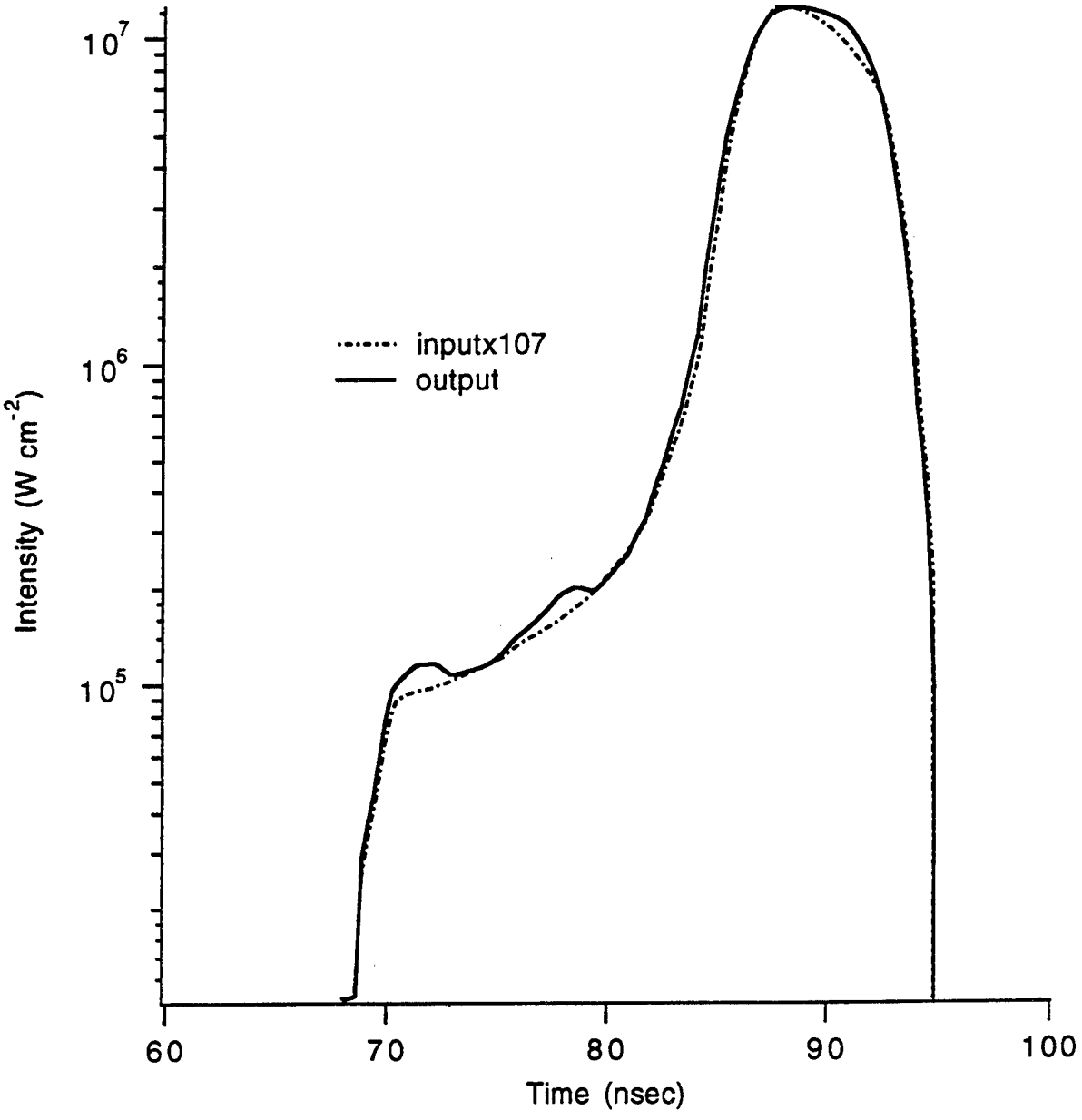


Figure 23 Propagation of laser pulse scaled to 1.8 MJ through segmented amplifier, at interpulse separation 6.5 nsec.

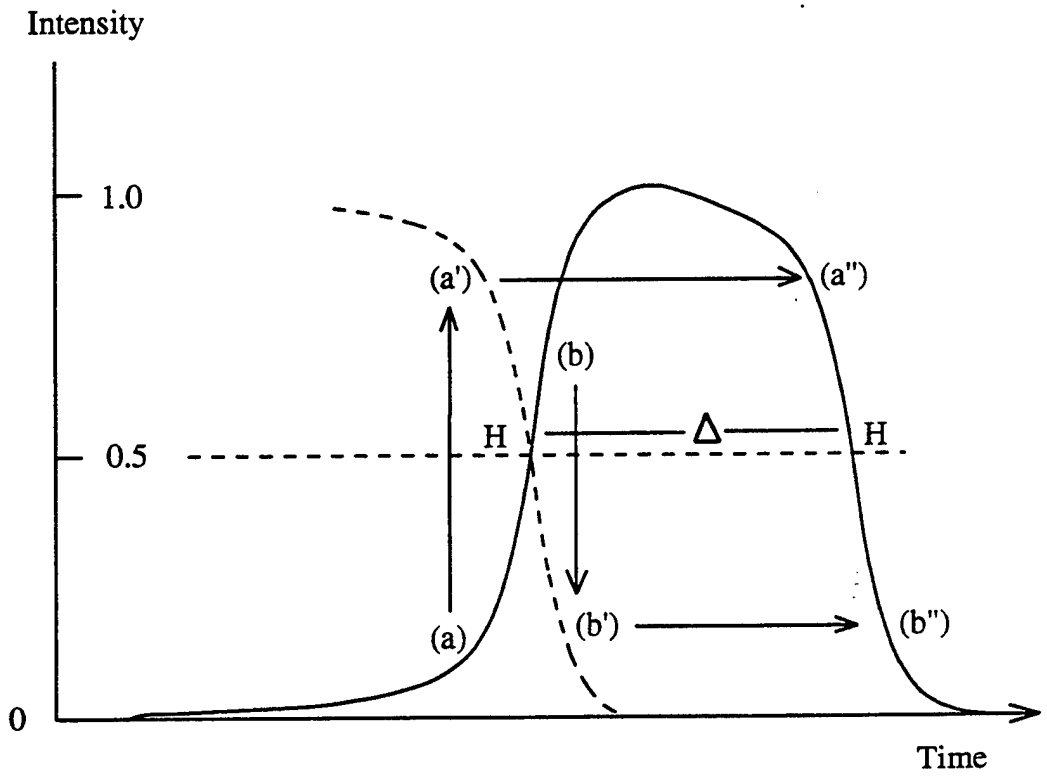


Figure 24 Illustration of construction of mod2 pulses for a given interpulse separation Δ

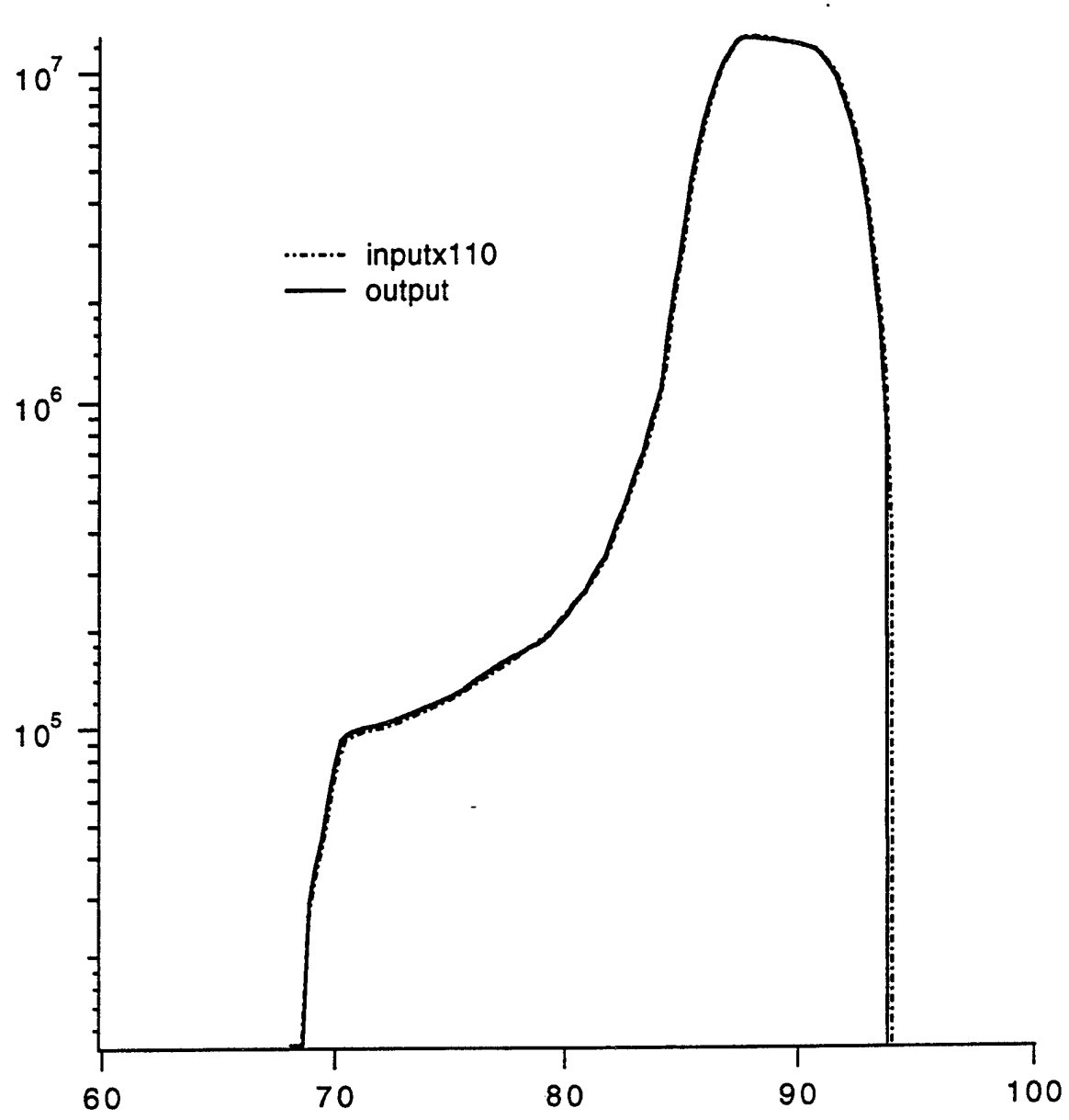


Figure 25 Propagation through segmented amplifier of mod2 pulse based on 6.5 nsec width and baseline shape at 1.8 MJ, separation 6.5 nsec.

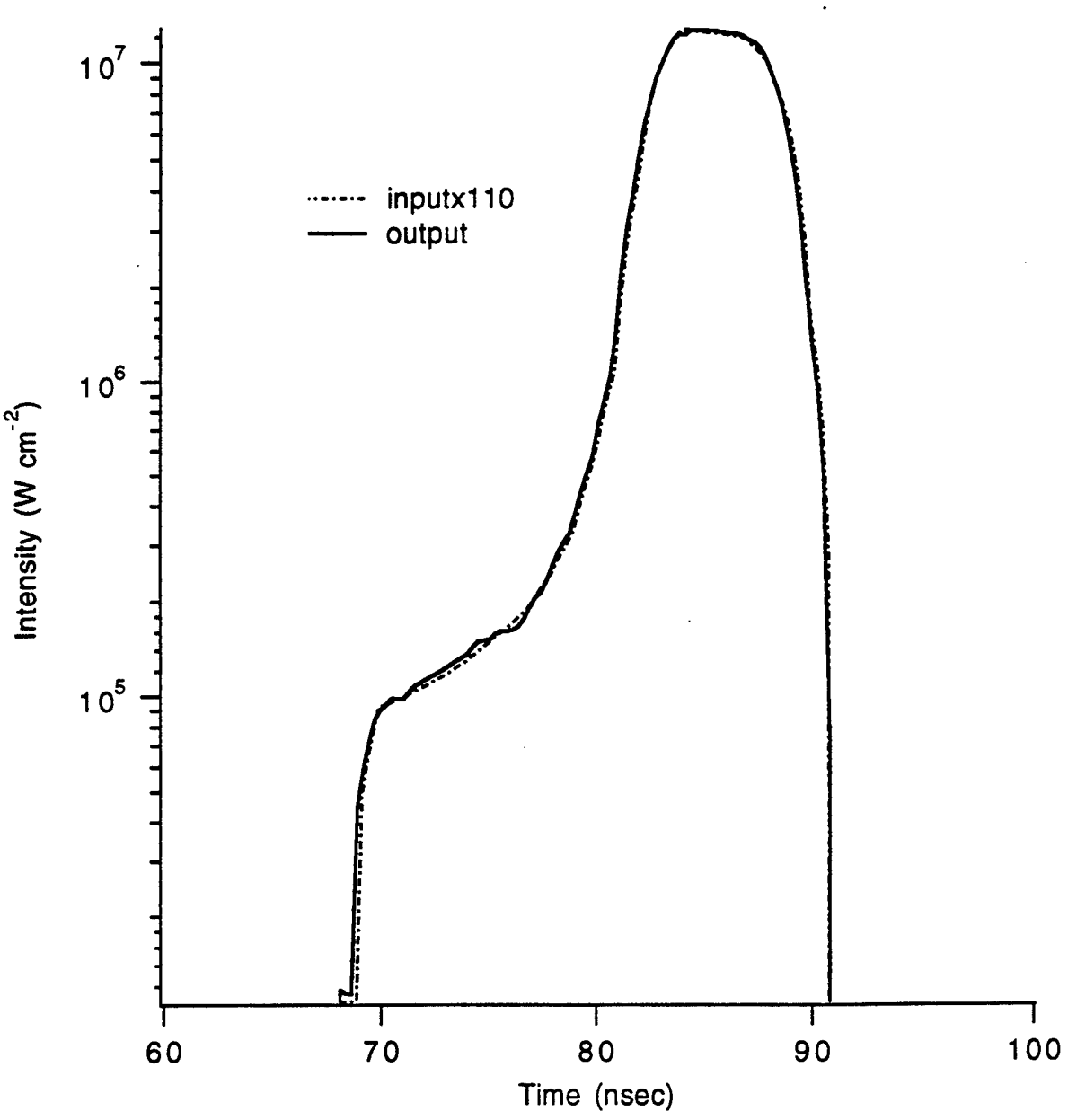


Figure 26 Propagation through segmented amplifier of mod2 pulse based on 6.5 nsec width and baseline shape at 1.0 MJ, separation 6.5 nsec.

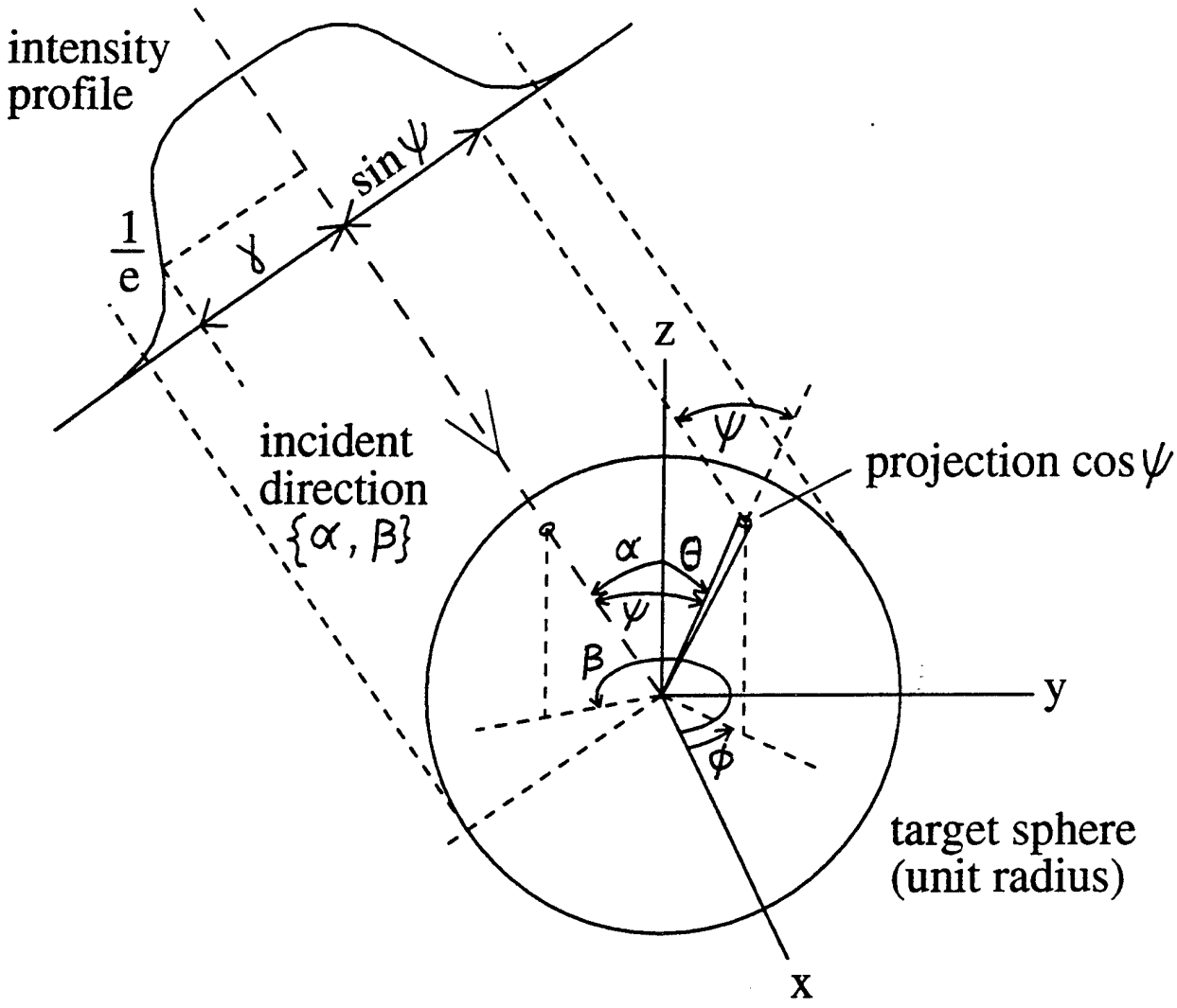


Figure 27 Illustration of illumination symmetry calculation.

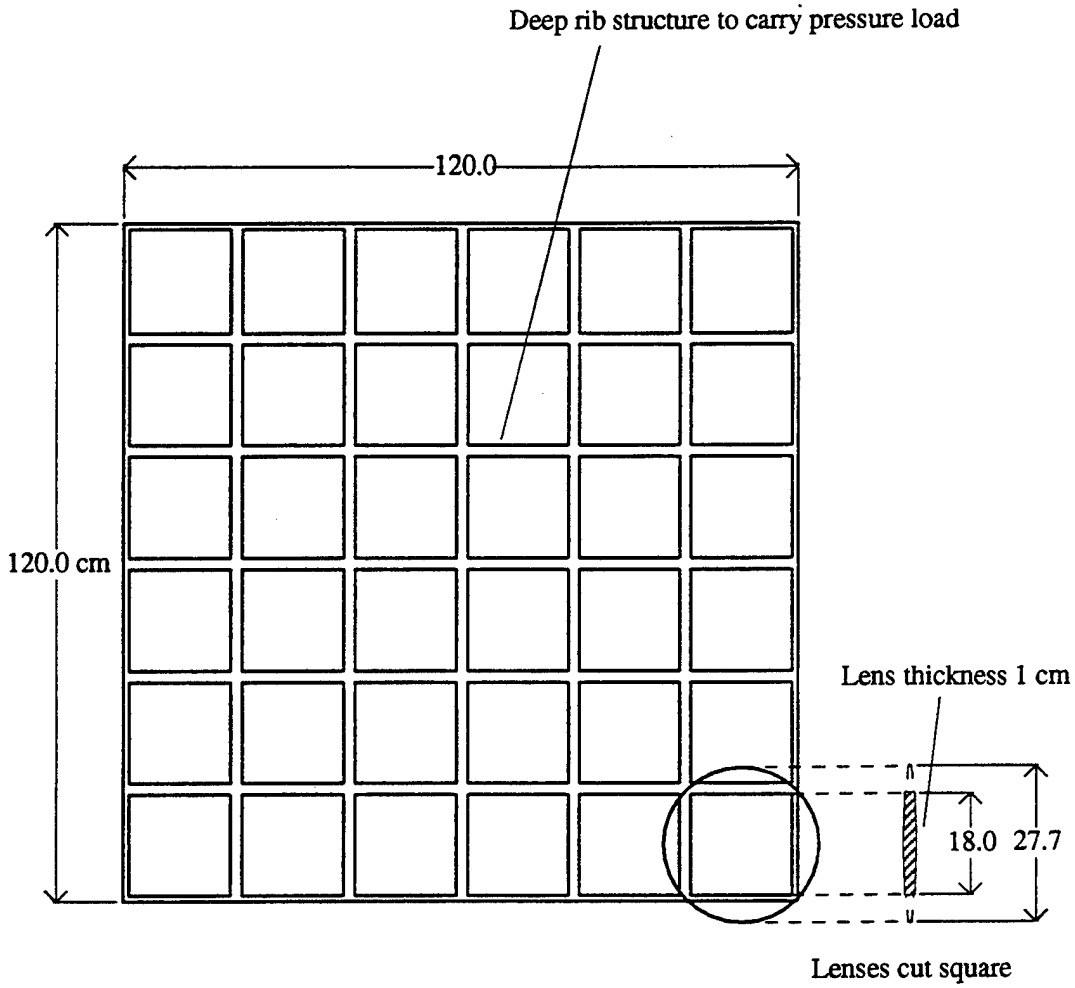


Figure 28 Unit array of 36 final focusing lenses. One of 64 similar arrays on target chamber. Target energy handled by one array = 31 kJ

Appendix J

X-Ray Emission from a 650-fs Laser-Produced Barium Plasma

X-ray emission from a 650-fs laser-produced barium plasma

W. H. Goldstein
L-Division, Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, California 94550

A. Zigler, P. G. Burkhalter, and D. J. Nagel
Naval Research Laboratory, Department of the Navy, Washington, DC 20375-5000

A. Bar-Shalom and J. Oreg
Nuclear Research Center of the Negev, Box 9001, Beer Sheva, Israel

T. S. Luk, A. McPherson, and C. K. Rhodes
University of Illinois at Chicago, Chicago, Illinois 60680

(Received 19 November 1992)

We have used x rays in the 9–15-Å band emitted from a solid target of BaF_2 irradiated by ~120 mJ of 248-nm radiation in a 650-fs full width at half maximum Gaussian laser pulse to characterize spectroscopically the resulting ultrashort-pulse laser-produced plasma. The emission was spectrally resolved but space and time integrated. By comparing the spectrum with unresolved-transition-array calculations and measurements of plasma emission from longer-pulse experiments, it is clear that ions as highly stripped as titaniumlike barium are present. We have successfully modeled the observed spectrum by assuming an optically thin source in local thermodynamic equilibrium (LTE) and using the super-transition-array theory [A. Bar-Shalom *et al.*, Phys. Rev. A **40**, 3183 (1989)] for emission from a hot, dense plasma. The model indicates that the emitting region is at comparatively low temperature (200–300 eV) and high electron density (10^{23}–10^{24} cm^{-3}). The degree of agreement between the model and the measured spectrum also suggests that the emitting plasma is near LTE. We conclude that the emission in this band arises from a solid-density plasma formed early in time, and is thus localized in both space and time. This interpretation is reinforced by LASNEX [G. B. Zimmerman and W. L. Kruer, Comments Plasma Phys. Controlled Fusion **11**, 51 (1975)] simulations that indicate that emission in this band closely tracks the laser pulse.

PACS number(s): 52.50.Jm, 32.30.Rj, 52.70.La

Ultrashort-pulse laser technology, providing subpicosecond pulse lengths and power densities of up to 10^{19} W/cm^2 from table-top systems, has significantly expanded the field of laser-produced plasma research. Applications presently under investigation include x-ray lasers [1], ultrashort-pulse x-ray flashlamps [2], and the study of the laser-matter interaction at high density and temperature [3]. Progress in these and other directions depends on understanding x-ray production in ultrashort-pulse laser-produced plasmas. Unfortunately, experimental investigations have generally been compromised by the lack of time- and space-resolved information. Broadband measurements have put the duration of the x-ray burst at less than 2 ps, for a 100-fs-long laser pulse [2], but, in general, time resolution has been sufficient only to put upper bounds on the duration of x-ray emission [4]. Pinhole cameras have yielded images of the emitting plasmas with resolutions down to about 5 μm [4,5], but this is insufficient to resolve a transient, solid-density plasma expected to extend over less than 500 nm, based on the absence of hydrodynamic motion and measurements of heat front penetration [6]. Prior attempts to characterize these plasmas based on spectroscopically resolved measurements have been limited to estimates of electron density using line broadening [4,7] in light elements. These analyses have encountered ambiguities owing to the long time scales involved in the atomic kinetics [4,7] and have involved short-pulse energy deposition in plasmas preformed by an amplified-spontaneous-emission (ASE) prepulse and consequently emitting at much lower than solid density [4,7,8].

In the absence of sufficient instrumental time and space resolution, it is important to explore spectral signatures that can be correlated with localized plasma conditions. Recently, Zigler *et al.* [5] measured intense x-ray emission from ultrashort-pulse laser-produced barium plasmas, formed without an ASE prepulse [9], at or near the maximum possible intensity. We have analyzed this spectrally resolved, but space- and time-integrated, emission in the 9–15-Å band created by 650-fs full width at half maximum (FWHM), 248-nm, KrF-laser pulses on solid targets of BaF_2. Time scales for the atomic kinetics of highly stripped barium should be short enough to remove atomic transients from the analysis. From the spectrum, we have been able to characterize the charge-state distribution, temperature, and density of the emitting plasma. For the latter two quantities, we have used the super-transition-array (STA) theory for heavy-element emission from plasma in local thermodynamic equilibrium (LTE) [10]. This method of spectral simulation is essential for

analyzing the quasicontinuum emission of many-electron systems at high density, in distinction to the simpler line spectra encountered in light-element plasmas even under ultrashort-pulse irradiation. Our conclusion is that this emission arises primarily from a cool, near-solid-density region, rather that from the hot, expanding corona, in contrast to the light-element emission seen in prepulse experiments [4,7,8]. This interpretation is supported by hydrodynamic simulations that further confirm that the emission is local in time.

The KrF laser for these experiments had a pulse width of 650 fs and an energy of 120 mJ, yielding a peak power of 200 GW and an average power density of $\sim 10^{17}$ W/cm^2. A 20-mJ prepulse was spread over ~ 20 ns. A detailed description of the laser system has been presented elsewhere [11]. The laser beam was focused onto a planar rotating target with a plano-convex $f/10$ CaF_2 lens. Several positions of the focusing lens were selected in order to vary the incident peak laser intensity. The x-ray emission from the target was collected by a flat potassium acid phthalate crystal ($2d = 26.6$ Å) spectrometers using Kodak DEF film. The calibration is described in Ref. [5]. Each spectrum was produced by 25 laser shots. Shot-to-shot variations in laser intensity were typically smaller than 10%, but could be as large as 50%. Ten shot averages always varied by less than 10%. The BaF_2 target was chosen since it is sufficiently transparent to low intensities of the uv laser that no prepulse plasma is formed [9]. Figure 1 shows the spectrum obtained with a laser intensity of 5×10^{15} W/cm^2. Figure 2 shows the spectrum, previously reported in Ref. [5], obtained at $\sim 10^{17}$ W/cm^2.

We turn now to analyzing the spectrum and plasma conditions. The lower trace in Fig. 1 shows an unresolved-transition-array (UTA) calculation [12] of the $3d$-$4f$ transitions in charge states of barium from scandiumlike through nickel-like. Comparing with this calculation, it is clear that the quasicontinuum emission in the

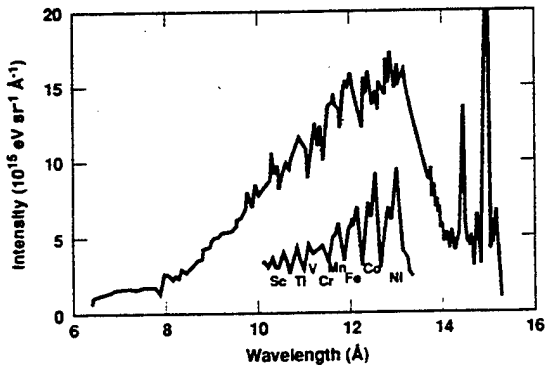


FIG. 1. Space- and time-integrated x-ray spectrum emitted by a barium plasma formed by ultrashort-pulse irradiation with 5×10^{15} W/cm^2 at $\lambda = 248$ nm. The vertical axis is intensity, in units of 5×10^{15} eV/sr Å. The narrow lines at long wavelength are fluorine He-β and H-α. The lower trace is a UTA calculation of the $3d$-$4f$ transition arrays in charge states of barium from scandiumlike through nickel-like. The calculated positions of these transition arrays indicates these same charge states contribute to the measured spectrum.

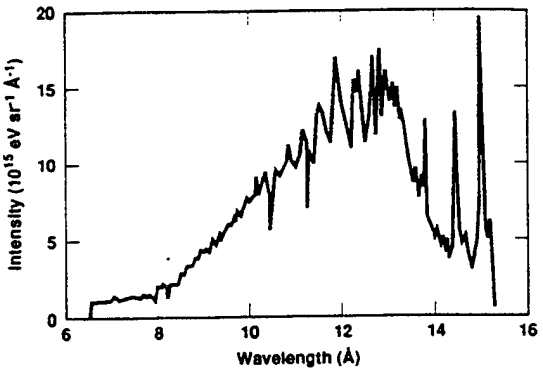


FIG. 2. Same as Fig. 1, but for an irradiation intensity of $\sim 10^{17}$ W/cm^2. (From Ref. [5].)

measured spectrum arises from the same charge states. If near-coronal conditions are obtained in the emitting plasma, this implies a fairly high temperature, on the order of 1 keV. On the other hand, if emission arises primarily from a near-solid-density plasma, conditions can be nearer LTE, and much lower temperatures are required. In fact, assuming Saha equilibrium, at solid density a temperature as low as 100 eV will strip barium beyond nickel-like. The almost complete absence of structure in the measured spectrum, as contrasted, for example, with the distinctive non-LTE, long pulse (4.5 ns) laser-produced plasma spectrum [13] in Fig. 3, supports the latter hypothesis, since, at high density, multistep collisional processes populate a huge array of states never sampled under non-LTE conditions, and these "satellites" fill in the spectrum. In Fig. 3, emission from specific charge states of barium have been identified using Cowan's atomic structure computer code [14].

We have modeled the barium spectrum shown in Fig. 1, assuming that it is emitted by an optically thin plasma in LTE. In Fig. 4(a), STA spectra are displayed for barium at normal density (3.5 g/cm^3) and a range of temperatures. The solid trace represents a weighted sum of the emission at the several temperatures shown. The weights were chosen to give a reasonably good fit to the data shown in Fig. 1. Though this is not a precise procedure, it conveys the essential point that the emission can be

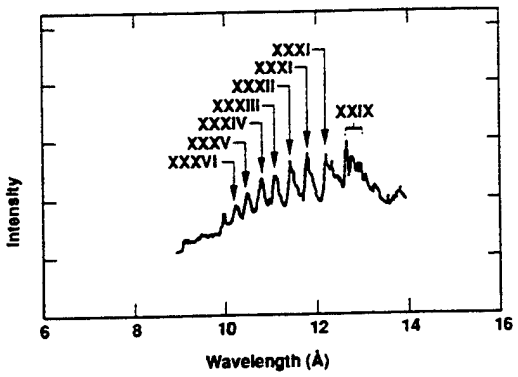


FIG. 3. Barium spectrum from a plasma formed by a 4.5 ns pulse. Contributing charge states between nickel-like Ba XXIX and scandiumlike Ba XXXVI are identified. Characteristically, the long-pulse laser-produced plasma is dominated by UTA's.

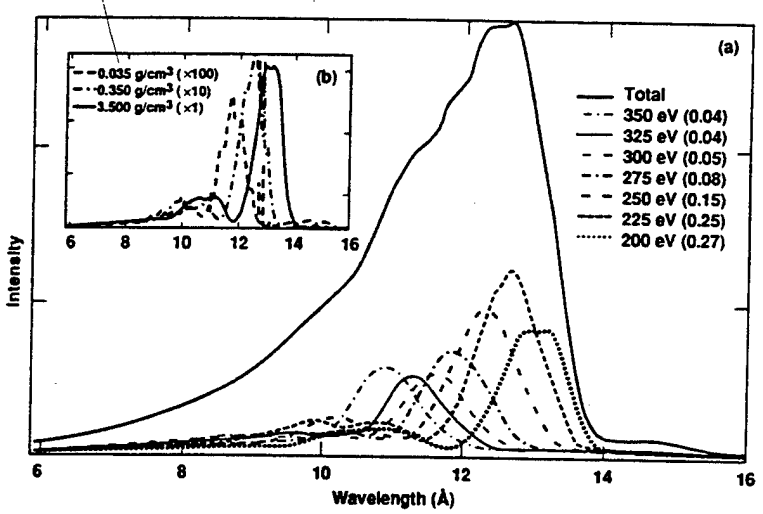


FIG. 4. (a) STA model for the spectrum in Fig. 1. Solid trace is the sum of the spectra obtained for each temperature, weighted by the fractions in parentheses. (b) Density dependence of STA $n=3-4$ peak position at 200 eV. (The intensities for 0.35 and 0.035 g/cm^3 have been scaled by factors of 10 and 100, respectively, relative to the normal density trace.)

well represented assuming a distribution of temperatures between 200 and 350 eV, and that the peak emission corresponds to a temperature around 225 eV. (Discrepancies between the measured and model spectra are attributable to emission from lower densities—between solid and coronal—and higher and lower temperatures than those shown.)

Note that the position of the peak emission, corresponding in this case to $n=3-4$ transitions, is a strong function of temperature. In Fig. 4(b), the peak is shown to vary with density as well, shifting to shorter wavelength with decreasing density. Since the model assumes LTE, this trend is easy to understand: three-body recombination is suppressed at lower densities, leading to a more highly ionized charge-state distribution and, thus, more energetic bound-bound transitions. The model, then, strongly suggests that LTE emission from lower density regions does not contribute appreciably, since these regions are almost certainly at *higher* temperature than the normal density plasma.

Based on general features of the measured spectrum and the fit obtained in the STA model, we conclude that emission in the 9–15-Å band is dominated by radiation from a normal density plasma at a temperature between 200 and 300 eV. The data is consistent with this plasma being in LTE, and optically thin. Emission from lower-density, higher-temperature regions is not apparent, and is not required to explain the observations.

A cool, dense emitting plasma is formed early in time by skin-depth absorption and a thermal wave propagating into the solid material. If the region does not ablate during the laser pulse, but remains at high density, it will rapidly recombine once the drive is off, effectively localizing the emission in space and time. If, on the other hand, the plasma ablates before it recombines through the M shell, the drop in density will "freeze" a charge-state distribution reflecting LTE at low temperature and high density, and emission will be nonlocal, persisting into the corona. Such coronal emission ought to contribute structure on top of the high-density background. Since the spectral measurement integrates over all time, even weak emission, if persistent, could easily swamp the early-time component. Although there are hints of this structure in the data, it clearly does not suggest appreciable late-time or coronal emission.

We must also consider the possibility that, though emission comes from a high-density plasma localized early in time, there is insufficient time for the charge-state distribution to equilibrate. Our LTE model is hardly applicable in this case, and the temperature inferred thereby will be incorrect. It is difficult to dispose of this possibility without supplementing the spectral evidence with a simulation of the experiment. But we can, at least, check the consistency of the LTE hypothesis by comparing collisional ionization times with the ~1-ps pump duration. Assuming $T_e \sim 250$ eV and solid density, and using the Lotz formula [15], the total LTE ionization rate for an $n=3$ electron of nickel-like barium is $\sim 6\times10^{-11}$ cm^3 sec^{-1}, indicating that electron densities of only $\sim 2\times10^{22}$ cm^{-3}, well below that of the solid, are needed at this temperature.

Our assumptions and conclusions are substantiated by simple hydrodynamic simulations. We simulated the experiment in one dimension, with an incident laser flux of 5×10^{15} W/cm^2 and 30% absorption, using the Lagrangian hydrodynamic code LASNEX [16], with hydrogenic, non-LTE average-atom atomic physics (XSN) [17]. Laser energy was deposited by both inverse bremsstrahlung and resonant absorption at the critical surface, $N_c = 1.6\times10^{22}$ cm^{-3}. We found that assuming 30% absorption at N_c best reproduced the observed position of the Ba emission maximum at about 12.5 Å. Resonant absorption is restricted to the ablated corona, where suprathermal electron temperatures of several keV, and fractions as high as 95% are generated [18]. The thermal electron temperatures, however, do not exceed 350 eV, as shown in Fig. 5(a), and peak just after the peak of the laser pulse at 1 ps. Note that peak temperatures are reached later for deeper zones, owing to thermal transport. Figure 5(b) shows that the electron density rises initially in each zone, as the plasma is heated and stripped, then drops as the material expands.

In Fig. 6, we plot integrated radiant energy output as a function of time in the band from 900 to 1100 eV. Evidently, over 80% of the emission into this band takes place in the first 4 ps following onset of the pump; at the

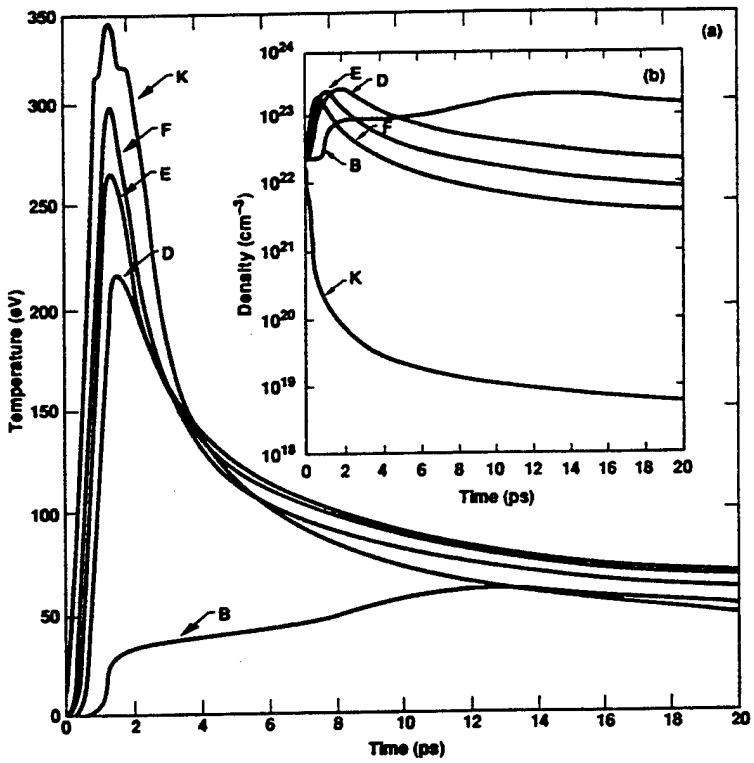


FIG. 5. As predicted by LASNEX, (a) the thermal electron temperature and (b) electron density, as functions of time, in a series of Lagrangian zones: *K* is the front surface of the 2 μm target; *F*, *E*, and *D* are, respectively, 0.02, 0.05, and 0.12 μm deep initially, and *B* is at 0.7 μm. These zones are chosen to represent three regions: the blow-off plasma (*K*); the emitting plasma (*F*,*E*,*D*) at ~100–1000 Å, heated by thermal transport; and a deeper region, > 1000 Å, that does not participate in the emission or ablation.

emission peak, ~1000 eV, the figure is over 90%. Localization of the emission in this band in both space and time is further demonstrated in Fig. 7. Finally, at the time of peak emission (2 ps), in zones *D*, *E*, and *F*, LASNEX predicts mean occupancies of the *M* shell of 17.8, 16.9, and 16.2 electrons, respectively. The corresponding LTE values, also calculated by LASNEX, are 17.6, 17.3, and 16.4.

In conclusion, we have observed x rays in the 9–15-Å band emitted from a solid target of BaF_2 irradiated by ~100 mJ of 0.25 μm, KrF, light in a 650-fs FWHM Gaussian laser pulse. We have successfully modeled the observed spectrum using the STA theory and concluded that the emitting region is at comparatively low temperature (200–300 eV) and high electron density (10^{23}–10^{24} cm^{-3}), and is localized in both space and time. This in-

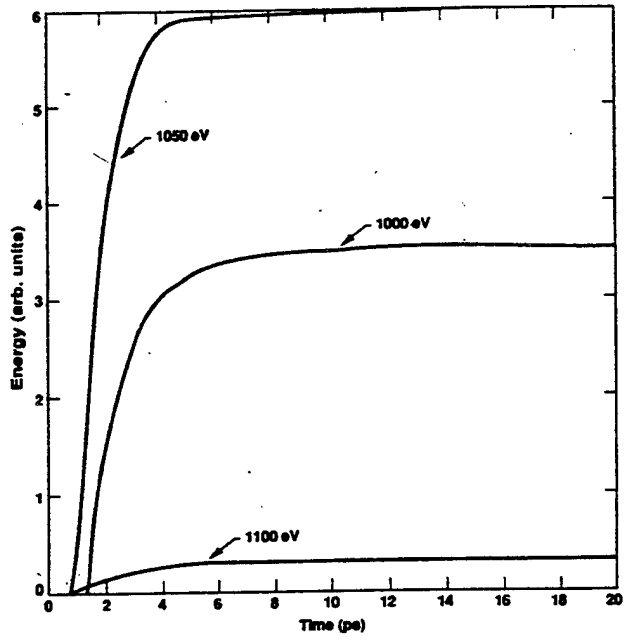


FIG. 6. Radiated energy (arbitrary units) as a function of time at 1000, 1050, and 1100 eV. The global emission peak is between 1000 and 1050 eV.

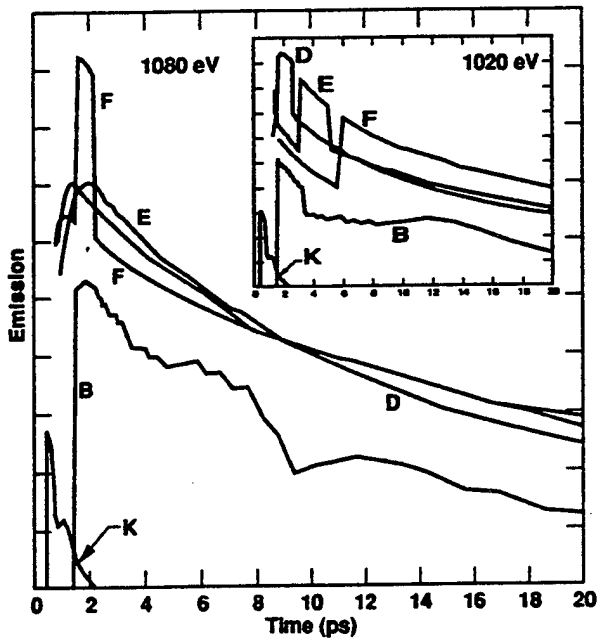


FIG. 7. Emission in arbitrary logarithmic units at 1080 eV and, inset, at 1020 eV, as a function of time in zones *K*, *F*, *E*, *D*, and *B*. The emission is highly localized in space and time. Emitting zones have temperatures of 200–300 eV, and electron densities of $\sim 3\times10^{23}$ cm^{-3}. The inset abscissa also covers the time range from 0 to 20 ps.

terpretation is reinforced by LASNEX simulations that indicate that emission in this band closely tracks the laser pulse. We expect to be able to apply an understanding of this localization to develop spectroscopic diagnostics for ultrashort-pulse laser-produced plasmas.

The authors thank Dr. M. Rosen for valuable assistance in simulating the experiment and Dr. A. Osterheld for providing atomic data. Work performed under the auspices of the U.S. Dept. of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48, the U.S. Air Force Office of Scientific Research under Grant No. AFOSR 89-0159, the Department of Energy under Grant No. DE-F602091ER12108, the U.S. Office of Naval Research and the Strategic Defense Initiative Organization.

[1] N. H. Burnett and P. B. Corkum, J. Opt. Soc. Am. B **6**, 1195 (1989); M. M. Murname, H. C. Kapteyn, and R. W. Falcone, IEEE J. Quantum Electron. **25**, 2417 (1989); C. J. Keane, J. N. Bardsley, L. da Silva, N. Landen, and D. Matthews, in *Femtosecond to Nanosecond High Intensity Lasers and Applications*, edited by E. M. Campbell, Proc. SPIE Vol. 1229 (SPIE, Bellingham, WA, 1990), pp. 190–195.

[2] M. M. Murname, H. C. Kapteyn, and R. W. Falcone, Phys. Rev. Lett. **62**, 155 (1989); M. M. Murname, H. C. Kapteyn, M. D. Rosen, and R. W. Falcone, Science **251**, 531 (1991).

[3] R. Fedosejevs, R. Ottmann, R. Sigel, G. Kühnle, S. Szatmari, and F. P. Schäfer, Phys. Rev. Lett. **64**, 1250 (1990); R. M. More, Opt. Soc. Am. Tech. Dig. Ser. **17**, 67 (1989); H. M. Milchberg, R. R. Freeman, S. C. Davey, and R. M. More, Phys. Rev. Lett. **61**, 2364 (1988).

[4] J. A. Cobble, G. A. Kyrala, A. A. Hauer, A. J. Taylor, C. C. Gomez, N. D. Delameter, and G. T. Schappert, Phys. Rev. A **39**, 454 (1989).

[5] A. Zigler, P. G. Burkhalter, D. J. Nagel, K. Boyer, T. S. Luk, A. McPherson, J. C. Solem, and C. K. Rhodes, Appl. Phys. Lett. **59**, 777 (1991).

[6] A. Zigler, P. G. Burkhalter, D. J. Nagel, M. D. Rosen, K. Boyer, G. Gibson, T. S. Luk, A. McPherson, and C. K. Rhodes, Appl. Phys. Lett. **59**, 534 (1991); A. Zigler, P. G. Burkhalter, D. J. Nagel, J. C. Solem, K. Boyer, T. S. Luk, A. McPherson, D. A. Tate, and C. K. Rhodes (unpublished).

[7] C. H. Nam, W. Tighe, S. Suckewer, J. F. Seely, U. Feldman, and L. A. Woltz, Phys. Rev. Lett. **59**, 2427 (1987).

[8] C. H. Nam, W. Tighe, E. Valco, and S. Suckewer, Appl. Phys. B **50**, 275 (1990).

[9] A. Zigler, P. G. Burkhalter, D. J. Nagel, M. D. Rosen, K. Boyer, T. S. Luk, and C. K. Rhodes, Opt. Lett. **16**, 1261 (1991).

[10] A. Bar-Shalom, J. Oreg, W. H. Goldstein, D. Shvarts, and A. Zigler, Phys. Rev. A **40**, 3183 (1989).

[11] T. S. Luk, A. McPherson, K. Boyer, and C. K. Rhodes, Opt. Lett. **14**, 1113 (1989).

[12] J. Bauche, C. Bauche-Arnoult, and M. Klapisch, Adv. At. Mol. Phys. **23**, 131 (1988).

[13] G. Mehlman, P. G. Burkhalter, D. A. Newman, and B. H. Ripin, J. Quant. Spectrosc. Radiat. Transfer **45**, 225 (1991).

[14] R. D. Cowan, *The Theory of Atomic Structure* (University of California, Berkeley, 1981).

[15] W. Lotz, Astrophys. J. Suppl. Ser. **14**, 207 (1967).

[16] G. B. Zimmerman and W. L. Kruer, Comments Plasma Phys. Controlled Fusion **11**, 51 (1975).

[17] W. A. Lokke and W. H. Grassberger, Lawrence Livermore National Laboratory Report No. UCRL-52276, 1977 (unpublished).

[18] D. W. Forslund, J. M. Kindel, and K. Lee, Phys. Rev. Lett. **39**, 284 (1977).

Appendix K

Effects of Configuration Widths on the Spectra of Local Thermodynamic Equilibrium Plasmas

Effect of configuration widths on the spectra of local thermodynamic equilibrium plasmas

A. Bar-Shalom and J. Oreg
Nuclear Research Center Negev, P.O. Box 9001, Beer Sheva 84190, Israel

W. H. Goldstein
Lawrence Livermore National Laboratory, Livermore, California 94550
(Received 31 August 1994)

We present the extension of the supertransition-array (STA) theory to include configuration widths in the spectra of local thermodynamic equilibrium (LTE) plasmas. Exact analytic expressions for the moments of a STA are given, accounting for the detailed contributions of individual levels within the configurations that belong to a STA. The STA average energy is shifted and an additional term appears in its variance. Various cases are presented, demonstrating the effect of these corrections on the LTE spectrum.

PACS number(s): 52.25.Nr, 32.90.+a

I. INTRODUCTION

The supertransition-array (STA) model for calculation of atomic spectra under local thermodynamic equilibrium plasma conditions was presented in Ref. [1] (hereafter denoted as I). The model accounts for *all possible* bound-bound and bound-free radiative transitions. It divides the set of all accessible configurations into subsets, or superconfigurations. Each superconfiguration comprises configurations that are similar in energy. The transition array between two superconfigurations is represented by a Gaussian distribution of spectral strength called a supertransition array. Thus a STA can be considered a collection of unresolved transition arrays (UTAs) [2,3] where, as is usually defined, each UTA represents the set of level-to-level transitions between a pair of configurations.

The first three moments of a STA—total intensity, average energy, and variance—can be calculated exactly using an analytic technique that bypasses the direct summation over all the UTAs it comprises. A convergence procedure is used [4,5] to successively better approximate the detailed structure of the spectral distribution. Each STA is iteratively divided into a number of smaller STAs as described below. The end point of this process, of course, is reached when each STA contains a single UTA. In practice, it is never necessary to carry the procedure past a few iterations to obtain an excellent approximation to the UTA spectrum.

In I, the STA moments were calculated neglecting the energy splitting within configurations, i.e., in the approximation of vanishing UTA widths. In this case each UTA is represented by a single line. Recently, several publications have appeared presenting STA results [4–6] and comparing with experiments [7,8]. These works have in fact been based on an improved model that already includes UTA widths, as well as additional improvements. However, these theoretical developments have not yet been presented.

In this paper we present the STA theory accounting for UTA widths and shifts. This involves deriving the analytical expressions of the STA moments that explicitly take into account individual level-to-level contributions rather than the configuration averages used in I.

In the next section we reintroduce the concepts and definitions required in the STA model. In Sec. III, we rederive the STA moments, taking into account the individual levels of each superconfiguration, and obtain additional terms representing the energy splitting within configurations. Examples with and without UTA width effects and comparisons with recent experiments are presented in Sec. IV, showing that the UTA widths and shifts significantly affect the spectrum. A discussion and a summary are given in Sec. V.

II. THE BOUND-BOUND SPECTRUM

The spectral distribution $S(E)$ of transition strength between bound atomic levels gives the probability per unit length for absorption (or emission) of a photon of energy E per unit of energy. In the following we will treat absorption: emission is obtained simply by multiplication with the Planck function.

The set of bound-bound transitions may be divided into subsets G such that

$$S(E)=\sum_G S_G(E)\ , \tag{1}$$

where

$$S_G(E)=\sum_{i,j\in G} N_i w_{ij} P_{ij}(E-E_{ij})\ . \tag{2}$$

N_i is the density of atoms in the lower level i and w_{ij} is the absorption transition probability to level j,

$$w_{ij}=\frac{\pi e^2 h}{mc} f_{ij}\ , \tag{3}$$

where f_{ij} is the absorption oscillator strength and $P_{ij}(E-E_{ij})$ is the normalized line shape of the transition, centered at the transition energy $E_{ij}=E_j-E_i$.

The total intensity, the average energy, and the variance of the group of transitions G are given by, respectively,

$$I_G = \int S_G(E)dE = \sum_{i,j\in G} N_i w_{ij} , \tag{4}$$

$$E_G = \frac{\int S_G(E)EdE}{I_G} = \frac{\sum_{i,j\in G} N_i w_{ij} E_{ij}}{I_G} , \tag{5}$$

$$(\Delta E_G)^2 = \frac{\int S_G(E)(E-E_G)^2 dE}{I_G} = \Delta_G^2 + \Delta_P^2 . \tag{6}$$

In Eq. (6),

$$\Delta_G^2 = \frac{\sum_{i,j\in G} N_i w_{ij}(E_{ij}-E_G)^2}{I_G} \tag{7}$$

and

$$\Delta_P^2 \equiv \int P(E-\bar{E})(E-\bar{E})^2 dE \tag{8}$$

is the variance of the single line shape profile assumed equal for all the transitions ij in G.

With the moments of S_G we could present the group of neighboring lines G as a Gaussian. This, however, has two drawbacks: (i) This approach hides the non-Gaussian profile of the individual line obtained by convolution of the Doppler and collisional profiles and (ii) the variance Δ_P^2 may be infinite, as is the case when a Lorentzian contribution is important. Instead, we use the moments I_G, E_G, and Δ_G to construct the distribution of the line centers in G as the Gaussian:

$$\Gamma(E-E_G) = \frac{I_G}{\sqrt{2\pi}\Delta_G} \exp\left[-\frac{1}{2}\left[\frac{E-E_G}{\Delta_G}\right]^2\right] . \tag{9}$$

In order to describe correctly the contribution of the individual lines to the spectrum we define $\bar{S}_G$ by convolution of this Gaussian with the individual line shape P (approximated in our code as a convolution of doppler and Lorentzian collisional profiles, i.e., a Voigt function)

$$\bar{S}_G(E) \equiv \int \Gamma(E'-E_G)P(E-E')dE' . \tag{10}$$

This representation of the spectrum of G keeps the non-Gaussian contribution of the individual line shape and has the same moments as S_G of Eq. (1).

Equations (4)–(6) for the moments include detailed summation over all the level-to-level transitions in G. The detailed transition energies (which were averaged over configurations in I) will be retained in the analytic derivation presented in the next section.

So far the theory relates to an unspecified group G of neighboring lines. In the STA model we are dealing with such well defined groups, i.e., the STAs, having the advantage that their moments can be derived analytically. In the next section we derive exact analytical expressions for the STA moments including energy splitting within configurations.

III. EVALUATION OF THE SPECTRAL MOMENTS BY SUMMATION OVER LEVELS

In the STA model one such group G is the collection of transitions originating from a specific superconfiguration by a specified one electron jump. A superconfiguration Ξ is a collection of ordinary configurations defined symbolically by the product over supershells σ,

$$\Xi \equiv \prod_{\sigma} \sigma^{Q_\sigma} . \tag{11}$$

A supershell, in turn, is the union of energetically adjacent ordinary atomic subshells $s \equiv \mathrm{j}_s \equiv n_s l_s j_s$. In Eq. (11) the superconfiguration is constructed by distributing the Q_σ electrons occupying supershell σ among the subshells in all possible ways subject to $\sum_{s\in\sigma} q_s = Q_\sigma$:

$$\sigma^{Q_\sigma} \equiv \sum_{\sum_{s\in\sigma} q_s = Q_\sigma} \prod_s \mathrm{j}_s^{q_s} . \tag{12}$$

Clearly, each partition of Q_σ in Eq. (11) is an ordinary configuration and the particular one-electron jump from this configuration results in another configuration. The transitions between the two configurations constitute an UTA.

The convergence procedure mentioned above splits supershells to smaller supershells according to their energy spread. For each superconfiguration in its turn, at each step, supershells that give rise to relatively well-separated configurations are preferentially split. The detailed structure of the spectrum is thus gradually revealed, yielding a converging spectrum or any other criteria, such as the Rosseland or the Planck means. The process terminates when a dictated change in these quantities is reached.

The STA model makes the essential approximation that the plasma is hot enough that the Boltzmann factor for the level population does not vary significantly over an ordinary configuration. This approximation is identical to that adopted in the UTA model of Bauche-Arnoult, Bauche, and Klapisch [2,3]. However, this would not be a good assumption for a superconfiguration and the variation of the Boltzmann factor from one configuration to another is accounted for in the STA model.

Another essential point is related to the dependence of the oscillator strength f_{ij} on the transition energy E_{ij} between the corresponding two levels. This transition energy can always be written as

$$E_{ij} = E_G + \Delta E_{ij} , \tag{13}$$

where ΔE_{ij} is the deviation from the average STA energy. We have found that the dependence of f_{ij} on ΔE_{ij} can be dealt with by modifying the Voigt function of Eq. (10). (However, since the derivation of this modification is nontrivial and its effect is negligible, it is not presented in this paper.) Thus in the calculation of the STA moments, f_{ij} depends only on E_G,

$$f_{ij}(E_{ij}) \approx f_{ij}(E_G) . \tag{14}$$

The explicit dependence of the moments on E_{ij} is of

course retained.

It can be shown from Eqs. (4)–(7) and (11) that the moments of a STA that contributes to the one-electron jump from orbital α to orbital $\beta(\alpha \Longrightarrow \beta,\ \alpha = j_\alpha \equiv n_\alpha l_\alpha j_\alpha \cdots)$ are given in terms of the moments of the UTAs it comprises with no approximation:

$$I_G = \sum_{C \in \Xi} N_C w_C^{(\alpha\beta)} , \tag{15}$$

$$E_G = \frac{\sum_{C \in \Xi} N_C w_C^{(\alpha\beta)} E_C^{(\alpha\beta)}}{I_G} , \tag{16}$$

$$\Delta_G^2 = \sum_{C \in \Xi} \sum_{\substack{i \in C \\ j \in C'}} N_i w_{ij} (E_{ij} - E_C^{(\alpha\beta)} + E_C^{(\alpha\beta)} - E_G)^2 / I_G$$

$$= \Delta_G'^2 + \delta[\Delta_G'^2] , \tag{17}$$

with

$$\Delta_G'^2 = \left[\frac{\sum_{C \in \Xi} N_C w_C^{(\alpha\beta)} (E_C^{(\alpha\beta)})^2}{I_G} - (E_G)^2 \right] , \tag{18}$$

$$\delta[\Delta_G'^2] \equiv \frac{\sum_{C \in \Xi} N_C w_C^{(\alpha\beta)} (\Delta_C^{(\alpha\beta)})^2}{I_G} . \tag{19}$$

In Eqs. (15)–(19) the summations are over UTAs represented by the initial and the final configurations C and C' connected by the orbital jump $\alpha \Longrightarrow \beta$. The quantities

$$w_C^{(\alpha\beta)} \equiv \frac{1}{N_C} \sum_{i \in C} N_i \sum_{j \in C'} w_{ij}, \quad N_C = \sum_{i \in C} N_i \tag{20}$$

$$E_C^{(\alpha\beta)} \equiv \frac{\sum_{i \in C} N_i \sum_{j \in C'} w_{ij} E_{ij}}{N_C w_C^{(\alpha\beta)}} , \tag{21}$$

and

$$(\Delta_C^{(\alpha\beta)})^2 \equiv \frac{\sum_{i \in C} N_i \sum_{j \in C'} w_{ij} (E_{ij} - E_C^{(\alpha\beta)})^2}{N_C w_C^{(\alpha\beta)}} , \tag{22}$$

are precisely the first three UTA moments. These were evaluated analytically by Bauche-Arnoult, Bauche, and Klapisch [2,3], assuming statistical distributions $N_i/N_C = g_i/g_C$, where g_i and g_C are the statistical weights of level i and configuration C, respectively. The results were obtained in terms of the *occupation numbers* of C and radial integrals. A compact presentation of these results is presented, for the case of orthogonal orbitals, in Appendix A. Equation (19) gives the correction to the width of the STA owing to the widths of the UTAs it comprises. The term $(\Delta_C^{(\alpha\beta)})^2$ was neglected in I. In addition to this correction in the STA width, it is important to note (see I) that the effect of individual lines in Eq. (21) includes the UTA shift of Eq. (A4) and by substitution in Eq. (16) these UTA shifts cause a global shift in the STA average energy.

Since the analytic expressions for the UTA moments do not depend on individual i,j levels, substituting them into Eqs. (15)–(17) replaces the sum over levels in Eqs. (4)–(6) with sums over configurations (i.e., occupation numbers). But the number of terms in these summations is, in general, still enormously high. To overcome this difficulty we use the same mathematical techniques developed in I. (i) Using binomial relations we first express the moments as generalized partition functions, which still include the summation *over occupation numbers.* (ii) Recursion formulas are then derived, without approximations, for the generalized partition functions that bypass the need for direct summation over occupation numbers. Details of this procedure are given in Appendix B; the results are summarized below.

A. Total STA intensity

The total intensity is not affected by the UTA widths and the result is as in I:

$$I_G = b(N/U_{\rm tot}) X_\alpha g_\alpha g_\beta \left[\prod_\sigma U_{Q'_\sigma}(g^{\alpha\beta}) \right] \times \left[\sum_\kappa (E_G)^{2\kappa-1} \begin{bmatrix} j_\alpha & \kappa & j_\beta \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}^2 (r_{\alpha\beta}^\kappa)^2 \right] , \tag{23}$$

where

$$b = \frac{8\pi^2 e a_0}{3h^4 c^3} ,$$

κ is the rank of the transition multipole (in most cases dipole $\kappa = 1$ is sufficient), and the statistical weight of shell s is $g_s = 2j_s + 1$. The Boltzmann factor X_s is

$$X_s = e^{-[(\varepsilon_s - \mu)/kT]} , \tag{24}$$

where ε_s is the energy of orbital s and μ is the chemical potential. The quantity $r_{\alpha\beta}^\kappa$ is the relativistic radial integral for the electron jump $\alpha \Longrightarrow \beta$,

$$Q'_\sigma = Q_\sigma - \tilde{\delta}_{\alpha\sigma}, \quad \tilde{\delta}_{\alpha\sigma} \equiv \begin{cases} 1, & \alpha \in \sigma \\ 0, & \alpha \notin \sigma , \end{cases} \tag{25}$$

$g^{\alpha\beta}$ is the set of modified orbital statistical weights with

$$g_s^{\alpha\beta} = g_s - \delta_{s\alpha} - \delta_{s\beta} , \tag{26}$$

for orbital s (here δ is the Kronecker delta), and

$$U_Q(g) \equiv \sum_{i \in \Xi} g_i e^{(E_i - Q\mu)/kT} \tag{27}$$

is the partition function of the superconfiguration Ξ occupied by Q electrons. N and $U_{\rm tot}$ are, respectively, the total number density and the partition function of the system.

B. Average STA transition energy

The average energy is given by

$$E_G = D_0 + \sum_\sigma \varepsilon_{\alpha\beta}^\sigma(D') , \tag{28}$$

where for the supershell σ

$$\varepsilon^{\sigma}_{\alpha\beta}(D')=\sum_{n=1}^{Q_\sigma}\phi^{\sigma}_{n}(D')U_{Q_\sigma-n}(g^{\alpha\beta})\ , \tag{29}$$

with

$$\phi^{\sigma}_{n}(D')\equiv-\sum_{s\in\sigma}g_s^{\alpha\beta}D'_s(-X_s)^n \tag{30}$$

and

$$D'_s=D_s+\left[\delta_{s\alpha}-\delta_{s\beta}\frac{2j_\alpha}{2j_\beta}\right]\delta E_{\alpha\beta}\ . \tag{31}$$

The explicit expressions for the radial one-body and two-body parts $D_0=\langle\beta\rangle-\langle\alpha\rangle$, $D_s=\langle s,\beta\rangle-\langle s,\alpha\rangle$, and for $\delta E_{\alpha\beta}$ are given in Appendix A.

C. STA variance

The result for the variance is

$$\Delta^2_G=\sum_{\sigma=1}^{N}\Delta^{\sigma}_{\alpha\beta}(D'')\ , \tag{32}$$

$$\Delta^{\sigma}_{\alpha\beta}(D'')=\sum_{n=1}^{Q_\sigma}\eta^{\sigma}_{n}(D'')U_{Q_\sigma-n}(g^{\alpha\beta})/U_{Q_\sigma}(g^{\alpha\beta})-[\varepsilon^{\sigma}_{\alpha\beta}(D')]^2\ , \tag{33}$$

$$\eta^{\sigma}_{n}(D'')\equiv\sum_{m=1}^{n-1}\phi^{\sigma}_{m}(D')\phi^{\sigma}_{n-m}(D')+n\phi^{\sigma}_{n}(D'^2+D''^2)\ , \tag{34}$$

where

$$\begin{aligned}&(D'^2+D''^2)_s\equiv(D'^2)_s+(D''^2)_s\ ,\\&(D'^2)_s\equiv D'^2_s\ ,\\&(D''^2)_s\equiv(g_s^{\alpha\beta}-1)\Delta^2_s\ ,\end{aligned} \tag{35}$$

where $\Delta^2_s\equiv(\Delta_s^{\alpha\beta})^2$ was derived in Refs. [2,3]. A compact and convenient presentation for this quantity appears for completeness in Appendix A as well. A comparison with the results of I shows that the UTA widths are included in the STA moments with the replacements $D_s\rightarrow D'_s$ in the working formulas for the average energy and $(D)^2_s\rightarrow(D')^2_s+(D'')^2_s$ in the second term of the variance.

The equations above are derived in Appendix B on the basis of the technique developed in I that assumes zeroth-order energies in the Boltzmann factors. This approximation can be removed by a common correction to the STA intensity that converges to the correct first-order results in the convergence procedure. This part is beyond the scope of the present subject and will be reported elsewhere.

IV. EXAMPLES OF THE UTA WIDTH IN PLASMA SPECTRA

The above analytical expressions were incorporated in the STA code and used to calculate the following typical spectra. The first example is the spectrum of the $2p_{3/2}$-$3d_{5/2}$ transition array in iron at $T=59$ eV and ion density 0.0113 g/cm^3, shown in Fig. 1. This transition

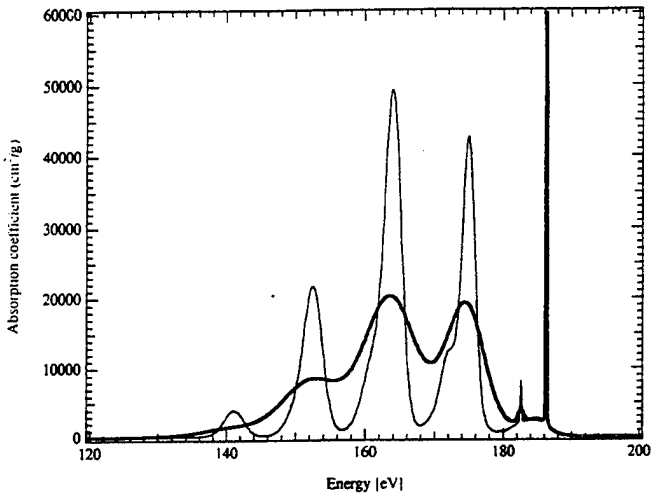


FIG. 1. Absorption spectrum of the $2p_{3/2}$-$3d_{5/2}$ transition array in iron at $T=59$ eV and an ion density of 0.0113 g/cm^3. The heavy and the thin lines describe the spectrum with and without the UTA widths, respectively.

array is the most intense under these conditions. In order to demonstrate the effect of the UTA widths we compare this spectrum with the one obtained without UTA widths as in I. The solid line includes the UTA widths and shifts while the thin line does not. We see in this example that the UTA width smears some of the structure.

In a recent experiment [8] the spectrum of the mixture 0.802 wt. % Fe+0.198 wt. % NaF was obtained at $T=59$ eV and an ion density of 0.0113 g/cm^3 (the same plasma conditions were used in Fig. 1). In Fig. 2 we show the calculated and the experimental transmission spectra. In this calculation, in addition to the bound-bound transitions we have also included the bound-free and free-free transitions as well as scattering. The main features are

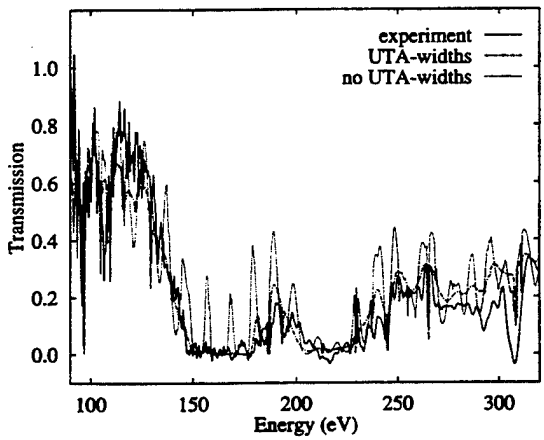


FIG. 2. Total transmission spectrum of the mixture 0.802 wt. % Fe+0.198 wt. % NaF at $T=59$ eV and an ion density of 0.0113 g/cm^3 (the same plasma conditions were used in Fig. 1). The dashed and the dotted lines describe the transmission with and without the UTA widths, respectively. The solid line is the experimental result.

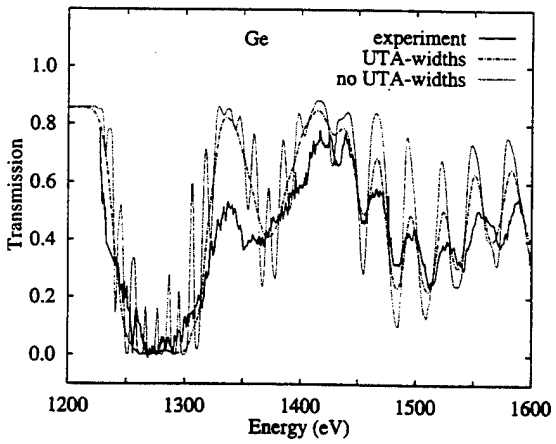


FIG. 3. Total transmission spectrum of Ge at $T=68$ eV and an ion density of 0.05 g/cm^3. The dashed and the dotted lines describe the transmission with and without the UTA widths, respectively. The solid line is the experimental result.

dominated by the bound-bound part. The dashed and the dotted lines present the calculations with and without UTA widths, respectively. The effect of UTA widths is obvious here. The third (solid) line in Fig. 2, which is the experimental result of Ref. [8], indicates the importance of UTA widths. In this case the total Rosseland opacity is also measured and within the experimental error agrees with the calculated result of the STA model including UTA widths, whereas the omission of these widths results in a discrepancy of about a factor of 2.

Another experiment reported recently [9] presents the transmission spectrum of Ge assumed at $T=76$ eV and an ion density of 0.05 g/cm^3. In Fig. 3 we compare the experimental results with calculations for these plasma conditions with and without the inclusion of the UTA widths. The calculated spectra here are for $T=68$ eV, which fits better to the experiment. Again the effect of the UTA widths is apparent and the agreement between the calculations with UTA widths and the experiment is very satisfactory.

V. DISCUSSION

In this work we have derived the analytical expressions for the STA moments including the effect of the detailed level structure within the configurations. In this respect the model is therefore equivalent to the detailed term accounting (DTA) approach. However, since it avoids dealing directly with levels, it is much faster. In fact, for heavier atoms where DTA calculations are impossible, the STA model is easily applied. It was shown above that the UTA width does not affect the formal form of the STA moments and with only an alternative definition of the radial quantities this extension, though fairly complicated, is easily incorporated in the numerical code and does not increase the required computer time.

The convergence procedure of the STA method leads in this case to the detailed UTA result where each UTA is a Gaussian with the UTA variance. This assumes that each UTA is completely unresolved. The results of I, on the other hand, ignore the UTA widths and are therefore equivalent to the detailed configuration accounting approach. The only assumption, still hidden in our corrected model, is that an UTA is completely unresolved. This approximation may overestimate the absorption when the number of lines within an UTA is very small so that the accumulated width of these lines is smaller compared to the UTA width. However, in all the experiments performed so far, as in the examples of Figs. 2 and 3, the effect of the UTA widths on the spectra is very important in reproducing the accurate absorption.

ACKNOWLEDGMENTS

We thank Dr. Springer from LLNL and Dr. Foster from AWE England and their experimental groups for providing us with their experimental data of references [8] and [9], respectively.

APPENDIX A: COMPACT FORM FOR THE ANALYTICAL EXPRESSIONS OF THE UTA MOMENTS

The results obtained by Bauche-Arnoult, Bauche, and Klapisch [2,3] for the UTA moments can be rewritten in a concise form. We present here the results for the case where the same set of orbitals is used for the initial and the final configurations. Furthermore, we extend the formulas to transitions of the more general case with electric multipoles of rank κ (not necessarily dipole). The working formulas are obtained as follows.

1. The UTA average energy

The UTA average energy

$$E_C^{(\alpha\beta)}=E_C^{\prime(\alpha\beta)}+\delta E_C^{\prime(\alpha\beta)} \tag{A1}$$

includes a shift $\delta E_C^{\prime(\alpha\beta)}$ from the difference between the corresponding configuration energy averages (first order),

$$E_C^{\prime(\alpha\beta)}\equiv E_{C'}-E_C\ , \tag{A2}$$

$$E_C^{\prime(\alpha\beta)}=D_0+\sum_s(q_s-\delta_{a\alpha})D_s\ . \tag{A3}$$

The shift is given by

$$\delta E_C^{\prime(\alpha\beta)}=\left[\frac{q_\alpha-1}{2j_\alpha}-\frac{q_\beta}{2j_\beta}\right]2j_\alpha\delta E(\mathrm{j}_\alpha^2\rightarrow\mathrm{j}_\alpha\mathrm{j}_\beta)\ . \tag{A4}$$

The radial quantities in (A3) and (A4) are

$$D_0\equiv D_0^{\mathrm{j}_\alpha\mathrm{j}_\beta}=\langle\mathrm{j}_\beta\rangle-\langle\mathrm{j}_\alpha\rangle\ , \tag{A5}$$

$$\langle\mathrm{j}_s\rangle\equiv\langle\mathrm{j}_s|h_D|\mathrm{j}_s\rangle\ , \tag{A6}$$

$$D_{\mathrm{j}_s}^{\mathrm{j}_\alpha\mathrm{j}_\beta}\equiv\langle\mathrm{j}_s,\mathrm{j}_\beta\rangle-\langle\mathrm{j}_s,\mathrm{j}_\alpha\rangle\ , \tag{A7}$$

$$\langle\mathrm{j}_r,\mathrm{j}_s\rangle=F^0(\mathrm{j}_r,\mathrm{j}_s)-\frac{g_{j_s}}{(g_{j_s}-\delta_{\mathrm{j}_r,\mathrm{j}_s})}\times\sum_k\begin{bmatrix}j_r & k & j_s\\ \frac{1}{2} & 0 & -\frac{1}{2}\end{bmatrix}^2G^k(\mathrm{j}_r,\mathrm{j}_s)\ , \tag{A8}$$

where h_D is the Dirac single-particle zeroth-order Hamiltonian, F^k and G^k are the Slater integrals, and

$$2j_\alpha \delta E(\mathrm{j}_\alpha^2 \rightarrow \mathrm{j}_\alpha \mathrm{j}_\beta) = F^{\mathrm{j}_\alpha,\mathrm{j}_\beta} + G^{\mathrm{j}_\alpha,\mathrm{j}_\beta} , \tag{A9}$$

with

$$F^{\mathrm{j}_\alpha \mathrm{j}_\beta} \equiv - \sum_{k \neq 0 \text{ even}} g_{j_\alpha} g_{j_\beta} \begin{Bmatrix} k & j_\alpha & j_\alpha \\ \kappa & j_\beta & j_\beta \end{Bmatrix} \begin{bmatrix} j_\alpha & k & j_\alpha \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \times \begin{bmatrix} j_\beta & k & j_\beta \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} F^k(\mathrm{j}_\alpha \mathrm{j}_\beta) ,$$

$$G^{\mathrm{j}_\alpha,\mathrm{j}_\beta} = \sum_k \frac{g_{j_\alpha} g_{j_\beta} \delta_{k,\kappa} - (2\kappa+1)}{2\kappa+1} \begin{bmatrix} j_\alpha & k & j_\beta \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}^2 G^k(\mathrm{j}_\alpha, \mathrm{j}_\beta) \tag{A10}$$

depending on the radial orbitals, but not on the shell occupation numbers.

We thus notice immediately from (A4) that the UTA shift has the same dependence on the occupation numbers as in (A3), i.e.,

$$E_C^{(\alpha\beta)} = D_0 + \sum_s (q_s - \delta_{s\alpha}) D_s' , \tag{A11}$$

with the following substitution:

$$D_s' = D_s + \left[\delta_{s\alpha} - \delta_{s\beta} \frac{2j_\alpha}{2j_\beta} \right] \delta E_{\alpha\beta} . \tag{A12}$$

2. The UTA variance

The compact form of the results of Bauche-Arnoult, Bauche, and Klapisch [2,3] for the UTA variance is

$$(\Delta_C^{(\alpha\beta)})^2 = \sum_s (q_s - \delta_{s\alpha})(g_s - q_s - \delta_{s\beta}) \Delta_s^2 , \tag{A13}$$

where Δ_s^2 depends only on radial integrals but not on occupation numbers. This general form of the occupation numbers dependence as written in Eq. (A13) is not apparent in Refs. [2,3]. The explicit expressions for Δ_s^2 of (A13) is

$$\Delta_s^2 = \frac{\Delta^2(\mathrm{j}_s \mathrm{j}_\alpha, \mathrm{j}_s \mathrm{j}_\beta)}{2j_s - \delta_{s\alpha} - \delta_{s\beta}} , \tag{A14}$$

where $\Delta^2(\mathrm{j}_s \mathrm{j}_\alpha, \mathrm{j}_s \mathrm{j}_\beta)$ for $s = \alpha, \beta$ and $s \neq \alpha, \beta$ are given below for simplicity in terms of the factored radial Slater integrals

$$\begin{aligned} \bar{F}^k(\mathrm{j}_a \mathrm{j}_b) &\equiv X^k(\mathrm{j}_a \mathrm{j}_b, \mathrm{j}_a \mathrm{j}_b) , \\ \bar{G}^k(\mathrm{j}_a \mathrm{j}_b) &\equiv X^k(\mathrm{j}_a \mathrm{j}_b, \mathrm{j}_b \mathrm{j}_a) , \end{aligned} \tag{A15}$$

where

$$X^k(\mathrm{j}_a \mathrm{j}_b, \mathrm{j}_c \mathrm{j}_d) \equiv \langle \mathrm{j}_a \| C^{(k)} \| \mathrm{j}_c \rangle \langle \mathrm{j}_b \| C^{(k)} \| \mathrm{j}_d \rangle \times R^k(\mathrm{j}_a \mathrm{j}_b, \mathrm{j}_c \mathrm{j}_d) , \tag{A16}$$

$$\langle \mathrm{j} \| C^{(k)} \| \mathrm{j}' \rangle = (-1)^{j+1/2} [j, j']^{1/2} \begin{bmatrix} j & k & j' \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \Pi(lkl') , \tag{A17}$$

and $\Pi(lkl') = 1$ when l, k, and l' obey the triangle condition and $l + l' + k$ is even, and zero otherwise:

$$\Delta^2(\mathrm{j}_s \mathrm{j}_\alpha, \mathrm{j}_s \mathrm{j}_\beta) = \mathcal{A}_s + \mathcal{B}_s + \mathcal{C}_s + \mathcal{D}_s + \mathcal{E}_s + \mathcal{F}_s(\alpha\beta) + \mathcal{F}_s(\beta\alpha) , \tag{A18}$$

where

$$\mathcal{A}_s = \sum_{\nu=\alpha,\beta} \sum_{k \,(\neq 0) \text{ even}} \frac{1}{(2k+1)(2j_s+1)(2j_\nu+1)} \times [\bar{F}^k(\mathrm{j}_s, \mathrm{j}_\nu)]^2 ,$$

$$\mathcal{B}_s = \sum_{\nu=\alpha,\beta} \sum_k \sum_{k'} \left[\frac{\delta_{kk'}}{2k+1} - \frac{1}{(2j_s+1)(2j_\nu+1)} \right] \times \frac{1}{(2j_s+1)(2j_\nu+1)} \times \bar{G}^k(\mathrm{j}_s, \mathrm{j}_\nu) \bar{G}^{k'}(\mathrm{j}_s, \mathrm{j}_\nu) ,$$

$$\mathcal{C}_s = \sum_{\nu=\alpha,\beta} \sum_{k \,(\neq 0)} \sum_{k'} 2(-1)^{k'} \begin{Bmatrix} j_\nu & j_s & k' \\ j_s & j_\nu & k \end{Bmatrix} \times \frac{1}{(2j_s+1)(2j_\nu+1)} \times \bar{F}^k(\mathrm{j}_s, \mathrm{j}_\nu) \bar{G}^{k'}(\mathrm{j}_s, \mathrm{j}_\nu) ,$$

$$\mathcal{D}_s = - \sum_k \frac{2}{2k+1} \begin{Bmatrix} j_\alpha & j_\alpha & k \\ j_\beta & j_\beta & \kappa \end{Bmatrix} \frac{1}{(2j_s+1)} \times (-1)^{j_\alpha + j_\beta + 1} \bar{F}^k(\mathrm{j}_s, \mathrm{j}_\alpha) \bar{F}^k(\mathrm{j}_s, \mathrm{j}_\beta) ,$$

$$\mathcal{E}_s = 2 \sum_k \sum_{k'} \left[- \begin{Bmatrix} k & k' & \kappa \\ j_\beta & j_\alpha & j_s \end{Bmatrix}^2 + \frac{1}{(2j_s+1)(2j_\alpha+1)(2j_\beta+1)} \right] \times \frac{1}{(2j_s+1)} (-1)^{j_\alpha + j_\beta + 1} \times \bar{G}^k(\mathrm{j}_s, \mathrm{j}_\alpha) \bar{G}^{k'}(\mathrm{j}_s, \mathrm{j}_\beta) ,$$

$$\mathcal{F}_s(\alpha\beta) = 2 \sum_{k \neq 0 \text{ even}} \sum_{k'} \left[-(-1)^{k'} \begin{Bmatrix} j_\beta & j_\beta & k \\ j_\alpha & j_\alpha & \kappa \end{Bmatrix} \times \begin{Bmatrix} j_\beta & j_\beta & k \\ j_s & j_s & k' \end{Bmatrix} \right] \times \frac{1}{(2j_s+1)} (-1)^{j_\alpha + j_\beta + 1} \times \bar{F}^k(\mathrm{j}_s, \mathrm{j}_\alpha) \bar{G}^{k'}(\mathrm{j}_s, \mathrm{j}_\beta) ,$$

and $\mathcal{F}(\beta\alpha)$ is the same as $\mathcal{F}(\alpha\beta)$ with the interchange $\alpha \leftrightarrow \beta$. For $s = \alpha$,

$$\Delta^2(\mathbf{j}_s\mathbf{j}_\alpha,\mathbf{j}_s\mathbf{j}_\beta)=\mathcal{A}'_s+\mathcal{B}'_s+\mathcal{C}'_s+\mathcal{D}'_s+\mathcal{E}'_s+\mathcal{F}'_s\ , \tag{A19}$$

$$\mathcal{A}'_s=\sum_{k\,(\neq 0)\ \text{even}}\ \sum_{k'\,(\neq 0)\ \text{even}}\left[\frac{\delta_{kk'}}{2k+1}-\frac{1}{2j_\alpha(2j_\alpha+1)}+\begin{Bmatrix} j_\alpha & j_\alpha & k\\ j_\alpha & j_\alpha & k'\end{Bmatrix}\right]\frac{1}{2j_\alpha(2j_\alpha+1)}\bar{F}^k(\mathbf{j}_\alpha,\mathbf{j}_\alpha)\bar{F}^{k'}(\mathbf{j}_\alpha,\mathbf{j}_\alpha)\ ,$$

$$\mathcal{B}'_s=\sum_{k\,(\neq 0)\ \text{even}}\ \sum_{k'\,(\neq 0)\ \text{even}}\begin{Bmatrix} j_\alpha & j_\alpha & k'\\ j_\beta & j_\beta & \kappa\end{Bmatrix}\left[-\frac{\delta_{kk'}}{2k+1}-\begin{Bmatrix} j_\alpha & j_\alpha & k'\\ j_\alpha & j_\alpha & k\end{Bmatrix}+\frac{1}{2j_\alpha(2j_\alpha+1)}\right]$$
$$\times\frac{1}{j_\alpha}(-1)^{j_\alpha+j_\beta+1}\bar{F}^k(\mathbf{j}_\alpha,\mathbf{j}_\alpha)\bar{F}^{k'}(\mathbf{j}_\alpha,\mathbf{j}_\beta)\ ,$$

$$\mathcal{C}'_s=\sum_{k\,(\neq 0)\ \text{even}}\ \sum_{k'}\left[-\begin{Bmatrix} k & k' & \kappa\\ j_\beta & j_\alpha & j_\alpha\end{Bmatrix}^2+\begin{Bmatrix} j_\beta & j_\beta & k\\ j_\alpha & j_\alpha & \kappa\end{Bmatrix}\begin{Bmatrix} j_\beta & j_\beta & k\\ j_\alpha & j_\alpha & k'\end{Bmatrix}-\frac{1}{2j_\alpha(2j_\alpha+1)}\left[\frac{\delta_{k'\kappa}}{2\kappa+1}-\frac{1}{2j_\beta+1}\right]\right]$$
$$\times\frac{1}{j_\alpha}(-1)^{j_\alpha+j_\beta+1}\bar{F}^k(\mathbf{j}_\alpha,\mathbf{j}_\alpha)\bar{G}^{k'}(\mathbf{j}_\alpha,\mathbf{j}_\beta)\ ,$$

$$\mathcal{D}'_s=\sum_{k\,(\neq 0)\ \text{even}}\ \sum_{k'\,(\neq 0)\ \text{even}}\left[\frac{\delta_{kk'}}{(2k+1)(2j_\beta+1)}-\begin{Bmatrix} j_\alpha & j_\alpha & k\\ j_\alpha & \kappa & j_\beta\\ k' & j_\beta & j_\beta\end{Bmatrix}\right.$$
$$\left.-\frac{1}{2j_\alpha}\begin{Bmatrix} j_\alpha & j_\beta & \kappa\\ j_\beta & j_\alpha & k\end{Bmatrix}\begin{Bmatrix} j_\alpha & j_\beta & \kappa\\ j_\beta & j_\alpha & k'\end{Bmatrix}\right]\frac{1}{2j_\alpha}\bar{F}^k(\mathbf{j}_\alpha,\mathbf{j}_\beta)\bar{F}^{k'}(\mathbf{j}_\alpha,\mathbf{j}_\beta)\ ,$$

$$\mathcal{E}'_s=\sum_k\sum_{k'}\left[\frac{\delta_{kk'}}{(2k+1)(2j_\beta+1)}-\begin{Bmatrix} j_\alpha & j_\beta & k\\ j_\beta & \kappa & j_\alpha\\ k' & j_\alpha & j_\beta\end{Bmatrix}-\frac{1}{2j_\alpha}\left[\frac{\delta_{k\kappa}}{2\kappa+1}-\frac{1}{(2j_\beta+1)}\right]\left[\frac{\delta_{k'\kappa}}{2\kappa+1}-\frac{1}{(2j_\beta+1)}\right]\right]$$
$$\times\frac{1}{2j_\alpha}\bar{G}^k(\mathbf{j}_\alpha,\mathbf{j}_\beta)\bar{G}^{k'}(\mathbf{j}_\alpha,\mathbf{j}_\beta)\ ,$$

$$\mathcal{F}'_s=\sum_{k\,(\neq 0)\ \text{even}}\ \sum_{k'}\left[\frac{-1}{2j_\beta+1}\begin{Bmatrix} j_\beta & j_\beta & k\\ j_\alpha & j_\alpha & k'\end{Bmatrix}-(-1)^{j_\alpha+j_\beta}\begin{Bmatrix} k & k' & \kappa\\ j_\alpha & j_\beta & j_\beta\end{Bmatrix}\begin{Bmatrix} k & k' & \kappa\\ j_\beta & j_\alpha & j_\alpha\end{Bmatrix}\right.$$
$$\left.+\frac{1}{2j_\alpha}\begin{Bmatrix} j_\alpha & j_\beta & \kappa\\ j_\beta & j_\alpha & k\end{Bmatrix}\left[\frac{\delta_{k'\kappa}}{2\kappa+1}-\frac{1}{2j_\beta+1}\right]\right]\frac{1}{j_\alpha}\bar{F}^k(\mathbf{j}_\alpha,\mathbf{j}_\beta)\bar{G}^{k'}(\mathbf{j}_\alpha,\mathbf{j}_\beta)\ .$$

APPENDIX B: STA MOMENTS INCLUDING UTA WIDTHS

Equation (13) for the STA average energy reads

$$E_G=\frac{\sum_C N_C w_C^{(\alpha\beta)}E_C^{(\alpha\beta)}}{I_G}\ . \tag{B1}$$

It is seen immediately that the shifted and the unshifted average energies (A3) and (A11), respectively, have exactly the same occupation number dependence. Since the derivation procedure of I operates only on occupation numbers, the result for the STA average energy has the same form as that of I:

$$E_G=D_0+\sum_\sigma \varepsilon^\sigma_{\alpha\beta}(D)\ , \tag{B2}$$

$$\varepsilon^\sigma_{\alpha\beta}(D)=\sum_{n=1}^{Q_\sigma}\phi^\sigma_n(D)U_{Q_\sigma-n}(g^{\alpha\beta})/U_{Q_\sigma}(g^{\alpha\beta})\ , \tag{B3}$$

with the substitution $D_s\rightarrow D'_s$ of (A12). The final results appear in Eqs. (28)–(31).

The derivation for the STA variance due to configuration widths is more complex. As shown in Eq. (17), the correction of the STA variance due to UTA variances is given by

$$\delta[\Delta'^2_G]=\frac{\sum_{C\in\Xi}N_C w_C^{(\alpha\beta)}(\Delta_C^{(\alpha\beta)})^2}{I_G}\ . \tag{B4}$$

Using the relations [1]

$$N_C=\frac{N}{U}\prod_{s\in C}\begin{bmatrix}g_s\\q_s\end{bmatrix}X_s^{q_s}\,, \tag{B5}$$

$$w_C^{(\alpha\beta)}=b(E_G^{(\alpha\beta)})^{2\kappa-1}q_\alpha(g_\beta-q_\beta)\times\begin{bmatrix}j_\alpha & \kappa & j_\beta\\ \frac{1}{2} & 0 & -\frac{1}{2}\end{bmatrix}^2(r_{\alpha\beta}^{(\kappa)})^2\,, \tag{B6}$$

$$I_G=\sum_{C,C'\in G}N_C w_C^{(\alpha\beta)} \tag{B7}$$

together with (A13) gives, by substitution in (B2),

$$\delta[\Delta_G'^2]=\frac{\sum_C\sum_r\Delta_r^2(q_r-\delta_{r\alpha})(g_r-\delta_{r\beta}-q_r)q_\alpha(g_\beta-q_\beta)\prod_{s\in C}\begin{bmatrix}g_s\\q_s\end{bmatrix}X_s^{q_s}}{\sum_C q_\alpha(g_\beta-q_\beta)\prod_{s\in C}\begin{bmatrix}g_s\\q_s\end{bmatrix}X_s^{q_s}}\,. \tag{B8}$$

It should be mentioned that the $3j$ symbol in Eq. (B6) contains the hidden triangular and the parity conditions $\Delta(l_\alpha,\kappa,l_\beta)$ and $l_\alpha+\kappa+l_\beta$ even. Using the binomial relations

$$q\begin{bmatrix}g\\q\end{bmatrix}=g\begin{bmatrix}g-1\\q-1\end{bmatrix},\quad (g-q)\begin{bmatrix}g\\q\end{bmatrix}=g\begin{bmatrix}g-1\\q\end{bmatrix} \tag{B9}$$

we can follow the steps below:

$$\delta[\Delta_G'^2]=\frac{\sum_{C\in\Xi}\sum_r\Delta_r^2 q_r^\alpha(g_r^\beta-q_r)\prod_s\begin{bmatrix}g_s^{\alpha\beta}\\q_s^\alpha\end{bmatrix}X_s^{q_s^\alpha}}{\sum_{C\in\Xi}\prod_s\begin{bmatrix}g_s^{\alpha\beta}\\q_s^\alpha\end{bmatrix}X_s^{q_s^\alpha}}=\frac{\sum_r\Delta_r^2 g_r^{\alpha\beta}(g_r^{\alpha\beta}-1)X_r\sum_{C\in\Xi}\prod_s\begin{bmatrix}g_s^{\alpha\beta rr}\\q_s^{\alpha r}\end{bmatrix}X_s^{q_s^{\alpha r}}}{\sum_{C\in\Xi}\prod_s\begin{bmatrix}g_s^{\alpha\beta}\\q_s^\alpha\end{bmatrix}X_s^{q_s^\alpha}}$$

$$=\frac{\sum_r\Delta_r^2 g_r^{\alpha\beta}(g_r^{\alpha\beta}-1)X_r\prod_{\sigma\in\Xi}\sum_{\sum_{s\in\sigma}q_s=Q_\sigma}\prod_{s\in\sigma}\begin{bmatrix}g_s^{\alpha\beta rr}\\q_s^{\alpha r}\end{bmatrix}X_s^{q_s^{\alpha r}}}{\prod_{\sigma\in\Xi}\sum_{\sum_{s\in\sigma}q_s=Q_\sigma}\prod_{s\in\sigma}\begin{bmatrix}g_s^{\alpha\beta}\\q_s^\alpha\end{bmatrix}X_s^{q_s^\alpha}}=\frac{\sum_r\Delta_r^2 g_r^{\alpha\beta}(g_r^{\alpha\beta}-1)X_r\prod_{\sigma\in\Xi}U_{Q_\sigma^{\alpha r}}(g^{\alpha\beta rr})}{\prod_{\sigma\in\Xi}U_{Q_\sigma^\alpha}(g^{\alpha\beta})}\,, \tag{B10}$$

where, in general,

$$g_s^{abcd\cdots}\equiv g_s-\delta_{as}-\delta_{bs}-\delta_{cs}-\delta_{ds}-\cdots\,,$$

$$q_s^{abcd\cdots}\equiv q_s-\delta_{as}-\delta_{bs}-\cdots\,,$$

$(g^{abcd\cdots})$ stands for the set of values for all s, and

$$Q_\sigma^{ab}\equiv Q_\sigma-\bar\delta_{a\sigma}-\bar\delta_{b\sigma},\quad \bar\delta_{a\sigma}=\begin{cases}1, & a\in\sigma\\ 0, & a\notin\sigma.\end{cases} \tag{B11}$$

In Eq. (B10) all the terms with the shell r, $r\notin\sigma$, cancel out from both the nominator and denominator and we obtain

$$\delta[\Delta_G'^2]=\sum_r\frac{\Delta_r^2 g_r^{\alpha\beta}(g_r^{\alpha\beta}-1)X_r U_{Q_r-1}(g^{\alpha\beta rr})}{U_{Q_r}(g^{\alpha\beta})}\,, \tag{B12}$$

where

$$Q_r\equiv Q_\sigma-\bar\delta_{a\sigma} \tag{B13}$$

and σ is the supershell containing the shell r. The appearance of r twice in the set of statistical weights $g^{\alpha\beta rr}$ means, according to (B9), that for the r shell the weight should be reduced by 2 (or 3 in the cases $r=\alpha,\beta$). We now use the identity [1]

$$U_{Q_\sigma}(g^\gamma)=\sum_{n=0}^{Q}(-X_\gamma)^n U_{Q_\sigma-n}(g)\,, \tag{B14}$$

leading to

$$U_{Q_r-1}(g^{\alpha\beta rr})=\sum_{n=0}^{Q_r-1}(-X_r)^n U_{Q_r-1-n}(g^{\alpha\beta r})$$

$$=\sum_{n=0}^{Q_r-1}(-X_r)^n\sum_{m=0}^{Q_r-1-n}(-X_r)^m\times U_{Q_r-1-n-m}(g^{\alpha\beta})\,, \tag{B15}$$

and defining $k=n+m+1$ we obtain

$$U_{Q_r-1}(g^{\alpha\beta rr})=\sum_{n=0}^{Q_r-1}(-X_r)^n\sum_{k=n+1}^{Q_r}(-X_r)^{k-n-1}\times U_{Q_r-k}(g^{\alpha\beta})\,. \tag{B16}$$

With substitution in (B12)

$$\delta[\Delta_G'^2]=\sum_r \frac{\Delta_r^2 g_r^{\alpha\beta}(g_r^{\alpha\beta}-1)X_r \sum_{k=1}^{Q_r} k(-X_r)^{k-1} U_{Q_r-k}(g^{\alpha\beta})}{U_{Q_r}(g^{\alpha\beta})}$$

$$=\sum_\sigma \frac{-\sum_{k=1}^{Q_\sigma} U_{Q_\sigma^\alpha-k}(g^{\alpha\beta})k \sum_{r\in\sigma}\Delta_r^2 g_r^{\alpha\beta}(g_r^{\alpha\beta}-1)(-X_r)^k}{U_{Q_\sigma^\alpha}(g^{\alpha\beta})}=\sum_\sigma \frac{\sum_{k=1}^{Q_\sigma} U_{Q_\sigma^\alpha-k}(g^{\alpha\beta})k\phi_k^\sigma(\{D''^2\})}{U_{Q_\sigma^\alpha}(g^{\alpha\beta})}, \tag{B17}$$

where

$$\phi_n^\sigma(D''^2)\equiv -\sum_{s\in\sigma} g_s^{\alpha\beta}\{D''^2\}_s(-X_s)^n \tag{B18}$$

and

$$(D''^2)_s\equiv(g_s^{\alpha\beta}-1)\Delta_s^2 , \tag{B19}$$

the results can be summarized as

$$\delta[\Delta_G'^2]=\sum_{\sigma\in G}\delta[\Delta_\sigma'^2] , \tag{B20}$$

where

$$\delta[\Delta_\sigma'^2]=\frac{\sum_{n=1}^{Q_\sigma} n\phi_n^\sigma(D''^2)U_{Q_\sigma-n}(g^{\alpha\beta})}{U_{Q_\sigma}(g^{\alpha\beta})} . \tag{B21}$$

The working formula of I for the variance where the UTA widths were neglected is

$$\Delta_G'^2=\sum_{\sigma\in G}\Delta_\sigma'^2 , \tag{B22}$$

where

$$\Delta_\sigma'^{\alpha\beta}=\Delta_\sigma'^{\alpha\beta}(D)$$

$$=\sum_{n=1}^{Q_\sigma}\eta_n^\sigma(D)U_{Q_\sigma-n}(g^{\alpha\beta})/U_{Q_\sigma}(g^{\alpha\beta})-(\varepsilon_{\alpha\beta}^\sigma(D))^2 , \tag{B23}$$

with

$$\eta_n^\sigma(D)=\sum_{m=1}^{n-1}\phi_m^\sigma(D)\phi_{n-m}^\sigma(D)+n\phi_n^\sigma(D^2) . \tag{B24}$$

It is seen immediately that the same working formula holds for the corrected variance with the replacement

$$(D^2)_s\rightarrow(D^2)_s+(D''^2)_s \tag{B25}$$

in the second term of (B24), i.e., the factor in (B23) should be replaced by

$$\eta_n^\sigma(D'')\equiv\sum_{m=1}^{n-1}\phi_m^\sigma(D)\phi_{n-m}^\sigma(D)+n\phi_n^\sigma(D^2+D''^2) , \tag{B26}$$

as presented in the text.

[1] A. Bar-Shalom, J. Oreg, W. H. Goldstein, D. Shvarts, and A. Zigler, Phys. Rev. A **40**, 3183 (1989).

[2] C. Bauche-Arnoult, J. Bauche, and M. Klapisch, Phys. Rev. A **25**, 2641 (1982).

[3] C. Bauche-Arnoult, J. Bauche, and M. Klapisch, Phys. Rev. A **31**, 2248 (1985).

[4] A. Bar-Shalom, J. Oreg, and W. H. Goldstein, in *Radiative Properties of Hot Dense Matter*, edited by W. H. Goldstein, C. Hooper, J. Gautier, J. Seely, and R. Lee (World Scientific, Singapore, 1990).

[5] A. Bar-Shalom, J. Oreg, and W. H. Goldstein, in *Atomic Processes in Plasmas*, edited by E. S. Marmar and J. L. Terry, AIP Conf. Proc. No. 257 (AIP, New York, 1991), p. 68.

[6] A. Bar-Shalom, J. Oreg, and W. H. Goldstein, J. Quant. Spectosc. Radiat. Transfer **51**, 27 (1994).

[7] P. T. Springer, T. S. Perry, D. F. Fields, W. H. Goldstein, B. G. Wilson, and R. E. Stewart, in *Atomic Processes in Plasmas* (Ref. [5]), p. 78.

[8] P. T. Springer, D. F. Fields, B. G. Wilson, J. K. Nash, W. H. Goldstein, C. A. Iglesias, F. J. Rogers, J. K. Swenson, M. H. Chen, A. Bar-Shalom, and R. E. Stewart, Phys. Rev. Lett. **69**, 3735 (1992).

[9] J. M. Foster, D. J. Hoarty, C. C. Smith, P. A. Rosen, S. J. Davidson, S. J. Rose, T. S. Perry, and F. J. D. Serduke, Phys. Rev. Lett. **67**, 3255 (1991).

Appendix L

Phase-Amplitude Algorithms for Atomic Continuum Orbitals and Radial Integrals

ELSEVIER

Computer Physics Communications 93 (1996) 21–32

Computer Physics Communications

Phase-amplitude algorithms for atomic continuum orbitals and radial integrals

A. Bar-Shalom [a], M. Klapisch [b,1], J. Oreg [a]

[a] *Nuclear Research Center Negev, P.O. Box 9001, Beer Sheva 84190, Israel*
[b] *ARTEP Inc., Columbia, MD 21045, USA*

Received 24 May 1995

Abstract

We present a new, fast and accurate phase-amplitude algorithm for the calculation of atomic continuum orbitals needed for cross sections computations of various atomic processes in plasmas. A coarse, energy independent, mesh is sufficient to achieve high accuracy. A straightforward application of a predictor-corrector procedure to the non-linear differential equations would fail, in particular for high energy free electrons in any atomic potential. The present algorithm overcomes this problem. In addition, we describe a novel method for calculating the radial integrals by integration over the phases instead of r. With the use of Gaussian trigonometric formulas over half periods, the integrals are expressed as alternating series. Levin's transform for convergence acceleration then provides the sum of the series with a few terms only. These methods are applicable in a relativistic framework as well as non-relativistic.

1. Introduction

The calculation of continuum orbitals is a bottleneck in many areas of central interest to atomic and plasma spectroscopy. An important application is the collisional radiative model which requires cross sections for the various atomic processes in plasmas such as the collisional ionization and excitation, autoionization and electron capture, photoionization and radiative recombination, etc. For LTE plasmas these continuum orbitals are essential for the calculation of the bound–free and free–free spectra. For all these calculations, a large amount of continuum orbitals of different energy in a general atomic potential are required. An obvious approach [1,2] is to solve the Schrödinger (or Dirac) equation directly, on a mesh of points that is dense enough to accommodate for the highest energy orbital considered in the problem, and also extends to values of r large enough to reach the asymptotic behavior of the functions necessary for the normalization and the convergence of the integrals. This means several thousands of points per wavefunction, and it results in computations that are expensive in time and memory. This is a serious drawback when large numbers of continuum orbitals are needed [3,4].

We give here an algorithm based on the phase-amplitude (PA) representation [5] that describes the highly oscillating orbitals by a product of two smooth functions, an amplitude and a phase. This representation was investigated previously [6–10] in the context of improving the WKB approximation and as a

[1] Mailing Address: Laser Plasma Branch, Code 6730, Naval Research Laboratory, Washington DC 20375-5320, USA.

SSDI 0010-4655(95)00100-X

tool for obtaining exact solutions for a variety of specific problems. For atomic systems in hot and dense plasmas however, one needs to solve the Schrödinger or Dirac equation for free electrons in a generally non-Coulomb atomic potential. This potential is assumed to include an approximation of exchange, so our differential equations are homogeneous. Since efficiency is an issue here, we wish to use predictor-corrector methods [11], which are known to be the fastest methods for solving differential equations [12]. It turns out, however, that in general direct application of the usual predictor-corrector solver, as well as other commonly used numerical methods, on the PA equations encounters severe numerical problems especially at large r values and in particular for high-energy free electrons. This may be one of the reasons why the PA approach has not been applied earlier to such problems, the other main reason being that these orbitals are needed for the calculations of integrals, and there was no method up to now that could take advantage directly of the PA representation. Therefore, it made no sense to try and solve an apparently more difficult differential equation, and then to expand or interpolate the outcome in order to compute the integrals.

In this work we identify the origin of these problems and present solutions within the PA approach. The result is an extremely fast, compact and accurate method which allows to calculate continuum orbitals on the same logarithmic mesh as bound wavefunctions. Furthermore, we present an new algorithm taking advantage of the PA for calculating the radial integrals by integration over the phases instead of r. This avoids the necessity of expanding the wavefunctions, and generating a large grid altogether. The method is applicable to both Schrödinger and Dirac continuum orbitals and could be used for high Rydberg bound states as well.

We have developed in the last decade two atomic codes for spectroscopic diagnostics of plasma systems: The HULLAC code on the one hand [4,13], calculates cross sections for all atomic processes and solves the collisional radiative model rate equations for the level populations, and on the other hand, the STA code computes absorption and emission spectra in LTE plasmas [14]. In both codes the PA algorithms presented in this paper have been used extensively and this contributes to their high efficiency.

In Section 2 we establish the PA equations for a logarithmic mesh. In Section 3 we explain the numerical failure of direct application of commonly used numerical schemes in the solution of the PA equations and present a new scheme overcoming these difficulties. In Section 4 we describe the method for calculating the radial integrals. Accuracy and efficiency are discussed in Section 5.

2. The phase-amplitude equations

2.1. The differential equations

Let us write the radial Schrödinger equation (in atomic units) for continuum orbitals [15] with angular momentum ℓ and energy $\varepsilon = k^2/2$, in a potential $U(r)$, as

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2} P_{\varepsilon\ell} + \omega(r)\ P_{\varepsilon\ell} = 0, \tag{1}$$

where

$$\omega(r) = k^2 + 2U(r) - \frac{\ell(\ell+1)}{r^2}, \tag{2}$$

with the boundary condition

$$P_{\varepsilon\ell}(0) = 0 \tag{3}$$

and the asymptotic behavior

$$P_{\varepsilon\ell}(r)\,|_{r\to\infty} \approx \sqrt{\frac{2}{\pi k}}\,\sin(\theta+\tau), \tag{4}$$

where

$$\theta = kr + \mathrm{O}(\ln(kr)). \tag{5}$$

τ is the phase shift, and the normalization of $P_{\varepsilon\ell}(r)$ (in atomic units) is defined by:

$$\int_0^\infty P_{\varepsilon\ell} P_{\varepsilon'\ell}\,\mathrm{d}r = \delta(\varepsilon - \varepsilon'). \tag{6}$$

It is shown in Appendix A that in the relativistic (Dirac) case, or in the "semi-relativistic" approximation, one can also obtain a second-order equation with an effective potential. Therefore, from now on, we consider a general equation of the form,

$$\frac{\mathrm{d}^2 P}{\mathrm{d}r^2} + \omega(r)\ P = 0. \tag{7}$$

The solution acquires an oscillatory character when $\omega(r) > 0$. For electrons in the continuum ($\varepsilon > 0$) which are the main interest of this paper, this leads to the definition of the turning point r_t:

$$\omega(r_t) \equiv 0. \tag{8}$$

Now, let us introduce the PA representation

$$P(r) = y(r) \sin \varphi(r). \tag{9}$$

By substitution in Eq. (7) we obtain the following separate equations for the phase and the amplitude

$$\frac{d^2 y}{dr^2} + \omega(r)\, y - \frac{a^2}{y^3} = 0, \tag{10}$$

$$\varphi(r) = \int_{r_0}^{r} \frac{a}{y^2}\, dr. \tag{11}$$

The non-linear form of Eq. (10), together with the choice

$$a = \frac{2}{\pi}, \tag{12}$$

ensures the correct normalization of $P(r)$. The choice of r_0 appearing in Eq. (11) is discussed in Section 3.4.

2.2. Logarithmic mesh

Since the functions y and φ are "smooth" (i.e. not oscillatory) we can use a very coarse mesh and in fact we opt for the same logarithmic mesh as used for the bound orbitals [16] i.e.

$$\rho = \ln r. \tag{13}$$

To keep the same form of equation without a first derivative, let us define a new function

$$F(r) = r^{-1/2} P(r). \tag{14}$$

We obtain for F an equation of the same form

$$\frac{d^2}{d\rho^2} F + WF = 0, \tag{15}$$

with

$$W = r^2 \omega - \tfrac{1}{4}. \tag{16}$$

The PA equations take now the form:

$$F(\rho) = Y(\rho) \sin \Phi(\rho), \tag{17}$$

$$\frac{d^2}{d\rho^2} Y + W(\rho)\, Y - \frac{a^2}{Y^3} = 0, \tag{18}$$

$$\Phi(\rho) = \int_{\rho_0}^{\rho} \frac{a}{Y^2}\, d\rho', \tag{19}$$

with the asymptotic behavior

$$F(\rho)\,|_{\rho \to \infty} = \left(\frac{2}{\pi k e^{\rho}}\right)^{1/2} \sin(\theta + \tau). \tag{20}$$

3. The numerical technique

The general strategy in the numerical solution is as follows: In order to use algorithms that are stable and efficient, we will distinguish three regions in r (see Table 1): From the asymptotic values of the solution at large r we integrate Eq. (18) for the amplitude from a sufficiently large $\rho = \rho_{max}$ inwards. However, when reaching small r, the irregular solution will dominate. Therefore, the equation is integrated from both ends in opposite directions and the solutions are matched around the turning point ρ_t defined by $W(\rho_t) = 0$ (see above, Eq. (8)).

Specifically, in region I defined by $-\infty < \rho < \rho_t$, $W(\rho) < 0$, F does not oscillate and we solve Eq.

Table 1
Algorithms for solving the phase-amplitude equations in the different ranges of the radial variables r and $\rho = \ln r$.

Region	Boundaries	Directions	Method
Origin	$r \approx 0$	Outwards	Power development or Eq. (25) for $\ell > 5$
I	$-\infty < \rho < \rho_t$	Outwards	Numerov
II	$\rho_t < \rho < \rho_{pc}$	Inwards	Pred.-Cor.
III	$\rho_{pc} < \rho < \rho_{max}$	Inwards	Mod. Pred.-Cor.
Asymptotic	$r \approx r_{max}$	Inwards	Eq. (26)

(15) directly using the Numerov method [16,17] outwards:

$$F_{n+1} = (a_{n+1})^{-1}\{2b_nF_n - a_{n-1}F_{n-1}\},$$

where

$$a_n = \left(1 - \frac{h^2}{12}W(\rho_n)\right), \quad b_n = \left(1 + \frac{5h^2}{12}W(\rho_n)\right).$$

Since Eq. (15) is homogeneous, the solution in this region is determined up to a multiplicative constant. The amplitude equation (Eq. (18)) is solved in the oscillatory regions where $W(\rho) > 0$, inwards. The turning point ρ_t determines the border between the two methods and matching the solutions at that point determines the arbitrary constant of region I, and the phase through ρ_0 of Eq. (19). We will show below that a usual predictor-corrector scheme works well for the amplitude equations (18) in region II defined by $\rho_t < \rho < \rho_{pc}$ such that

$$W(\rho_{pc}) \equiv \frac{3}{h^2}, \tag{21}$$

where h is the mesh size.

In region III defined by $\rho > \rho_{pc}$ however, the usual predictor-corrector scheme fails and a modified scheme is necessary. Table 1 summarizes the different schemes.

3.1. Initial values

In addition, one needs starting values for the above algorithms at the boundaries. These are obtained from the a priori known behavior of the solutions:

3.1.1. Starting values for Numerov

In order to start the Numerov integrator near the origin we use a power series expansion [16]. However, for high ℓ values the orbitals near the origin are very small, behaving, in the non-relativistic case like

$$F \approx r^{\ell+1/2} \tag{22}$$

and one can save time by starting at a larger r. A series expansion around $r = 0$ is then inappropriate. In this case we use for the initial points a modified phase-amplitude representation (where the phase would be imaginary):

$$F(\rho) = Z(\rho)\, e^{\theta(\rho)} \tag{23}$$

yielding

$$Z'' + WZ + \frac{1}{Z^3} = 0, \qquad \theta' = \frac{1}{Z^2}. \tag{24}$$

$Z(\rho)$ is represented by a polynomial of degree four, the coefficients of which are obtained by approximating the above equation. We can now obtain the ratio between two initial points needed to start the Numerov procedure,

$$\frac{F_1}{F_2} = \frac{Z_1}{Z_2}\exp\left(\int_{\rho_1}^{\rho_2}\frac{d\rho}{Z^2}\right). \tag{25}$$

3.1.2. Large r values

At large r values we represent the solution by the asymptotic solution multiplied by a fourth-order polynomial:

$$Y|_{\rho\to\infty} = \left(\frac{\pi k e^{\rho}}{2}\right)^{-1/2}\sum_{n=0}^{4} a_n \rho^n \tag{26}$$

and using the differential equation (18) we determine the values of a_n and values of Y on five points.

3.2. The problem of the predictor-corrector method in region III

The usual predictor-corrector [11] method is used in region II for the amplitude equation (18). In this scheme the prediction at point $n+1$ (where n counts grid points from ρ_{max} *inwards*) is obtained from the previous points by the commonly used differences approximation for a second-order differential equation:

$$Y^{p}_{n+1} = 2Y_{n-1} - Y_{n-3} + \frac{4h^2}{3}(Y''_n + Y''_{n-1} + Y''_{n-2}) \tag{27}$$

and the correction is given by

$$Y^{c}_{n+1} = 2Y_n - Y_{n-1} + \frac{h^2}{12}(Y''^{p}_{n+1} + 10Y''_n + Y''_{n-1}), \tag{28}$$

using the differential equation (18) for Y''^{p}_{n+1}.

In region III, however, any method that uses Eq. (18) to obtain the derivative Y''^{P} from a given prediction of Y^{P} fails. In order to clarify this point we note the following characteristics of the solutions in that region:

For large r we have the asymptotic behavior

$$Y(r) \approx \sqrt{\frac{2}{\pi kr}}$$

and

$$W \approx (kr)^2 = \frac{a^2}{Y^4}.$$

For large kr, Y is small and both WY and a^2/Y^3 are very large and behave similarly with r. Thus, using Eq. (18), Y'' is obtained as a difference between two large numbers and even a small error in Y results in a disproportionate error in Y''. Specifically, if we look at the approximate equation for the difference $\Delta Y \equiv Y - Y^{\mathrm{P}}$ (assuming a good prediction, i.e. $Y \approx Y^{\mathrm{P}}$):

$$\Delta Y'' = -\left(\frac{3a^2}{Y^4} + W\right)\Delta Y \approx -4(kr)^2 \Delta Y \qquad (29)$$

and for typical values: $r = 10$ and $k = 30$ we have $(kr)^2 \approx 10^5$. A very good prediction is therefore followed by a huge error in the correction.

It should be emphasized that this failure is not a property of the predictor-corrector method but is an inherent property of the differential equation and will occur with any method that uses it to obtain the derivative when the function is given. Note that with a non-logarithmic mesh, $w \approx k^2$, and the absolute error would be smaller, but the relative error will still be of the same order.

Using relations (29) and (28) between Y''^{P} and Y^{c} we obtain the error in Y^{c} due to ΔY:

$$\Delta Y^{\mathrm{c}} \approx -\frac{h^2}{12}\left(\frac{3a^2}{Y^4} + W\right)\Delta Y \approx -\frac{h^2 W}{3}\Delta Y. \qquad (30)$$

Equating the coefficient of ΔY in the r.h.s. to 1 defines the boundary ρ_{pc} of Eq. (21) for the applicability of the predictor-corrector method.

3.3. The modified predictor-corrector method

We have developed a modified predictor corrector scheme which reverses the role of Y and Y''. Using Eq. (18) we now obtain Y from Y'':

$$Y = \sqrt{a}\left(W + \frac{Y''}{Y}\right)^{-1/4}. \qquad (31)$$

Since $W \gg Y''/Y$ in region III, the previous flaw turns out to be an advantage where now even a large error in the predicted derivative Y'' results in a small error in Y. Now, it is tempting to use Eq. (28) for obtaining Y'' from Y. However, this reverse approach is unstable and leads to large accumulated errors. Specifically, in order to obtain a prediction for the derivative we define

$$Y_n^{\mathrm{H}} \equiv Y_n - \tfrac{1}{3}h^2 Y_n'' \qquad (32)$$

and use the approximate relation

$$Y_{n+1} = 2Y_n - Y_{n-1} + \tfrac{1}{3}h^2(Y_{n+1}'' + Y_n'' + Y_{n-1}'') \qquad (33)$$

to obtain

$$Y_{n+1}^{\mathrm{H}} = 3Y_n - \left(Y_n^{\mathrm{H}} + Y_{n-1}^{\mathrm{H}}\right). \qquad (34)$$

With this prediction of Y^{H} we write Eq. (18) as

$$\left(1 + \tfrac{1}{3}h^2 W\right)Y - \frac{\frac{1}{3}h^2 a^2}{Y^3} - Y^{\mathrm{H}} = 0 \qquad (35)$$

and solve for Y using one step of a Newton–Raphson iteration with initial estimate for Y^{P} obtained from the previous points by a polynomial expansion of order three. We obtain:

$$Y^{\mathrm{c}} = \frac{\left(\frac{4}{3}a^2 h^2 Y^{\mathrm{P}} + Y^{\mathrm{H}}(Y^{\mathrm{P}})^4\right)}{\left(1 + \frac{1}{3}h^2 W\right)(Y^{\mathrm{P}})^4 + a^2 h^2}. \qquad (36)$$

To improve the result, after calculating the values of Y_n for all n in region III, we evaluate the second derivative numerically by

$$Y_n''^{\mathrm{c}} = \frac{-Y_{n-2}^{\mathrm{c}} + 16Y_{n-1}^{\mathrm{c}} - 30Y_n^{\mathrm{c}} + 16Y_{n+1}^{\mathrm{c}} - Y_{n+2}^{\mathrm{c}}}{12h^2} \qquad (37)$$

and again use one Newton–Raphson iteration to solve Eq. (18) for our final result for Y, namely

$$Y^{\mathrm{cc}} = \frac{4a^2 Y^{\mathrm{c}} - Y''^{\mathrm{c}}(Y^{\mathrm{c}})^4}{W(Y^{\mathrm{c}})^4 + 3a^2}. \qquad (38)$$

3.4. Matching the solutions at the turning point

The matching of the two solutions – i.e. PA and Numerov – is done for convenience near the turning point where the functions are most smooth. Equating the two solutions at two nearby grid points we can write

$$F_1 = cY_1 \sin \Phi_1, \tag{39}$$

$$F_2 = cY_2 \sin(\Phi_1 + \Delta\Phi), \tag{40}$$

where

$$\Delta\Phi = \int_{\rho_1}^{\rho_2} \frac{a}{Y^2}\, d\rho. \tag{41}$$

Dividing (40) by (39), we determine the phase Φ_1 by:

$$\tan(\Phi_1) = \sin(\Delta\Phi)\left(\frac{F_2 Y_1}{F_1 Y_2} - \cos(\Delta\Phi)\right)^{-1} \tag{42}$$

using this value and inserting into (39) the normalization constant of region I is obtained.

3.5. Calculation of the phase function

Having the phase Φ_1 and the amplitude for $\rho > \rho_t$ the remaining phases are obtained by integrating Eq. (19) numerically outwards. Since this is a quadrature, any method will do. In order to avoid accumulated errors we subtract from the integrand of Eq. (19) its asymptotic form, $kr = ke^{\rho}$, which can be integrated analytically.

4. Calculation of the radial integrals

The radial integrals encountered in the various atomic processes can always be reduced to the form

$$I = \int_{\rho_1}^{\rho_2} X(\rho) \cos \phi(\rho)\, d\rho, \tag{43}$$

where $X(\rho)$ is not oscillatory. For processes involving a single continuum orbital the representation (43) is immediate. Let us show that it is the case also when two continuum orbitals are present. We describe here in details the case of collisional excitation.

The Slater integrals appearing in this case in the expression for the cross section, involve two bound and two continuum orbitals. The exchange integrals do not present an essential problem since each continuum function overlaps a bound orbital, and is cut off when its bound counterpart vanishes. Paradoxically, however, they will take longer to compute, since we will not be able to use the method described in the following paragraph. The only way that we see to evaluate exchange integrals is to create a mesh of points adequate for the continuum function with less oscillations and treat this function and its bound counter part as two "bound" orbitals thus leaving only a single continuum orbital. In the direct type integrals on the other hand, the two continuum orbitals belong to the same electron and the integration limits are extended to infinity.

Let us recall that the wavefunctions are computed by the Numerov algorithm until the turning point, on the same grid as the bound electron functions. For that part of the integrals, therefore, there is no difference with the case where all electrons are bound, and we apply a variant of Simpson's rule.

For the proper PA part of the wavefunctions, we need then to calculate an integral of the form

$$R = \int_{\rho_t}^{\infty} B(\rho)\, F_1(\rho)\, F_2(\rho)\, d\rho, \tag{44}$$

where $B(\rho)$ contains the effect of the bound orbitals and $F_i(\rho) = Y_i(\rho) \sin \phi_i(\rho)$ are the continuum orbitals. By a simple trigonometric manipulation we can write

$$\begin{aligned} R &= \int_{\rho_t}^{\infty} B(\rho)\, F_1(\rho)\, F_2(\rho)\, d\rho \\ &= \frac{1}{2}\int_{\rho_t}^{\infty} B(\rho)\, Y_1(\rho)\, Y_2(\rho) \\ &\quad \times [\cos \eta(\rho) - \cos \zeta(\rho)]\, d\rho, \end{aligned} \tag{45}$$

where

$$\begin{aligned} \eta(\rho) &= [\phi_1(\rho) - \phi_2(\rho)], \\ \zeta(\rho) &= [\phi_1(\rho) + \phi_2(\rho)]. \end{aligned} \tag{46}$$

The integrals in Eq. (45) have now the form of Eq. (43).

Let us now describe a method [18] that takes advantage of this form where the oscillatory charac-

ter of the cosine is explicit. The phase-amplitude relation

$$d\rho = 1/aY^2(\rho)\, d\phi \tag{47}$$

allows the transformation

$$I = \int_{\rho_1}^{\rho_2} X(\rho) \cos \phi(\rho)\, d\rho$$

$$= 1/a \int_{\phi_1}^{\phi_2} X(\rho(\phi))\, Y^2(\rho(\phi)) \cos \phi\, d\phi. \tag{48}$$

Let us now introduce Gauss trigonometric formulas over half periods

$$I_n = \int_{(n-\frac{1}{2})\pi}^{(n+\frac{1}{2})\pi} P(\phi) \cos \phi\, d\phi \cong \sum_{i=1}^{N} w_i P(\phi_i), \tag{49}$$

for any function P that can be represented by a polynomial of degree N. As usual for Gauss formulas, P is evaluated at some predetermined abcissas ϕ_i and is multiplied by weights w_i. The abcissas and the weights are shown in Table 2, reproduced from Ref. [18]. Using Eq. (49) avoids evaluating the cosines altogether, but requires the values of P at specified phases. Therefore, when applying to (48), we interpolate XY^2, using cubic splines [19].

The integration in half periods produces an alternating series with terms $a_n = I_n$. For summing this series, we use the Levin t-transformation [7,20]. This gives a sequence of terms t_n defined by

$$t_n = \frac{\sum_{j=0}^{n} (-1)^j \binom{n}{j} \left(\frac{j+1}{n+1}\right)^{n-1} \frac{A_{j+1}}{a_{j+1}}}{\sum_{j=0}^{n} (-1)^j \binom{n}{j} \left(\frac{j+1}{n+1}\right)^{n-1} \frac{1}{a_{j+1}}}, \tag{50}$$

where

$$A_j = \sum_{i=0}^{j} a_i, \tag{51}$$

which converges rapidly to A_∞.

Specifically for the collisional excitation integral (45) we obtain

$$2R_{\rho_1}^{\rho_2} \equiv 2\int_{\rho_1}^{\rho_2} B(\rho)\, F_1(\rho)\, F_2(\rho)$$

Table 2
Abcissas and weights for Gaussian trigonometric formulas of order N, taken from [18]. cos φ is defined on $[-\frac{1}{2}, +\frac{1}{2}]$. Only positive φ_i are shown. (n) means $\times 10^n$.

Order	φ_i	w_i
2	6.8366739009(−1)	1.0000000000(0)
3	0.0000000000(0)	1.0883191839(0)
	1.01260124000(0)	4.5584040804(0)
4	4.3928746686(−1)	7.7592938187(−1)
	1.1906765639(0)	2.2407061813(−1)
5	0.0000000000(0)	7.5221097881(−1)
	7.2598673794(−1)	5.0378251239(−1)
	1.2964402800(0)	1.2011199821(−1)
6	3.2385211421(−1)	6.05813700129(−1)
	9.1979066552(−1)	3.2479855138(−1)
	1.3639113021(0)	6.9387748500(−2)
7	0.0000000000(0)	5.7554040031(−1)
	5.6350196618(−1)	4.5607388993(−1)
	1.0555399634(0)	2.1353015976(−1)
	1.4094168673(0)	4.2625750151(−2)
8	2.5649650742(−1)	4.9199579660(−1)
	7.4346864788(−1)	3.3626447785(−1)
	1.1537256455(0)	1.4420409203(−1)
	1.4414905402(0)	2.7535633514(−2)

$$= \int_{\eta_1}^{\eta_2} \frac{B(\eta)\left[Y_1(\eta)\ Y_2(\eta)\right]^3}{Y_2(\eta)^2 - Y_1(\eta)^2} \cos\ \eta\ \mathrm{d}\eta$$

$$- \int_{\zeta_1}^{\zeta_2} \frac{B(\zeta)\left[Y_1(\zeta)\ Y_2(\zeta)\right]^3}{Y_2(\zeta)^2 + Y_1(\zeta)^2} \cos\ \zeta\ \mathrm{d}\zeta. \tag{52}$$

The first integral in Eq. (52) may contain a pole at some point ρ_p where

$$Y_1(\eta_p) = Y_2(\eta_p). \tag{53}$$

In this case Eq. (45) is integrated directly – i.e. using functions $P_{\varepsilon\ell}$ (Eq. (1)) – up to $\rho_1 \geq \rho_p$, and the PA method (48)–(49) is used for the region $\rho > \rho_p$. The occurrence of a pole indicates that the phases η are varying slowly and direct integration over ρ is most efficient in this region with no need for extra grid points.

In cases where the upper limit of the integral in Eq. (48) is finite (e.g. for the photoionization or collisional exchange integrals), by having the spline coefficients we can integrate directly using Eq. (49) without using Levin's transformation since we have an integral of polynomial times a cosine function.

5. Discussion

5.1. Accuracy

The advantage of solving for smooth functions is the ability to achieve high precision with a relatively coarse mesh. In particular it allows to use the same logarithmic mesh for both bound and continuum orbitals. Specifically, we have used:

$$\rho = \ln\ r, \qquad r_{\max} = 50,$$

$$\rho_n = \rho_0 \mathrm{e}^{nh}, \quad \rho_0 = \frac{\mathrm{e}^{-10}}{z}, \tag{54}$$

The step used for bound orbitals is $h = 1/n$ with $n = 32$ (i.e. r is multiplied by e every 32 points). The value of $r_{\max}$ is determined by two requirements: the normalization of the orbitals and convergence of the radial integrals.

In order to check the accuracy of our continuum wavefunction F on that mesh we have used a Coulomb potential for which exact solutions F^c are available [21] (for energy ε and angular momenta ℓ). It should be mentioned that our program is built for relativistic calculations and for the comparisons presented below we have taken the non-relativistic limit $c \Rightarrow \infty$ (practically $c = 10^6$ a.u. is more than sufficient).

We can measure the accuracy by the relative quantity

$$\Delta F = \max\left|\left[F(r) - F^c(r)\right]/Y^c(r)\right|. \tag{55}$$

Where the amplitude $Y^c(r)$ was used to avoid division by zero.

In Table 3 we compare calculations of ΔF obtained by the PA and the Numerov methods for various energies ε and angular momenta ℓ. For the Numerov method we present two sets of results obtained for two sets of grid points. In one set, the number of points N was chosen to yield similar accuracy to our PA results (calculated here with $h = 1/16$), and the second takes, as a standard, 12 points for each half cycle following Sampson et al. [1]. The comparison in Table 3 refers to the accuracy, number of points, and time required on a personal computer equipped with a 486 microprocessor running at 50 Mhz. The computation times are given in ms. It is clear that the PA results are always more accurate than the Numerov method, with these conditions.

It is less easy to check radial integrals. As an example we have calculated the radial Slater integral

$$R_n \equiv R^1\left(1\mathrm{s}\ \varepsilon_{\mathrm{in}}\ell_{\mathrm{in}}, 2\mathrm{p}\ \varepsilon_{\mathrm{out}}\ell_{\mathrm{out}}\right),$$

using a logarithmic mesh with step size $h = 1/n$. Here 1s and 2p are Hydrogenic orbitals calculated for $Z = 20$ and ℓ_{out} and ℓ_{in} are continuum orbitals with the respective energies $\varepsilon_{\mathrm{out}}$ and

$$\varepsilon_{\mathrm{in}} = \varepsilon_{\mathrm{out}} - \varepsilon_{1\mathrm{s}} + \varepsilon_{2\mathrm{p}}.$$

In order to evaluate the accuracy of our results we have repeated the calculations for four grids with $n = 8$, 16, 32 and 64. In Table 4 we present the relative quantity $\Delta R_n \equiv |[R_n - R_{64}]/R_{64}|$ for $n =$ 8, 16, 32 and for various energies and angular momenta of continuum electrons having relatively large contribution. It is seen that even taking $n = 16$ is satisfactory.

In fact an accuracy of 2–3 digits for cross sections is, in most cases, sufficient. A question to be

asked therefore is: why is high accuracy in the orbitals themselves required? The answer is twofold:

(1) As described above the calculation of the radial integral extends to infinity (the Levin formula). In addition for all atomic processes the continuum is described by an infinite sum over partial waves. In both convergence procedures, high precision in each individual term is required in order to achieve reasonable accuracy in the final result.

(2) Accurate methods allow the use of a coarse mesh. Specifically for atomic processes the same logarithmic mesh can now be used with adequate accuracy for both oscillatory (continuum) and non-oscillatory (bound) orbitals. Furthermore the accuracy does not depend on the continuum orbital energy, thus the same mesh applies to all continuum orbitals.

5.2. *Efficiency*

In spite of the seemingly involved algorithm presented for solving the continuum orbitals, the use of

Table 3
Comparison between the phase-amplitude and Numerov methods. "Numerov" uses a number of points chosen to obtain accuracy similar to the PA method. "Numerov-12" uses 12 points per half period, see [1]. ΔF: Maximum relative error compared to exact Coulomb functions [21]. N: Number of points required. t: time in msec on a PC with a 486 micropocessor running at 50 Mhz.

ε (a.u.)	ℓ	Item	Phase Amp.	Numerov	Numerov-12
1	0	ΔF	1×10^{-8}	1×10^{-7}	1×10^{-3}
		N	272	2100	315
		t	2.7	19.8	2.8
1	5	ΔF	1×10^{-6}	1×10^{-5}	1×10^{-3}
		N	133	1050	315
		t	1.6	9.6	2.8
10	0	ΔF	1×10^{-8}	5×10^{-6}	5×10^{-4}
		N	272	3300	990
		t	4.9	31.6	9.4
10	5	ΔF	1×10^{-6}	5×10^{-6}	1×10^{-3}
		N	136	3300	792
		t	2.5	31.6	7.2
100	0	ΔF	1×10^{-8}	5×10^{-6}	1×10^{-3}
		N	272	10550	3165
		t	4.9	112	30.4
100	5	ΔF	5×10^{-7}	5×10^{-6}	1×10^{-3}
		N	147	10550	2532
		t	3.1	112	24
100	10	ΔF	5×10^{-6}	1×10^{-5}	1×10^{-3}
		N	104	10550	2532
		t	2.7	112	24
1000	0	ΔF	1×10^{-8}	1×10^{-5}	1×10^{-3}
		N	272	33450	8028
		t	4.9	370	84
1000	5	ΔF	1×10^{-7}	1×10^{-5}	1×10^{-3}
		N	162	33450	8028
		t	3.4	376	86
1000	10	ΔF	1×10^{-6}	1×10^{-5}	1×10^{-2}
		N	121	33450	8028
		t	3.1	376	86
1000	20	ΔF	1×10^{-6}	1×10^{-5}	1×10^{-2}
		N	93	33450	8028
		t	2.7	376	86

Table 4
Accuracy of the radial integrals, as a function of the step size $h = 1/n$. ΔR_n is the relative difference with $n = 64$.

ε_{out} (a.u.)	ℓ_{out}	ℓ_{in}	ΔR_8	ΔR_{16}	ΔR_{32}
1	0	1	2.4×10^{-4}	5.2×10^{-5}	2.9×10^{-6}
10	0	1	9.0×10^{-5}	2.4×10^{-5}	2.8×10^{-6}
100	0	1	5.7×10^{-4}	8.6×10^{-5}	2.4×10^{-5}
1000	0	1	6.1×10^{-4}	8.8×10^{-5}	2.2×10^{-5}
1	4	5	1.4×10^{-2}	2.6×10^{-3}	5.4×10^{-4}
10	4	5	9.6×10^{-3}	2.9×10^{-3}	2.8×10^{-4}
100	4	5	5.6×10^{-4}	4.4×10^{-4}	5.1×10^{-5}
1000	4	5	1.2×10^{-3}	6.0×10^{-5}	1.0×10^{-5}
100	9	10	1.2×10^{-2}	4.4×10^{-3}	1.4×10^{-3}
1000	9	10	2.0×10^{-3}	4.4×10^{-5}	2.3×10^{-5}
1000	14	15	1.1×10^{-2}	8.4×10^{-4}	2.5×10^{-5}

the coarse mesh makes it very fast and efficient. For high ℓ continuum orbitals the calculation begins at relatively large r and in many cases the orbital is described with a few tens of grid points only. It is obvious from Table 3 that the PA method reduces drastically the computation time and memory requirements for comparable accuracy (for high energies by more than a factor of 100). For the Numerov calculations, we have started with $r = 3$ and constrained two initial accurate values. (The number of points with $r > 3$ in our PA mesh is only 45). In addition we have skipped the normalization step in the Numerov calculations. When these points are appropriately handled, the comparison will favor even more the PA method. Because of the logarithmic mesh there is no problem in reaching very high r values, since the addition of $1/h$ points after ρ_{max} multiplies the radius by e. This is very convenient to reach the asymptotic amplitude of the wavefunction.

Furthermore, thanks to the small number of points the calculation of the radial integral is very fast as well. The calculation of the spline coefficients is relatively slow, but the time needed is proportional to the number of points which in our case is small, and the calculation on the intermediate points is very fast. Another advantage of the proposed method is that in the convergence procedure for the integrals there is no problem with very large r values and the convergence method itself is much faster and more accurate then the asymptotic series expansion [22]. Computation times on a personal computer using a 486 microprocessor running at 50 Mhz for the integrals of Table 4 are given in Table 5 for the mesh steps $h = 1/n$, $n = 16, 32$. For this comparison the

Table 5
Calculation times t_n (ms) of radial integrals, same conditions as in Table 3.

ε_{out} (a.u.)	ℓ_{out}	ℓ_{in}	t_{16}	t_{32}
1	0	1	6.4	11.9
10	0	1	6.4	11.9
100	0	1	6.5	12.2
1000	0	1	7	13
1	4	5	3.2	5.7
10	4	5	3.5	5.9
100	4	5	3.8	6.7
1000	4	5	4.7	8.4
100	9	10	3	5.2
1000	9	10	3.8	6.9
1000	14	15	3.4	6

analytic expressions of hydrogenic bound orbitals where used for both methods. The times vary from 3 to 7 ms for $n = 16$ and twice as long for $n = 32$. In order to evaluate the computation time needed by the Numerov method for these integrals at the same degree of accuracy, we have calculated the same integrals using Simpson integration with varying steps [1] and number of points which meets the accuracy requirements. For simplicity we have taken $r_{max} = 10$.a.u. A typical computation time was 180 ms. Of course r_{max} depends strongly on the problem. For transitions between energetically near lying levels, with transition energy ΔE, like $\Delta n = 0$ transitions or transitions within a configuration, the required r_{max} is much higher. It can be seen from Eqs. (5), (19) and (52), that the phase of the "slow frequency" component of the integrand at large r is approximately $(\Delta E/2\sqrt{\varepsilon})r$. However, the accuracy and convergence depend on the number of points per period and on the number of periods. As already mentioned above the number of points in the PA method is logarithmic with r whereas the direct Numerov integration requires a number of points linear with r and therefore converges very slowly. In these cases the gap between the two methods is even much higher.

Finally, as a by-product, the normalization of the continuum orbitals is achieved naturally in the PA method because of the non homogeneity of the amplitude equation (18).

Appendix

In the relativistic theory [23] the Dirac spinor orbitals with the upper and lower components P and Q respectively obey the following coupled equations:

$$\frac{dP}{dr} + \frac{\kappa}{r}P = \eta Q,$$

$$\frac{dQ}{dr} + \frac{\kappa}{r}Q = -\chi P,$$

where

$$\eta = \frac{1}{c}(\varepsilon - U(r) + 2c^2),$$

$$\chi = \frac{1}{c}(\varepsilon - U(r)),$$

$$\kappa = \begin{cases} -\left(j + \frac{1}{2}\right), & j = l + \frac{1}{2} \\ \left(j + \frac{1}{2}\right), & j = l - \frac{1}{2} \end{cases}.$$

There are many ways to write Schrödinger type uncoupled equations. The one presented here is similar to [1]. Let us define

$$F = \sqrt{\frac{2c}{\eta}}P, \qquad G = \sqrt{\frac{2c}{\chi}}Q. \tag{A.1}$$

After some algebra, we obtain

$$\frac{d^2}{dr^2}F + \varpi_F(r)\, F = 0, \tag{A.2a}$$

$$\frac{d^2}{dr^2}G + \varpi_G(r)\, G = 0, \tag{A.2b}$$

with

$$\varpi_F = \eta\chi - \frac{3}{4}\left(\frac{\eta'}{\eta}\right)^2 + \frac{1}{2}\frac{\eta''}{\eta} - \frac{\kappa}{r}\frac{\eta'}{\eta} - \frac{\kappa(\kappa+1)}{r^2}, \tag{A.3a}$$

$$\varpi_G = \eta\chi - \frac{3}{4}\left(\frac{\chi'}{\chi}\right)^2 + \frac{1}{2}\frac{\chi''}{\chi} + \frac{\kappa}{r}\frac{\chi'}{\chi} - \frac{\kappa(\kappa+1)}{r^2}. \tag{A.3b}$$

We have also used the semi-relativistic scheme [1] ignoring the small component Q and averaging over the two j values $j = \ell \pm \frac{1}{2}$ yielding Eq. (A.3a) with

$$\varpi_F = \eta\chi - \frac{3}{4}\left(\frac{\eta'}{\eta}\right)^2 + \frac{1}{2}\frac{\eta''}{\eta} + \frac{\eta'}{\eta}\frac{1}{r} - \frac{\ell(\ell+1)}{r^2}.$$

In Eq. (A.3) the following relations hold:

$$\eta' = \chi' = -\frac{1}{c}U', \qquad \eta'' = \chi'' = -\frac{1}{c}U''$$

and

$$\eta\chi = \frac{1}{c^2}\left[(\varepsilon - U + 2c^2)(\varepsilon - U)\right]$$

$$= 2(\varepsilon - U) + \frac{(\varepsilon - U)^2}{c^2}.$$

The derivatives of the potential are easily obtained if one uses a parametric potential representation [24,25]. After solving for F in the semi-relativistic scheme we must transform back to P. However in order to retain normalization, the replacement

$$\sqrt{\frac{2c}{\eta}} \Rightarrow \sqrt{\frac{2c}{\eta + \chi}}$$

must be made in Eq. (A.1).

References

[1] H.L. Zhang, D.H. Sampson and A.K. Mohanty, Phys. Rev. A 40 (1989) 616.
[2] P.L. Hagelstein and R.K. Jung, Atom. Dat. Nucl. Data Tables 37 (1987) 121.
[3] J. Oreg, W.H. Goldstein, M. Klapisch and A. Bar-Shalom, Phys. Rev. A 44 (1991) 1750.
[4] A. Bar-Shalom, M. Klapisch and J. Oreg, Phys. Rev. A 38 (1988) 1773.
[5] W.E. Milne, Phys. Rev. 35 (1930) 863.
[6] J.A. Wheeler, Phys. Rev. 52 (1937) 1123.
[7] D.A. Smith and W.F. Ford, Math. Comput. 38 (1982) 481.
[8] W.I. Newman and W.R. Thorson, Phys. Rev. Let. 29 (1972) 1350.
[9] H.J. Korsch and H. Laurent, J. Phys. B (At. Mol.) 14 (1981) 4213.
[10] F. Robicheaux, U. Fano, M. Cavagnero and D.A. Harmin, Phys. Rev. A 35 (1987) 3619.
[11] R.W. Hamming, Numerical Methods for Scientists and Engineers (McGraw-Hill, New York, 1973).
[12] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, Numerical Recipes (Cambridge Univ. Press, Cambridge, 1989).
[13] A. Bar-Shalom, M. Klapisch, W.H. Goldstein and J. Oreg, HULLAC-Hebrew University Lawrence Livermore Atomic Code, Jerusalem (1989).
[14] A. Bar-Shalom, J. Oreg, W.H. Goldstein, D. Shvarts and A. Zigler, Phys. Rev. A 40 (1989) 3183.
[15] R.D. Cowan, The Theory of Atomic Structure and Spectra (Univ. California Press, Berkeley, 1981).
[16] C. Froese-Fischer, The Hartree-Fock Method for Atoms: a Numerical Approach (Wiley, New York, 1977).
[17] D.R. Hartree, The Calculation of Atomic Structures (Wiley, New York, 1957).
[18] M. Blakemore, G.A. Evans and J. Hyslop, J. Comput. Phys. 22 (1976) 352.
[19] J.H. Ahlberg, E.N. Nilson and J.L. Walsh, The Theory and Splines and their Applications (Academic Press, New York, 1967).
[20] D. Levin, Int. J. Comput. Math. B 3 (1973) 371.
[21] A.R. Barnett, D.H. Feng and L.J.B. Goldfarb, Comput. Phys. Commun. 8 (1974) 377.
[22] J.A. Belling, J. Phys. B (At. Mol.) 1 (1968) 136.
[23] I.P. Grant, Adv. Phys. 19 (1970) 747.
[24] M. Klapisch, Comput. Phys. Commun. 2 (1971) 239.
[25] M. Klapisch, J.L. Schwob, B.S. Fraenkel and J. Oreg, J. Opt. Soc. Am. 67 (1977) 148.

Appendix M

Effects of Heavy Metals on Astrophysical Opacities

THE ASTROPHYSICAL JOURNAL, 445:855–860, 1995 June 1

EFFECTS OF HEAVY METALS ON ASTROPHYSICAL OPACITIES

C. A. IGLESIAS, B. G. WILSON, F. J. ROGERS, AND W. H. GOLDSTEIN
Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94550

AND

A. BAR-SHALOM AND J. OREG
Nuclear Research Center of the Negev, P.O. Box 9001, Beer Sheva, Israel
Received 1994 October 17; accepted 1994 December 7

ABSTRACT

Recent reexaminations of astrophysical opacities resulted in large opacity increases. The increases were mostly due to improvements in the atomic physics used in the calculations; in particular, the treatment of bound-bound transitions of partially ionized Fe. The new opacities were also found to depend on other Fe group elements even though in typical stars these elements have a combined abundance that is more than an order of magnitude lower than Fe. It is, therefore, important to examine the contributions from even heavier elements. These elements are much lower in abundance still, but have many more bound electrons and could, in principle, impact astrophysical opacities. It is shown that although the heavy elements are stronger photon absorbers their small abundance leads to a marginal effect on the Rosseland mean opacities of stars with solar metal distributions.

Subject headings: atomic processes — stars: interior

1. INTRODUCTION

The Rosseland mean opacity is an important physical input for stellar models. In recent years a reexamination of the opacity has led to the resolution of several long-standing problems in astrophysics (e.g., Rogers & Iglesias 1994). The most significant change was an increase in the opacity of approximately a factor of 3 for Population I stars near temperatures of a few hundred-thousand degrees. The major source of the opacity increase has been the use of improved atomic physics in the calculations of partially ionized Fe ions. Consequently, the new opacity results display a strong dependence on the metal mass fraction not present in earlier calculations; hence, the new so-called Z-bump.

The new Z-bump was also found to depend on other Fe group elements (Rogers & Iglesias 1992, 1993; Seaton et al. 1994). Even though elements such as Cr, Mn, and Ni typically have a combined abundance that is more than an order of magnitude smaller than that of Fe, their explicit inclusion in the calculations can increase the Z-bump by as much as 40%. It is important then to examine what effect, if any, heavier elements may have on the opacities. The study here concentrates on normal compositions. That is, it assumes the relative metal distribution in the Sun and ignores processes that may enhance or redistribute the heavier metals.

2. METHOD

The OPAL opacities (Iglesias, Rogers, & Wilson 1992, hereafter IRW) are assumed as the standard results. These opacities explicitly include photon absorption cross sections from 14 elements; the heaviest being Fe. The results in IRW treat in detail the bound-bound transitions using LS coupling for these elements except Fe for which it uses full intermediate coupling. However, this detailed line-by-line approach is not practical for the heavier elements.

The myriad spectral lines present in photoabsorption calculations of heavier elements makes unresolved transition array (UTA) approaches attractive (Bauche-Arnoult, Bauche, & Klapisch 1988 and references therein). Here, the Super Transition Array (STA) code is used (Bar-Shalom, Oreg, & Goldstein 1994) which includes, in a statistical manner, all transitions between states of the various ions and electrons in the plasma. The atomic calculations are based on fully relativistic, quantum mechanical theory using a parametric potential approach (Klapisch 1971; Klapisch et al. 1977). It also includes configuration interaction effects between neighboring jj configurations. Unlike detailed line accounting methods such as OPAL, the STA method can handle cases where the number of relevant configurations is immense. Also, unlike average atom approximations (Rozsnyai 1972), the STA method reveals the UTA spectrum (Bar-Shalom et al. 1991).

The STA can be viewed as an extension of the UTA concept where the latter accounts for all the spectral lines connecting two configurations. In the STA method the central idea is to introduce superconfigurations that are collections of configurations constructed by distributing the electrons occupying a supershell among its constituent electron shells or subshells in all possible ways. The STA represents all transitions between two superconfigurations. Similar to the UTA method, the main results provide analytical expressions for the STA's moments (total intensity, average transition energy, and variance) that allow the representation of an STA by a single Gaussian. A more detailed spectrum is revealed by increasing the number of STAs where each new STA includes fewer configurations located within a narrower energy range. The procedure is stopped when convergence of a chosen quantity is reached (e.g., Rosseland mean opacity).

The comparison between OPAL and STA given in Figure 1 helps clarify the STA method. The calculations are for Ga at a temperature and density near the peak of the opacity Z-bump. The figure shows that STA provides an envelope to the OPAL results computed in detail with full intermediate coupling, but since the former is a UTA-type approach it is not intended to resolve individual lines in the spectrum. For the low densities encountered in stellar envelopes, where the individual spectral lines do not significantly overlap (see Fig. 1), statistical

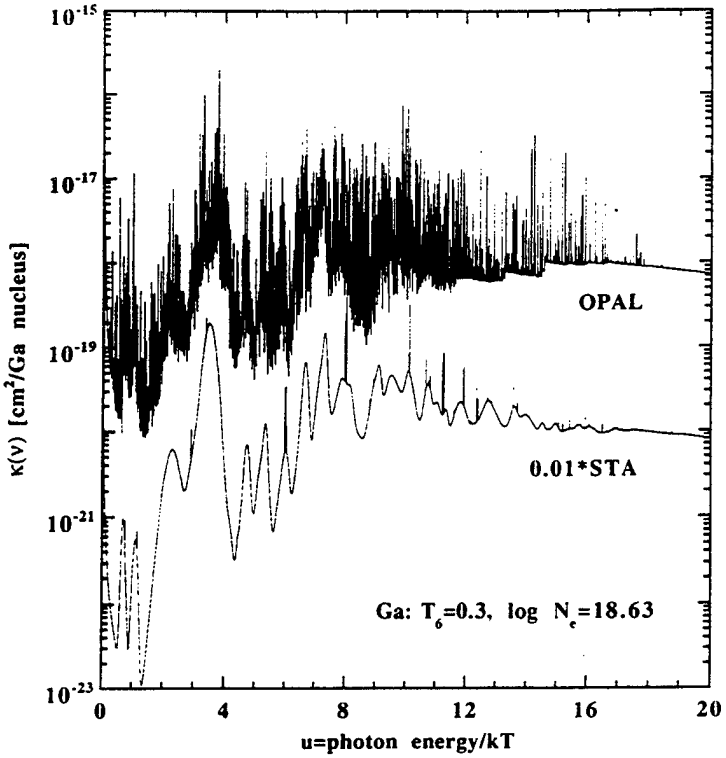


FIG. 1.—Comparison of the OPAL and STA gallium photoabsorption cross sections as a function of u = (photon energy/kT), where kT is the temperature in energy units and N_e is the free electron number density corresponding to the mixture with log $R = -3.5$ shown in Figure 3. Note that for clarity the STA result has been reduced by 0.01.

approaches can overestimate the Rosseland mean opacity (Iglesias, Rogers, & Wilson 1990; Iglesias & Wilson 1994). (Recall that the Rosseland opacity is a harmonic mean that weighs the low absorption regions.) The advantage, however, is that the STA method is computationally fast and can be used for the heavier elements where the detailed line accounting methods become impractical. Furthermore, for much heavier elements the number of lines increases dramatically and thus improves the chances for lines to overlap making the UTA-type methods more reliable.

For the lowest temperatures there were numerical problems with the STA code. For these cases, where there are relatively few excited configurations forming the plasma ensemble, the photoabsorption cross sections for the heavy metals were generated with the MCDCA code (Monte Carlo Detailed Configuration Accounting) developed by Wilson, Albritton, & Liberman (1991). Briefly, the MCDCA code computes the state of the gas by tracking the configuration history of a test ion in the ambient plasma and radiation field on an atomic scale using Monte Carlo techniques. The absorption spectra is then computed in the UTA approximation for each individual dipole allowed configuration-to-configuration electron transition.

3. RESULTS

The metal composition used in the present work is based on recent solar observations (Grevesse & Noels 1993). However, only elements that are listed in Table 1 are explicitly included in the opacity calculations. To obtain this reduced list, particle number was conserved so that the abundance of elements omitted from Table 1 was added to nearby included elements. The first 14 elements in Table 1 are the most abundant in the Sun, and their number fractions have been adjusted in order to account for all elements through Ni where the reduction also conserves the mixture molecular weight. For elements heavier than Ni the abundance of five neighboring elements were combined into one. For example, the abundance of Cu, Zn, Ge, and As were added to that of Ga and only the latter is explicitly included in Table 1.

Photoabsorption cross sections are displayed in Figures 2, 3, and 4 which compare the OPAL calculations (IRW) for the first 14 elements to the remaining heavy metals where the

TABLE 1

FRACTIONAL ABUNDANCE FOR A MIXTURE WITH $X = 0.7$ AND $Z = 0.02$[a]

Atomic Number	Number Fraction
1	9.0710 (−1)
2	9.1367 (−2)
6	3.769 (−4)
7	9.915 (−5)
8	7.876 (−4)
10	1.278 (−4)
11	2.271 (−6)
12	4.039 (−5)
13	3.135 (−6)
14	3.778 (−5)
16	1.764 (−5)
18	3.732 (−6)
20	2.701 (−6)
26	3.626 (−5)
31	6.328 (−8)
36	5.900 (−9)
41	7.246 (−10)
46	2.205 (−10)
51	3.843 (−10)
56	3.929 (−10)
61	5.439 (−11)
66	4.332 (−11)
71	2.980 (−11)

[a] Parentheses indicate power of 10.

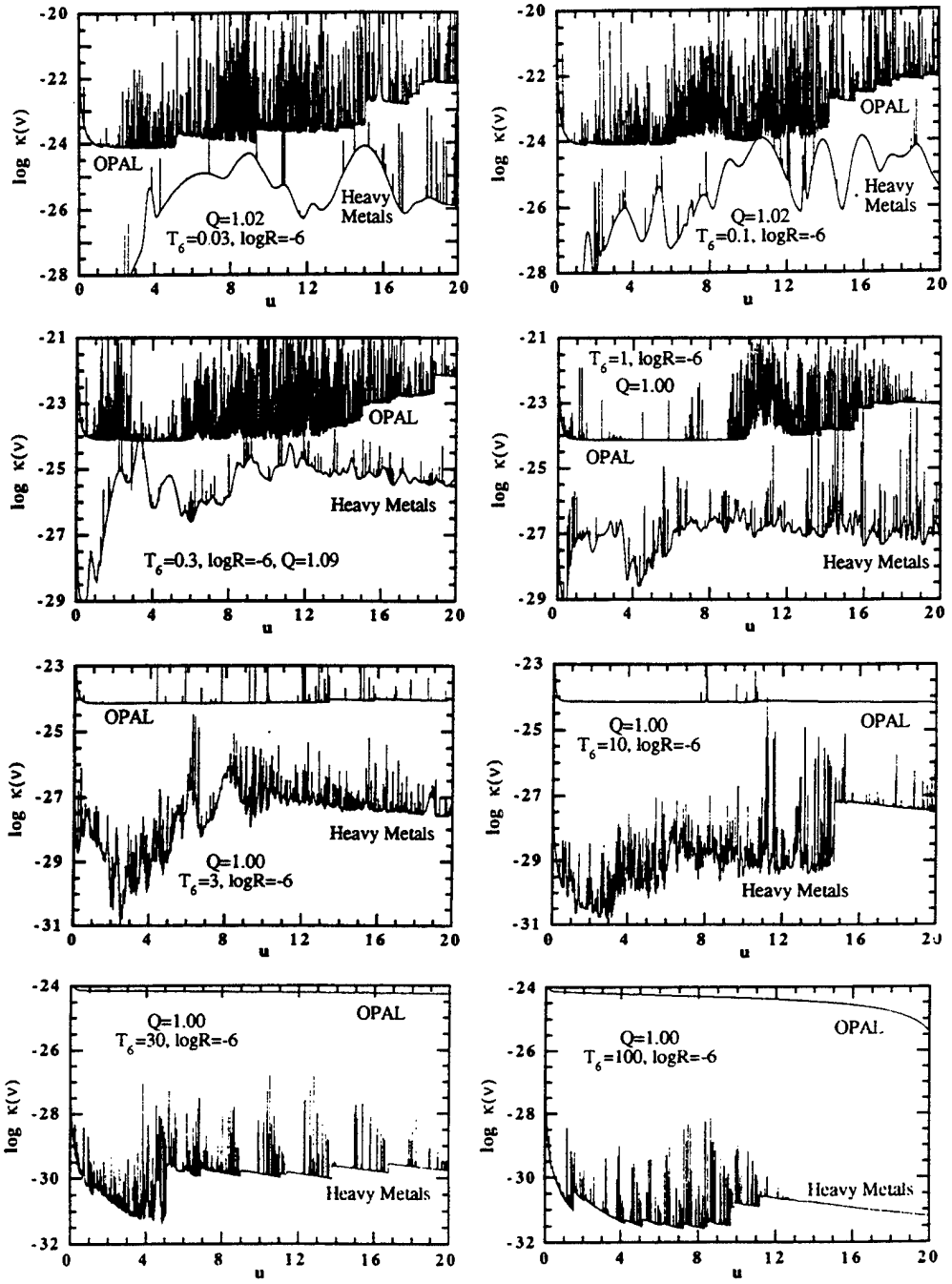


FIG. 2.—The OPAL photoabsorption cross sections compared to the heavy element contribution at various temperatures and constant log $R = -6$. The photoabsorptions are given per nucleus in the mixture and have been weighted by the number fractions in Table 1 corresponding to $X = 0.7$ and $Z = 0.02$.

curves are weighted by the number fractions in Table 1. The calculations are at various temperatures and R tracks for a composition with hydrogen mass fraction $X = 0.7$ and metal mass fraction $Z = 0.02$. Here, $R = \rho/T_6^3$ with T_6 the temperature in million degrees and ρ the matter density in g cm^{-3} where the values in the figures cover most of the temperatures and densities in which Rosseland mean opacities are relevant to stellar models. Indicated in Figure 2 are values of Q which is defined as the ratio of the Rosseland mean opacity for the 23 element mixture to the 14 element OPAL results. The figures show that the elements should have little effect on the total opacity. The possible exceptions are at the lower temperatures and lowest R tracks.

Since UTA-type methods could overestimate the impact of the heavier elements on the mixture Rosseland mean opacity, calculations at $T_6 = 0.3$ and $\log R = -6$ for Ga were done with the OPAL code using full intermediate coupling to compute the detailed line spectrum. This Ga calculation, weighted by the Ga abundance, was then added to the OPAL 14 element results to compute a new Rosseland mean giving a

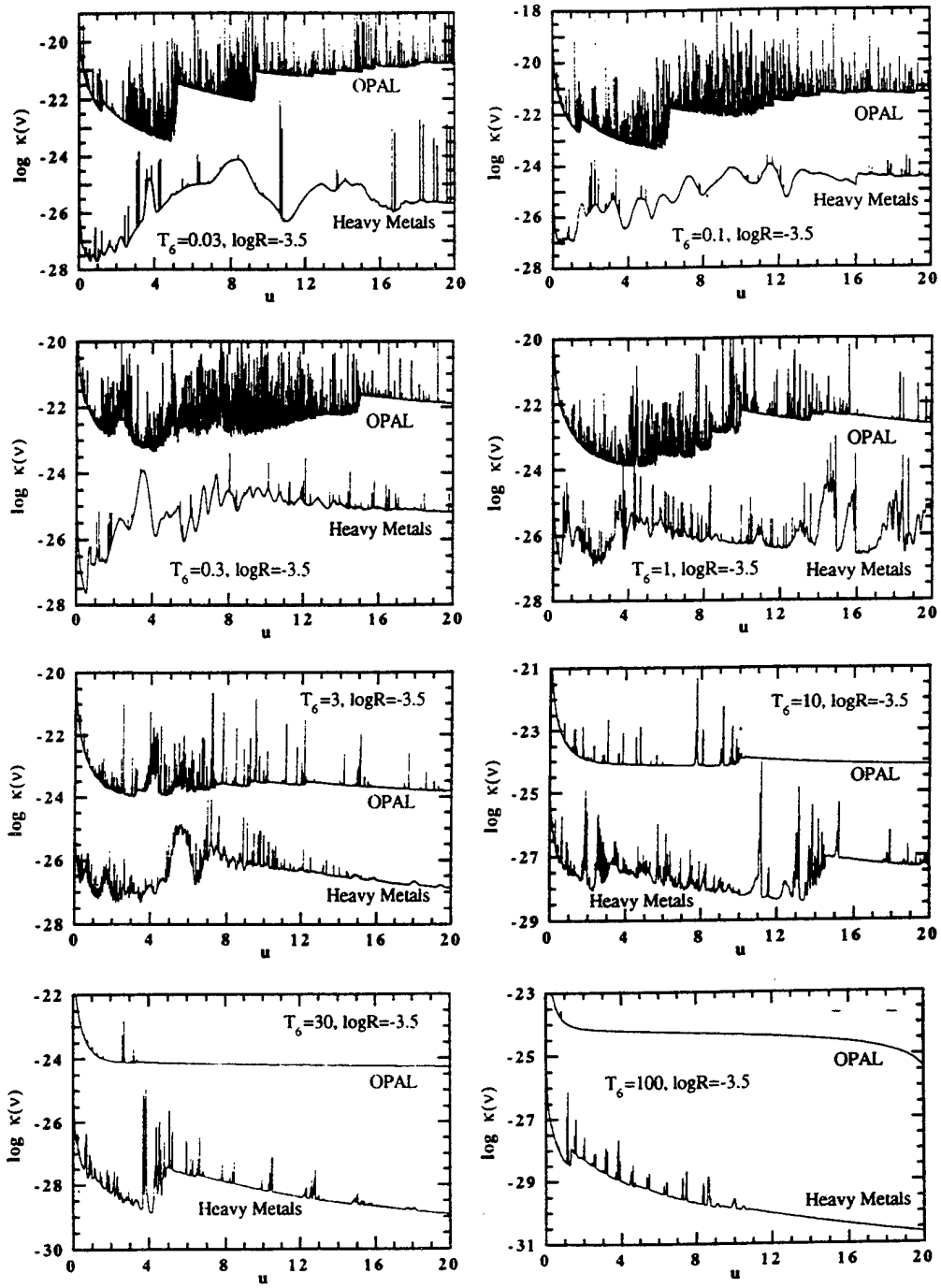


FIG. 3.—Same as Fig. 2 for log $R = -3.5$

5% increase over the IRW opacity. Consequently, both OPAL and STA calculations show that including the heavier elements can slightly increase the total mixture opacity at these matter conditions. In this temperature-density region, previous work (Rogers & Iglesias 1992, 1993; Seaton et al. 1994) found the largest opacity increases due to the weakly abundant iron-group elements. Therefore, it is not surprising that in this same region Ga would also make the largest impact. Note that the present calculations have combined the abundance of Cu, Zn, Ge, and As with that of Ga and a calculation with all these elements explicitly included would broaden the STA feature at $u \approx 3$. This broadening would not significantly affect the overall strength of this feature, but would reduce the portion that emerges above the 14 element results. Thus, it should further reduce the effect of the heavy elements on the total opacity.

It follows from these figures that at the lowest temperature the neutral hydrogen K-shell absorption dominates the spectrum near $u = 4$; thus, minimizing the heavy element contribution. Photoabsorption cross sections for H-depleted mixtures,

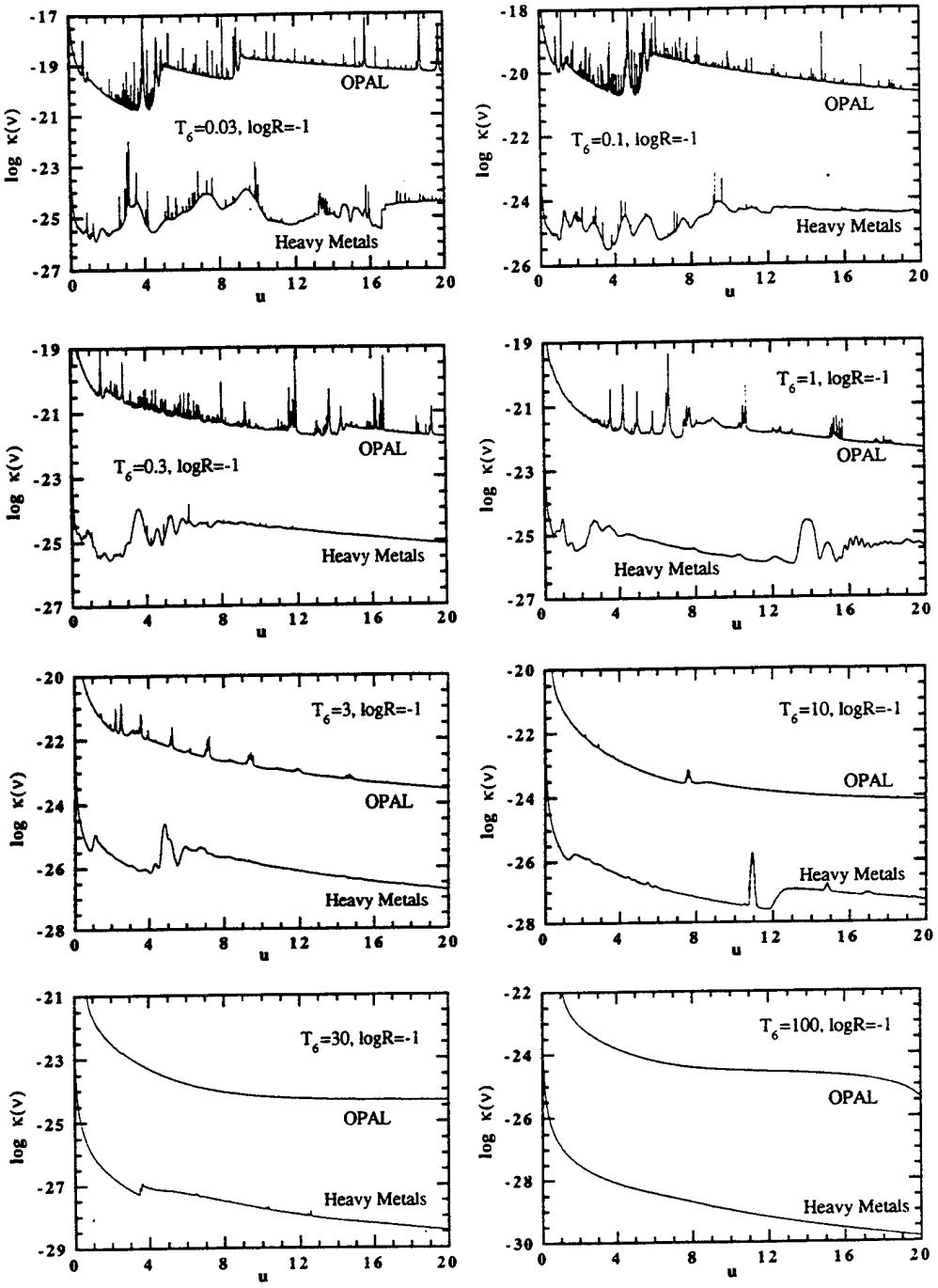


FIG. 4.—Same as Fig. 2 for log $R = -1$

$X = 0.0$ and $Z = 0.02$, are displayed in Figure 5. As before, the heavy elements clearly have no effect at the higher temperatures. The missing H absorption does make the heavy metal contribution relatively more important at $T_6 = 0.03$, but still without significant impact.

4. CONCLUSION

The effect of the heavier metals (atomic number >28) on astrophysical opacities has been shown to be small—at least for solar-mixtures. The largest opacity enhancement due to these heavy metals were found at low to medium temperatures along tracks of the smaller R values relevant to stellar models. The results show that the heavy elements are considerably better photon absorbers than the metals usually included in stellar opacity calculations, but their very small abundance dilutes their impact on the mixture opacities. There are, however, situations where some of the heavy element abundance can be significantly enhanced. For example, Ga overabundance of over 1000 times the solar value have been measured in peculiar A stars (e.g., Takada-Hidai, Sadakane, &

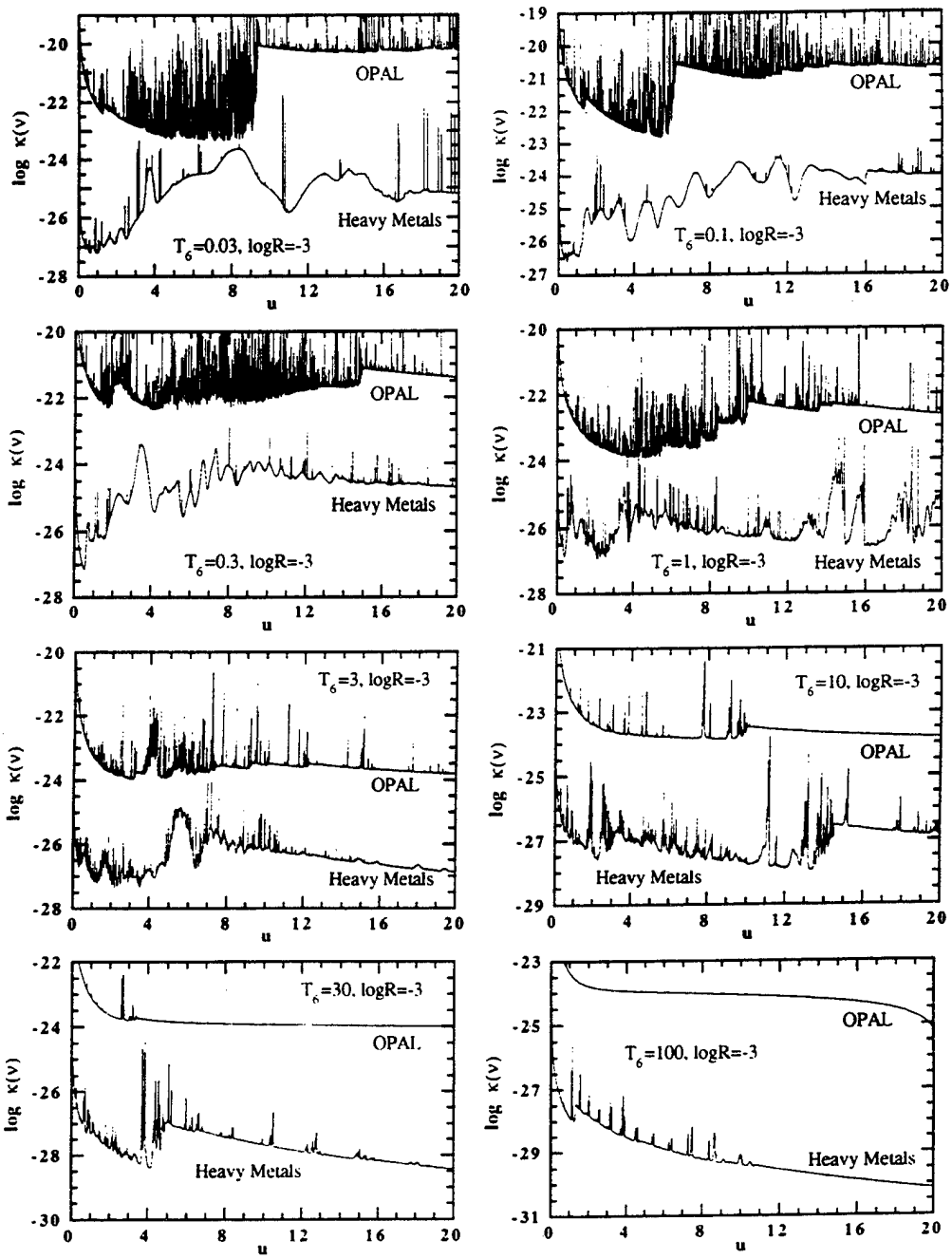


FIG. 5.—Same as Fig. 2 for $X = 0.0$ and $Z = 0.02$ at constant $\log R = -3$

Jugaku 1986). Even though in many cases these anomalies may be restricted to regions near the surface, the impact of a potentially large opacity increase due to an enhanced heavy element abundance should be considered in stellar models of such stars.

Work by C. A. I., B. G. W., F. J. R., and W. H. G. performed under the auspices of the Department of Energy by the Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

REFERENCES

Bar-Shalom, A., Oreg, J., & Goldstein, W. H. 1991, Radiative Properties of Dense Matter, ed. W. Goldstein, C. Hooper, J. Gauthier, J. Seely, & R. Lee (Singapore: World Scientific), 163
———. 1994, J. Quant. Spectros. Rad. Transf., 51, 27
Bauche-Arnoult, C., Bauche, J., & Klapisch, M. 1988, Adv. Atomic Molec. Phys., 23, 131
Grevesse, N., & Noels, A. 1993, in Origin and Evolution of the Elements, ed. N. Prantzo, E. Vangioni-Flam, & M. Casse (Cambridge: Cambridge Univ. Press)
Iglesias, C. A., Rogers, F. J., & Wilson, B. G. 1990, ApJ, 360, 221
———. 1992, ApJ, 397, 717 (IRW)
Iglesias, C. A., & Wilson, B. G. 1994, J. Quant. Spectros. Rad. Transf., 52, 127
Klapisch, M. 1971, Computer Phys. Comm., 2, 239
Klapisch, M., Schwob, J. L., Fraenkel, B. S., & Oreg, J. 1977, J. Opt. Soc. Am., 61, 148
Rogers, F. J., & Iglesias, C. A. 1992, ApJS, 79, 507
———. 1993, New Prespectives on Stellar Pulsation and Pulsating Variable Stars, ed. J. M. Nemec & J. Mathews (Cambridge: Cambridge Univ. Press), 221
———. 1994, Science, 263, 50
Rozsnyai, B. F. 1972, Phys. Rev. A, 5, 1137
Seaton, M. J., Yan, Mihalas, D., & Pradhan, A. K. 1994, MNRAS, 226, 805
Takada-Hiadai, M., Sadakane, K., & Jugaku, J. 1986, ApJ, 304, 425
Wilson, B., Albritton, J., & Liberman, D. 1991, Radiative Properties of Dense Matter, ed. W. Goldstein, C. Hooper, J. Gauthier, J. Seely, & R. Lee (Singapore: World Scientific), 189

Appendix N

Interpretation of Hot and Dense Absorption Spectra of a Near-Local-Thermodynamic-Equilibrium Plasma by the Super-Transition-Array Method

PHYSICAL REVIEW E VOLUME 52, NUMBER 6 DECEMBER 1995

Interpretation of hot and dense absorption spectra of a near-local-thermodynamic-equilibrium plasma by the super-transition-array method

A. Bar-Shalom and J. Oreg
Nuclear Research Center, Negev, P.O. Box 9001, Beer-Sheva, Israel 84190

J. F. Seely, U. Feldman, and C. M. Brown
E. O. Hulburt Center for Space Research, Naval Research Laboratory, Washington, D.C. 20375-5352

B. A. Hammel, R. W. Lee, and C. A. Back
Lawrence Livermore National Laboratory, Livermore, California 94550
(Received 30 May 1995)

The super-transition-array model is shown to be a very convenient tool for the interpretation of near-local-thermodynamic-equilibrium hot and dense plasmas. Specifically, we interpret here the absorption spectra of the CH-Ni-CH foil experiment performed at Lawrence Livermore National Laboratory using the backlighter technique. In this experiment a laminar foil composed of 200-Å Ni with 1000-Å CH on both sides was radiatively heated by the x-ray continuum from a nearby gold plasma and was backlit by the x-ray continuum from a distant gold plasma that could be time delayed with respect to the heating pulse. This setup was designed to achieve a uniform density and heating of the Ni middle layer. It is found that the Ni absorption features depend very weakly on the density of the foil but are quite sensitive to the foil temperature. Remarkably good agreement between the theory and the experiment is obtained for the Ni $2p$-$3d$ spectrum. The detailed features indicate that the plasma temperature is confined to a narrow range between 14 and 18 eV, demonstrating that the foil design, aiming to create a homogeneous Ni plasma, was successful. These results represent an alternative temperature diagnostic for high-Z plasma.

PACS number(s): 52.70.−m

I. INTRODUCTION

In this work we apply the super-transition-array (STA) method [1–5] to interpret the absorption spectra obtained from hot and dense plasmas in the near-local-thermodynamic-equilibrium (LTE) regime, created by irradiation of matter by x-ray emission of a gold plasma. For relatively high-Z materials (e.g., Ni) such plasmas contain a huge number of transitions even when LTE conditions do not exactly hold, and the spectrum has many unresolved structures. The application of the collisional radiative model in these cases is a tremendous effort if possible at all. It is shown that the STA model, though based on LTE conditions, is a very convenient tool for the interpretation of such plasmas. For given temperature and density, it produces the exact wavelengths and widths of *all* the contributing clusters, including the individual linewidths and the entire cluster widths constructed from the many overlapping lines. The deviation in the intensities reveals the degree of departure from LTE conditions.

Specifically, we interpret here the absorption spectrum obtained in the CH-Ni-CH foil experiments performed at Livermore [6,7]. In these experiments a laminar foil composed of 200-Å Ni with 1000-Å CH on both sides was radiatively heated by the x-ray continuum from a nearby gold plasma, and was backlit by the x-ray continuum from a distant gold plasma that could be time delayed with respect to the heating pulse [8]. The tamped CH-Ni-CH foil was designed to achieve a uniform density and heating of the Ni middle layer. As we shall see, the Ni absorption features depend very weakly on the density of the foil but are quite sensitive to the foil temperature. The comparison between theory and experiment shows remarkable agreement for the Ni $2p$-$3d$ spectrum. The detailed features in this region indicate that the plasma temperature is indeed confined to a narrow range between 14 and 18 eV. These results represent a temperature diagnostic for high-Z plasmas. In Sec. II we briefly review the STA model and the experimental setup. In Sec. III we compare the experimental and STA theoretical results. A discussion and summary are given in Sec. IV.

II. REVIEW OF THE THEORY AND EXPERIMENT

A. STA model

The most complex contribution to the spectrum emitted from hot and dense plasmas arises from the huge number of bound-bound and bound-free transitions. Unresolved clusters of many neighboring partially overlapping transition lines are created, constructing a complex intensity profile. The STA model reveals this complicated structure by a convergence procedure which increases the resolution of the calculated spectrum, until the re-

1063-651X/95/52(6)/6686(6)/$06.00

quired accuracy is achieved. Particularly, in each step the entire bulk of transitions is divided into groups G of neighboring lines, and each group is described as a Gaussian having the exact group moments, i.e., total intensity, average energy, and variance. The resolution is thus increased with the number of groups.

Specifically, the total spectrum can be written as

$$S(E)=\sum_G S_G(E)\ , \tag{1}$$

where

$$S_G(E)=\sum_{i,j\in G} N_i w_{ij} P_{ij}(E-E_{ij})\ . \tag{2}$$

In Eq. (2) the summation is over all the transitions $i\rightarrow j$ in G, where i and j indicate the corresponding initial and final levels. N_i is the population of the initial level, w_{ij} is the transition probability, and P_{ij} is the corresponding line shape centered on the transition average energy E_{ij}.

For normalized symmetric line profiles, the group moments are the following:

intensity

$$I_G\equiv\int S_G(E)dE=\sum_{i,j\in G} N_i w_{ij}\ , \tag{3}$$

average energy

$$E_G\equiv\frac{\int S_G(E)E\,dE}{I_G}=\frac{\sum_{i,j\in G} N_i w_{ij} E_{ij}}{I_G}\ , \tag{4}$$

and variance

$$(\Delta E_G)^2\equiv\frac{\int S_G(E)(E-E_G)^2dE}{I_G}=\Delta_G^2+\Delta_P^2\ , \tag{5}$$

where

$$\Delta_G^2=\frac{\sum_{i,j\in G} N_i w_{ij}(E_{ij}-E_G)^2}{I_G}\ , \tag{6}$$

and

$$\Delta_P^2=\int P(E-\bar{E})(E-\bar{E})^2dE \tag{7}$$

is the variance of the individual line shape, assumed to be equal for all lines in group G.

The central achievement of STA theory is the ability to obtain, under certain conditions detailed below, analytical formulas for the moments, bypassing the impractical need to account for the huge number of transitions one by one [1,5]. The only two assumptions made are that (1) the plasma is in LTE conditions, yielding Boltzmann populations N_i; and (2) The configuration widths are smaller than kT.

In addition, the definition of STA groups described below has two advantages [1,5]: (1) it enables one, using these assumptions, to derive analytic expressions for the group moments; and (2) it allows a group splitting strategy which speeds up the convergence. For bound free transitions the final level j belong to the continuum, and the moments of G are obtained by integration over the continuum [2,3].

In order to account for the non-Gaussian nature of P, we first construct a Gaussian from the moments I_G, E_G, and Δ_G,

$$\Gamma(E-E_G)=\frac{I_G}{\sqrt{2\pi}\Delta_G}\exp\left[-\frac{1}{2}\left[\frac{E-E_G}{\Delta_G}\right]^2\right]\ , \tag{8}$$

and then construct the spectrum $S_G(E)$ by the convolution with the individual line shape

$$S_G(E)\equiv\int\Gamma(E-E_G)P(E-E)dE \tag{9}$$

having the same moments as the original spectrum of G defined by Eq. (2).

In order to complete this brief description of the theory, we now define the STA groups. A STA group (termed STA) is the collection of all transitions between two superconfigurations. A superconfiguration Ξ is a collection of ordinary configurations defined symbolically by the product over supershells σ,

$$\Xi\equiv\prod_\sigma \sigma^{Q_\sigma}\ . \tag{10}$$

A supershell, in turn, is the union of energetically adjacent ordinary atomic subshells $s\equiv\not{s}_s\equiv n_s l_s j_s$. In Eq. (10), the superconfigurations are constructed by distributing the Q_σ electrons occupying the supershell σ among the subshells s in all possible ways subject to $\{\sum_{s\in\sigma} q_s=Q_\sigma\}$:

$$\sigma^{Q_\sigma}\equiv\sum_{\{\sum_{s\in\sigma} q_s=Q_\sigma\}}\prod_s \not{s}_s^{q_s}\ . \tag{11}$$

Clearly each partition of Q_σ is an ordinary configuration. The transitions between two configurations constitute an unresolved transition array (UTA) [9,10], and a STA is thus a collection of energetically near UTA's.

The convergence procedure mentioned above splits supershells into smaller supershells according to their energy spread. For each superconfiguration in its turn, at each step, supershells that give rise to relatively well-separated configurations, are preferentially split. The detailed structure of the spectrum is thus gradually revealed, yielding a converging spectrum. This procedure converges to the UTA spectra where each UTA is completely unresolved.

It is appropriate to collect separately all STA's belonging to a specific one electron jump $\not{\alpha}\rightarrow\not{\beta}$, e.g., $2p_{3/2}\rightarrow 3d_{5/2}$, which fall in the same region. This characterization will be used in comparison with the experiment given in Sec. II B.

Finally, three essential points should be emphasized here.

(1) The convergence procedure also allows the use of first order energies in the Boltzmann factor for the level configurations [2,3].

(2) Orbital relaxation is applied by taking different potentials for different superconfigurations [2,3]. This relaxation is carried out for different STA's as well as for the lower and upper superconfigurations of the same STA. As we shall see, this orbital relaxation significantly improves the agreement between the calculated and experimental spectra.

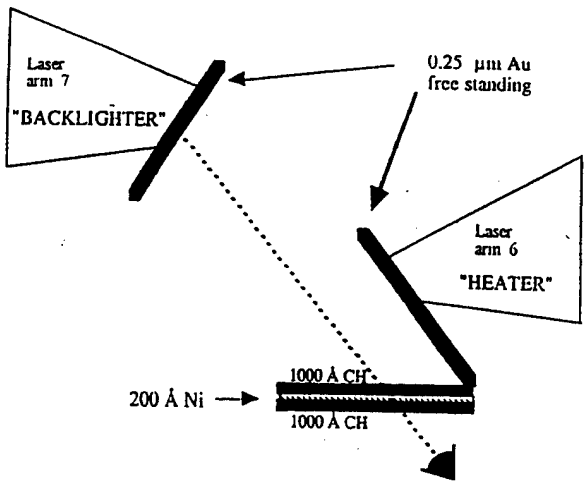
(3) The STA model was extended [4] to include the im-

portant part of the configuration interaction, i.e., the interaction between *j-j* configurations belonging to the same *LS* configuration. As will be seen below, these developments significantly improve the agreement with experimental results. The STA code is a very convenient tool for spectral diagnostics. Given the material density and temperature, it produces the entire spectra in a single run. It allows fast identification of the various arrays and lines. The deviation in intensities may help to understand the experimental conditions, i.e,. the degree of departure from LTE conditions or the occurrence of stimulated emission as detailed below.

B. Experiment

The experiment and its motivation were discussed in detail in Refs. 6 and 7. A brief summary of the experiment is given below for convenience.

The experimental setup is shown schematically in Fig. 1. Two arms of the NOVA laser were used. One irradiates a thin gold foil emitting, in turn, x rays which heat the CH/Ni/CH foil target. The second laser arm heats a second gold foil which emits an x ray directly on the already heated target. The spectrally resolved absorption of this backlight continuum is detected at the back of the target. The tamped CH-Ni-CH foil, composed of 200-Å Ni with 1000-Å CH on both sides, was designed to achieve a uniform density and heating of the middle layer. The carbon on both sides of the Ni, on the other hand, blows off and is expected to have a much wider range of temperature and density. The spectrum in the wavelength range 6–100 Å of the backlighter continuum transmitted through the radiatively heated foil was photographically recorded by a high-resolution grazing incidence spectrograph.



The Experimental Setup

FIG. 1. The experimental setup: two arms of the NOVA laser are used. One irradiates a thin gold foil, emitting, in turn, x rays which heat the CH-Ni-CH foil target. The second laser arm heats a second gold foil which emits an x ray directly on the already heated target. The absorption of this second laser light vs frequency is detected at the back of the target.

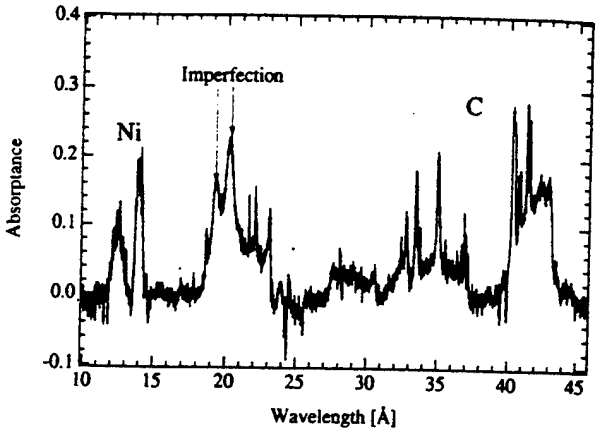


FIG. 2. The experimental spectrum includes contributions from C II to C V in the wavelength range 32–42 Å and a few ionization stages of Ni at wavelengths 12–14 Å. The features near 22 Å are of oxygen existing in the plastic layer not considered in this work. The features at 19–21 Å are imperfections in the photographic emulsion that appear in this particular lineout.

The Ni absorption features depend very weakly on the density of the foil, but are quite sensitive to the foil temperature. We will show that Ni transitions of the type $2p$-$3d$ originate from plasma of temperatures in the narrow range 14–18 eV, indicating that a homogeneous Ni plasma was indeed achieved. This analysis represents an alternative temperature diagnostic for high-Z plasmas. In comparing the experimental and theoretical results, we have found deviations in intensity ratios between transition arrays originating from the same initial states. This is attributed to stimulated emission from highly populated final levels due to the photon flux from a much higher temperature gold plasma.

The experimental spectrum presented in Fig. 2 includes contributions from C II to C V in the wavelength range 32–42 Å and a few ionization stages of Ni at wavelengths 12–14 Å. The identification of the spectral features and the calibration of the photographic plates were discussed in Ref. [6]. The cold Ni L edge at 14.5 Å is absent in the spectrum. The features near 22 Å (the cold edge at 23.3 Å) are of oxygen existing in the plastic layer not considered in this work. The features at 19–21 Å are imperfections in the photographic emulsion that appear in this particular lineout.

III. INTERPRETATION OF THE SPECTRA BY THE STA MODEL

The experimental *absorption* spectrum is related to the absorption coefficient $\sigma(\lambda,\rho,T)$ by

$$A(\lambda)=(1-e^{-\sigma(\lambda)\rho\Delta x}) , \tag{12}$$

where $\rho\Delta x$ is the plasma depth, estimated to be 2.10^{-5} g/cm^2 in our experiment. The STA model described in Sec. II calculates $S=\sigma\rho$ of Eq. (1) from which we obtain $A(\lambda)$.

The experimental and theoretical spectra are shown in Figs. 3–5. We have found that the resolved structures

originate from the superconfigurations

$$\Xi=(1s^2 2s^2 2p^6)(3s3p_{1/2}3p_{3/2}3d_{3/2}3d_{5/2})^N \times(4s4p_{1/2}\ldots)^M , \quad (13)$$

where $N=8-11$ and the M's (mainly between $M=0$ and 2) are not resolved. The observed transition arrays belong to one electron jumps $2p\rightarrow nd$ ($2p_{1/2}\rightarrow nd_{3/2}+2_{3/2}\rightarrow nd_{3/2}+2p_{3/2}\rightarrow nd_{5/2}$) with $n=3$, 4, and 5.

In Fig. 3 we examine the $2p\rightarrow 3d$ spectra. Four spectra are shown for four temperatures from 10 to 22 eV. In Fig. 3(a) it is seen immediately that the calculated spectra of the two extreme temperatures (belonging to $T=10$ and 22 eV) extend beyond the experimental spectral structure. This limits the temperature spread to about 4 eV only ($T=14-18$ eV) as seen in Fig. 3(b). We have found that due to its narrow range the contribution of the entire continuous temperature profile can be well represented by the two limiting values ($T=14$ and 18 eV). The addition of intermediate temperatures alters the spectrum only slightly.

In Figs. 4(a) and 4(b) the two temperature spectra of $2p\rightarrow 3d$ are shown. The "total" curve is the "best" combination of the two temperatures (40% $T=14$ eV+60% $T=18$ eV). In Fig. 4(a) we present the theoretical result without configuration interaction (CI) [11], whereas Fig. 4(b) includes CI in agreement with the experiment. In Fig. 4(b) we also present a complete identification of all the resolved arrays using the compact notation N^- and N^+ for arrays $2p_{1/2}\rightarrow nd_{3/2}$ and $2p_{3/2}\rightarrow nd_{5/2}$, respectively. As expected, array $2p_{3/2}\rightarrow nd_{3/2}$ is negligibly weak, and only N^- and N^+ are observed. The agreement is strikingly good, and all arrays $N=9-11$ (denoted by $9^{\pm}-11^{\pm}$) are identified. In pure j-j coupling (without CI) the ratio between the + and − transitions is in favor of the + transitions, and the departure from j-j already seen in Fig. 3(a). It should be emphasized here that this excellent agreement was obtained only after improving the orbital relaxation, using a sufficient number of opti-

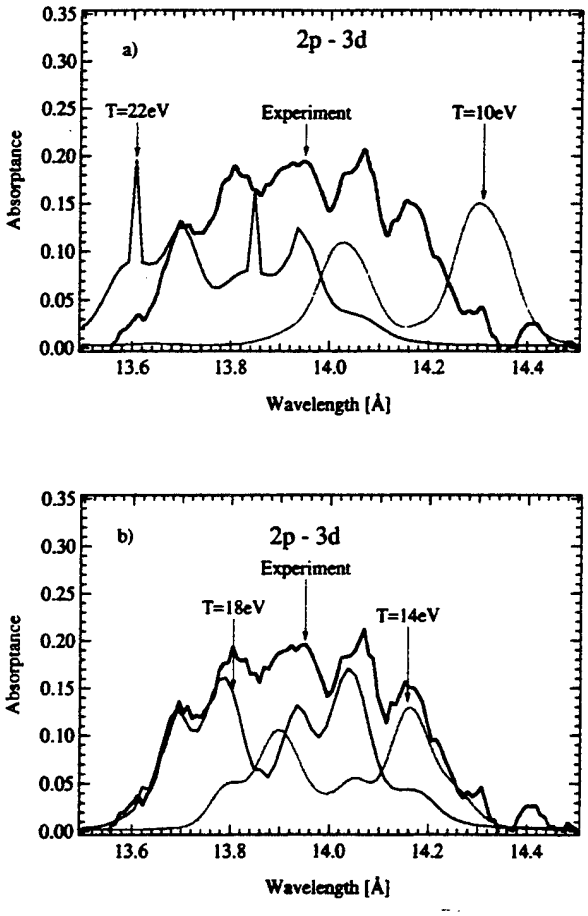


FIG. 3. Comparison between the theoretical and the experimental $2p\rightarrow 3d$ spectrum. (a) Two temperature (10 and 22 eV) spectra beyond the experimental limit. (b) The theoretical spectra of the temperature bounds (14 and 18 eV) reproducing the experimental spectrum.

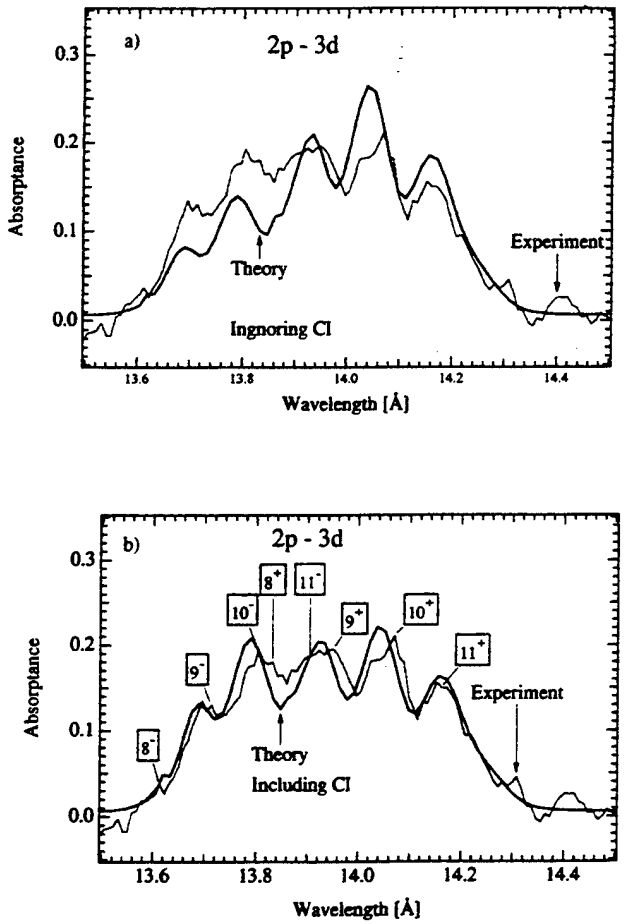


FIG. 4. The total two temperature (40% $T=14$ eV+60% $T=18$ eV) spectra of $2p\rightarrow 3d$: (a) without CI, and (b) with CI. N^+ and N^- indicate the $2p_{1/2}\rightarrow nd_{3/2}$ and $2p_{3/2}\rightarrow nd_{5/2}$ transitions, respectively.

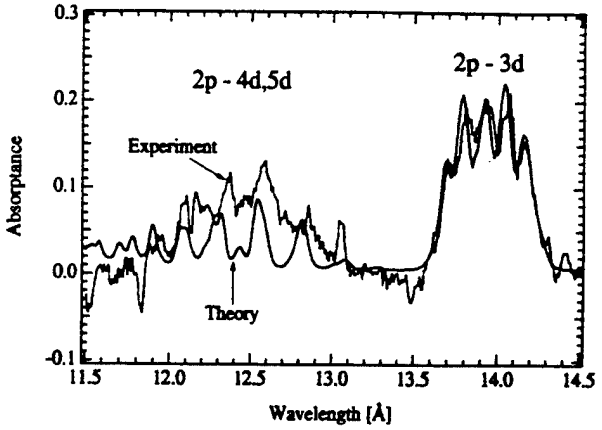


FIG. 5. The complete Ni LTE spectrum at $T=14$ eV$+18$ eV and density 0.01 g/cc.

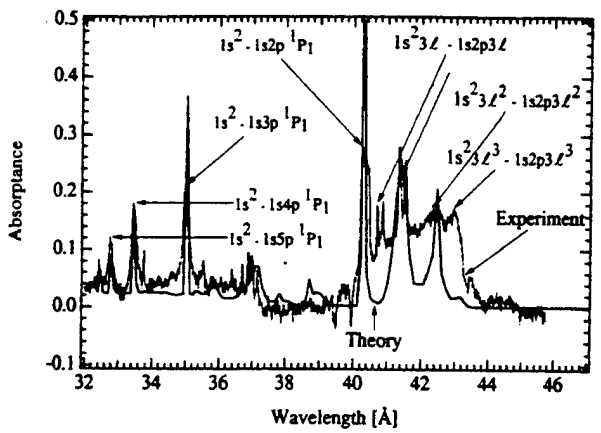


FIG. 7. Identification of the carbon LTE spectrum at $T=14$ eV$+18$ eV and density 0.01 g/cc. The noisy and smooth lines are the experimental and theoretical results, respectively.

mized potentials for both initial and final superconfigurations. This comparison thus demonstrates the importance of both orbital relaxation and CI for the interpretation of plasma experiments.

The fact that the $2p$-$3d$ features are modeled so well is the most interesting result of this work, indicating that only three Ni ionization states contribute to this spectrum, and that the temperature range is extremely narrow (14–18 eV). These results, together with the absence of the Ni cold L edge at 14.5 eV, imply that the Ni plasma was uniformly heated by the x-ray continuum without large temperature gradients. This confirms the success of the foil design to avoid fast expansion of the middle layer and to allow a uniform heating.

In Fig. 5 the complete two temperature (40% $T=14$ eV$+60\%$ $T=18$ eV) Ni spectrum is presented. The comparison between the experiment and STA are not so good for the transitions $2p\rightarrow 4d$, $5d$. Whereas shifts in wavelength can be improved by better optimization of the potential, we can see significant deviations is intensity ratios between transition arrays $2p\rightarrow 4d$, $5d$ and $2p\rightarrow 3d$. These deviations cannot be attributed to a departure from LTE *absorption* since all these arrays originate from the same initial states. Departure from the LTE may affect these intensity ratios only through stimulated emission. Whereas under LTE conditions stimulated emission is obtained by simply multiplying the absorption spectra by the Plank function at the plasma temperature (in our case 14 eV $< T <$ 18 eV), a non-LTE effect can originate here from *the photon flux from the much higher temperature (keV) gold plasma (in the spectral region of these arrays)*. The discrepancy in the intensities can therefore be explained as follows: We could adjust the plasma depth Δx of Eq. (12) to fit the total intensity at the $2p\rightarrow 4d$, $5d$ region. In such a normalization the $2p$-$3d$ calculated spectrum will still follow the detailed structure of the experimental spectrum but will be globally intensified. The lower experimental intensity in this region can now be attributed to the *non-LTE stimulated emission from the highly populated final levels of the 2p-3d array due to the photon flux from the much higher temperature (keV) gold and to the strong 2p-3d transition probability (relative to the $2p\rightarrow 4d$, $5d$)*. A more quantitative explanation can be obtained by solving the detailed collisional radiative model. Such a calculation is beyond the scope of the present work.

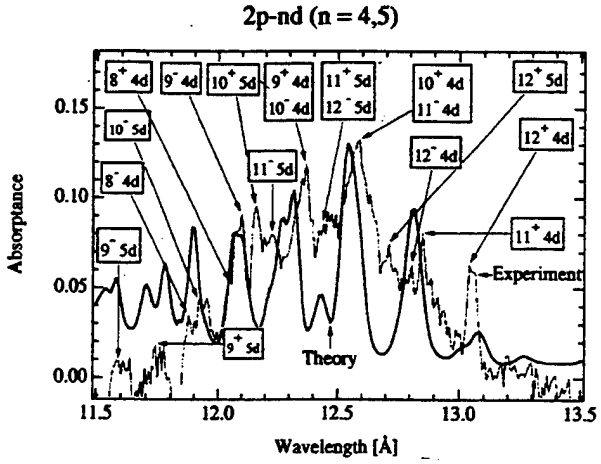


FIG. 6. Identification of the $2p\rightarrow 4d,5d$ transitions. (N,nl) indicates the $2p\rightarrow nl$ one-electron jump origination from the bulk of superconfigurations denoted in the text by N, $(N,M$: summed over all $M)$ of Eq. (13).

In spite of these deviations, almost all the unresolved arrays of the transitions $(8^{\pm},4d)$ $(12^{\pm}$-$4d)$ and $(9^{\pm}$-$12^{\pm},5d)$ could be identified, as is specified in Fig. 6. The carbon on both sides of the Ni layer blows off, and is expected to have a much higher range of temperatures and density. It is thus not surprising that the modeling of carbon is not so good. The results for carbon are presented in Fig. 7. Again the strongest features could be identified in spite of the temperature spread which could not be accounted for accurately in the calculation.

IV. SUMMARY AND DISCUSSION

In this work we have demonstrated how the STA model can be used for the interpretation of near-LTE plasma

experiments. We have shown that for sufficiently relaxed orbitals the wavelengths of the various arrays in the entire spectral region are accurately calculated. The deviations from the LTE level populations may, in general, change the entire spectrum. Dominant LTE arrays may vanish and others may become significant. In this case a detailed collisional radiative model must be solved to obtain the level populations and the resulting array intensities. This is an enormous task. In contrast the STA model is a very simple tool, yielding the entire spectrum in a single run and allowing fast identification of all resolved arrays and lines. Of course STA spectra can be used quantitatively only if the departure from the LTE is not too large so that the LTE features are still recognizable. In the present experiments we have distinguished only small non-LTE effects. The remarkable reconstruction of the experimental $2p$-$3d$ spectra by the theory with only two temperatures in a range of only 4 eV confirmed the success of the foil design aiming at homogeneous plasma in both density and temperature. The deviations in relative intensities of $2p \rightarrow nd$ arrays, between transitions with $n=4$ and 5 and with $n=3$, are attributed to spontaneous and stimulated emission from highly populated final levels due to the photon flux from the much higher temperature gold plasma. This work introduces an alternative temperature diagnostic for high-Z plasma.

[1] A. Bar-Shalom, J. Oreg, W. H. Goldstein, D. Shvarts, and A. Zigler, Phys. Rev. A **40**, 3183 (1989).

[2] A. Bar-Shalom, J. Oreg, and W. H. Goldstein, in *Radiative Properties of Hot Dense Matter,* edited by W. H. Goldstein, C. Hopper, J. Gautier, J. Seely and R. Lee (World Scientific, Singapore, 1990).

[3] A. Bar-Shalom, J. Oreg, and W.H. Goldstein, in *Atomic Processes in Plasmas,* edited by E. S. Marmar and J. L. Terry, AIP Conf. Proc. No. 257 (AIP, New York, 1991).

[4] A. Bar-Shalom, J. Oreg, and W. H. Goldstein, J. Quantum Spectrosc. Radiat. Transfer **51**, 27 (1994).

[5] A. Bar-Shalom, J. Oreg, and W.H. Goldstein, Phys. Rev. E **51**, 4882 (1995).

[6] J. F. Seely *et al.* J. Quantum Spectrosc. Radiat. Transfer **51**, 349 (1994).

[7] J. D. Bauer *et al.* Phys. Rev. E (to be published).

[8] D. R. Kania *et al.*, Phys. Rev. A **46**, 7853 (1992).

[9] C. Bauche-Arnoult, J. Bauche, and M. Klapisch, Phys. Rev. A **25**, 2641 (1982).

[10] C. Bauche-Arnoult, J. Bauche, and M. Klapisch, Phys. Rev. A **31**, 2248 (1995).

[11] A. Bar-Shalom, J. Oreg, and W. H. Goldstein, J. Quant. Spectrosc. Radiat. Transfer **51**, 27 (1994).

Appendix O

Opacity of Dense, Cold, and Strongly Coupled Plasmas

Opacity of Dense, Cold, and Strongly Coupled Plasmas

A. N. Mostovych, L. Y. Chan, and K. J. Kearney
Laser Plasma Branch, Plasma Physics Division, U.S. Naval Research Laboratory, Washington, D.C. 20375

D. Garren,[1] C. A. Iglesias,[2] M. Klapisch,[3] and F. J. Rogers[2]
[1]*Science Applications International Corporation, McLean, Virginia 22102*
[2]*Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, California 94550*
[3]*ARTEP, Inc., Columbia, Maryland 20899*
(Received 22 March 1995)

The physics of dense [$(1-5) \times 10^{19}$ cm^{-3}], cold (1–15 eV), strongly coupled ($\Gamma \sim 0.7$) plasmas is probed with 0.351, 0.527, and 1.054 μm opacity measurements in well characterized, laser-heated, aluminum plasmas. Current opacity models are tested, for the first time, in the regime where the probing photon energies are of the same order as the average interparticle interaction energies in the plasma. Predicted enhancements of the opacity at low temperatures are not observed, but overall agreement between experiment and theory is within a factor of 2.

PACS numbers: 52.25.Qt, 52.25.Rv, 52.50.Jm

At low temperatures, the electrons and ions of dense plasmas interact strongly to produce highly nonideal gas systems. The interaction to thermal energy ratio Γ is no longer small as in "weakly coupled" ideal gas plasmas. The opacity of strongly coupled plasmas (SCP) is of fundamental interest because it is sensitive to nonideal effects through its strong dependence on the equation of state, particle collisions, plasma microfields, and atomic line shapes in the plasma. The largest nonideal effects are expected for photons with energies comparable to the plasma interaction energies, i.e., $h\nu/kT \leq \Gamma$. At these energies the photon-plasma interaction is primarily determined by electron-ion collisions and transitions between high-lying atomic levels which are subject to strong perturbations from the plasma. The opacity and transport properties of dense plasmas are important issues in astrophysics, laser-fusion, and shock wave research, as well as in the physics of high current discharges. The physics of SCP's [1] and the opacity [2] of dense plasmas have been the subject of vigorous theoretical and numerical investigations with relatively few experiments. Comparisons between opacity models show reasonable agreement between models for high temperature plasmas with high-charge-state ions, but rather poor agreement for low temperature plasmas with low-charge-state ions [3]. Recently, the hot, high-charge-state plasma regime has been investigated by several soft x-ray and XUV opacity experiments [4]. Previous opacity work with low temperature plasmas was primarily with plasma arcs [5], shock waves in high density gases [6], or laser vaporization and heating of thin metal films [7]. Typically, the arc plasma experiments were at low densities such that $\Gamma < 0.2$, whereas in the shock wave experiments the opacity was not measured directly but was inferred from plasma emission measurements. In our early work [7] with laser-heated vapors, the plasma expansion from laser heating was only measured in two dimensions (2D), making computer modeling of the plasma hydrodynamics necessary to fully analyze the data. In this letter, we report on new multi-wavelength opacity experiments with laser-heated metallic vapors. New diagnostics and instrumentation measure the plasma expansion in 3D, removing previous uncertainties. The UV to IR emission spectrum and the 0.351, 0.527, and 1.054 μm opacity of strongly coupled plasmas for $n_e \approx (1-5) \times 10^{19}$ cm^{-3}, $T_e \approx 1-15$ eV, and $\Gamma = 0.3-0.8$ are measured under controlled and well diagnosed conditions. Comparisons with calculations show agreement, within a factor of 2, with the STA [8] and OPAL [9] opacity models.

Similar to Ref. [7], a laser-produced plasma is used as the source for these experiments; see Fig. 1. A glass substrate is coated with an Al film, and a laser beam vapor-

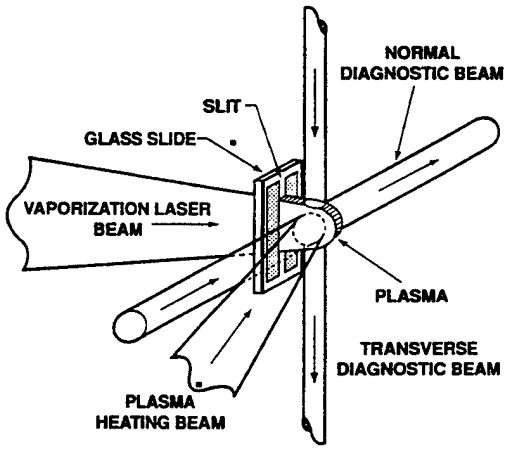


FIG. 1. Schematic of experimental setup. A laser vaporizes a thin aluminum film from a glass substrate. The aluminum vapor flows through a thin slit (100 μm) and is heated and ionized by a second laser. Multiwavelength probes measure the plasma absorption coefficient while transverse interferometer beams measure the plasma density profile.

 0031-9007/95/75(8)/1530(4)$06.00

izes this film by irradiation through the glass. The expanding supersonic vapor, with a diameter comparable to the laser focal spot ($d \sim 1.0$ mm), flows through a slit ($\Delta x \approx 100$ μm) which limits its transverse extent and total mass. After the vapor slab has reached the appropriate expansion length ($l \approx 1000$ μm at $t \approx 100$ ns) and desired vapor density, a second laser ($\lambda = 1.054$ μm, 10^{10}–10^{11} W/cm^2) is used to ionize the Al vapor. The heating occurs over a time of about 5–8 ns and produces a fully ionized plasma (roughly 1000 μm by 1000 μm) with an average ionization $\langle Z \rangle \approx 1$–4. Three short pulse (700 ps) probe beams (1.054, 0.527, and 0.351 μm) measure the transmission through the plasma, normal to the slab geometry. The beams are focused to a small spot ($d \approx 100$ μm) in the center of the plasma to provide spatial resolution and to ensure a uniform density in the focal plane. The degree of transmission is measured with a set of fast ($\tau \approx 350$ ps) incident and transmitted photodiodes. The transmission probes are much brighter than the plasma, and the transmission measurements are not affected by plasma emission. Two unfocused interferometry beams, one normal and one transverse, measure the line integrated density from the two orthogonal orientations. These measurements are unfolded to give the density profile along the line of sight of the transmission measurement. The emission of the plasma, in the focal volume of the transmission probes, is measured with an absolutely calibrated 0.5 m monochromator with a temporal resolution of 0.5 ns. The plasma temperature is determined from the absolute emission and the degree of transmission at 0.527 μm. For plasmas in local thermodynamic equilibrium (LTE), these quantities are related to the average temperature through the radiation transfer equation and Kirchhoff's law [10], i.e., $I(\nu, \Gamma) = I_p(\nu, \Gamma)(1 - I/I_0)$, where $I_p(\nu, \Gamma)$ is the Planck distribution, I/I_0 is the transmission fraction, and ν is the frequency of the absorbing radiation. The plasma emission spectra (2000–8000 Å) are recorded with a 0.33 m spectrometer coupled to a streak camera or gated microchannel plate detector.

The initial plasma geometry is slablike; however, the density profile of the plasma along the transmission-probe line of sight evolves into a Gaussian-like profile as a result of heating and expansion; see inset in Fig. 2. The profiles are symmetric as long as the vapor areal-mass-density remains below some critical value. In these experiments, the maximum Al mass density is restricted (to approximately 4.5×10^{-4} g/cm^3, $N_{\text{ion}} \sim 10^{19}$ cm^{-3}) to ensure symmetry in the density profile and to ensure full fringe visibility in the highly absorbed transverse interferometer beam. The average mass density (i.e., ion density) is inferred from the measured electron density and temperature and the degree of ionization given by a Saha equation calculation for Al. The Saha equation and the equations of state in the OPAL and STA codes give ion densities which are within 2%–3%. Typical electron densities in the experiment, averaged over the line-of-sight profile, are plotted in the solid circles of Fig. 2. The uncertainty in the temperature for these data is typically 10%–20%. The solid curve through the data is a Saha calculation for an average mass density of 3.6×10^{-4} g/cm^3. The spread in the data is due to shot-to-shot variations in the mass density of the aluminum vapor, resulting from variations in the film thickness, laser power, beam alignments, and slit width. The solid triangles in Fig. 2 correspond to the strength of the ion coupling parameter $[\Gamma = e^2 Z^{5/3}(4\pi N_e/3)^{1/3}/T_e]$. Coupling between ions is the strongest ($\Gamma \approx 0.75$) at temperatures (4–5 eV), where the population of triply ionized aluminum peaks. For the most part the plasmas are in LTE, with collisional excitation and relaxation rates dominating over radiative rates [11]. However, for Al IV, the rate of collisional excitation between the ground and 1st excited states (excitation time $\approx$ 10 ns at 10 eV) is too slow to bring the excited and higher ionization states above the Al IV ground state into LTE in the 5–8 ns heating time of the experiment. As a result, for temperatures in the 10–15 eV range, where there is significant excitation and ionization of Al IV, the degree of ionization may not reach its LTE value. The worst case deviation from LTE is estimated from a Saha calculation (dashed curve in Fig. 2) for which all states above the Al IV ground state are assigned a statistical weight of zero.

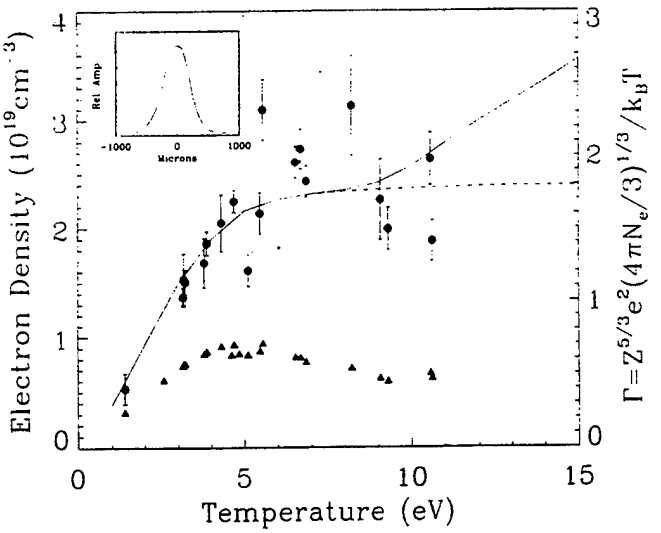


FIG. 2. Measured electron densities as a function of plasma temperature. Solid curve due to Saha equation calculation for an aluminum mass density of $\rho_m = 3.6 \times 10^{-4}$ g/cm^3. Dashed curve estimates the worst case deviation from LTE. Triangular data points correspond to the strength of the ion coupling parameter. Curve in inset is a typical transverse density profile of the plasma.

Opacity measurements are taken in the region where deviations from LTE are small. The transmission of the laser probes (1.054, 0.527, and 0.351 μm) is measured at the peak of the heating pulse and in tempo-

ral and spatial synchronization with the interferometry and absolute emission measurements. The transmission T through the plasma is related to the absorption coefficient κ by $T = \exp[-\int \kappa(x)dx]$. For comparison between experiment and theory it is convenient to define an average absorption coefficient $\langle\kappa\rangle = -\ln(T)/L$, where L is an average plasma thickness calculated from the measured density profile $n_e(x)$ such that $L\int n_e^2(x)dx = \{\int n_e(x)dx\}^2$. With this choice, processes which depend on $n_e(x)^2$ are well described by average quantities even though the plasma profile is not flat. This model is useful for our conditions because the STA and OPAL opacity codes and the data show a $n_e(x)^2$ dependence for the absorption coefficient. In practice, L is also used to calculate the average electron densities. L is comparable to the thickness containing 90% of the plasma mass (0.06–0.1 cm) and is about 1.5 times the full width at half maximum (FWHM) of the density profile.

The 0.351, 0.527, and 1.054 μm absorption coefficients are plotted in Fig. 3. These opacity data are a subset corresponding to only those measurements having an initial mass density of $\rho_0 = (3.6 \pm 1.8) \times 10^{-4}$ g/cm^3. In addition, the data are scaled by $(\rho_0/\rho_m)^2$ to account for the finite distribution of actual mass densities ρ_m and the n^2 dependence of the opacity. Error in the data is minimized by averaging over multiple data in temperature steps of 0.5 eV. The error bars include instrumental, alignment,

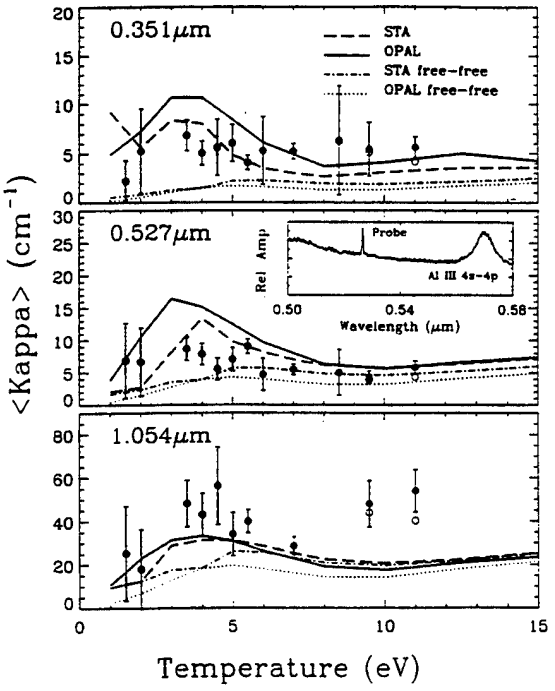


FIG. 3. The measured absorption coefficient at 0.351, 0.527, and 1.054 μm compared to OPAL and STA model calculations for an aluminum mass density of $\rho_m = 3.6 \times 10^{-4}$ g/cm^3. Open circles indicate impact of possible non-LTE density distribution on the opacity. Inset contains typical ($T \sim 5$ eV) spectra in the vicinity of the 0.527 μm probe.

and calibration uncertainties of the transmission measurements, as well as the effects of the uncertainties in density and temperature on the absorption coefficient. The effect of non-LTE density distributions on the opacity is to reduce the absorption coefficient around 10 eV, as indicated by the open data points.

To better characterize the conditions under which the opacity was measured, simultaneous time-resolved emission spectra were recorded in the 0.25–0.70 μm range. Detectors sensitive in the 1 μm range were not available for this experiment. The 0.351 and 0.527 μm probing wavelengths are in smooth regions of the emission spectra between very broad emission lines; see inset in Fig. 3. The spectra also revealed some narrow but very strong resonance absorption lines (transitions from the ground state) originating from neutral aluminum and from neutral slit material impurities. These features are far from the probing wavelengths and are the result of emission from the plasma interior being absorbed in the low-density wings of the plasma profile. Here, the temperature is sufficiently low ($\approx$1 eV) to permit finite populations of ground-state neutral atoms, and the lines are not strongly broadened by high density. The population of impurities is estimated from the relative strength of the impurity-to-aluminum lines and is found to be less than 1% of the aluminum population. At this concentration, even in the plasma interior, changes in the opacity from impurities are not expected at wavelengths outside the narrow resonance lines.

The data are compared to predictions of the STA and OPAL opacity models. They calculate the complete absorption coefficient $\kappa = \kappa_{ff} + \sum\kappa_{bf} + \sum\kappa_{bb}$, including free-free, bound-free, and bound-bound contributions. These are LTE models that include contributions from all ionization stages in the plasma, as well as from all LTE populated bound and free states of each ionization stage. In the STA model the large number of bound states that are populated in LTE are grouped into smaller supertransition arrays of similar energy-level configurations. The energy, strength, and variance of transitions between arrays are determined from solutions of the Dirac equation in parametric potentials, self-consistently optimized for each super configuration. Energies and strengths are corrected to first order to account for departure from pure jj coupling. The population of levels follows Fermi statistics; dense plasma effects are incorporated with an ion-sphere radius model of ionization potential lowering and by including degeneracy effects for the continuum. The OPAL model uses parametric potentials, but it differs in that all terms are accounted for in detail, and LS coupling determines the configuration term structure. In OPAL, dense plasma effects are included in the equation of state and in the occupation number of bound states through systematic expansions of the grand canonical partition function for a system of electrons and nuclei interacting in a Coulomb potential. In both models, bound transition energies do not have spectroscopic accuracy ($\Delta E \approx 1\%$), and pertur-

bation of states by the plasma microfields is taken into account with a semiempirical and semiclassical electron-impact broadening theory [12]. Free-free transition cross sections for the two models are calculated from their respective plasma-screened parametric potentials with quantum mechanical partial-wave expansions. These models have been successful in reproducing previous x-ray ultraviolet opacity measurements in higher temperature experiments [4] but have not been tested in the low-temperature UV-visible-IR region. This low-temperature, low-photon-energy regime is different because free-free transitions now form a significant fraction of the opacity, the broadening of bound transitions by plasma perturbations is greater than configuration splitting, and many of the contributing bound transitions are near the plasma-perturbed continuum edge.

Initial comparisons of the models to the data differed by factors of 2 to 5. Spontaneous emission reductions to the opacity had to be included (STA), and approximate free-free calculations had to be replaced with full dipole matrix evaluations (STA and OPAL). In Fig. 3, the OPAL and STA calculations are displayed alongside the data. The calculations are based on the average density and temperature measured in the experiment. In practice, the plasmas have finite density and temperature gradients. The sensitivity of the calculations to these gradients was checked by integrating the local OPAL opacities over the measured electron-density profile and several physically reasonable temperature profiles. The set of temperature profiles was chosen to span the range from the ideal, flat profile, to the worst-case, highly peaked profile. The temperature profiles are constrained by the measured density profile, via the equation of state, and by the requirement to reproduce the measured plasma emission. The difference in opacity between calculations using average values vs full profile averages increases as the temperature profile is peaked, limiting to about 15%–20% for the worst case where the concomitant mass density, as dictated by the equation of state, becomes flat. The actual deviation will be less than 20% because an expanding plasma can not have a flat mass-density profile.

Both STA and OPAL predict enhancement of bound state contributions in the 2–6 eV temperature range, where the Al III ionization state and its large set of spectral lines in the UV-visible-IR range are well populated. Even though the probing wavelengths do not coincide with aluminum absorption lines, the wings of many spectral lines contribute to the absorption as a result of spectral line broadening by the dense plasma. The measured opacity, while greater than the free-free opacity, is nearly constant and does not show significant enhancement in this temperature range. The 0.351 and 0.527 μm data indicate that bound states contribute less than is calculated by STA and OPAL. The sensitivity of the opacity to line broadening issues is illustrated by the difference in STA and OPAL opacities. The OPAL bound state contributions are larger, primarily as a result of broader (2×) line profiles in OPAL calculations. The role of spectral line broadening is being investigated in a new set of experiments and will be published elsewhere. The overall differences between experiment and theory imply theoretical uncertainties on the order of 20%–50% in this dense, low temperature regime.

The opacity (0.351, 0.527, and 1.054 μm) of dense, cold, and strongly coupled plasmas ($\Gamma \approx 0.5-0.7$) has been measured under well characterized conditions. Comparison of the data with current opacity codes such as OPAL and STA show that both bound-bound and free-free absorption processes are important in this parameter regime, but that their relative contributions are not yet certain. Overall agreement between experiment and theory is within a factor of 2.

The NRL authors were supported by the U.S. Office of Naval Research.

[1] Good reviews are found in S. Ichimaru, H. Iyetomi, and S. Tanaka, Phys. Rep. **149**, 91 (1987); S. Ichimaru, *Statistical Plasma Physics* (Addison-Wesley, Reading, MA, 1992); *Strongly Coupled Plasma Physics*, edited by H. M. Van Horn and S. Ichimaru (University of Rochester, Rochester, 1993), and references within.

[2] See, for example, *Radiative Properties of Hot Dense Matter*, edited by W. Goldstein, C. Hooper, J. Gauthier, J. Seely, and R. Lee (World Scientific, Singapore, 1991).

[3] S. J. Rose, J. Quant. Spectrosc. Radiat. Transfer **51**, 317 (1994).

[4] P. T. Springer *et al.*, Phys. Rev. Lett. **69**, 3735 (1992); L. B. da Silva *et al.*, Phys. Rev. Lett. **69**, 438 (1992); J. M. Foster *et al.*, Phys. Rev. Lett. **67**, 3255 (1991); T. S. Perry *et al.*, Phys. Rev. Lett. **67**, 3784 (1991).

[5] V. E. Gavrilov *et al.*, Sov. J. Opt. Technol. **49**, 403 (1982); M. M. Popović, and D. S. Dordević, in *Strongly Coupled Plasma Physics*, edited by H. M. Van Horn and S. Ichimaru (University of Rochester, Rochester, 1993).

[6] V. E. Bespalov, V. K. Gryaznov, and V. E. Fortov, Sov. Phys. JETP **49**, 71 (1979); D. Erskine, B. Rozsnyai, and M. Ross, J. Quant. Spectrosc. Radiat. Transfer **51**, 97 (1994).

[7] A. N. Mostovych *et al.*, Phys. Rev. Lett. **66**, 612 (1991).

[8] A. Bar-Shalom *et al.*, Phys. Rev. A **40**, 3183 (1989).

[9] F. J. Rogers and C. A. Iglesias, Astrophys. J. Suppl. Ser. **79**, 507 (1992).

[10] Ya. B. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Academic Press, New York, 1966).

[11] H. R. Griem, *Plasma Spectroscopy* (McGraw-Hill, New York, 1964).

[12] M. S. Dimitrijevic and N. Konjevic, J. Quant. Spectrosc. Radiat. Transfer **29**, 451 (1980).

Appendix P

Improved Two-Moment Radiation Transport

To apppear in: *Proc. 23rd European Conference on Laser Interaction with Matter*

Improved two-moment radiation transport

David Alan Garren†[12], Marcel Klapisch‡, John Gardner§ and Denis Colombant‖

† Science Applications International Corporation, 1710 Goodridge Dr., P.O. Box 1303, McLean, Virginia 22102, USA

‡ ARTEP, Columbia, Maryland 21045, USA

§ Laboratory for Computational Physics and Fluid Dynamics, Naval Research Laboratory, Washington, DC 20375, USA

‖ Plasma Physics Division, Naval Research Laboratory, Washington, DC 20375, USA

Abstract. A two moment model of radiative transport is developed, based on the *ad hoc* assumption that the specific intensity of radiation has the angular dependence of an ellipsoid with the origin lying at one of the foci. This model reduces to the standard diffusion approximation in the limit $\delta \equiv F/cU \ll 1$, with F and U the radiant flux and energy density. However, unlike standard diffusion theory, the present model gives the correct radiative transport in the limit in which all the radiative energy streams in one direction, i.e., $\delta = 1$. This property is crucial in predicting accurately the radiation transport in two and three-dimensional laser-plasma simulations.

1. Introduction

1.1. Radiation transport equation

Computer simulations of laser-plasma interactions require an accurate model for the transport of radiation. The radiation transfer equation [1, 2, 3]

$$\left[\frac{1}{c}\frac{\partial}{\partial t} + \mathbf{\Omega}\cdot\nabla\right] I_\nu(\mathbf{\Omega}) = S_\nu - \kappa'_\nu I_\nu(\mathbf{\Omega}) \tag{1}$$

determines the radiation intensity $I_\nu(\mathbf{\Omega})$ at frequency ν in the direction of the unit vector $\mathbf{\Omega}$. The effective absorption coefficient κ'_ν includes the contribution due to induced emission and S_ν gives the remaining source part of the emission.

[1] Mailing Address: Code 6730, Naval Research Laboratory, Washington, DC 20375, USA
[2] E-mail (INTERNET): GARREN@THIS.NRL.NAVY.MIL.

For a plasma in approximate local thermodynamic equilibrium (LTE), the source is $S_\nu = \kappa'_\nu B_\nu$ and the effective absorption coefficient is $\kappa'_\nu \equiv \kappa_\nu[1 - \exp(-h\nu/kT)]$. Here κ_ν is the standard absorption coefficient, T is the electron temperature, and

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp(h\nu/kT) - 1} \quad . \tag{2}$$

is the Planck radiation intensity. For a non-LTE plasma [4], κ'_ν and S_ν in Eq. (1) are modified from their LTE values.

1.2. Moments of radiation transport equation

One method of solving the radiation transport equation (1) is through moments in the radiation direction $\mathbf{\Omega}$. The zeroth and first moments of Eq. (1) yield [2, 3]

$$\frac{\partial U_\nu}{\partial t} + \nabla \cdot \mathbf{F}_\nu = 4\pi S_\nu - c\kappa'_\nu U_\nu \ , \tag{3}$$

$$\frac{1}{c}\frac{\partial \mathbf{F}_\nu}{\partial t} + c\nabla \cdot \vec{\mathbf{P}}_\nu = -\kappa'_\nu \mathbf{F}_\nu \ , \tag{4}$$

for a spherically symmetric source function S_ν. The first three moments of the radiation intensity in Eqs. (3) and (4) are defined by

$$U_\nu \equiv (1/c)\int d\Omega \, I_\nu(\mathbf{\Omega}) \ , \tag{5}$$

$$\mathbf{F}_\nu \equiv \int d\Omega \, \mathbf{\Omega} \, I_\nu(\mathbf{\Omega}) \ , \tag{6}$$

$$\vec{\mathbf{P}}_\nu \equiv (1/c)\int d\Omega \, \mathbf{\Omega}\,\mathbf{\Omega} \, I_\nu(\mathbf{\Omega}) \ , \tag{7}$$

in terms of the solid angle $d\Omega \equiv \sin\theta \, d\theta \, d\phi$ and the radiation unit vector $\mathbf{\Omega} \equiv \hat{\mathbf{z}}\cos\theta + [\hat{\mathbf{x}}\cos\phi + \hat{\mathbf{y}}\sin\phi]\sin\theta$, with θ and ϕ standard spherical coordinates. The sequence of moments in $\mathbf{\Omega}$ of Eq. (1) does not close since there are always moments through one higher order than there are equations. Therefore, some assumption on either the highest moment or equivalently the form of $I_\nu(\mathbf{\Omega})$ must be made in order to close the system of equations. The closure of the two moment equations (3) and (4) is often expressed in terms of a tensor Eddington factor defined by

$$\vec{\mathbf{f}}_\nu \equiv \vec{\mathbf{P}}_\nu / U_\nu \ . \tag{8}$$

2. Ellipsoidal model for radiation intensity

2.1. Radiative flux reference frame

To close the moment equations (3) and (4) we make the *ad hoc* assumption that the specific intensity of radiation has the angular distribution of an ellipsoid having its origin lying at one of the foci. Mathematically, this has the form

$$I_\nu(\mu) = \frac{(1 - \epsilon_\nu^2)a_\nu}{1 - \epsilon_\nu \mu} \ . \tag{9}$$

with a_ν and ϵ_ν depending upon space $\mathbf{x}$, time t, and spectral frequency ν. The variable $\mu \equiv \cos(\theta)$ is defined so that $\theta = 0$ corresponds to the direction of the spectral flux $\mathbf{F}_\nu$. Since the assumed intensity function is independent of an azimuthal angle ϕ, i.e., $I_\nu = I_\nu(\mu)$, we choose first to define $\mathbf{F}_\nu$ and $\vec{\mathbf{P}}_\nu$ in terms of some coordinate system $(\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}')$ that is arbitrary with the restriction that $\hat{\mathbf{z}}'$ lie in the local $\theta = 0$ direction. The angular integrals in Eqs. (5)-(7) then can be expressed in the form

$$U_\nu = \frac{2\pi}{c} \frac{(1-\epsilon_\nu^2) a_\nu}{\epsilon_\nu} \ln\left(\frac{1+\epsilon_\nu}{1-\epsilon_\nu}\right) , \tag{10}$$

$$\mathbf{F}_\nu = \hat{\mathbf{z}}' \, 2\pi \frac{(1-\epsilon_\nu^2) a_\nu}{\epsilon_\nu^2} \left[\ln\left(\frac{1+\epsilon_\nu}{1-\epsilon_\nu}\right) - 2\epsilon_\nu \right] , \tag{11}$$

$$[\vec{\mathbf{P}}_\nu]_{(\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}')} = \begin{pmatrix} P_\nu & 0 & 0 \\ 0 & P_\nu & 0 \\ 0 & 0 & P_\nu \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 3P_\nu - U_\nu & 0 & 0 \\ 0 & 3P_\nu - U_\nu & 0 \\ 0 & 0 & 0 \end{pmatrix} , \tag{12}$$

in terms of a scalar radiation pressure

$$P_\nu = \frac{2\pi}{c} \frac{(1-\epsilon_\nu^2) a_\nu}{\epsilon_\nu^3} \left[\ln\left(\frac{1+\epsilon_\nu}{1-\epsilon_\nu}\right) - 2\epsilon_\nu \right] . \tag{13}$$

The values of radiant energy density U_ν and the magnitude of the radiative flux $F_\nu \equiv \| \mathbf{F}_\nu \|$ in Eqs. (10) and (11) can be expressed equivalently in terms of a_ν and ϵ_ν. The value of ϵ_ν is found from the ratio of Eqs. (10) and (11), i.e.,

$$\delta_\nu \equiv \frac{F_\nu}{cU_\nu} = \frac{1}{\epsilon_\nu} - \frac{2}{\ln\Big((1+\epsilon_\nu)/(1-\epsilon_\nu)\Big)} = \frac{1}{\epsilon_\nu} - \frac{1}{\operatorname{arctanh}(\epsilon_\nu)} , \tag{14}$$

which is a monotonically-increasing function from zero to unity defined over the same range in ϵ_ν. The value of a_ν then determines the effective magnitude of the radiation intensity via Eq. (10). Thus a_ν and ϵ_ν can be found easily from values of U_ν and F_ν.

2.2. Simulation reference frame

The only aspect remaining to be determined is the rotation from the coordinates $(\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}')$ of the spectral flux to those of the simulation $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$. This is required in order to find an expression for the radiation stress tensor $[\vec{\mathbf{P}}_\nu]_{(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})}$ in the simulation reference frame in terms of $[\vec{\mathbf{P}}_\nu]_{(\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}')}$ of Eq. (12). The radiative flux in the simulation frame is found using the standard Euler angles:

$$[\mathbf{F}_\nu]_{(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})} = \vec{\mathbf{R}}_y(\theta) \, \vec{\mathbf{R}}_z(\phi) \, [\mathbf{F}_\nu]_{(\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}')} \tag{15}$$

with the rotation matrices $\vec{\mathbf{R}}_y(\theta)$ and $\vec{\mathbf{R}}_z(\phi)$ defined by

$$\vec{\mathbf{R}}_y(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} , \quad \vec{\mathbf{R}}_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{16}$$

The angles θ and ϕ are determined by the angle of the spectral flux in the simulation frame, i.e., $F_\nu = F_{\nu,x}\hat{\mathbf{x}} + F_{\nu,y}\hat{\mathbf{y}} + F_{\nu,z}\hat{\mathbf{z}}$ via

$$\theta \equiv \arccos\left(\frac{F_{\nu,z}}{\sqrt{F_{\nu,x}^2 + F_{\nu,y}^2 + F_{\nu,z}^2}} \right) , \quad \phi \equiv \arctan\left(\frac{F_{\nu,y}}{F_{\nu,x}} \right) . \tag{17}$$

The transformation of the radiative tensor in the primed frame $(\hat{x}', \hat{y}', \hat{z}')$ to the simulation frame $(\hat{x}, \hat{y}, \hat{z})$ is found via

$$[\vec{\mathbf{P}}_\nu]_{(\hat{x},\hat{y},\hat{z})} = \vec{\mathbf{R}}_y(\theta)\,\vec{\mathbf{R}}_z(\phi)\,[\vec{\mathbf{P}}_\nu]_{(\hat{x}',\hat{y}',\hat{z}')}\vec{\mathbf{R}}_z^T(\phi)\vec{\mathbf{R}}_y^T(\theta)\,, \tag{18}$$

with the superscript T denoting the matrix transpose. Using Eq. (18) for $\vec{\mathbf{P}}_\nu$ then closes the moment equations (3) and (4).

For spatial regions in which the intensity is nearly isotropic, i.e., $\epsilon_\nu, \delta_\nu \ll 1$, the tensor Eddington factor of Eq. (8) reduces to the optically thick limit with $\vec{\mathbf{f}}_\nu = \vec{\mathbf{I}}/3$. Then $\nabla\cdot\vec{\mathbf{P}}_\nu$ in Eqs. (3) and (4) becomes $\nabla U_\nu/3$, yielding the standard radiative diffusion equations. Equations (3) and (4) also give the correct result for the streaming limit in an optically thin plasma. That is, if all photons are traveling in a single direction, then $\nabla\cdot\vec{\mathbf{P}}_\nu$ becomes ∇U_ν.

3. Discussion

We have modified the radiative diffusion subroutine within a multigroup hydrodynamics code to apply for a tensor Eddington factor given by $\vec{\mathbf{f}}_\nu \equiv [\vec{\mathbf{P}}_\nu]_{(\hat{x},\hat{y},\hat{z})}/U_\nu$. There is very little difference between 1D runs using our new Eddington tensor compared with that of ordinary diffusion theory.

More dramatic differences are expected between the tensor and scalar Eddington factors for 2D and 3D simulations, since this tensor Eddington factor allows for "shadows" that are impossible to obtain with ordinary diffusion theory. We are currently investigating the effect of the shadows on the ablative Rayleigh-Taylor instability in the nonlinear regime. This formulation will not only improve current LTE modeling of laser matter interactions, but it will also be a valid extension of radiative diffusion theory when non-LTE effects are taken into account using the method of Busquet [4].

Acknowledgments

Work supported by the U.S. Department of Energy and the Office of Naval Research. The authors are grateful to Jill Dahlburg for useful suggestions.

References

[1] Zel'dovich Y B and Raizer Y P 1966 *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (New York: Academic) pp 107-72

[2] Minerbo G N 1978 *J. Quant. Spectrosc. Radiat. Transfer* **20** 541-5

[3] Mihalas D and Mihalas B W 1984 *Foundations of Radiation Hydrodynamics* (New York: Oxford) pp 311-6

[4] Busquet M 1993 *Phys Fluids* B **5** 4191

Appendix Q

Variable Eddington Radiative Transport

VARIABLE EDDINGTON RADIATIVE TRANSPORT

DAVID A. GARREN
Science Applications International Corp., McLean, VA 22102

MARCEL KLAPISCH
ARTEP, Inc., Columbia, MD 21045

DENIS COLOMBANT, JOHN GARDNER
Naval Research Laboratory, Washington, DC 20375

Abstract

- A multi-group hydrodynamics code is used to compute the evolution of a planar aluminum target irradiated by a $1/4\,\mu$ pulse.

- For diagnostic purposes, it is desirable to obain a much more detailed spectrum than that produced by the multi-group hydro code.

- We have developed a post-processor that uses the temperature and density profiles from the hydro code to compute a detailed emission spectrum based on STA opacities.

- We apply these results to examine the validity of the variable Eddington approximation that is used to determine the transfer of radiation within the hydro code.

Radiative Transfer Equation

- Radiative Transfer Equation for an LTE plasma:

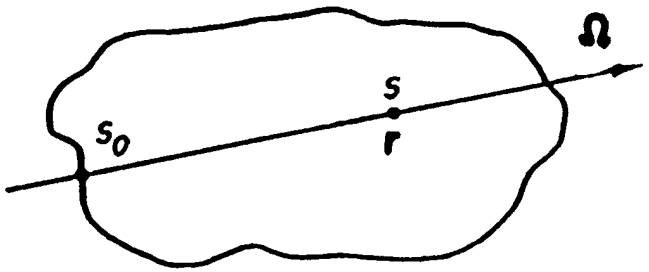
$$\frac{dI_\nu}{ds} = \left[\frac{1}{c}\frac{\partial}{\partial t} + \mathbf{\Omega}\cdot\nabla\right] I_\nu = \kappa'_\nu\left[I_{\nu p} - I_\nu\right]$$

  - $I_\nu(\mathbf{r}, \mathbf{\Omega}, t)$ = specific intensity of radiation
  - $\mathbf{\Omega}$ = propagation direction of radiation
  - s = path length along $\mathbf{\Omega}$
  - $I_{\nu p}(T) = (2h\nu^3/c^2)[\exp(h\nu/kT) - 1]^{-1}$ = Planck function.
  - $\kappa'_\nu \equiv \kappa_\nu\left[1 - \exp(-h\nu/kT)\right]$ = effective absorption coefficient
  - κ_ν = absorption coefficent

General Solution of Radiative Transfer Equation

- General solution can be found via ray-tracing :

$$I_\nu(s) = \int_{s_0}^{s} \exp\Big(-\int_{s'}^{s} \kappa'_\nu(s')\,ds''\Big)\, I_{\nu p}\big(T(s)\big)\, \kappa'_\nu(s)\,ds'$$
$$+ I_{\nu 0} \exp\Big(-\int_{s_0}^{s} \kappa'_\nu(s')\,ds''\Big)$$



- ◇ $I_{\nu 0}$ = radiation intensity at the boundary S_0
- ◇ This exact solution can be generated for any plasma
- ◇ However, it is numerically prohibitive $\Longrightarrow$ other approach ???

Moments of Radiative Transfer Equation

- Solution also can be found via moments in $\mathbf{\Omega}$:

$$\frac{\partial U_\nu}{\partial t} + \nabla \cdot \mathbf{F}_\nu = \kappa' \left[4\pi I_{\nu p} - c U_\nu \right]$$

$$\frac{1}{c}\frac{\partial \mathbf{F}_\nu}{\partial t} + c\nabla \cdot \vec{\mathbf{P}}_\nu = \kappa'_\nu \mathbf{F}_\nu$$

$$\vdots$$

  - $U_\nu \equiv (1/c) \int d^2\Omega \, I(\mathbf{\Omega}) =$ radiant energy density
  - $\mathbf{F}_\nu \equiv \int d^2\Omega \, \mathbf{\Omega} \, I(\mathbf{\Omega}) =$ radiant energy flux
  - $\vec{\mathbf{P}}_\nu \equiv (1/c) \int d^2\Omega \, \mathbf{\Omega}\,\mathbf{\Omega} \, I(\mathbf{\Omega}) =$ radiant energy stress tensor

Variable Eddington Approximation

- Define variable tensor Eddington factor: $\vec{\mathbf{f}}_\nu \equiv \vec{\mathbf{P}}_\nu / U_\nu$
  - $\vec{\mathbf{f}}_\nu$ depends on $I_\nu(\mathbf{\Omega})$ profile
  - Value for $\vec{\mathbf{f}}_\nu$ closes first two moment eqns.
- Variable scalar Eddington factor in planar geometry:

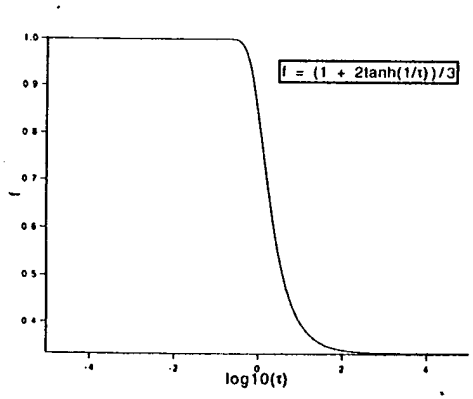
$$\nabla \cdot \vec{\mathbf{P}}_\nu = \nabla \cdot \left[\vec{\mathbf{f}}_\nu U_\nu\right] \rightarrow \nabla \left[f_\nu U_\nu\right]$$

$$f_\nu \equiv \frac{\int_{-1}^{1} I_\nu \mu^2 \, d\mu}{\int_{-1}^{1} I_\nu \, d\mu}$$

  - $\mu \equiv \cos(\theta)$, with $\theta =$ angle w.r.t plane normal.

Specific Values for Scalar Eddington

- Optically Thick Plasmas: Standard Diffusion Theory
  - ◇ Isotropic dist. $I_\nu(\mu) \to I_\nu \implies f_\nu \to 1/3$
- Optically Thin Plasmas:
  - ◇ Streaming dist. $I_\nu(\mu) \to I_\nu\, \delta(\mu \pm 1) \implies f_\nu \to 1.$
  - ◇ Thin slab dist. $I_\nu(\mu) \to I_\nu\, \delta(\mu) \implies f_\nu \to 0.$
- Mixed Thick/Thin ???
  - ◇ *Ad hoc* Eddington: $f_\nu = (1 + 2\tanh(1/\tau_\nu))/3$



Spectral Surface Brightness

- Radiant Energy Flux from a Surface:

$$S_\nu \equiv \int_{hemisphere} \cos(\theta)\, I_\nu(\mathbf{\Omega})\, d\mathbf{\Omega}$$

- For planar geometry, S_ν can be found analytically:

$$S_\nu = 2\pi \int_0^d E_2\big(\tau_\nu(z)\big)\, I_{\nu p}\big(T(z)\big)\, \kappa'_\nu(z)\, dz$$

  - $\tau_\nu(z) \equiv \int_0^z \kappa'_\nu(z')\, dz' =$ optical depth.
  - $E_2(x) = \int_1^\infty w^{-2} \exp(-xw)\, dw =$ second exponential integral.

Post-Processing

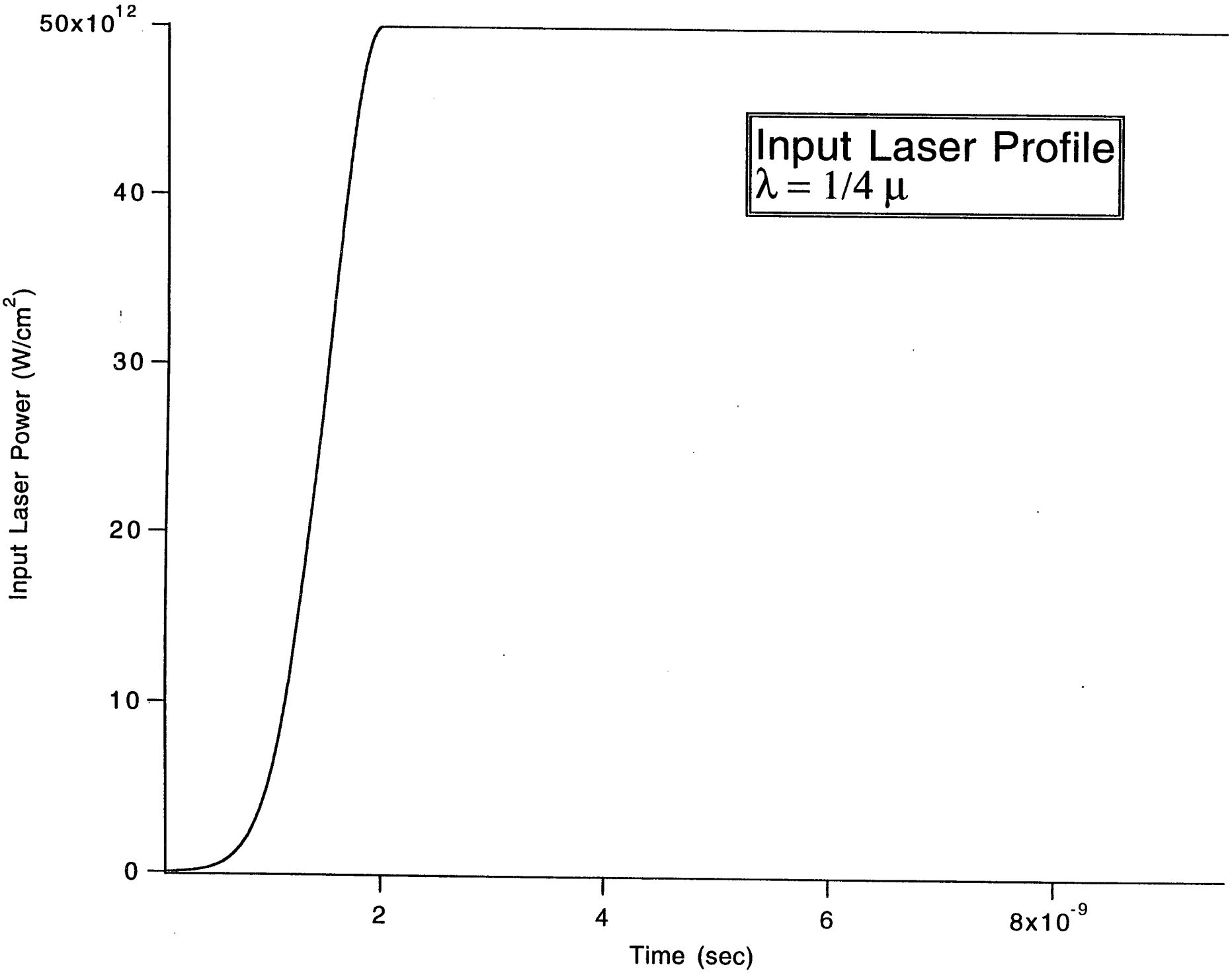
- We ran the hydrocode RAD2D using group-averaged STA opacities to simulate an aluminum solid target irradiated with a KrF (0.248μ) laser.

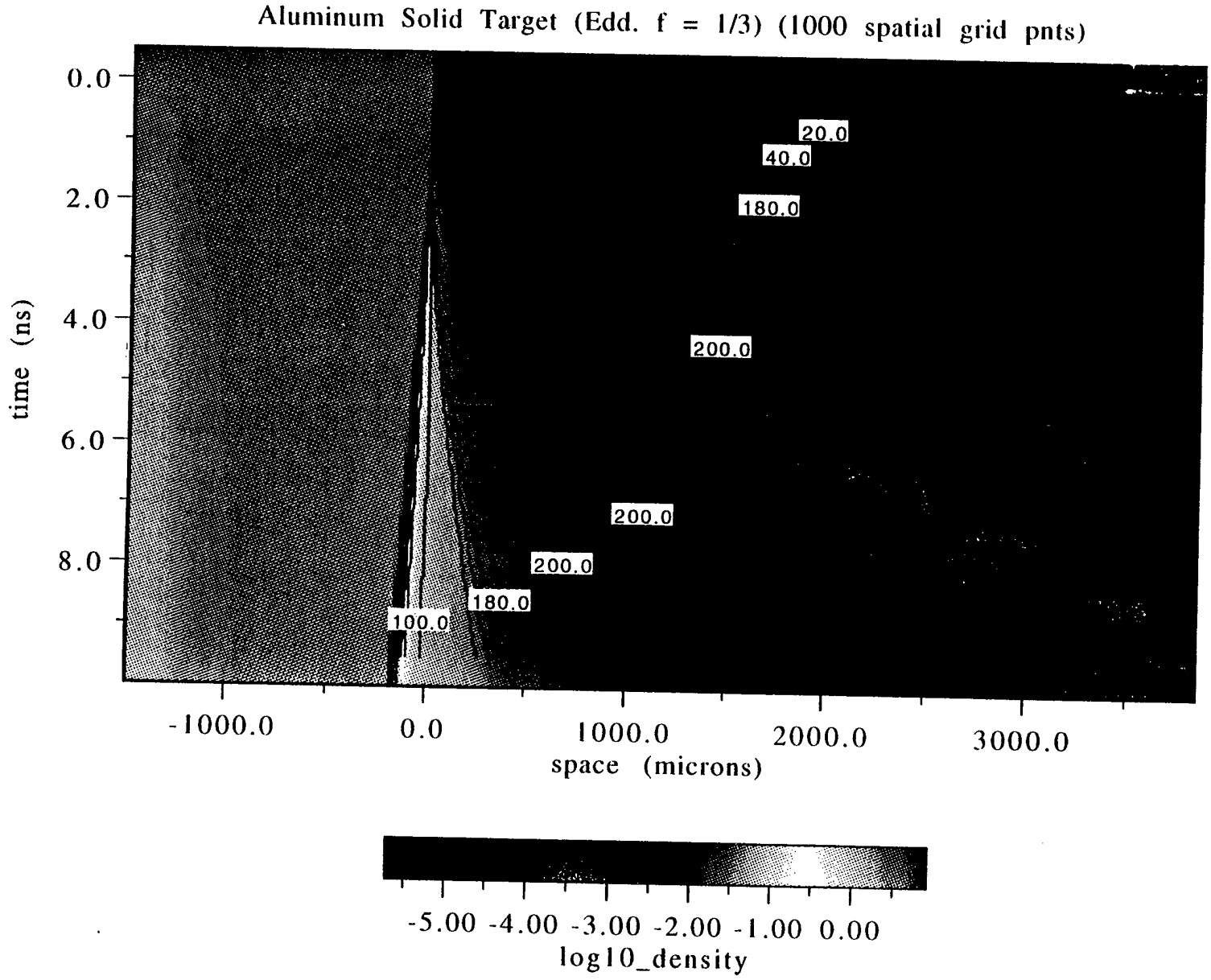
- The resulting 1-D temperature and density profiles as a function of position and time are dumped to the post-processor.

- We use the exact analytical expression for S_ν to calculate the detailed spectrum emitted by the irradiated target.

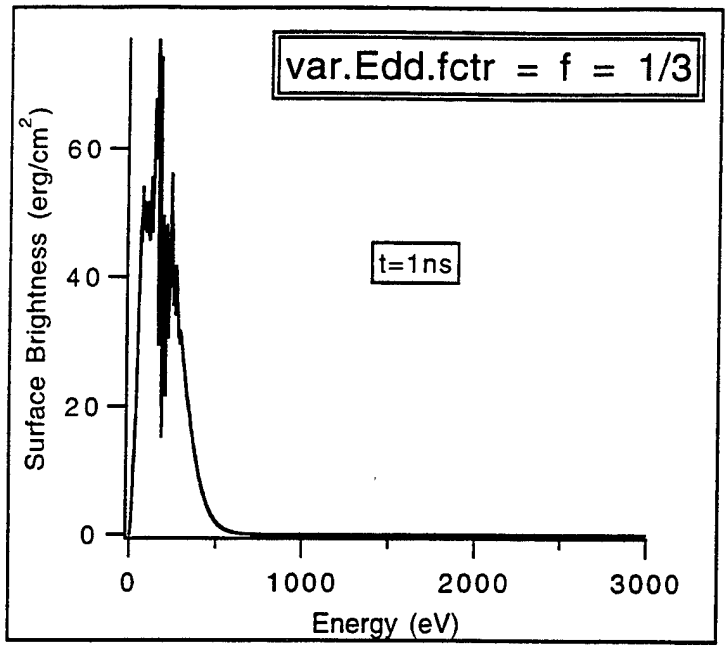
- The detailed spectra of the STA-generated absorption coefficients κ_ν are in used in finding S_ν.

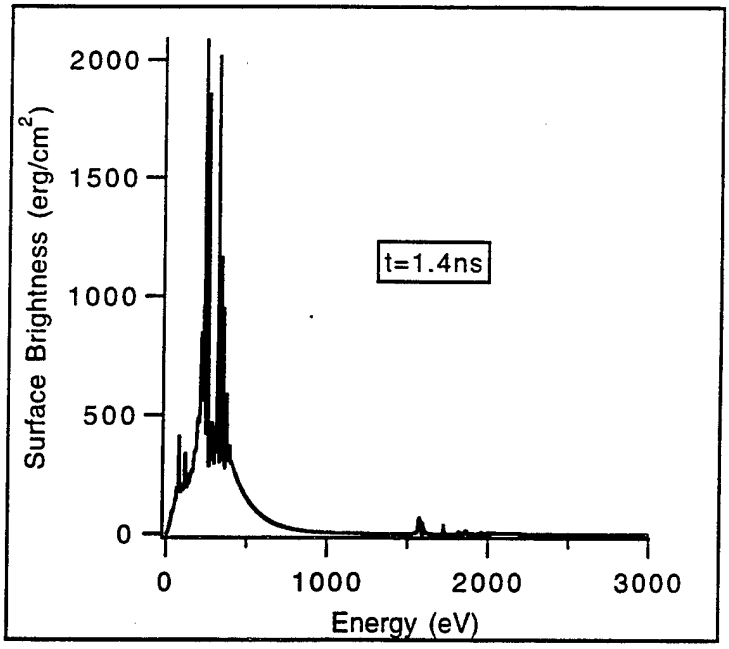
- The resulting post-processed value of S_ν is compared with the hydrodynamics accounting, providing a means of analyzing the validity of approximate Eddington factors.

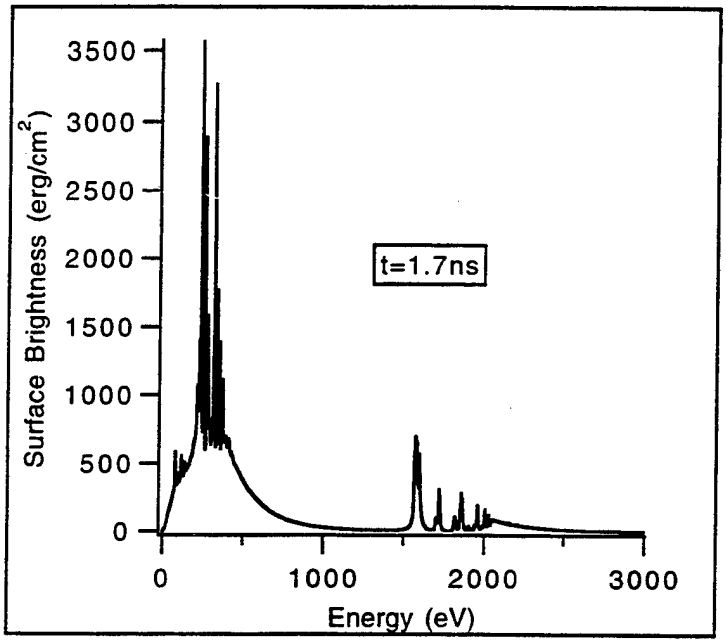

Input Laser Profile
λ = 1/4 μ
50x10^12
40
30
20
10
0
Input Laser Power (W/cm^2)
2
4
6
8x10^-9
Time (sec)

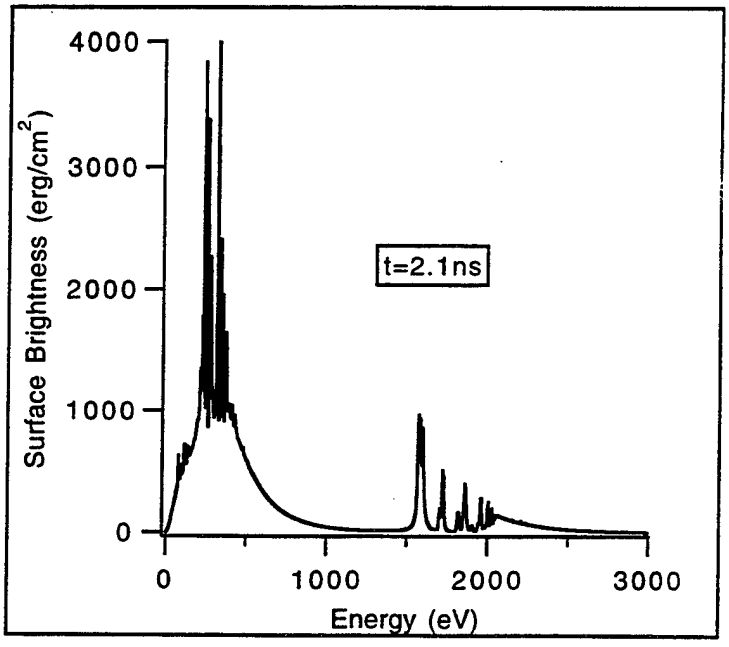


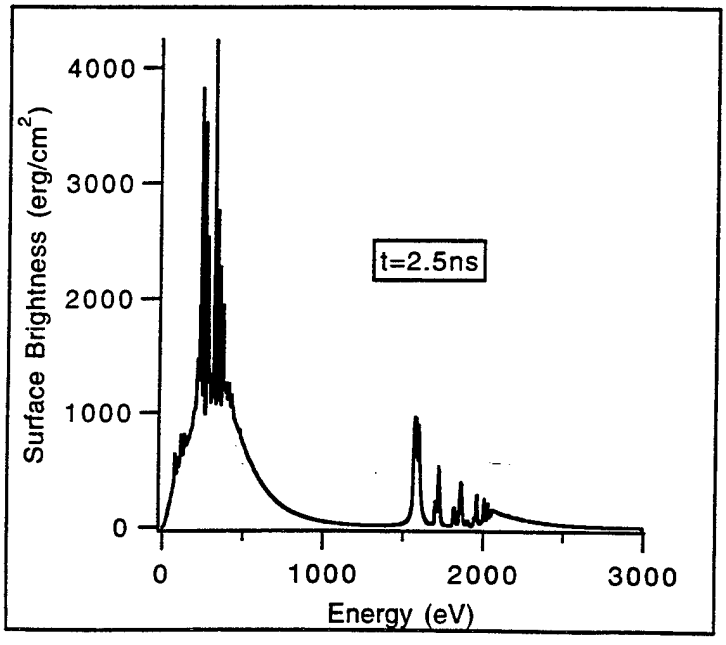
avgr_d vs. (row, col)

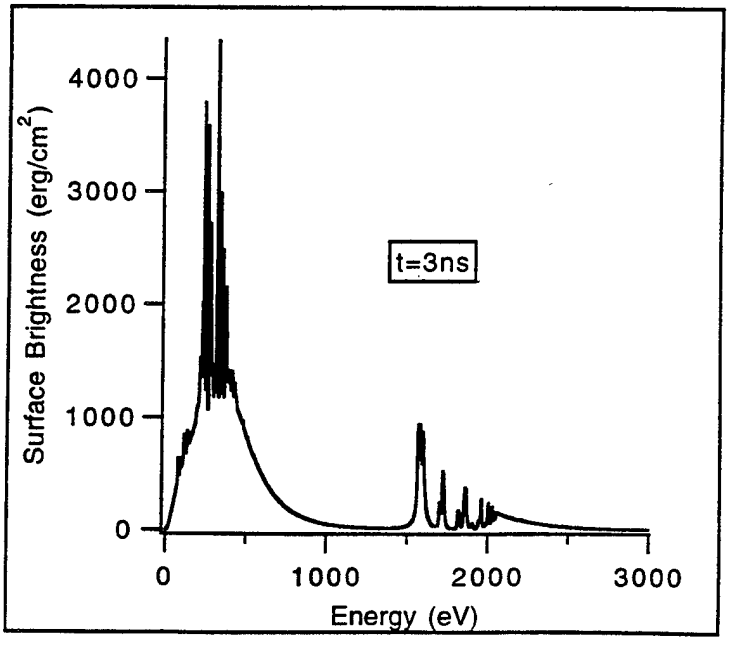

var.Edd.fctr = f = 1/3
t=1ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
20
40
60
0
1000
2000
3000

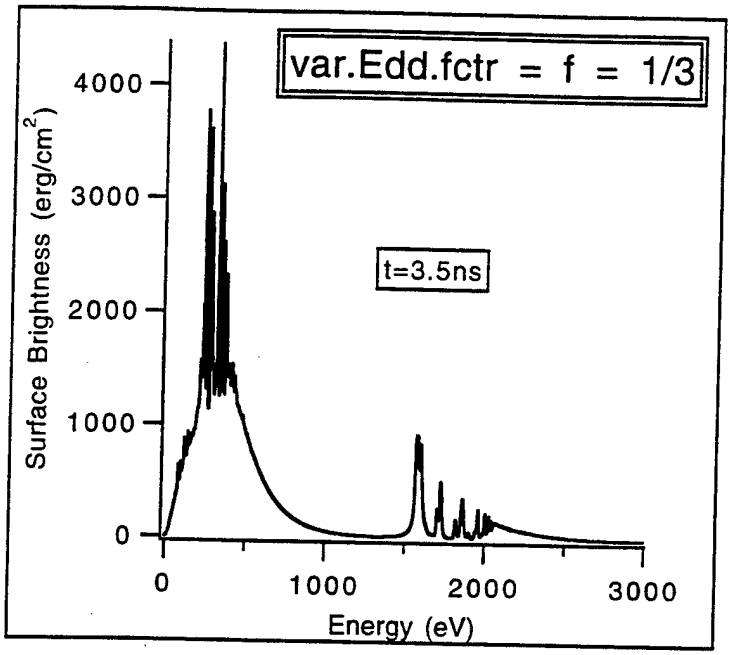

t=1.4ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
500
1000
1500
2000
0
1000
2000
3000

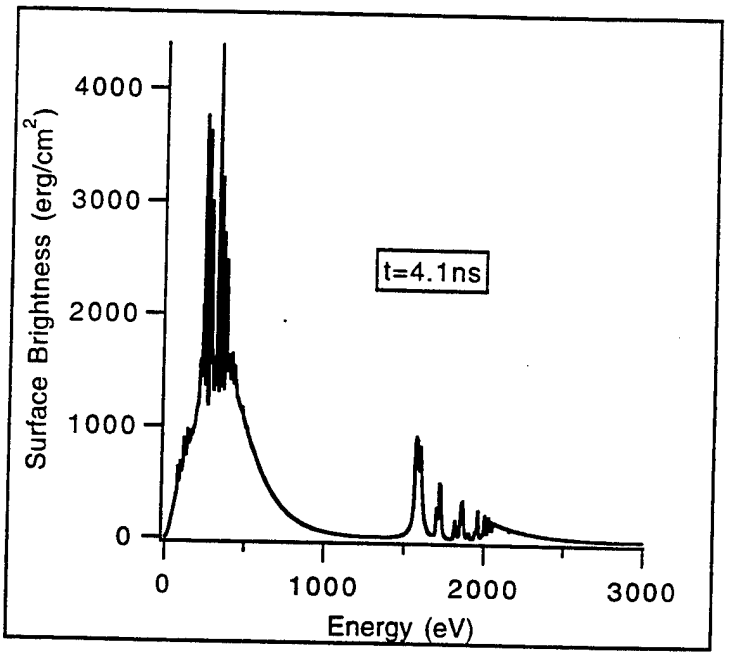

t=1.7ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
500
1000
1500
2000
2500
3000
3500
0
1000
2000
3000

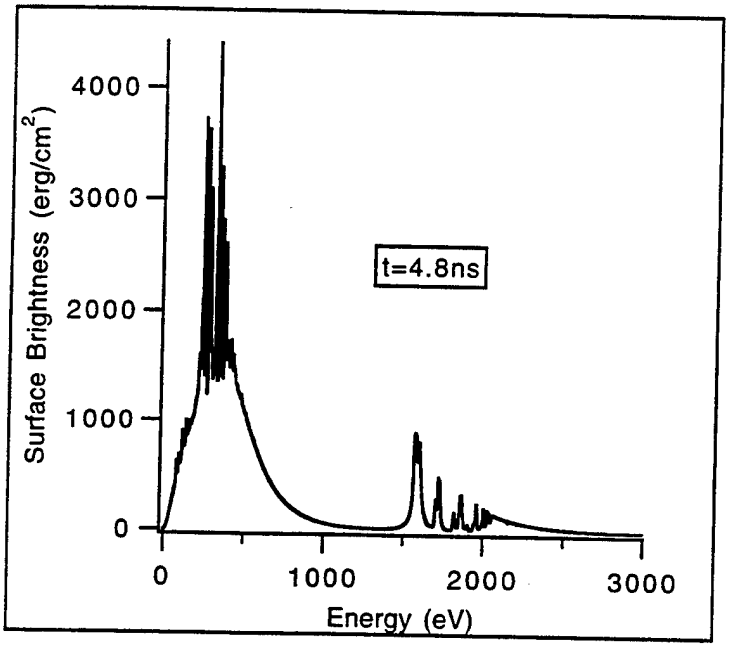

t=2.1ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000
0
1000
2000
3000

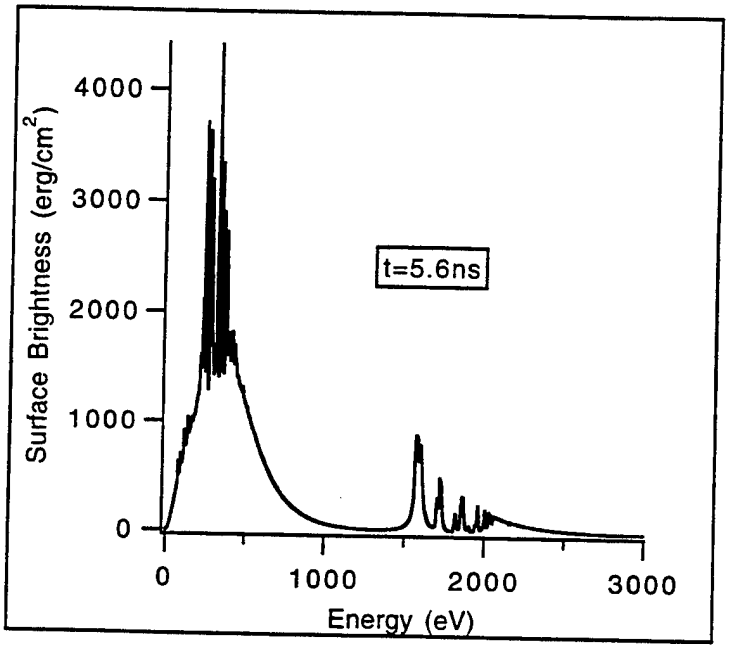

t=2.5ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000
0
1000
2000
3000

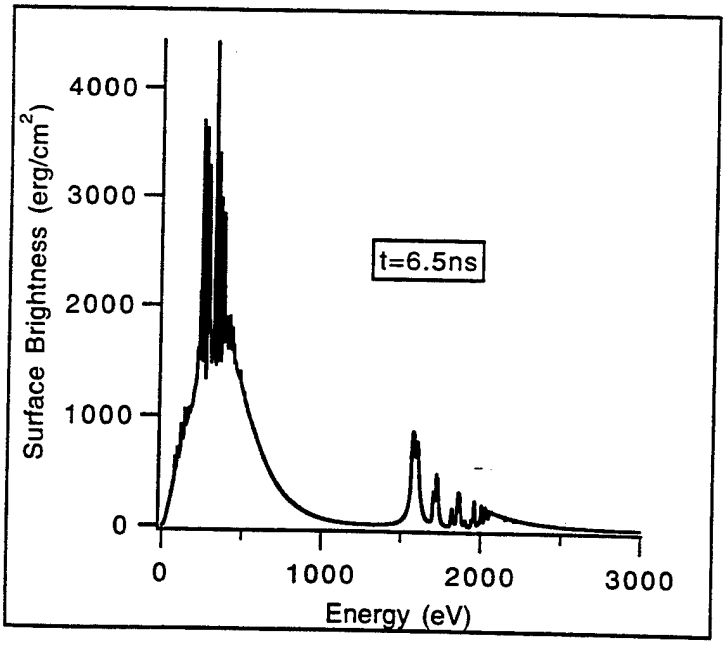

t=3ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000
0
1000
2000
3000

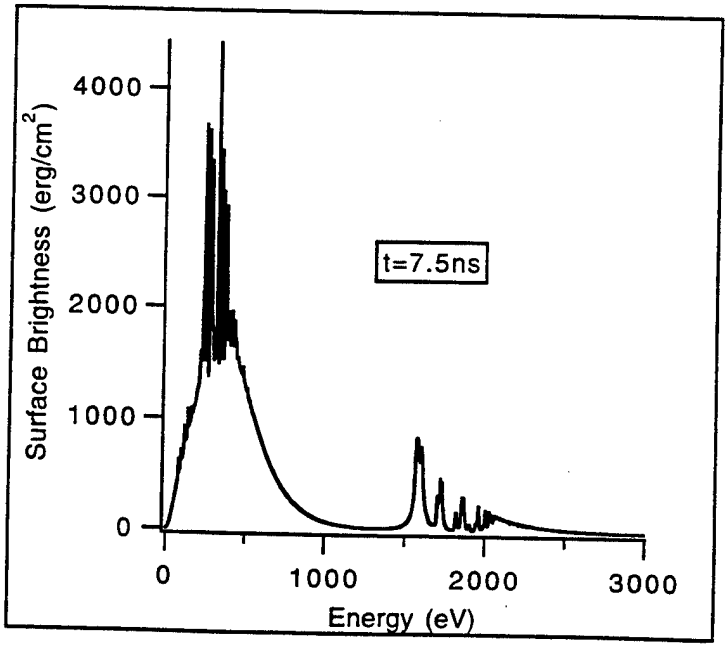

var.Edd.fctr = f = 1/3
t=3.5ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

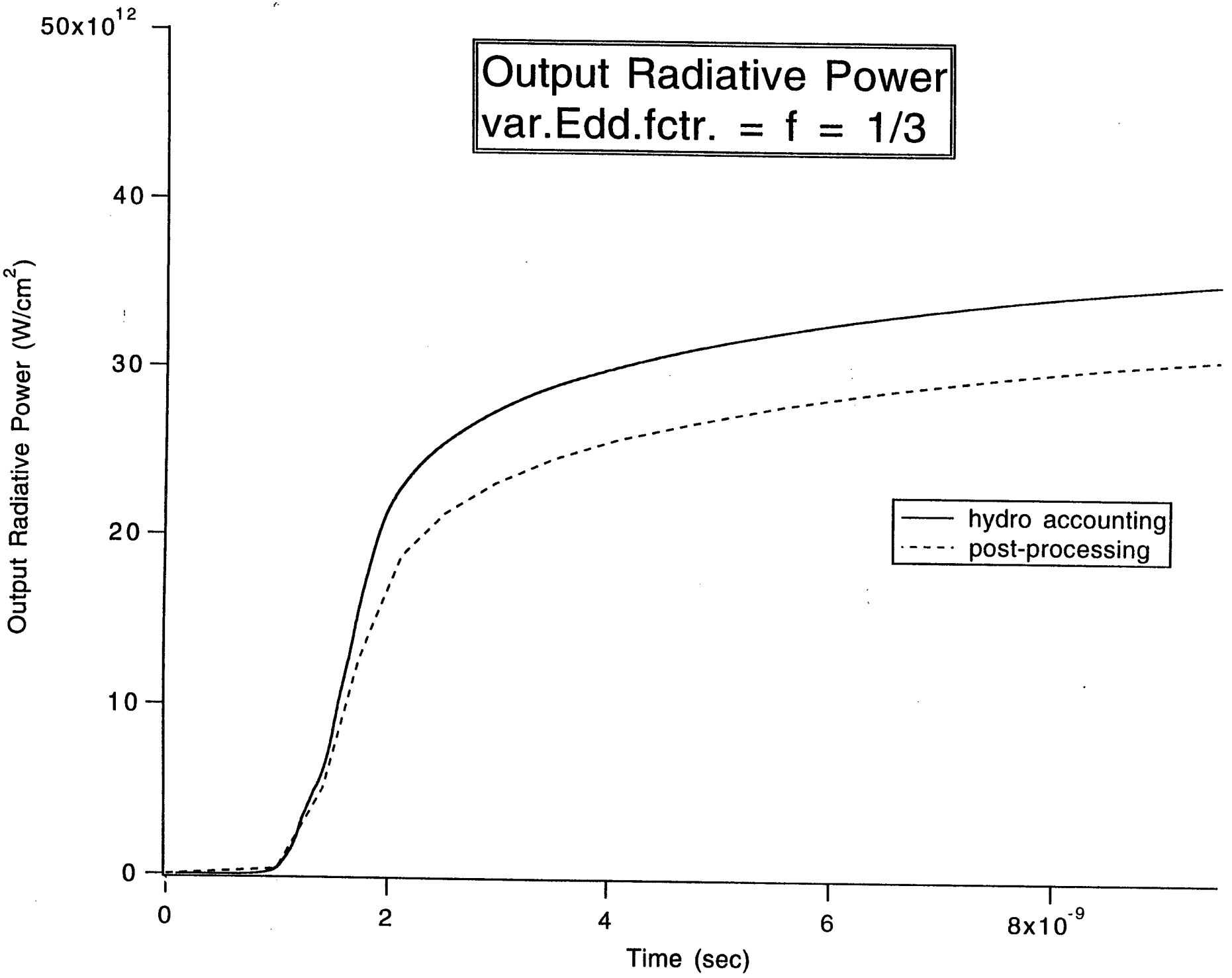

t=4.1ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

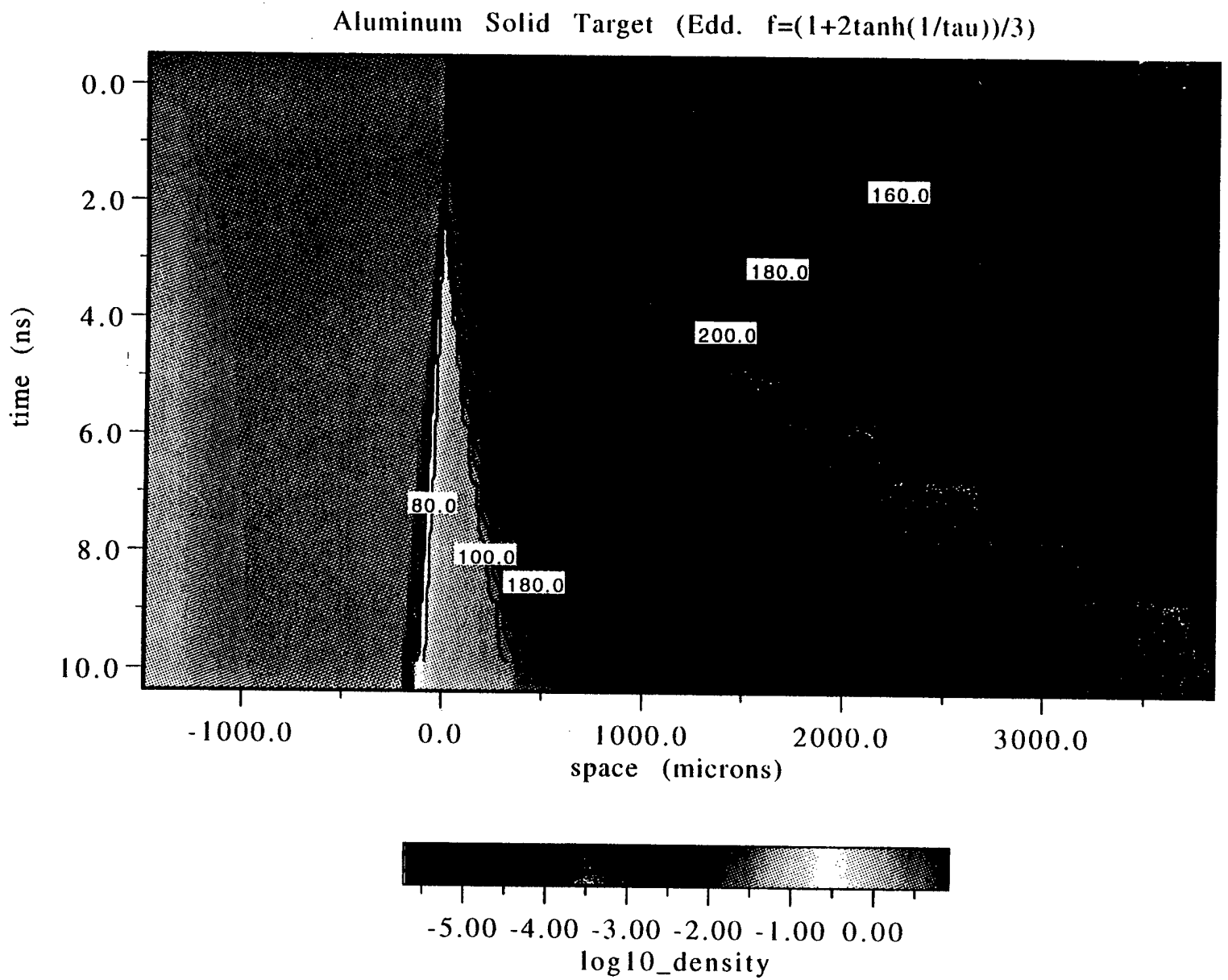

t=4.8ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000


t=5.6ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

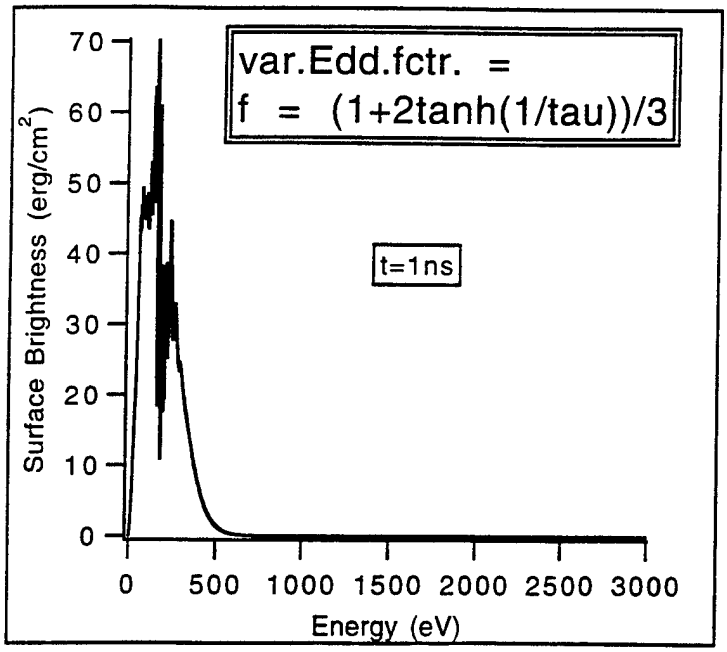

t=6.5ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

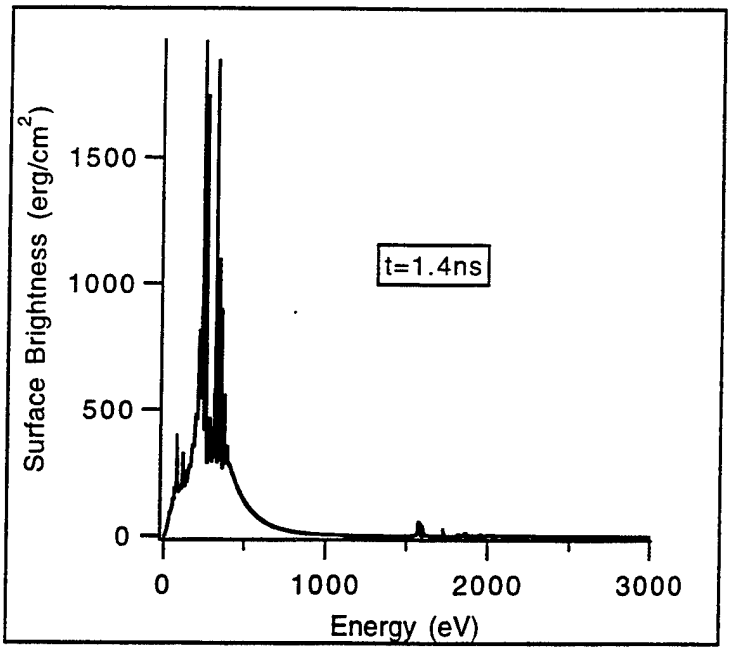

t=7.5ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

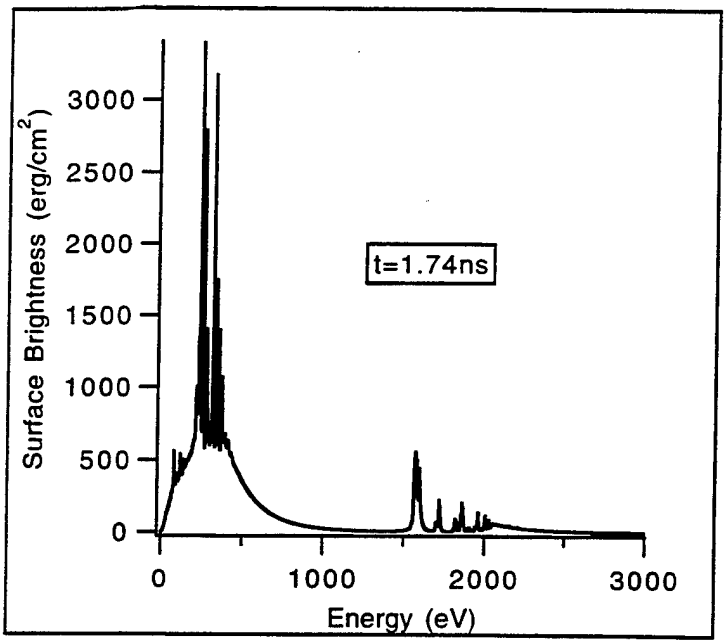

Output Radiative Power
var.Edd.fctr. = f = 1/3
50x10^12
40
30
20
10
0
Output Radiative Power (W/cm^2)
0
2
4
6
8x10^-9
Time (sec)
hydro accounting
post-processing

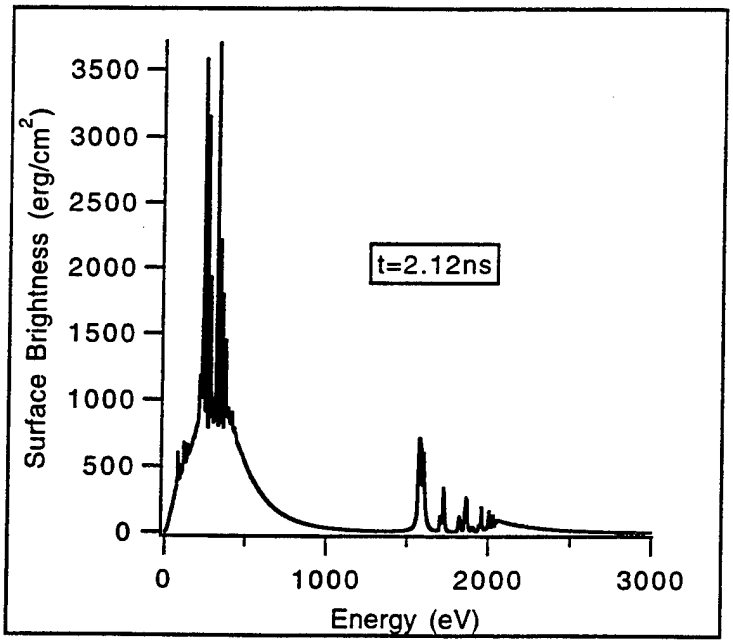


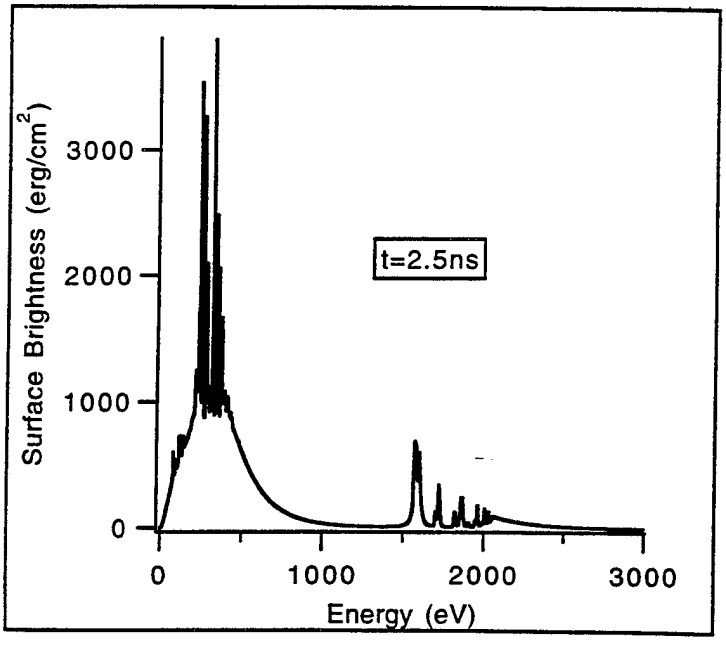
avgr_d vs. (row, col)

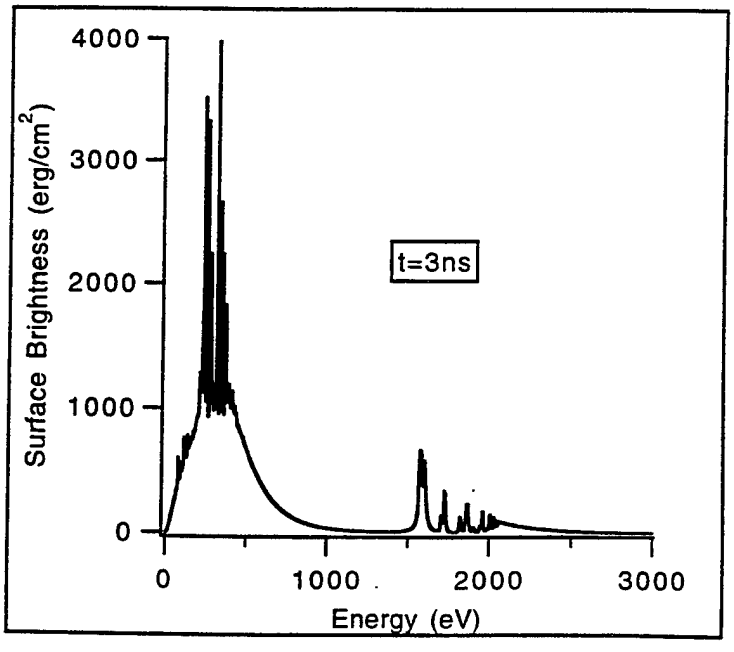

var.Edd.fctr. =
f = (1+2tanh(1/tau))/3
t=1ns
Surface Brightness (erg/cm^2)
Energy (eV)

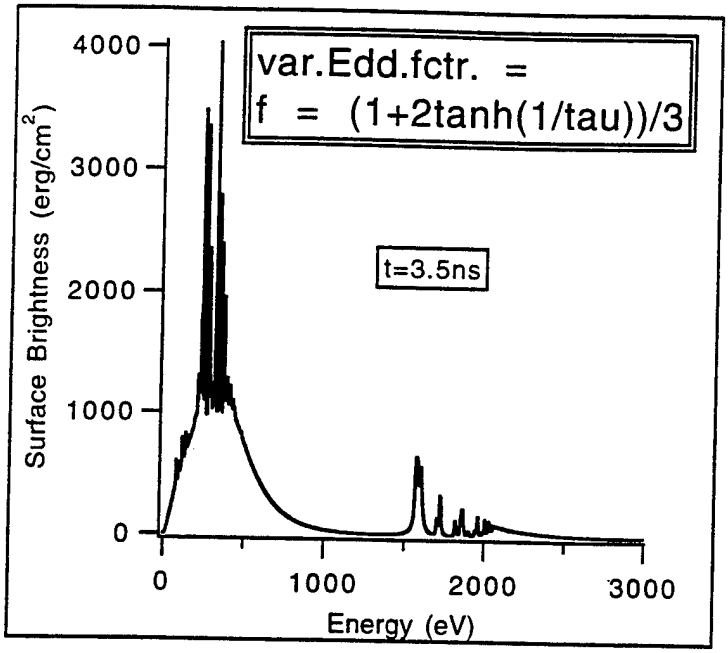

t=1.4ns
Surface Brightness (erg/cm^2)
Energy (eV)

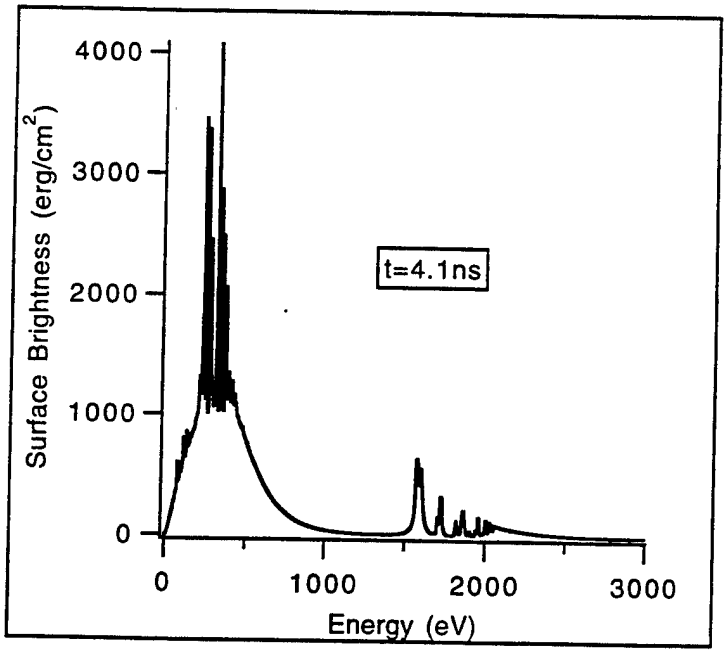

t=1.74ns
Surface Brightness (erg/cm^2)
Energy (eV)

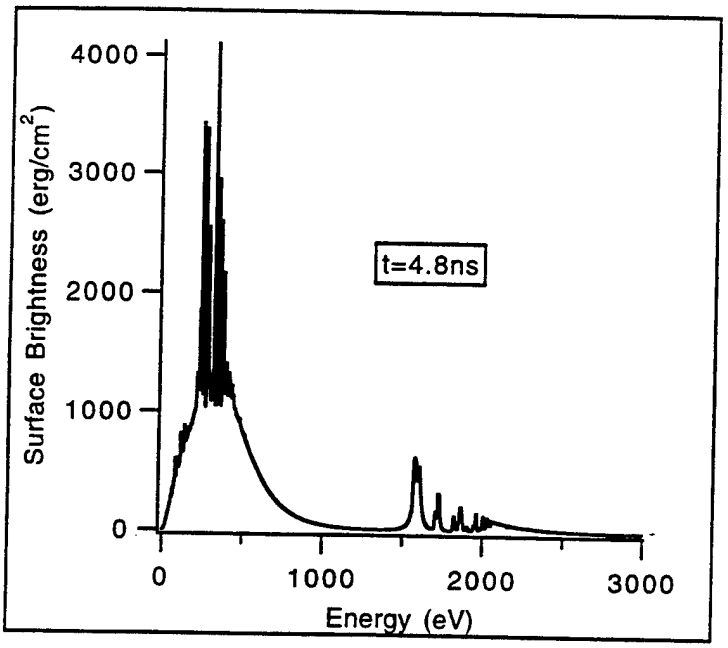

t=2.12ns
Surface Brightness (erg/cm^2)
Energy (eV)

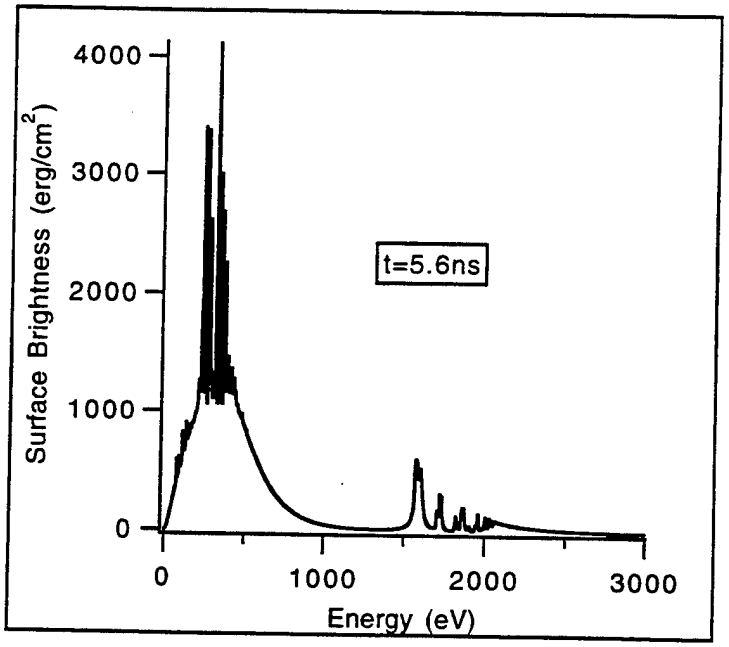

t=2.5ns
Surface Brightness (erg/cm^2)
Energy (eV)

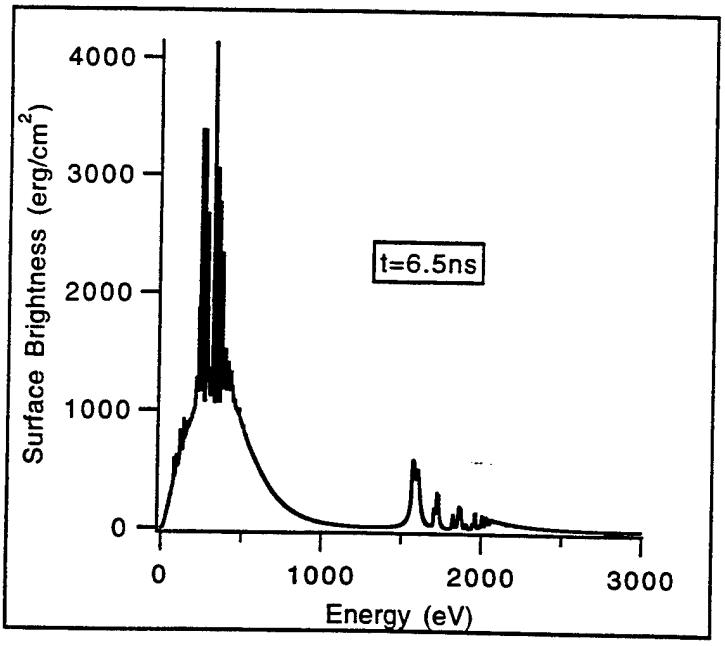

t=3ns
Surface Brightness (erg/cm^2)
Energy (eV)

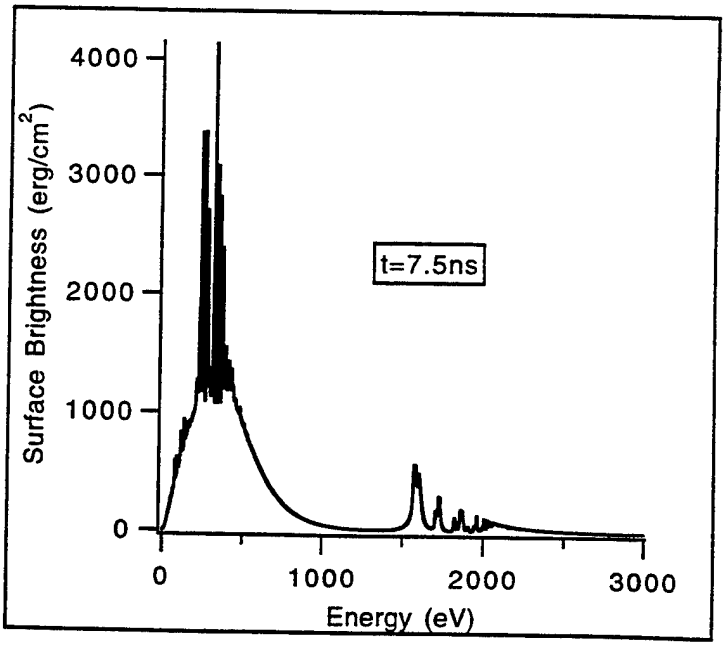

var.Edd.fctr. =
f = (1+2tanh(1/tau))/3
t=3.5ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

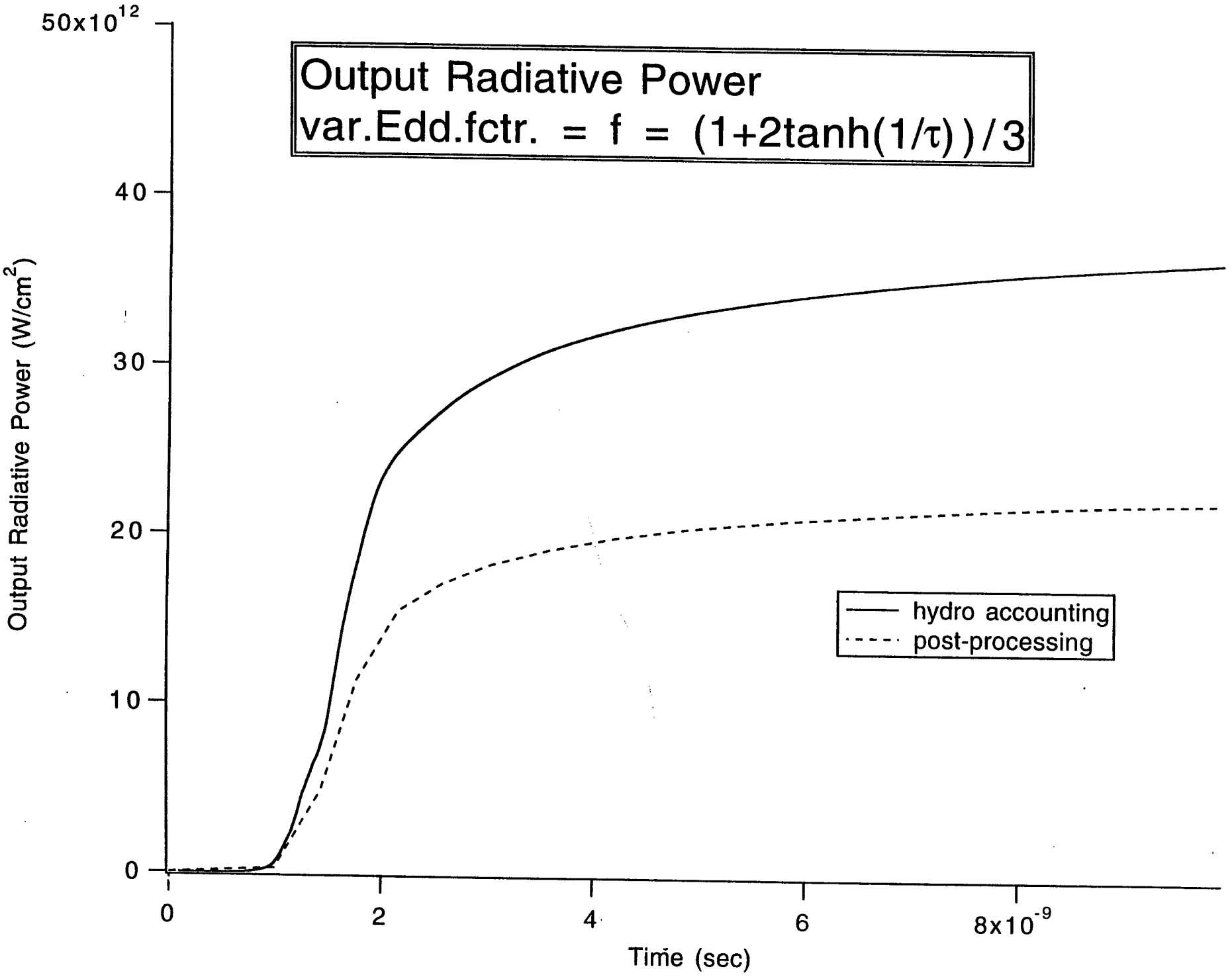

t=4.1ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

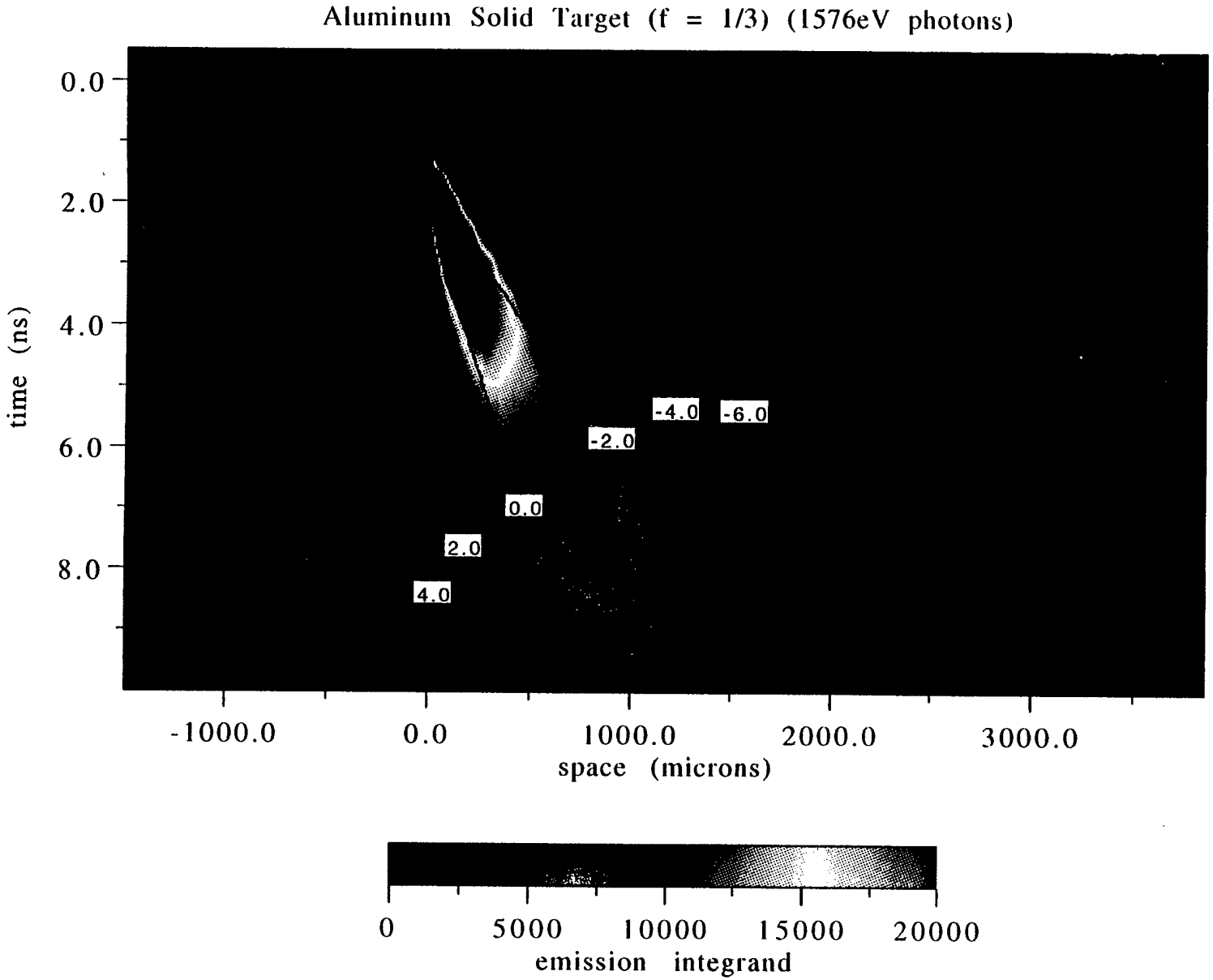

t=4.8ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000


t=5.6ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000

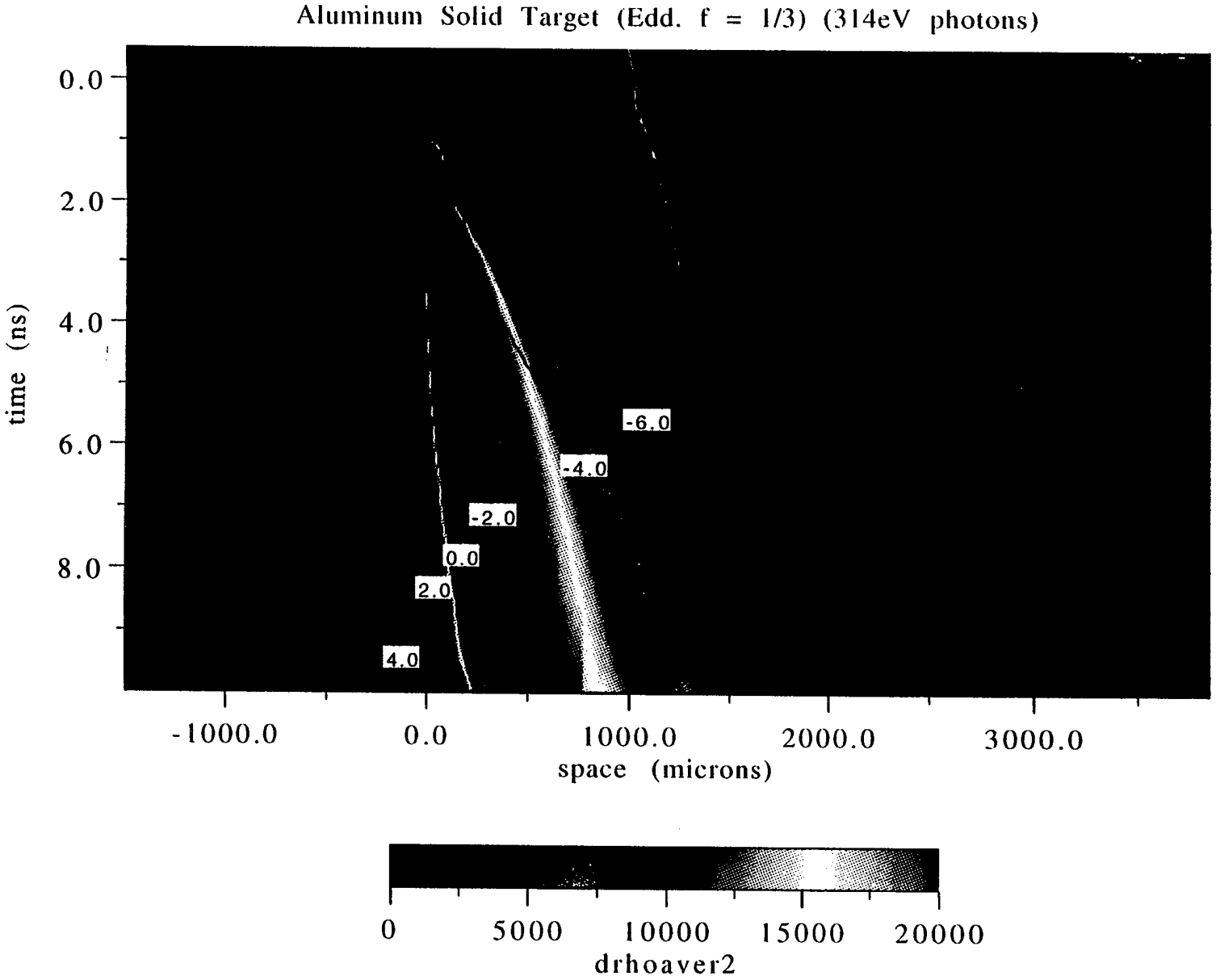

t=6.5ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000


t=7.5ns
Surface Brightness (erg/cm^2)
Energy (eV)
0
1000
2000
3000
4000


Output Radiative Power
var.Edd.fctr. = f = (1+2tanh(1/τ))/3
50x10^12
40
30
20
10
0
Output Radiative Power (W/cm^2)
0
2
4
6
8x10^-9
Time (sec)
hydro accounting
post-processing



drhoaver2 vs. (row, col)



drhoaver2 vs. (row, col)

Summary

- We obtain high-accuracy emission spectra by dumping the temperature and density profiles obtained from a 1-D hydro code to a post-processor that uses STA opacities.

- The standard diffusion model ($f_\nu = 1/3$) for the radiation transport of optically thick plasmas is compared with an *ad hoc* Eddington factor ($f_\nu = (1 + 2\tanh(1/\tau_\nu))/3$) that attempts to model the plasma in both optically thick and optically thin regions.

- The post-processor is more consistent with the hydrodynamics accounting if $f = 1/3$ rather than $f_\nu = (1 + 2\tanh(1/\tau_\nu))/3$ is used, even though the hydro simulation has regions of both thick and thin optical depth.

- It is surprising that the simpler model has such a broad range of validity.

Appendix R

Electronmagnetohydrodynamic Response of a Plasma to an External Current Pulse

Electronmagnetohydrodynamic response of a plasma to an external current pulse

H. B. Zhou,[a] K. Papadopoulos, and A. S. Sharma
University of Maryland, College Park, Maryland 20742

C. L. Chang
Science Applications International Corp., McLean, Virginia 22101

(Received 30 June 1995; accepted 1 February 1996)

In this paper we examine the dynamic response of a magnetoplasma to an external time-dependent current source in the context of electronmagnetohydrodynamics (EMHD). A combined analytic and numerical technique is developed to address this problem. The set of cold electron plasma and Maxwell's equations are first solved analytically in the $(\mathbf{k},\omega)$ space. Inverse Laplace and three-dimensional complex Fast Fourier Transform techniques are used subsequently to numerically transform the radiation fields and plasma currents from the $(\mathbf{k},\omega)$ space to the $(\mathbf{r},t)$ space. The results show that the electron plasma responds to a time-varying current source imposed across the magnetic field by exciting whistler/helicon waves and forming an expanding local current loop, driven by field-aligned plasma currents. The current loop consists of two antiparallel field-aligned current channels concentrated at the ends of the imposed current and a cross-field Hall current region connecting these channels. The characteristics of the current closure region are determined by the background plasma density, the magnetic field, and the time scale of the current source. The results are applied to the ionospheric generation of extremely low-frequency (ELF) and very low-frequency (VLF) radiation using amplitude modulated high-frequency heating. It is found that contrary to previous suggestions the dominant radiating moment of the ELF/VLF ionospheric source is an equivalent horizontal magnetic dipole. [S1070-664X(96)01805-4]

I. INTRODUCTION

The transient response of a magnetoplasma to externally imposed stationary or moving current pulses is a fundamental plasma physics problem with a wide range of applications. The formal solution of this problem for a cold, collisionless, isotropic plasma is well known.[1] Kuehl[2] used the far-zone dyadic Green's function to calculate the radiation of an electric dipole in a two-dimensional (2-D) magnetoplasma. A solution was found only for the case when the frequency ω is much higher than the electron plasma frequency ω_e and the electron cyclotron frequency Ω_e, viz. $\omega/\omega_e \gg 1$ and $\omega/\Omega_e \gg 1$. For low frequencies, $\omega < \Omega_e$, the integral equation reduces to a transcendental equation that cannot be solved analytically. Vidmar[3] used a saddle point method to study the delta function excitation of waves in the Earth's ionosphere in one dimension. He found an asymptotic solution, valid for the far-zone field and long after the source turn-on. Transient effects were lost through the use of the saddle point method. Furthermore, saddle point methods are very difficult to use[4] for multidimensional cases. Because of the mathematical difficulty in calculating the integrals analytically in two or three dimensions, numerical solutions are required.

The objective of this paper is to study the transient response of a magnetoplasma to a current pulse, in the parameter range where the ratio of the electron cyclotron to electron neutral collision frequency (ν_e) is larger than unity, while the ratio of the ion cyclotron to the ion neutral collision frequency (ν_i) is smaller than unity. This is the case for the plasma in the D and E regions of the ionosphere between 70–130 km altitude range. At high latitudes this region is penetrated by electric fields and currents. As a result, transient precipitation events that cause local conductivity modifications induce current pulses, to which the plasma responds by generating electromagnetic (EM) waves and currents. Of particular interest is the plasma response to periodic heating of the D region by ionospheric high-frequency (HF) radio wave heaters. This phenomenon is extremely important, since it leads to the generation of low-frequency EM waves in the ultra low-frequency (ULF), extremely low-frequency (ULF), and very low-frequency (VLF) ranges.[5–8] A new technique that combines analytic and numerical methods is developed in this paper. The electron plasma and Maxwell's equations are solved analytically in Fourier space; then inverse complex fast Fourier transform (FFT) technique is used to transform the radiation fields and plasma currents from Fourier space into real space and time. A general form of time-varying current source is used, and collisional effects on the plasma response are retained. This paper is organized as follows. In the next section we discuss the dielectric properties of the D region and present the model used in the

[a]Current address: GE Capital Spacenet Services, Inc., 1750 Old Meadow Road, McLean, Virginia 22102-4300.

 1070-664X/96/3(5)/1484/11/$10.00

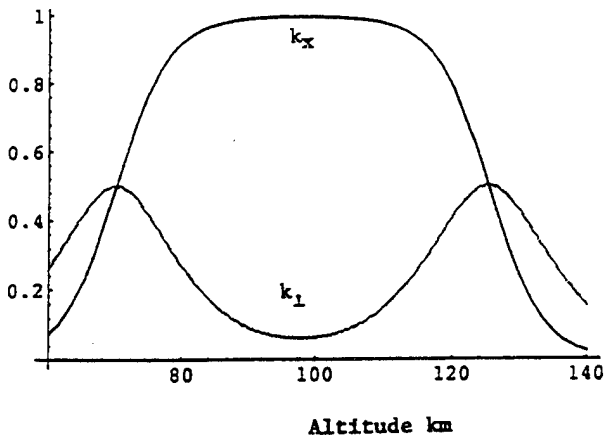


FIG. 1. $K_\perp(z)$ and $K_x(z)$ as functions of altitude for typical ionospheric parameters.

computations. In Sec. III we present the basic set of equations and the methods used to obtain their solutions. The computational results are discussed in Sec. IV. In Sec. V we discuss the application of the model to the generation of ELF/VLF waves in ionospheric heating. In the final section we summarize the results and present suggestions for future studies.

II. EIGENMODES OF THE LOWER IONOSPHERE

The linear wave equation describing electromagnetic wave propagation excited by a current $\mathbf{J}_s(\mathbf{r},t)$ in a vertically stratified ionosphere is given by[9]

$$-\nabla\times(\nabla\times\mathbf{E})+\frac{\omega_e^2(z)}{c^2}\frac{1}{\Omega_e}\mathbf{G}(z)\cdot\frac{\partial\mathbf{E}}{\partial t}=\frac{4\pi}{c^2}\frac{\partial\mathbf{J}_s(z,t)}{\partial t}, \tag{1}$$

where

$$\mathbf{G}(z)=\begin{bmatrix} K_\perp(z) & K_x(z) & 0 \\ -K_x(z) & K_\perp(z) & 0 \\ 0 & 0 & K_z(z) \end{bmatrix}, \tag{2}$$

$$K_\perp=\frac{\epsilon}{1+\epsilon^2},\quad K_x=\frac{1}{1+\epsilon^2},\quad K_z=\frac{\Omega_e}{\nu_e}, \tag{3}$$

and

$$\epsilon=\frac{\nu_e}{\Omega_e}+\frac{\Omega_i}{\nu_i}. \tag{4}$$

In deriving Eqs. (1)–(4) a vertical magnetic field $\mathbf{B}=\hat{e}_z B_0$ was assumed and $\nu_e(\nu_i)$ is the electron (ion) neutral collision frequency. Figure 1 shows plots of $K_\perp(z)$ and $K_x(z)$ as a function of altitude for typical ionospheric parameters. The important aspect of Fig. 1 is that for $z>130$ km the diagonal elements of $\mathbf{G}(z)$ dominate, giving rise to the traditional low-frequency Alfvén waves. However, for 70 km$<z<$130 km, the off-diagonal elements dominate, even for frequencies approaching zero. Since the value of $\epsilon\ll 1$, only the electron dynamics is important and the magnetized plasma modes resemble the well-known helicon modes in solid state plasmas.[10,11] The importance of this mode for the lower ionosphere can be seen by referring to the dispersion relation of the plasma in the 70–130 km altitude range for parallel propagation ($k=\hat{e}_z k$). This is given by

$$\left(\frac{kc}{\omega}\right)^2=1-\frac{\omega_e^2}{\omega(\omega-i\nu_e-\Omega_e)}-\frac{\omega_i^2}{\omega(\omega-i\nu_i-\Omega_i)}. \tag{5}$$

For $\omega<\Omega_i<\nu_i$, $\nu_e\ll\Omega_e$, and neglecting the displacement current, we obtain

$$\omega=k^2c^2\frac{\Omega_e}{\omega_e^2}\left[1-i\left(\frac{\nu_e}{\Omega_e}+\frac{\Omega_i}{\nu_i}\right)\right]. \tag{6}$$

This is the helicon mode and suffers weak attenuation, even for $\omega\ll\nu_e,\nu_i$. The important aspect of the above analysis is that in the 70–130 km range the plasma response is controlled by electron dynamics, not only in the usual whistler range ($\Omega_i<\omega<\Omega_e$), but also in the low-frequency range ($\omega<\Omega_i$). As a result, when the plasma response to externally induced perturbations is considered, it is sufficient to retain only the electron dynamics and ignore the motion of the ions. Such a model is referred to as the electronmagnetohydrodynamic (EMHD) model and is described in the next section.

III. IONOSPHERIC PLASMA MODEL

We study below the response of an electron plasma to a time-dependent and bounded current source $\mathbf{J}_s(\mathbf{r},t)$. The plasma is modeled by the EMHD equations, which, as discussed in Sec. II, are the equations that describe the plasma in the range between 70–130 km in the ionosphere. Assuming a homogeneous, cold plasma with the ions forming a stationary neutralizing background, the EMHD equations are[12]

$$\nabla\times\mathbf{B}=-\frac{4\pi}{c}en_0\mathbf{V}_e+\frac{4\pi}{c}\mathbf{J}_s, \tag{7}$$

$$\frac{d\mathbf{V}_e}{dt}=-\frac{e}{m_e}\left(\mathbf{E}+\frac{1}{c}\mathbf{V}_e\times\mathbf{B}_0\right)-\nu_e\mathbf{V}_e, \tag{8}$$

$$\nabla\times\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}, \tag{9}$$

where n_0 is the electron density, $\mathbf{V}_e$ is the electron fluid velocity, $\mathbf{E}$ and $\mathbf{B}$ are the radiation fields, and $\mathbf{J}_s$ is the externally driven current source, whose form will be specified later. The dominant plasma response enters through the first term on the right-hand side of Eq. (7). The displacement current has been neglected since we are interested only in low frequencies. From Eqs. (7)–(9) we obtain for the evolution of the magnetic field $\mathbf{B}(\mathbf{r},t)$,

$$\frac{\partial}{\partial t}\nabla^2\mathbf{B}+\nu_e\nabla^2\mathbf{B}-\frac{\omega_e^2}{c^2}\frac{\partial\mathbf{B}}{\partial t}-\Omega_e(\mathbf{b}\cdot\nabla)(\nabla\times\mathbf{B})$$
$$=\frac{4\pi}{c}\left(-\frac{\partial}{\partial t}\nabla\times\mathbf{J}_s+\Omega_e(\nabla\cdot\mathbf{J}_s)b-\Omega_e(\mathbf{b}\cdot\nabla)\right.$$
$$\left.\times\mathbf{J}-\nu_e\nabla_s\times\mathbf{J}_s\right), \tag{10}$$

where $\Omega_e=eB_0/m_ec$, $\mathbf{b}=\mathbf{B}_0/|\mathbf{B}_0|$, and $\omega_e^2=4\pi n_0e^2/m_e$. This is the key equation of the paper. Before proceeding with its

solution, it is instructive to examine the evolution of the electric field $\mathbf{E}(\mathbf{r},t)$, for an inertialless ($m_e=0$) and collisionless ($\nu_e=0$) plasma. From Eqs. (7)–(9) we find

$$\frac{\partial \mathbf{E}}{\partial t}-\frac{c^2\Omega_e}{\omega_e^2}\nabla^2\mathbf{E}\times\mathbf{b}=\frac{4\pi\Omega_e}{\omega_e^2}\mathbf{b}\times\frac{\partial \mathbf{J}_s}{\partial t}. \tag{11}$$

We solve Eq. (10) by using a spatial Fourier and temporal Laplace transform (the Appendix). This yields

$$\mathbf{D}(\mathbf{k},\omega)\cdot\mathbf{B}(\mathbf{k},\omega)=\mathbf{J}_{\mathrm{ext}}(\mathbf{k},\omega). \tag{12}$$

where $\mathbf{D}(\mathbf{k},\omega)$ is the dielectric tensor of magnetized plasma.

$$\mathbf{D}(\mathbf{k},\omega)=i\frac{4\pi}{c}\begin{pmatrix} -ik^2\omega+i\dfrac{\omega_e^2}{c^2}\omega+k^2\nu_e & -k_z^2\Omega_e & k_yk_z\Omega_e \\ k_z^2\Omega_e & -ik^2\omega-i\dfrac{\omega_e^2}{c^2}\omega-k^2\nu_e & -k_xk_z\Omega_e \\ -k_yk_z\Omega_e & k_xk_z\Omega_e & -ik^2\omega-i\dfrac{\omega_e^2}{c^2}\omega-k^2\nu_e \end{pmatrix}, \tag{13}$$

and $\mathbf{J}_{\mathrm{ext}}(\mathbf{k},\omega)$ is the transform of the external source and is given by

$$\mathbf{J}_{\mathrm{ext}}(\mathbf{k},\omega)=\frac{4\pi}{c}[\omega\mathbf{k}\times\mathbf{J}_s+i\Omega_e(\mathbf{k}\cdot\mathbf{J}_s)\hat{z}-ik_z\Omega_e\mathbf{J}_s -i\nu_e\mathbf{k}\times\mathbf{J}_s]. \tag{14}$$

The dispersion relation is

$$\mathbf{D}(\mathbf{k},\omega)=0. \tag{15}$$

From Eqs. (13) and (15) we find

$$\omega=\frac{kk_z\Omega_e}{k^2+\omega_e^2/c^2}\left(1-i\frac{k\nu_e}{k_z\Omega_e}\right), \tag{16}$$

which represents the propagation mode at low frequencies. For quasiparallel propagation, $k_z\approx k$, we recover Eq. (6) corrected for electron inertia and with ν_e/Ω_e corresponding to $\nu_e/\Omega_e[1+(\nu_i/\nu_e)(\Omega_e/\Omega_i)]$, we have

$$\omega=\frac{\Omega_e}{1+\omega_e^2/k^2c^2}\left(1-i\frac{\nu_e}{\Omega_e}\right). \tag{17}$$

In the following sections, Eq. (10) and its transform will be solved first for a two-dimensional current source (current sheet), given by

$$\mathbf{J}_s(x,z,t)=\begin{cases}\mathbf{e}_xJ_0\delta(z)(1-e^{-t/\tau}), & |x|<L,\\ 0, & |x|>L,\end{cases} \tag{18}$$

and then for a three-dimensional source,

$$\mathbf{J}_s(x,y,z,t)=\begin{cases}\mathbf{e}_xJ_0\delta(z)\delta(y)(1-e^{-t/\tau})u(t), & |x|<L,\\ 0, & |x|>L,\end{cases} \tag{19}$$

where $u(t)$ is a step function, and τ is the rise time of the source.

IV. COMPUTATIONAL RESULTS

We now present the results obtained by numerically solving Eq. (10) for the sources given by Eqs. (18) and (19).

A. Response to a current sheet: Collisionless case

Figures 2–6 show the results for the case $\nu_e=0$, $\omega_e/\Omega_e=1$, $\tau=45\Omega_e^{-1}$, and length $L=42c/\omega_e$ for the source given by Eq. (18). The results refer to a current source with $J_0=1$ mA/m. Figures 2 and 3 show the temporal evolution of the magnetic field $\mathbf{B}(\mathbf{r},t)$ as it propagates away from the source.

A pulse induced by the current propagates away on both sides of the source. It is generated during the switch-on and is characterized by strong dispersion with the perturbations with shorter wavelengths running ahead of those with longer wavelengths, as shown in Figs. 2(b) and 2(c) . The wave packet exhibits characteristics of the whistler wave and propagates with a group velocity $c/3$. Since the group velocity of the parallel propagating whistler is $V_g=2c\sqrt{\omega\Omega_e/\omega_e^2}$, this observed group velocity ($c/3$) corresponds to a frequency $\omega=\Omega_e/36$, comparable to the switch-on time $\tau=45\Omega_e^{-1}$. The wave numbers in the wave packets are in the range, $k\approx 5-6\omega_e/c$. The identification of the precursor waves as whistlers is clear from the wave polarization (Fig. 3), which shows that the B_x and B_y fields in the $x=0$ plane are 90° out of phase, as expected for these right-hand polarized waves.

The dispersion curve ω vs k_z obtained by computing wavelengths at different frequencies, is displayed in Fig. 4. The computed points fall on the theoretical dispersion curve for whistlers,[13] shown as a solid line in Fig. 4. When the source is turned off, the wave packets disconnect from the source and propagate as isolated wave packets (Fig. 5). This is quite different from the behavior during the switch-on (Fig. 3), during which the radiation field is connected to the source. By turning on and off the cross-field current source, we can generate isolated low-frequency wave packets propagating away on both sides of the current source.

Of particular significance is the identification of the currents induced in the plasma and of the current closure path. This aspect has not been emphasized in the previous studies. The current carried by the excited waves at $t=300\Omega_e^{-1}$ is shown in Fig. 6. The initial source current with $L=75c/\omega_e$ is

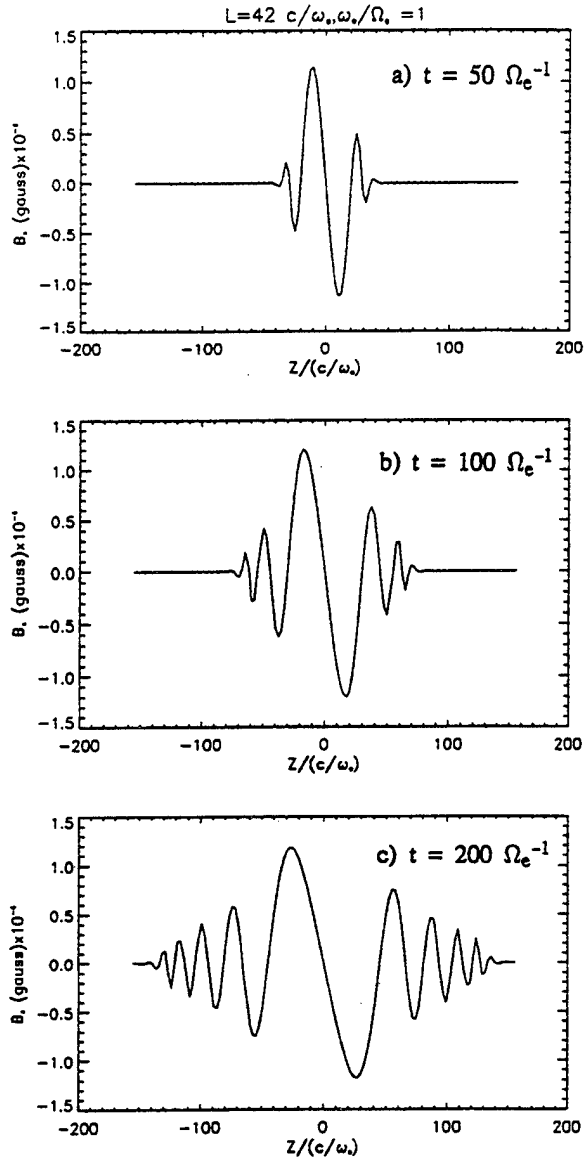


FIG. 2. The amplitudes of the magnetic field component B_x as a function of z along the external magnetic field line in the $x=0$ plane at different times: (a) $50\Omega_e^{-1}$, (b) $100\Omega_e^{-1}$, and (c) $200\Omega_e^{-1}$. The pulse propagates at the group velocity of whistler waves.

shown with a large arrow in the middle, while the plasma currents are proportional to the lengths of the arrows. The complete current path consists of (1) the outgoing portion of the closure current, as represented by the J_z flowing from the top of the current source outward along the magnetic field connecting the top; (2) the crossover portion of the closure current, as represented by the J_x flowing across the field and the midplane; and (3) the return portion of the closure current, as represented by the J_z flowing along the magnetic field toward the bottom of the current source, thus completing the circuit. The two current loops expand along the z direction, preceded by whistler radiation. The expanding loop has a whistler structure and the front expands in time as

$$Z(t)\approx\frac{c}{\omega_e}\sqrt{\Omega_e t}. \tag{20}$$

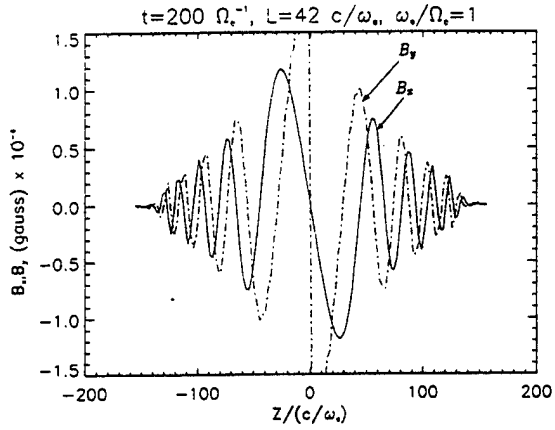


FIG. 3. The magnetic field components B_x and B_y, which are 90° out of phase in z. The excited waves propagate as right-hand circularly polarized waves along the background magnetic field.

The current closure region expands along the external magnetic field with the whistler, the group velocity, and is consistent with the time scale given by Eq. (20). However, the process is weakly dissipative.

We have performed a comprehensive study of the scaling of the field aligned length of the current loop as a function of the plasma parameters, and the results are shown in Fig. 7. The current front moves with the whistler group velocity, which is frequency dependent, and the typical group velocity is computed using a fixed frequency, $\omega=\Omega_e/170$. The size of the current closure region is defined by the location of the current reversal away from the source. For example, in Fig. 6, the closure current reverses direction at $z\simeq 45c/\omega_e$, defining the boundary of the region. The size of the loop r at a time $\Omega_e t=200$ as a function of the electron plasma frequency is shown in Fig. 7(a), and it scales as $\omega_e^{-1.1}$. The current closure size at time $\omega_e t=200$, but with ω_e constant and Ω_e varying, is shown in Fig. 7(b) and the scaling is $r\sim\Omega_e^{0.64}$. Finally, Fig. 7(c) shows the closure range as a function of time for the case $\omega_e/\Omega_e=1$, giving a $t^{0.6}$ scal-

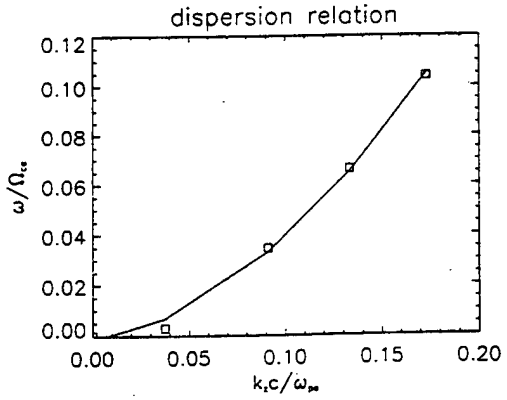


FIG. 4. The normalized dispersion relation, obtained numerically from the wavelengths corresponding to different frequencies (squares) and that given by $\omega=\Omega_e/(1+\omega_e^2/k^2c^2)$ for the given plasma parameter $\omega_e/\Omega_e=1$ (solid line). Both polarization and dispersion measurements show that the current perturbations excite whistlers.

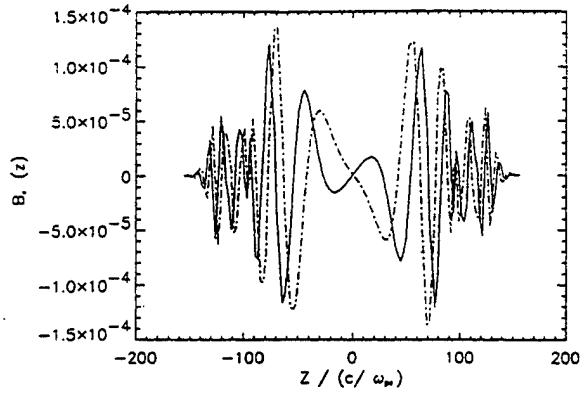


FIG. 5. The magnetic field components B_x (solid line) and B_y (dotted line) fields at $t=200\Omega_e^{-1}$. The current pulse is turned on at $t=0$ and turned off at $t=50\Omega_e^{-1}$. The plasma parameters are the same as in Fig. 3.

ing. These results confirm the validity of Eq. (20) in describing the loop expansion.

The above results show that the current closure is accomplished at early times. Besides the two field aligned currents, the current closure is accomplished by an electron cross-field Hall current and the closure region expands with the whistler group velocity. The electron Hall current is driven by an electric field E_y, whose temporal evolution in the x-z plane is shown in Fig. 8. Notice that the plasma response remains well confined in the transverse direction to a size comparable to $2L$, while propagating along B. We should remark that this transverse confinement was observed in another set of runs (not shown here), with $L=84c/\omega_e$ and $168c/\omega_e$.[14] The cross-field currents at the switch-off showed that current closure pattern is approximately the same as in the switch-on case with the current directions reversed.

An interesting effect related to long time propagation following the switch-off of the source current for the collisionless case can be seen in Fig. 9. It shows contours of the electric field at $t=2000\Omega_e^{-1}$, for a source with $\tau=100\Omega_e^{-1}$, which is turned off at $100\Omega_e^{-1}$. It can be seen that the two

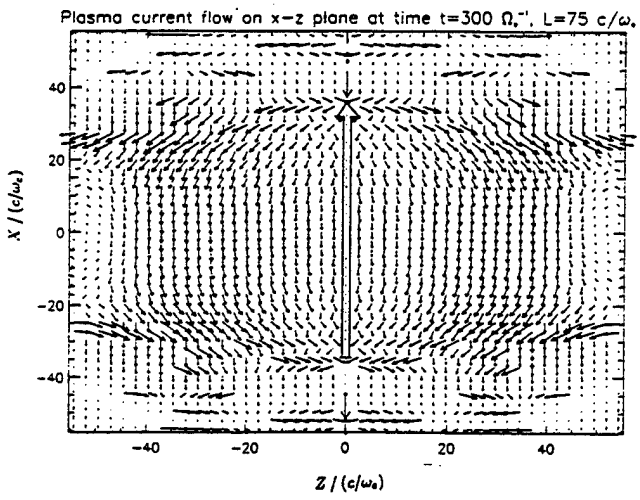


FIG. 6. The plasma current flow in the x-z plane at time $t=300\Omega_e^{-1}$, with $\Omega_e/\omega_e=1$ and $\tau=45\Omega_e^{-1}$. The current source size is $L=75c/\omega_e$ and the initial current is shown (not to scale) in the middle for comparison.

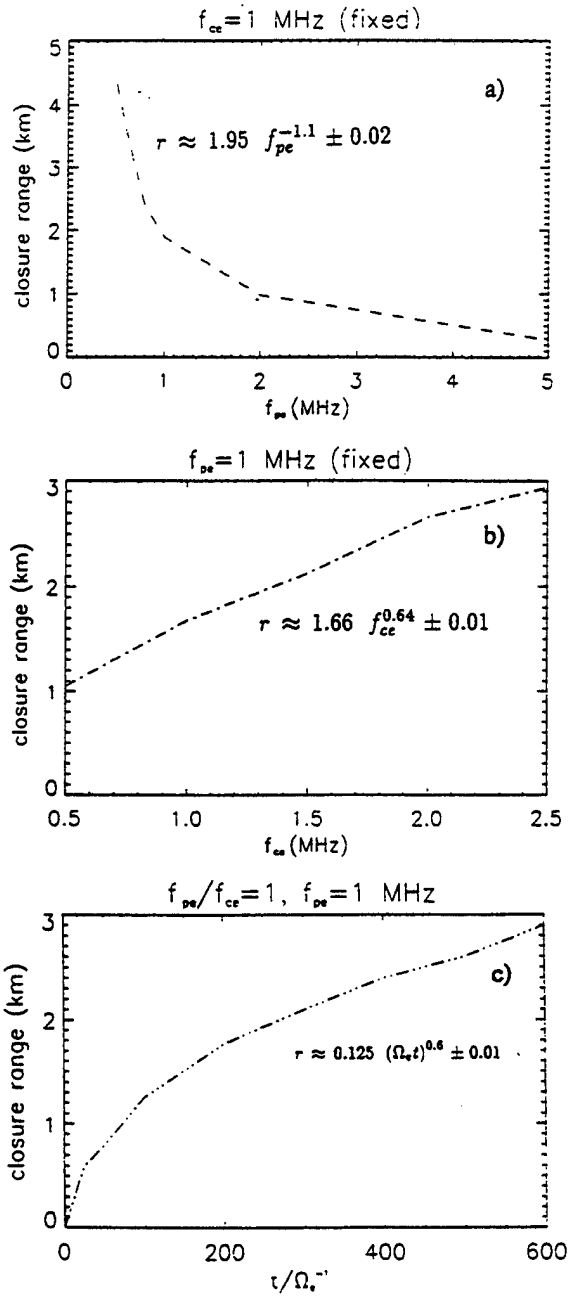


FIG. 7. The scaling of the current closure, (a) closure range as a function of plasma frequency for fixed $f_{ce}=1$ MHz, (b) closure range as a function of cyclotron frequency for fixed $f_{pe}=1$ MHz, and (c) closure range as a function of time for $f_{pe}/f_{ce}=1$, and $f_{ce}=1$ MHz.

wave packets disconnect and propagate uncoupled with the characteristic "Story" structure over a 19° angle.[13]

The previous results were constrained to relatively short values of τ. On the basis of the physics we expect similar behavior for longer time scales. To confirm this we performed a set of runs with $\tau=10^4\Omega_e^{-1}$, the other parameters being $\omega_e/\Omega_e=2$, and $L=84c/\omega_e$ and $1600c/\omega_e$. A summary of the results for the evolution of B_y is shown in Fig. 10 in a different format. The helicon wave packet is highly dispersive, exhibiting characteristics similar to the whistler wave discussed previously. The current closure structure is similar to the one shown in Fig. 6. It should be noted that in the

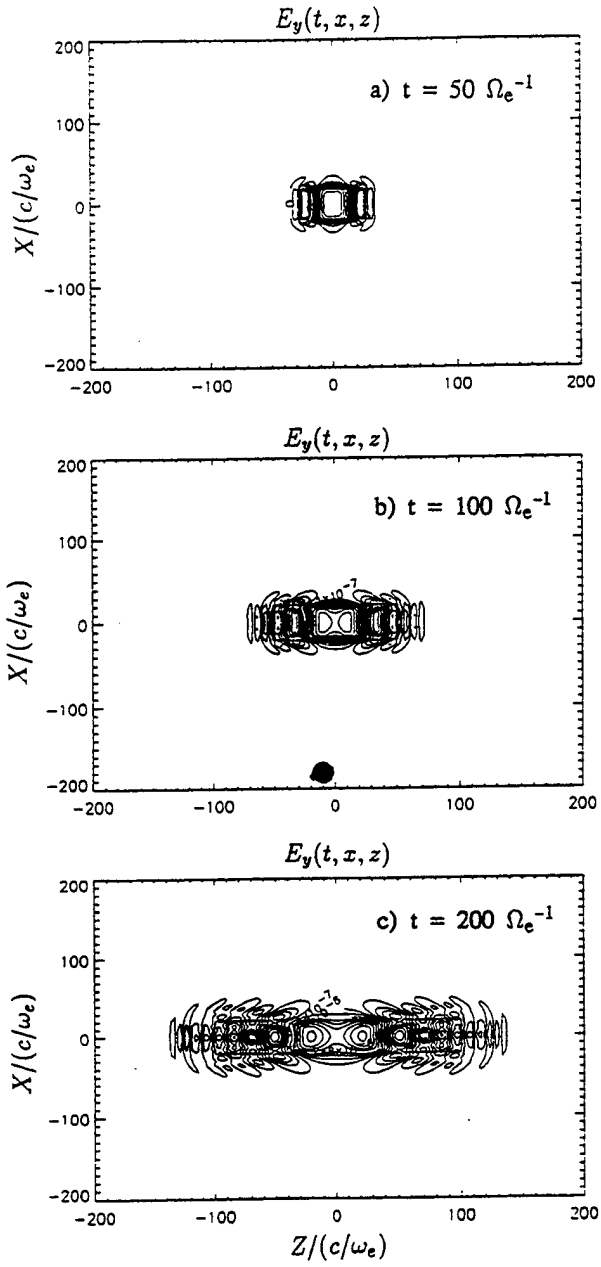


FIG. 8. The contours of E_y at different times after applying a current pulse across B_0: (a) $50\Omega_e^{-1}$, (b) $100\Omega_e^{-1}$, and (c) $200\Omega_e^{-1}$.

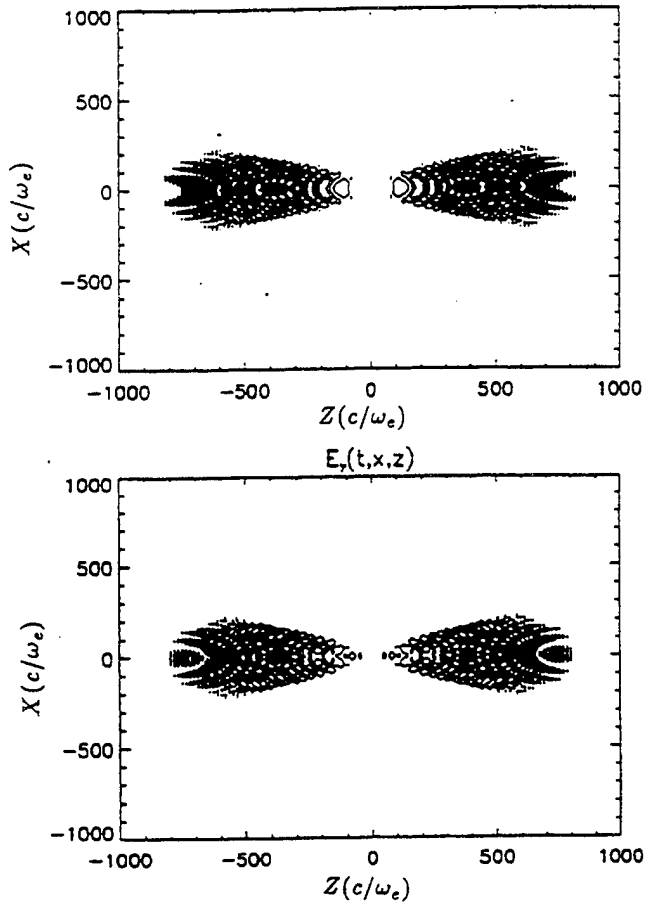


FIG. 9. The contours of $E_x(x,z)$ and $E_y(x,z)$ at time $t=2000\Omega_e^{-1}$ after applying a current across B_0. The current pulse is turned on at $t=0$ and turned off at $t=100\Omega_e^{-1}$. Note the wavefront spreads in a cone at an angle $\alpha\sim 20°$ with respect to B_0.

absence of ion dynamics, these results do not represent the complete physical picture. However, the long time effects of dispersion and dissipation on the propagating wave packets is shown by these results.

In concluding, we note that the observed field and current structures can be understood qualitatively with the following simple physical model. The time-varying current drives an inductive electric field $E_x=-\partial A_x/\partial t$ that is antiparallel to it when $\partial I/\partial t>0$ and parallel when $\partial I/\partial t<0$. Since the electrons are strongly magnetized, this electric field generates only a small polarization current. However, the electrons perform an $\mathbf{E}_x\times\mathbf{B}_0$ drift that gives rise to a space-charge electric field $\mathbf{E}_y$ perpendicular to the source current. The consequence of this nonuniform space charge separation is twofold: (1) It gives rise to field-aligned currents $J_\parallel$ that flows so as to neutralize the excess charges. These field-aligned currents give rise to the observed magnetic field component B_x shown in Figs. 2 and 3; and (2) the electrons perform an $\mathbf{E}_y\times\mathbf{B}_0$ drift that gives rise to cross-field currents $J_x=necE_y/B_0$, antiparallel to the imposed current. Although this Hall current has the appearance of an induced current, it is not directly driven by the inductive electric field but only indirectly via the space-charge separation.[15] As the current J_x moves along $\mathbf{B}_0$ into the plasma, the above processes repeat at the wave front, although with reversed signs: an induced electric field $-\partial A_x/\partial t$ gives rise to an electron drift v_y, resulting in a space-charge electric field E_y and Hall current $J_x=-\sigma_H E_y$, with $\sigma_H=\omega_e^2/4\pi\Omega_e$. The electron Hall currents, which are not balanced by ion Hall currents, form the cross-field currents.

B. Collisional and 3-D effects

Although the collective motion of magnetized electrons in a plasma dominates the dynamic response, collisional effects are important. On including collisional effects in the EMHD model, the dispersion relation of whistler/helicon waves in magnetized electron plasma becomes $\omega=\Omega_e(1-i\nu_e/\Omega_e)/(1+\omega_e^2/k^2c^2)$. The damping rate of the radiation fields due to collisions is controlled by the factor ν_e/Ω_e. We have performed a series of computations with the

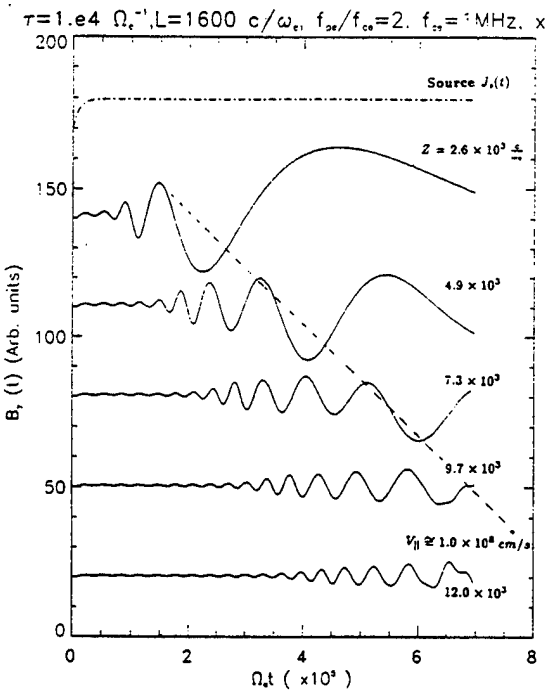


FIG. 10. The time variation of the magnetic field associated with a current switch-on (1 mA, rise time $\tau=10^4\Omega_e^{-1}$) across B_0 through a magnetoplasma. The transient field B_y at different z from the source exhibiting wave characteristics along B_0 are shown.

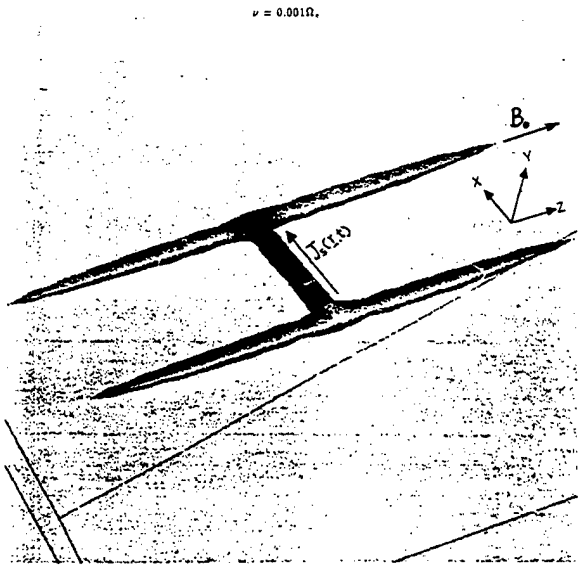


FIG. 12. The isosurface plot of the current density in three dimensions corresponding to a current pulse $\mathbf{J}_s(\mathbf{r},t)$ imposed across the ambient magnetic field B_0 for a strongly magnetized plasma. $\nu_e=0.001\Omega_e$. The outgoing (return) portion of the closure current flowing from the top (bottom) along the ambient magnetic field lines. The cross-field plasma currents are not shown due to their small values in this isosurface plot corresponding to $J=0.1J_s$.

values of the parameters corresponding to those of Figs. 2–6, but with ν_e varying between $(0.001-1)\Omega_e$. The results for the evolution of B_x are shown in Fig. 11, at $t=200\Omega_e^{-1}$. For $\nu_e=0.01\Omega_e$, the wave pattern is similar to the collisionless case [Fig. 2(c)]. When the collision frequency is increased, the waves are gradually damped. For $\nu_e=0.1\Omega_e$, the oscillating part of B_x is damped, while for $\nu_e=\Omega_e$ the collective mode ceases to exist and the field diffuses into the plasma in a fashion similar to a conventional conductor. The structure of the current loop follows a similar pattern.

For a 3-D time-dependent cross-field current source [Eq. (19)], the current closure pattern is similar to the 2-D case. For the strongly magnetized case, the field-aligned current is shown in Fig. 12 and is closed by electron Hall current (not shown). In three dimensions the cross-field plasma current is more distributed than in the 2-D case shown in Fig. 6. Consequently the current density is much smaller and does not show in the isosurface plot of the current density, shown in Fig. 12, for $|J|=0.1J_s$. The current distribution in the x-z plane (at $y=0$) for the 3-D distribution (Fig. 12) is shown in Fig. 13. The cross-field current is displayed here, but not shown in Fig. 12. The results obtained using the 3-D time-dependent current source given by Eq. (19) follow the same

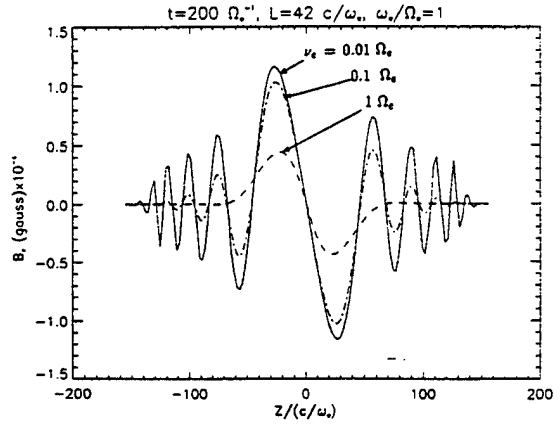


FIG. 11. The effect of collisions on radiation field $B_x(0,z,t)$ at $t=200\Omega_e^{-1}$ for collisional frequencies $\nu_e/\Omega_e=0.01$, 0.1, and 1.0.

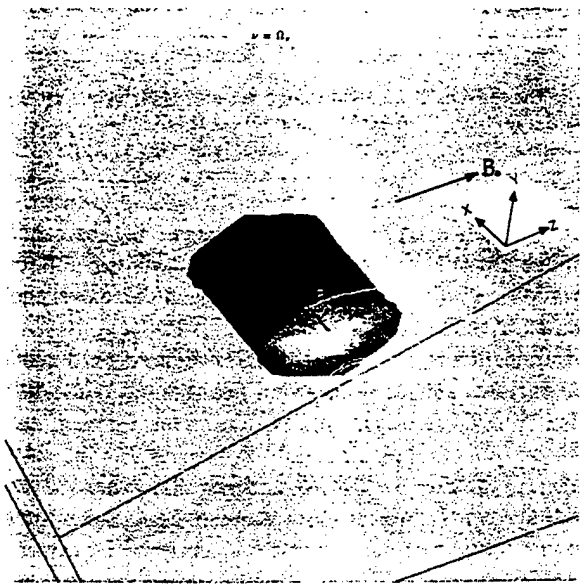


FIG. 13. The three-dimensional plasma current distribution generated by a current pulse $\mathbf{J}_s(\mathbf{r},t)$ imposed across the ambient magnetic field B_0 for a plasma with $\nu_e=\Omega_e$.

general physics of the previous results, although the current path appears more complex. An important result of the 3-D runs is the striking difference of the current path between the weakly collisional ($\nu_e=0.001\Omega_e$) and the strongly collisional ($\nu_e=\Omega_e$) cases shown in Figs. 12 and 13. Figure 12 shows a strong field-aligned response similar to the two-dimensional case, and the current is again closed by a Hall current (not shown). The three-dimensional features are confined in narrow regions about the field-aligned currents. On the contrary, the strongly collisional case shows a cylindrical distribution of currents about the imposed current, similar to the diffusive response expected when a current pulse is applied on an isotropic conductor. Also, 3-D results show the scale sizes of the current in the y direction. In Fig. 12 the field-aligned current channels has similar extents in the x and y directions. In Fig. 13 the current is distributed in a region that is narrower in the y-direction.

V. GENERATION OF ELF/VLF WAVES BY IONOSPHERIC HEATING

A fascinating and important property of the active ionosphere is its potential to act as a frequency transformer that converts HF power injected from the ionospheric heater to coherent VLF/ELF/ULF waves.[5-8] Waves between 10^{-3} Hz and 30 kHz have been generated in the ionosphere by amplitude modulated HF heating in the auroral zones. Ionospheric heating modulates the ambient conductivity, redistributing the ionospheric currents. This acts as an effective ionospheric antenna radiating waves back to the ground or to the magnetosphere at the low modulation frequency. This antenna has often been referred to as the Polar Electrojet Antenna (PEJ). A key issue on the subject is the radiative moment of the PEJ.

Early analysis[7,16] assumed that the radiative moment was a Horizontal Electric Dipole (HED) with moment M_E given by

$$M_E=E_0\Delta\sigma L_z L^2, \tag{21}$$

where L_z is the absorption length of the HF waves in the z direction and L is the horizontal dimension of the heated region. In the ionosphere (70–90 km), where the modification takes place, $\Omega_e>\nu_e$ but $\Omega_i<\nu_i$, as a result the EMHD model applies. The results of Sec. IV showed that the plasma responds to a cross-field current, such as the one that creates the assumed HED, by forming a current loop that includes a Hall current in the opposite direction to the applied current. We, therefore, expect that the HED model is incorrect, and should be replaced by a horizontal magnetic dipole (HMD) type source. The objective of this section is to use the EMHD model to determine the type of radiative source and its scaling with frequency and plasma parameters.

A. Current source by modulated HF heating

In the region of the polar electrojet, a time-varying current source is produced by modulated ionospheric heating. The absorption of a HF wave in the lower ionosphere results in the variation of the electron temperature (ΔT_e) and, to a lesser extent, of the electron density (ΔN_e) at the modulation frequency, and hence to a current modulation. The physics of the current source generation by modulated ionospheric heating is as follows. At high latitudes the solar wind interaction with the Earth's magnetosphere results in the creation of an electromotive force. Since the magnetic field lines are equipotentials, the high latitude electric field $\mathbf{E}_0=E_0\hat{e}_x$ maps into the lower ionosphere, where collisional processes allow for the generation of cross-field currents. Two types of currents flow across the magnetic field $\mathbf{B}=B_0\hat{e}_z$. The Pedersen current,

$$\mathbf{J}_P=\sigma_P E_0\hat{e}_x, \tag{22}$$

in the direction of the electric field $\mathbf{E}_0$, and the Hall current,

$$\mathbf{J}_H=\sigma_H E_0\hat{e}_y. \tag{23}$$

In Eqs. (22)–(23), σ_P and σ_H are the Pedersen and Hall conductivities, defined as

$$\sigma_P=\frac{ne^2}{m\Omega_e^2}\frac{\nu_e}{1+\nu_e^2/\Omega_e^2} \quad \text{and} \quad \sigma_H=\frac{ne^2}{m\Omega_e^2}\frac{\Omega_e}{1+\nu_e^2/\Omega_e^2}, \tag{24}$$

respectively. Since ν_e varies linearly with the electron temperature T_e, amplitude modulated heating at a low-frequency ω induces a modulation on the values of the conductivities through ν_e. As is clear from Eq. (24), for heating at altitudes with $\nu_e\ll\Omega_e$, the dominant modulation is in the Pederson conductivity, and this results in a low-frequency modulated cross-field current.

The relevant current density is the height integrated current density J_0 due to the modulation in the temperature. For the case that modification of the Pedersen conductivity dominates, the current generated by modulated HF heating can be written as a series of pulses, with each pulse represented by

$$\mathbf{J}_s(\mathbf{x},t)=\hat{e}_x S(\mathbf{x},t)J_0, \tag{25a}$$

$$S(\mathbf{x},t)=\Theta\left(|x|-\frac{L}{2}\right)\Theta\left(|y|-\frac{\Delta}{2}\right)\delta(z) \times\begin{cases}(1-e^{-t/\tau}), & t\le t_0,\\ e^{-(t-t_0)/\tau}-e^{-t/\tau}, & t>t_0,\end{cases} \tag{25b}$$

$$J_0=j_0L_z, \tag{25c}$$

where $\Theta(x)$ is the Heavyside step function, L and Δ are the extensions of the current region in the x and z directions, respectively, j_0 is the modified current density, and L_z the absorption length of the HF power. The values of j_0 and L_z are given by[10,11]

$$j_0=\frac{\omega_e^2}{4\pi\Omega_e^2}\frac{\Delta T}{T_0}\nu_e(T_0)E_0, \tag{26a}$$

$$L_z=\frac{c}{\nu_e(T_0)}\frac{\omega_{HF}^2}{\omega_e^2}, \tag{26b}$$

where T_0 is the ambient temperature, ΔT is the modification, and ω_{HF} is the frequency of the HF wave. In applying Eq. (26b), care should be exercised, in that it is valid only if

$$\frac{\omega_{HF}^2}{\omega_e^2}<\frac{\nu_e}{c}L_N,$$

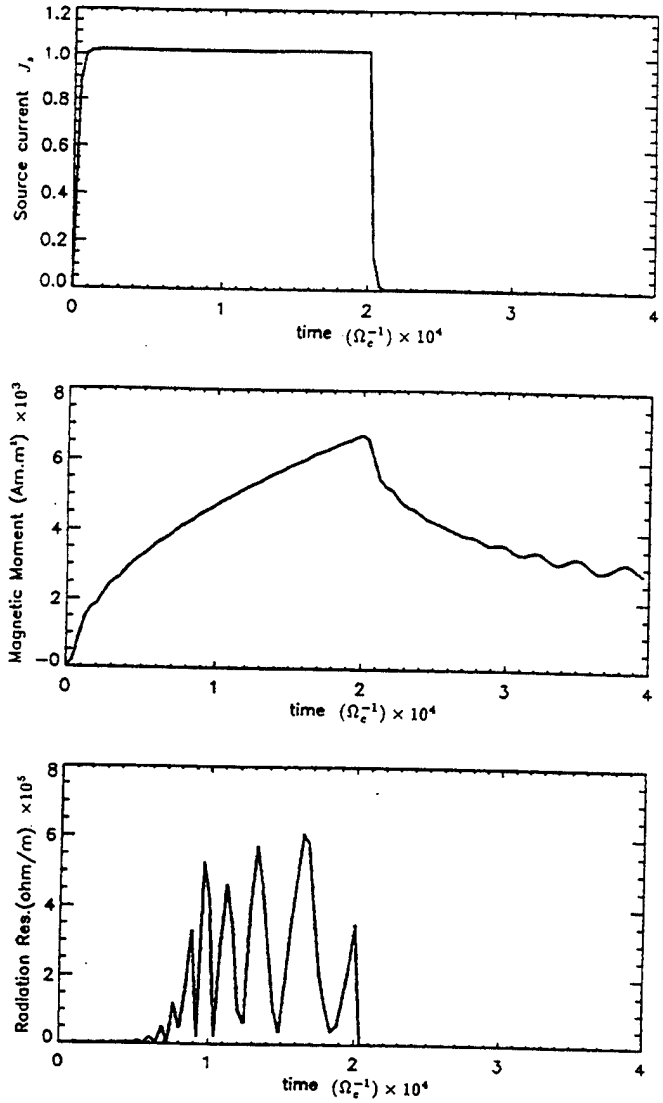


FIG. 14. The current source $J_s(t)$ generated by modulated HF heating (upper panel), the magnetic moment $M(t)$ generated by $J_s(t)$ (middle panel), and the radiation resistance due to $M(t)$ (lower panel).

where L_N is the plasma density gradient. From Eq. (26a) and (26b), we find

$$J_0=\frac{c}{\Omega_e}\left(\frac{\omega_{\mathrm{HF}}^2}{\Omega_e^2}\right)\frac{\Delta T}{T_0}(\Omega_e E_0). \tag{27}$$

B. Simulations of the PEJ structure

A set of simulations was performed using the code, to determine the expected structure of the PEJ. The parameters were taken as representative of 80–90 km of the ionosphere, with $f_{ce}=1$ MHz and $f_{pe}=2$ MHz. The simulation box was set in the x-z plane and covered a region of 120 km in each dimension. A current pulse whose temporal behavior is given by Eq. (25), with $J_0=1$ mA/m, $L=20$ km, and $\Delta=1$ m, was placed in the middle of the box. The value of the dipole moment was found by the numerical integration of the current moment in the upper half-plane,

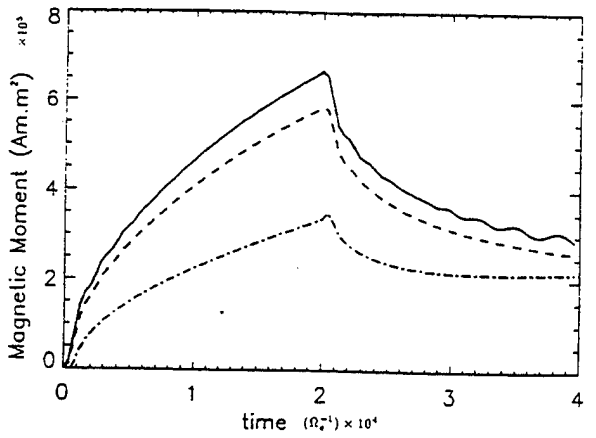


FIG. 15. The integrated magnetic moment as a function of time for different collision frequencies: $\nu_e=0.01\Omega_e$ (solid line), $\nu_e=0.1\Omega_e$ (dashed line), and $\nu_e=0.5\Omega_e$ (dot–dashed line). The current pulse is turned on at $t=0$ and turned off at $t=2\times10^4\Omega_e^{-1}$.

$$\mathbf{M}=\frac{1}{2c}\int_{-L/2}^{L/2}dx\int_{-\Delta/2}^{\Delta/2}dy\int_0^{\infty}dz\ \mathbf{x}\times\mathbf{J}(x,z,t). \tag{28}$$

For the current source shown in the top panel of Fig. 14, the value of

$$m\equiv\frac{My}{\Delta}=\frac{1}{2c}\int_{-L/2}^{L/2}dx\int_0^{\infty}dz[zJ_x(x,z,t)-xJ_z(x,z,t)], \tag{29}$$

as a function of time is shown in the middle panel of Fig. 14. In this particular case $t_0=2\times10^4\Omega_e^{-1}$, corresponding to an ELF frequency of 300 Hz, while $\nu=0.01\Omega_e$. After an initial transient the magnetic moment increases as $\sqrt{t}$ until the current is turned off. The peak value of m is 7×10^5 A m^2/m, for $J_0=1$ mA/m, and the minimum value is 3×10^5 A m^2/m, corresponding to a total net radiating moment of 4×10^5 A m^2/m. The temporal variation of $m(t)$ is due to the radiation of whistler waves propagating upward. A radiation resistance can be computed by integrating the Poynting flux across a detector located, e.g., at $z=38$ km. The temporal behavior of the radiation resistance $R(t)$ is shown in the lower panel of Fig. 14. It is expected that PEJ will radiate whistler waves with instantaneous power $P(t)=R(t)(J_0\Delta)^2$. The role of the collisionality in the value of m is shown in Fig. 15. It can be seen that the radiative value of m is relatively insensitive to ν_e for $\nu_e\ll\Omega_e$. However, the PEJ radiation becomes negligible when $\nu_e/\Omega_e\approx1$. We finally note that varying t_0, which corresponds to varying the ELF frequency, results in a scaling of $m\sim\sqrt{t_0}\sim1/\sqrt{f}$, where f is the ELF frequency.

C. Structure and scaling of the PEJ antenna

From the results of Secs. VI A and VI B we can determine the magnetic moment expected of the PEJ antenna and the resultant scaling by using a three layer model, such as shown in Fig. 16. The upper layer has low collisionality ($\nu_e/\Omega_e\leq0.5$), whereas the lower layer has high collisionality

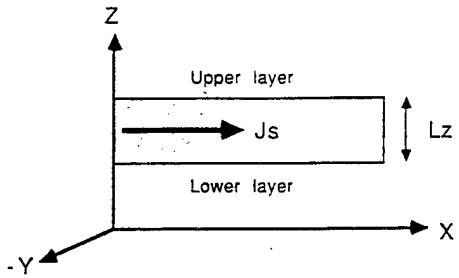


FIG. 16. The three layer model of the magnetic moment.

($\nu_e/\Omega_e \geq 0.5$), and the current flows in the middle layer with $0.5 \leq \nu_e/\Omega_e \leq 5.0$. On the basis of the results presented in Sec. V B,

$$M(t) \approx m_u(t) - m_l(t) \approx m_u(t), \tag{30}$$

where $m_u(m_1)$ are the values of M_y/Δ in the upper and lower layers.

The above analysis shows that the horizontal magnetic dipole moment produced by modulated HF heating can be written in mks units as

$$M_y(f) = 6\times 10^9 \left(\frac{\Delta}{20\ \text{km}}\right)\left(\frac{L}{20\ \text{km}}\right) \times \left(\frac{f_{\text{HF}}}{1.4\ \text{MHz}}\right)^2 \left(\frac{\Delta T}{T_0}\right) \sqrt{\frac{f_0}{f}} \left(\frac{E_0}{\text{V/m}}\right) \text{A m}^2, \tag{31}$$

where $f_0 = 300$ Hz. In Eq. (30), we note that M_y scales as $1/\sqrt{f}$, and the corresponding power as $1/f$. Such behavior is clearly seen in the near-field results ($f \leq 1$ Hz) of the Tromso experiments.[17] Similar behavior was observed in the transition from the near- to the far field for frequencies 78 and 154 Hz during high power active stimulation (HIPAS) experiments. Furthermore, during the HIPAS experiments with $\Delta \sim L \sim 20$ km, $f_{\text{HF}} \approx 2.8$ MHz and $f = 154$ Hz, M_y varied between 3×10^8–2×10^9 A m^2. This is consistent with Eq. (32) for values of E_0 varying between 10 and 60 mV/m.

VI. CONCLUSIONS

The dynamic response of a magnetoplasma to an external time-dependent current source is analyzed within the framework of the EMHD equations. It is shown that, besides short time scales, the EMHD model is a good physical model for the D and E regions of the ionosphere, due to the predominance of ion–neutral collisions. Following a transient emission of whistler/helicon waves, a current loop composed of two field-aligned currents and a transverse Hall current, in addition to the source current, is established. All the dynamics are controlled by the electron flow. The currents are parallel to the perturbed components of the magnetic field and all three components are important. The size of the loop scales as $\sqrt{\Omega_e t/\omega_e^2}$. Following the termination of the source, the current loop detaches from the source and damps at a rate controlled by collisions. When $\nu_e > \Omega_e$ the wave motion disappears and the plasma responds as an isotropic conductor. In a homogeneous plasma two symmetric current loops form on each side of the source. The presence of inhomogeneities breaks the symmetry, so that the magnetic moments of the loops do not cancel each other.

The results of the study have implications[10,11] in the response of electrodynamic tethers, short term perturbations of the magnetopause and the magnetotail by the solar wind, and on the physics of lower ionosphere when perturbed by natural or artificial sources at any frequency range. In this paper we emphasized the application of the theory to the generation of ELF/VLF radiation by modulated HF heating of the ionosphere. It is shown that, contrary to previous claims, the source region has a magnetic moment consistent with a horizontal magnetic dipole. The importance of the inhomogeneity in the vertical profile of the collision frequency is emphasized. The scaling properties and radiation power generated are determined and shown to be consistent with the experimental data.

Before closing, we should comment on some limitations of the model, arising from the neglect of the wave fields in the electron equation of motion [Eq. (8)]. The results found are linear in the strength of the source and obey the principle of superposition. They are valid up to current densities, such that $B \ll B_0$. This condition is easily satisfied for most ionospheric applications, since B/B_0 is at most 10^{-3}. A most stringent condition emerges in the very low collisionality regime and is related to the presence of the parallel electric field E_z. The value of this field can be estimated as

$$E_z = \frac{(\nu + i\omega)}{\omega_e^2} 4\pi j_z. \tag{32}$$

Since j_z is proportional to the current level at which phenomena associated with runaway electrons and collective modes would become important. Such phenomena could be important in magnetospheric applications and some tether applications, and will be explored in a future paper.

ACKNOWLEDGMENTS

This research is supported by the National Aeronautics and Space Administration under Contract No. NAS-51101 and the Office of Naval Research under Grant No. ONR-N0014-93-k-2019.

We would like to thank Dr. Mark Mandt for helpful discussions.

APPENDIX: EMHD MODEL (STATIONARY SOURCE)

Applying Fourier (with respect to x,z) and Laplace transforms (with respect to t) on Eq. (10), we obtain the radiation fields in ($\mathbf{k}.\omega$) space ($s = i\omega$), as discussed in Sec. III,

$$E_x(\mathbf{k},s) = i\,\frac{4\pi\Omega_e}{c}\,\frac{J_s}{s}\,\frac{k_x^2(1+\nu_e/s)+\omega_e^2/c^2}{\Delta}\,k_z, \tag{A1}$$

$$B_y(\mathbf{k},s)=i\frac{4\pi}{c}J_s \times\frac{k^2(1+\nu_e/s)+k^2\Omega_e^2/s^2+(\omega_e^2/c^2)(1+\nu_e/s)}{\Delta}k_z, \tag{A2}$$

$$B_z(\mathbf{k},s)=-i\frac{4\pi\Omega_e}{c}\frac{J_s}{s}\frac{k_x^2(1+\nu_e/s)+\omega_e^2/c^2}{\Delta}k_x, \tag{A3}$$

where $\Delta=[k^2(1+\nu_e/s)+\omega_e^2/c^2]^2+(k_zk\Omega_e/s)^2$. The following current source is introduced into a spatial region that is filled with stationary cold plasma and uniform magnetic field in the z direction,

$$\mathbf{J}_s(x,z,t)=\begin{cases}\mathbf{e}_xJ_0\delta(z)(1-e^{-t/\tau}), & |x|<L,\\ 0, & |x|>L.\end{cases} \tag{A4}$$

The Fourier–Laplace transforms of $J_s(x,z,t)$ is $J_s(k_x,k_z,s)=2J_0\sin k_xL/[k_xs(s\tau+1)]$, and the inverse Laplace transforms of Eqs. (A1)–(A3) yields

$$B_x(\mathbf{k},t)=i\frac{8\pi J_0\Omega_e}{c\tau}\frac{k_z\sin(k_xL)}{k_x(k^2+\omega_e^2/c^2)^2}\left[\left(m_{x1}\tau+\frac{m_{x2}}{A_1}+\frac{m_{x3}}{A_2}\right)-m_{x1}\tau e^{-t/\tau}-\frac{m_{x2}}{A_1}e^{-A_1t}-\frac{m_{x3}}{A_2}e^{-A_2t}\right], \tag{A5}$$

$$B_y(\mathbf{k},t)=i\frac{8\pi J_0}{c\tau}\frac{k_z\sin(k_xL)}{k_x(k^2+\omega_e^2/c^2)^2}\left[\left(m_{y1}\tau+\frac{m_{y2}}{A_1}+\frac{m_{y3}}{A_2}\right)-m_{y1}\tau e^{-t/\tau}-\frac{m_{y2}}{A_1}e^{-A_1t}-\frac{m_{y3}}{A_2}e^{-A_2t}\right] \tag{A6}$$

$$B_z(\mathbf{k},t)=-k_x/k_zB_x(\mathbf{k},t). \tag{A7}$$

The triads (m_{x1},m_{x2},m_{x3}) and (m_{y1},m_{y2},m_{y3}) are the solutions of the following linear equations.

$$\mathbf{T}\cdot\begin{pmatrix}m_{x1}\\m_{x2}\\m_{x3}\end{pmatrix}=\begin{pmatrix}0\\k_x^2+\dfrac{\omega_e^2}{c^2}\\k_x^2\nu_e\end{pmatrix},$$

$$\mathbf{T}\cdot\begin{pmatrix}m_{y1}\\m_{y2}\\m_{y3}\end{pmatrix}=\begin{pmatrix}k^2+\omega_e^2/c^2\\\nu_e(2k^2+\omega_e^2/c^2)\\k^2(\nu_e^2+\Omega_e^2)\end{pmatrix}, \tag{A8}$$

where

$$T=\begin{pmatrix}1 & 1 & 1\\ A_1+A_2 & 1/\tau+A_2 & 1/\tau+A_1\\ A_1A_2 & A_2/\tau & A_1/\tau\end{pmatrix}, \tag{A9}$$

and $A_1=k(-k\nu_e+ik_z\Omega_e)/(k^2+\omega_e^2/c^2)$, $A_2=k(-k\nu-ik_z\Omega_e)/(k^2+\omega_e^2/c^2)$

We can now transform $\mathbf{B}(\mathbf{k},t)$, given by Eqs. (A5)–(A7), numerically to obtain radiation fields and plasma currents in $(\mathbf{r},t)$ space by using two-dimensional inverse FFT techniques.

[1] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (Prentice–Hall, Englewood Cliffs, NJ, 1973).
[2] H. H. Kuehl, Phys. Fluids **5**, 1095 (1962).
[3] R. J. Vidmar, F. W. Crawford, and K. J. Harker, Radio Sci. **18**, 1337 (1983).
[4] K. C. Budden, *The Propagation of Radio Waves* (Cambridge University Press, Cambridge, 1985).
[5] G. G. Getmantsev, N. A. Zuikov, D. S. Kotik, L. F. Mironenko, N. A. Mityakov, V. O. Rapoport, Yu. A. Sazonov, V. Yu. Trakhtengerts, and V. Ya. Eidman, JETP Lett. **20**, 229 (1974).
[6] R. Barr and P. Stubbe, Radio Sci. **19**, 1111 (1984).
[7] P. Stubbe and H. Kopka, J. Geophys. Res. **82**, 2319 (1977).
[8] K. Papadopoulos, A. S. Sharma, and C. L. Chang, Comments Plasma Phys. Controlled Fusion **13**, 1 (1989).
[9] V. L. Ginzburg, *Propagation of Electromagnetic Waves in Plasmas* (Gordon & Breach, New York, 1964).
[10] K. Papadopoulos, H. B. Zhou, and A. S. Sharma, Comments Plasma Phys. Controlled Fusion **15**, 321 (1994).
[11] K. Papadopoulos, H. B. Zhou, and C. L. Chang, Geophys. Res. Lett. **21**, 1767 (1994).
[12] A. S. Kingsep, K. V. Chukbar, and V. V. Yankov, in *Reviews of Plasma Physics*, edited by M. A. Leontovich (Consultants Bureau, New York, 1990), Vol. 16, p. 243.
[13] R. A. Helliwell, *Whistlers and Related Ionospheric Phenomena* (Stanford University Press, Stanford, 1965).
[14] H. B. Zhou, "Dynamic response of a magnetized plasma to an external source: Application to space and solid state plasmas," Ph.D. thesis, 1994, University of Maryland, College Park, MD.
[15] R. L. Stenzel, J. M. Urrutia, and C. L. Rousculp, Phys. Plasmas B **2**, 325 (1993).
[16] C. L. Chang, V. Tripathi, K. Papadopoulos, J. Fedder, P. J. Palmadesso, and S. L. Ossakow, in *Effect of the Ionosphere on Radio Systems*, edited by J. M. Godman (U.S. Government Printing Office, Washington, DC, 1981), p. 91.
[17] R. Barr and P. Stubbe, J. Atmos. Terr. Phys. **46**, 315 (1984).
[18] C. L. Chang, A. S. Lipatov, A. T. Drobot, K. Papadopoulos, and P. Satyanarayana, Geophys. Res. Lett. **21**, 1015 (1994).
[19] D. Kahner, C. Moler, and S. Nash, *Numerical Methods and Software* (Prentice–Hall, Englewood Cliffs, NJ, 1989).

Appendix S

Direct Cerenkov Excitation of Waveguide Modes by a Mobile Ionospheric Heater

Direct Cerenkov Excitation of Waveguide Modes by a Mobile Ionopsheric Heater

N. Borisov**, A. Gurevich**,K. Papadopoulos* and C. L. Chang*

University of Maryland
College Park, MD 20742

and

Science Applications International Corporation*
McLean, VA 21102

and

**Lebedev Institute of Physics,
53 Leninsky Aveune, Moscow V-333, RUSSIA

3 February 1995

Abstract

A comprehensive theoretical analysis of direct Cerenkov excitation of the earth ionosphere waveguide using ionospheric heating is presented. The model relies in transient ionospheric heating with a heater spot moving horizontally at the bottom of the waveguide with speed close to the speed of light. The cases of isotropic ionospheric conductivity, corresponding to heating altitudes below 70 km, and of anisotropic conductivity, corresponding to heating higher altitude are examined separately. It is found that enhanced radiation coupling requires that the speed of the heater approaches the speed of light. For the anisotropic case such enhancement occurs independently of the direction of motion, while for the isotropic case motion parallel to the ambient electric field is required.

1. Introduction

The controlled generation of coherent low frequency waves in the ULF/ELF/VLF range using amplitude modulated ionospheric heating with an HF transmitter in the auroral zones, has been investigated extensively experimentally and theoretically [Getmantsev et al., 1974; Stubbe et al., 1981; Barr and Stubbe, 1984; Chang et al., 1981; Beljaev, 1987 and Papadopoulos et al., 1989]. The basic physics of the generation is as following. Absorption of the HF energy in the lower ionosphere modifies locally the ambient conductivity, leading to the generation of a polarization electric field and current inside the modified region. The surrounding magnetized plasma responds to the temporal modification by driving two field aligned currents carried by helicon or whistler waves [Papadopoulos et al., 1994] which close across the magnetic field by a Hall current. A current loop forms which expands upwards for the duration of the HF pulse. The loop damps quickly following the termination of the HF pulse. As a result amplitude modulated heating generates an equivalent horizontal magnetic moment M in the lower ionosphere given by

$$M = E_a \, \Delta\Sigma A \exp(i\omega_o t) \tag{1}$$

where E_a is the ambient electric field, $\Delta\Sigma$ the value of the modified conductance, A the area of the loop, and ω the frequency of the HF amplitude modulation. The power radiated by the oscillatory moment M couples to the waveguide formed by the conducting ground and the ionosphere and propagates over very large distances with small attenuation [Kotik et al., 1978; Papadopoulos et al., 1990]. It should be noted here that a current loop with magnetic moment in the opposite direction forms in the downward direction. However, the area of the loop is much smaller than A due to the larger collision frequency at the lower height, and its effect in the radiation can be neglected [Papadopoulos et al., 1993, 1994a].

The objective of this paper is to present a comprehensive analytical model of an alternative concept leading to injection of low frequency power in the Earth Ionosphere Waveguide (EIW) using controlled ionospheric heating. The concept is based on direct excitation of the waveguide by moving the ionospheric heater in the horizontal direction with a speed that matches the phase velocity of the desired waveguide mode. Its mathematical description follows the well known Cerenkov emission [Landau and Livshitz, 1960] which occurs when a charge moves in a medium with speed exceeding the phase speed of the waves the medium supports. The wave emission is due to the polarization electric field induced in the medium by the motion of the charge. As mentioned above ionospheric heating in regions penetrated by currents or electric fields, induces polarization fields and currents. One would, therefore, expect that polarization fields induced by a heater beam moving horizontally with speeds matching or exceeding the phase speed of the waveguide modes, will couple energy to the EIW by a process analogous to Cerenkov emission. The possibility for such an effect in the VLF range was previously noted by Kotik et al. [1986] and Borisov et al. [1991]. In fact, Kotik et al. [1986] performed an experiment in which they used a heater as an interferometer with a baseline d and frequency difference of the heating source Δf. By adjusting the values of d and Δf they created a supraluminal motion of the heated region. They noted that there was significant directional gain for frequencies between 8–10 kHz over the power produced by HF amplitude modulation in the absence of motion.

We present below a comprehensive theoretical analysis of Cerenkov excitation of the EIW by a moving ionospheric source. Emphasis is placed on the elucidation of the regimes of operation, the requirements for the validity of the theory and the expected observables. The aim of the study is to provide guidelines for the conduct of proof of principle experiments. This is particularly opportune since the opening of the HAARP heater in early 1995, will provide an

excellent facility for the conduct of such experiments [Papadopoulos et al., 1995]. The paper is organized as following. Section 2 presents a brief review of the fundamentals of HF heating of the lower ionosphere as a function of the heater parameters and the local ionospheric properties. The form of the polarization current generated by a moving heater is also described. Section 3 presents an analysis of the low frequency power coupled in the EIW by the moving polarization current. The EIW is modeled by a simple two layer model with a sharp boundary between the vacuum and the plasma. The heating is assumed to occur at the boundary. Two cases are examined separately. In the first, the layer boundary is taken much lower 70 km, where $\Omega_e << \nu_e$, resulting in isotropic conductivity. In the second, the layer boundary is taken above 70 km, where $\Omega_e >> \nu_e$ resulting in anisotropic conductivity. Section 4 examines the region of validity of the theory and its impact on the design of experiments. The last section summarizes the results and discusses some practical applications of the concept.

2. Ionospheric Heating Considerations

The auroral electrojet current system is driven by the solar wind flow past the geomagnetic field. The currents flow along the magnetic field and close by cross-field currents at altitudes between 60–110 km. For an ambient electric field $\underline{E}_a$ imposed by the solar wind across the magnetic field $\underline{B}_o$, there are two dominant currents, the Hall and Pedersen, given by

$$\underline{J}_H = -\frac{1}{4}\frac{\omega_e^2}{\Omega_e^2}\frac{\Omega_e}{1+\nu_e^2/\Omega_e^2}\frac{\underline{E}_a}{|\underline{B}_o|}\times\underline{B}_o \equiv \underline{\underline{\sigma}}_H\cdot\underline{E}_a \qquad (2)$$

$$\underline{J}_P = -\frac{1}{4\pi}\frac{\omega_e^2}{\Omega_e^2}\frac{\nu_e}{1+\nu_e^2/\Omega_e^2}\underline{E}_a \equiv \sigma_p\underline{E}_a \qquad (3)$$

The collision frequency ν_e is a function of the electron temperature T_e, which, in the parameter region of interest, can be approximated by

$$\nu_e(T_e) = \nu_{eo} \frac{T_e}{T_o} \tag{4}$$

where ν_{eo}, T_o are the ambient values. Under the action of HF radio waves with rms amplitude E_o, and frequency ω_o, the local electron temperature is given by [Gurevich, 1978]

$$\frac{\partial T_e}{\partial t} = \frac{1}{3} \frac{e^2 E_o^2 \nu_e}{m(\omega_o^2 + \nu_e^2)} - \nu_e \delta (T_e - T_o) \tag{5}$$

where δ is the average energy lost in one collision between electrons and neutrals. For strong heating with $T_e - T_o > T_o$ and assuming $\omega_o >> \nu_e$, Ω_e, the value of ν_e as a function of time following sudden switch-on of the HF power at t = 0, is given by

$$\nu_e/\nu_{eo} = \frac{1 + E_o^2/E_p^2}{1 + E^2/E_p^2 \exp\left[-\delta \nu_{eo} t (1 + E_o^2/E_p^2)\right]} \tag{6}$$

where E_p is the plasma field defined by [Gurevich, 1978]

$$E_p = \left(\frac{3 T_o \delta m \omega_o^2}{e^2}\right)^{1/2} \tag{7}$$

From eq. (6) follows that ν_e reaches a stationary value $\nu_e/\nu_o = 1 + \frac{E_o^2}{E_p^2}$ on a time scale $\tau \approx \left[\delta \nu_{eo}(1 + E_o^2/E_p^2)\right]^{-1}$. Notice that this times is shorter for $E_o/E_p > 1$ and approaches $1/\delta\nu_{eo}$ for $E_o/E_p < 1$. Following the end of the HF pulse the temperature returns to its ambient values on a time $1/\delta\nu_{eo}$.

We consider next an HF pulse whose energy is absorbed at an altitude z_o in the lower ionosphere over a region taken as a Gaussian with radial extent a and vertical extent Lz. If the HF beam remains on over a time scale $t \geq \tau$, the modification in the collision frequency will induce currents given by

$$\frac{\Delta J_H(\underline{r})}{|J_H|} = 1 - \frac{1 + \nu_{eo}^2/\Omega_e^2}{1 + (\nu_{eo}^2/\Omega_e^2) g^2} S(\underline{r}) \tag{8a}$$

$$\frac{\triangle J_p(\underline{r})}{|J_p|} = 1 - g\frac{1+\nu_{eo}^2/\Omega_e^2}{1+(\nu_{eo}^2/\Omega_e^2)g^2}S(\underline{r}) \tag{8b}$$

$$g \equiv 1 + \frac{E_p^2}{E_o^2}, \quad S(\underline{r}) = \exp\left(-\frac{r^2}{a^2}\right)\exp\left[-\frac{(z-z_o)^2}{L_z^2}\right] \tag{8c}$$

Consider next the case that the heater moves with velocity $\underline{u}$. In this case the modified current will be again given by eqs. (8), but with $S(\underline{r})$ replaced with $S(\underline{r} - \underline{u}t)$. These currents form the external source in the wave equation which excites the waveguide. We define thus as $\underline{J}_{ext}(\underline{r}, t)$

$$\underline{J}_{ext}(\underline{r},t) = \triangle\underline{J}_H(\underline{r}-\underline{u}t) + \triangle\underline{J}_p(\underline{r}-\underline{u}t). \tag{9}$$

3. Cerenkov Emission in an Isotropic Ionosphere

The physics of the interaction is best illustrated by considering the case of isotropic conductivity. In practice this is the case for heating altitude $z_o << 70$ km. corresponding to $\nu_e << \Omega_e$. We, furthermore, consider a simplified model with the Earth as an infinite conducting layer, followed by vacuum up to $z = z_o$. Above z_o dielectric constant is taken as homogeneous with

$$\epsilon = 1 - \frac{4\pi i\sigma_o}{\omega} = 1 - i\frac{\omega_e^2}{\omega\nu} \tag{10}$$

where ω_e is the plasma frequency. Since $\nu_e >> \Omega$

$$\underline{J}_{ext}(\underline{r},t) = \underline{J}_p\left(1-\frac{1}{g}\right)S(\underline{r}-\underline{u}t). \tag{11}$$

For $\underline{E} = E_o\hat{e}_y$ and $\underline{u} = u\,\hat{e}_y$, the wave equation for the TE and TM modes can be written as

$$\frac{d}{dz}\frac{k_o^2\epsilon}{k_o^2\epsilon - k_\perp^2}\frac{d}{dz}\psi_1 + k_o^2\epsilon\psi_1 = \frac{4\pi i\omega}{c^2}k_y J_o(z,k_\perp)\delta(\omega - k_y u) \tag{12a}$$

$$\left(\frac{d}{dz} - k_\perp^2\right)\ \psi_2 + k_o^2 \epsilon \psi_2 \ = \ \frac{4\pi i \omega}{c^2} k_x J_o(z, k_\perp)\delta(\omega - k_y u) \tag{12b}$$

$$J_o(z, k_\perp) \ \equiv \ E(\Delta\Sigma)\frac{a^2}{4\pi} \exp\left(-\frac{a^2 k_\perp^2}{4}\right)\delta(z - z_o) \tag{13a}$$

$$\Delta\Sigma \ = \ \Delta\sigma L_z \tag{13b}$$

Consider first the generation of TH waves. We look for solutions with

$$\psi_1 \ = \ b_1 \ \sin\sqrt{k_o^2 - k_\perp^2}\, z \qquad\qquad 0 < z < z_o \tag{14a}$$

$$\psi_2 \ = \ b_2 \ \sin\sqrt{k_o^2 \epsilon - k_\perp^2}\ \ (z - z_o) \qquad z \geq z_o \tag{14b}$$

From eqs. (12) and (14) the vertical component of the electric field on the ground $E_z(z = 0, y) \equiv E_z(0)$ is given by

$$E_z(z = 0) \ = \ i a^2 \int\int\int \frac{d\omega dk_x dk_y}{\epsilon(\omega) - \omega}$$

$$\frac{\sqrt{k_o^2 \epsilon - k_\perp^2}\ \Delta\,\Sigma\, k_y \delta(\omega - k_y V) e^{i[\omega t - (k_y y + k_x x)]}}{\sqrt{k_o^2 - k_\perp^2}\ \sin\sqrt{k_o^2 - k_\perp^2}\, z_o - i\frac{\sqrt{k_o^2\epsilon - k_\perp^2}}{\epsilon}\ \cos\sqrt{k_o^2 - k_\perp^2}\, z_o} \exp\left(-\frac{a^2 k_\perp^2}{4}\right) \tag{15}$$

The zeros of the denominator in eq. (15) correspond to the different modes of the waveguide.

It is well known that in the frequency range we are interested, upto several kHz, the main contribution to the field in the wave zone is due to the TH_o mode. For $\sigma/\omega >> 1$ the eigenvalue of this mode is determined by the equation

$$(k_o^2 - k_\perp^2)z_o + \frac{\omega}{4\pi\sigma}\sqrt{k_o^2\epsilon - k_\perp^2} = 0 \qquad (16)$$

After integration in (15) with respect to k_z and neglecting small terms, we obtain for the radiating field

$$E_z(z=0) = -\frac{i\pi^{1/2}E_a v^{1/2}}{\sigma^{1/2}z_o\beta\, c}\, a^2 \int dk_y e^{ik_y(y\pm\beta x - vt)}\sqrt{-ik_y}\ \Delta\Sigma\ \exp\left(-\frac{a^2k_\perp^2}{4}\right) \qquad (17)$$

where $\beta = \left(\frac{v^2}{c^2} - 1\right)^{1/2}$, and the signs $\pm$ before βx correspond to the regions $x > 0$ and $x < 0$. The meaning $\sqrt{-ik_y}$ is chosen in the same way as mentioned earlier (see eq. 14b). Assuming that the perturbation of the conductivity is in the horizontal plane we obtain for the radiating field

$$E_z(z=0) \approx -\frac{\sqrt{2\pi a}\ \Delta\Sigma_o v^{1/2}}{z_o\sigma^{1/2}c\beta}E_a\left[1.7\phi\left(\frac{3}{4};\frac{1}{2};\frac{\xi^2}{a^2}\right) - 2.6\frac{\xi}{a}\phi\left(-\frac{3}{4};\frac{3}{2};\frac{\xi^2}{a^2}\right)\right] \qquad (18)$$

Here $\xi = vt - y \mp \beta x$, the sign $\pm$ before βx corresponds to the regions $x > 0$ and $x < 0$, and $\phi(\alpha,\beta,x)$ is the confluent hypergeometric function. According to eq. (18) the radiation propagates in the horizontal plane as two temporally localized wave packets making angles $\alpha = \pm$ arctgβ to the y axis. In the limit $v >> c$ the field amplitude diminishes as a power law with increasing v. On the contrary radiation grows as $\beta \to 0$; that is if the speed v approaches the speed of light. In this case it is necessary to retain the small imaginary terms in the equation (16). The result is valid up to $\beta = \left(ca/\sigma z_o^2\right)^{1/4}$.

We consider next the case when the electric field $\underline{E}_a = \hat{e}_x E_a$ that is perpendicular to the direction of motion. In this case after a similar calculations we obtain for the field in the wave zone for |x| >> a as

$$E_z(z=0) \approx -\frac{\sqrt{2\pi a}\ \triangle\ \Sigma v^{12}}{z_o \sigma^{1/2} c} E_a \left[1,7\phi\left(\frac{3}{4};\frac{1}{2}-\frac{\xi^2}{a^2}\right) - 2.6\xi\phi\left(-\frac{3}{4};\frac{3}{2};\frac{\xi^2}{a^2}\right)\right] \qquad (19)$$

The main difference from the previous case is that E_z does not depend on β. As a result the field enhancement for $v \rightarrow c$ is absent. This implies that for isotropic media only the motion along $\underline{E}_a$ leads to enhanced radiation amplitude. According to eqs. (18), (19) the spectral amplitude diminishes for frequencies $\omega > 2c/a$. For $a \approx 20\text{km}$, and $v \approx c$ the characteristic frequency is 5 kHz and the main mode excited in the waveguide is the TH_o mode. Later on we consider only the contribution of this mode to the radiating field in the wave zone.

4. Cherenkov Radiation in Anisotropic Media

The anisotropy of the medium due to the external magnetic field makes the problem of ELF generation more complicated. For anisotropic media it is necessary to take into account the interaction of the waves with different polarization. In this case the external electric field $\underline{E}_a$ generates currents both in the direction and across the direction of the electric field. We write the external source as

$$\underline{J}_{ex} = \{\Delta\sigma_p E_a,\ \Delta\sigma_H E_a\} \qquad (20)$$

and proceed to investigate the solution of the wave equation

$$\nabla^2 \underline{E} - \frac{1}{c^2}\frac{\partial^2 \underline{E}}{\partial t^2} - \frac{4\pi}{c^2}\frac{\partial}{\partial t}\underline{\underline{\sigma}} \cdot \underline{E} = \frac{4\pi}{c^2}\frac{\partial}{\partial t}\underline{J}_{ex} \qquad (21)$$

The vacuum solution for the electric field is given again by eq. (12a). To solve eq. (21) for an anisotropic medium it is convenient to introduce instead of the components E_x, E_y their linear combinations corresponding to the ordinary and extraordinary modes,

$$\phi_1 = E_x + iE_y, \phi_2 = E_x - i\tilde{E}_y \tag{22}$$

where the $\sim$ denotes temporal transform. If we assume that the longitudinal component of theconductivity in the ionosphere $\sigma_{\parallel} = \omega_{pe}^2/4\pi\nu_e$ is infinite, we obtain for $\phi_{1,2}$

$$\frac{d^2\phi_1}{dz^2} + \left(k_o^2 - k_\perp^2 - \frac{4\pi i\omega\sigma_p}{c^2} + \frac{4\pi i\omega\sigma_H}{c^2}\right)\phi_1 + \frac{1}{2}(k_x + ik_y)\phi_2 = \frac{4\pi i\omega}{c^2}E_a(\triangle\sigma_p - i\triangle\sigma_H) \tag{23}$$

$$\frac{d^2\phi_1}{dz^2} + \left(k_o^2 - k_\perp^2 - \frac{4\pi i\omega\sigma_p}{c^2} + \frac{4\pi i\omega\sigma_H}{c^2}\right)\phi_2 + \frac{1}{2}(k_x + ik_y)\phi_1 = \frac{4\pi i\omega}{c^2}E_a(\triangle\sigma_p + i\triangle\sigma_H) \tag{24}$$

In the general case the system of eqs. (23–24) for $\sigma_p = \sigma_p(z)$, $\sigma_H = \sigma_H(z)$ can be solved only numerically. To obtain an approximate analytical solution we assume the following.

1. The Hall conductivity outside the waveguide is large enough so that, for the relevant frequencies ω and scales k, there is a small parameter

$$c^2|k_x \pm ik_y|^2/4\pi\omega\sigma_H << 1 \tag{25}$$

2. The thickness of the region where the source is located Lz is relatively thin

$$\left(4\pi\omega\sigma_H/c^2\right)^{1/2}L_z < 1 \tag{26}$$

3. Outside the waveguide the conductivities σ_H, σ_p change smoothly with height

$$\frac{4\pi\omega\sigma_H h_H^2}{c^2} > 1, \quad \frac{4\pi\omega\sigma_H h_p^2}{c^2} > 1 \ . \tag{27}$$

where $h_H = \left|\frac{d\ln\sigma_H}{dz}\right|^{-1}$, $h_p = \left|\frac{d\ln\sigma_p}{dz}\right|^{-1}$. We discuss the correctness of the above conditions for the frequency range of interest ($f \sim (1-5)$kHz). In the anisotropic ionosphere the formation of the upper boundary of the waveguide can be assumed as the height $z = z_o$ where the Hall conductivity becomes sufficiently large, $\sigma_H > \omega/4\pi$. For daytime conditions it corresponds to the heights $z_o \approx 65 - 70$km (depending on the frequency ω and parameters of the media). At such heights the Pedersen conductivity σ_p is considerably smaller than the Hall conductivity σ_H. The ordinary wave according to (24) is reflected to the Earth, while the extraordinary wave penetrates in the ionosphere. For frequencies $f \sim (1-5)$kHz and scales $a \sim (10-30)$km at heights $z > z_o$ condition 1 is fulfilled. This, to first order, allows us to split the equation for the ordinary and the modes.

The second condition demands that the effective thickness Lz where the source is localized, be smaller than the thickness of the skin depth. Otherwise the thickness of the source is limited to the skin depth. This condition allow us to use the approximation of source localized at a fixed height

$$J_{ext}(\omega, k, z) = \{\Delta\Sigma_p, -\Delta\Sigma_H\} E_o \delta(\omega - k_y v)\delta(z - z_o) \tag{28}$$

where $\Delta\Sigma_{p,H} = \int \Delta\sigma_{p,H}(z')dz'$. It should be mentioned that for frequencies of the order of several kHz the effective current becomes weaker than (28) due to screening outside the

waveguide. The third conditions allows us to use the approximation of geometric optics in computing the radiation outside the waveguide.

We choose for $\omega > 0$ as one of the solutions the function ϕ_1 which describes the wave propagating upward, and as the second solution the function ϕ_2, which corresponds to the attenuated wave. As a result

$$\phi_1 = \frac{b_1}{\sqrt{q_1}} \exp\left(-i\int_{z_o}^{z} q_1 dz\right), \quad \phi_2 = \frac{b_2}{\sqrt{q_2}} \exp\left(-i\int_{z_o}^{z} q_2 dz\right) \tag{29}$$

where $q_1 = \sqrt{\frac{4\pi\omega}{c^2}(\sigma_H - i\sigma_p)}$, $q_2 = \sqrt{\frac{4\pi\omega}{c^2}(\sigma_H + i\sigma_p)}$, and b_1, b_2 are arbitrary constants. Taking for $\omega < 0$ the meaning $\sqrt{\omega} = -i\sqrt{|\omega|}$ we find that if $\omega < 0$ the function ϕ_1 corresponds to the attenuated wave and ϕ_2 corresponds to the propagating upward wave.

We next consider the boundary conditions for the horizontal component at $z = z_o$. For this purpose we use equations similar to eq. (12) in which we take into account the tensor character of the dielectric permeability $\underline{\underline{\epsilon}}$

$$\begin{aligned}
&\frac{d}{dz}\frac{k_o^2\epsilon_{\|}}{k_o^2\epsilon_{\|} - k_\perp^2}\frac{d\psi_1}{dz} + k_o^2\epsilon_\perp\psi_1 - k_o^2\epsilon_{12}\psi_2 = \frac{4\pi i\omega}{c^2}E_a(k_x \triangle \sigma_p - k_y \triangle \sigma_p)\delta(\omega - k_y V), \\
&\left(\frac{d^2}{dz^2} - k_\perp^2\right)\psi_2^2 + k_o^2\epsilon_\perp\psi_2 - k_o^2\epsilon_{12}\psi_1 = \frac{4\pi i\omega}{c^2}E_o(k_x \triangle \sigma_p + k_y \triangle \sigma_p)\delta(\omega - k_y V)
\end{aligned} \tag{30}$$

Here $\epsilon_\perp$, $\epsilon_{\|}$ are the diagonal, and $\epsilon_{12} = -\epsilon_{21}$ the eff diagonal components of the tensor $\underline{\underline{\epsilon}}$. For the selected model the conductivities change abruptly at the upper boundary. We obtain with

the help of (30) the boundary conditions in the form

$$\frac{d\psi_1}{dz}|_{z_o+0} - \frac{k_o^2}{k_o^2 - k_\perp^2}\frac{d\psi_1}{dz}|_{z_o+0} = \frac{4\pi i\omega}{c^2}E_a(k_x \triangle E_p - k_y \triangle E_H)\delta(\omega - k_y v)$$

$$\frac{d\psi_2}{dz}|_{z_o+0} - \frac{d\psi_2}{dz}|_{z_o+0} = \frac{4\pi i\omega}{c^2}E_a(k_x \triangle E_p + k_y \triangle E_H)\delta(\omega - k_y v) \qquad (31)$$

At the same time the horizontal components E_x, E_y are continuous at the boundary $z = z_o$. Substituting in (31) the solution (29) for the field outside the waveguide and eliminating constants $b_{1,2}$ it is easy to obtain for the vertical component of the electric field the approximate expression

$$E_x(z=0) = -\frac{E_a}{z_o^2\sqrt{\sigma_p^2 + \sigma_H^2}}\int\int\int d\omega dk_z dk_y \delta(\omega - k_y v)e^{i\omega t - i(k_x x + k_y y)}$$

$$\frac{z_o(\triangle\Sigma_p k_y + \triangle\Sigma_H k_x)(iq_1 - q_2) - i2(k_x \triangle \Sigma_p + k_y \triangle \Sigma_H)}{k_o^2 - k_\perp^2} \qquad (32)$$

$$-\frac{iz_o(\triangle\Sigma_p k_x + \triangle\Sigma_H k_y)(iq_1 + q_2)}{k_o^2 - k_\perp^2}$$

After integrating eq. (32) with respect to ω and wave vectors $\underline{k}$ we obtain the expression for E_z in the wave zone. For $x > 0$ it takes the form

$$E_z(0) = \frac{\sqrt{\pi}(1+i)a^2E_o}{\beta z_o^2\sqrt{\sigma_p^2+\sigma_H^2}cv}\left(\frac{4v^2}{a^2(1+\beta^2)}\right)^{3/4}[(\Delta\Sigma_p^o+\beta\Delta\Sigma_H^o)$$

$$z_o\left(i\sqrt{\sigma_H-i\sigma_p}-\sqrt{\sigma_H+i\sigma_p}\right)-i(\beta\Delta\Sigma_p^o-\Delta\Sigma_H^o)z_o\left(i\sqrt{\sigma_H-i\sigma_p}+\sqrt{\sigma_h+i\sigma_p}\right)$$

$$\left[0.6\phi\left(\frac{3}{4};\frac{1}{2};-\frac{(\Delta t)^2v^2}{(1+\beta^2)a^2}\right)-0.9\frac{v\Delta t}{a(1+\beta^2)^{1/2}}\phi\left(\frac{9}{4};\frac{1}{2};-\frac{(\Delta t)^2v^2}{a(1+\beta^2)^2}\right)\right]$$

$$-i2(\beta\Delta\Sigma_p^o+\Delta\Sigma_H^o)\frac{cv}{a\sqrt{1+\beta^2}}e^{\frac{(\Delta t)^2v^2}{(1+\beta^2)e^2}} \tag{33}$$

Here $\Delta t = t\frac{y+\beta v}{v}$ and $\Delta\Sigma_{p,H}$ are the changes in the integrated Pedersen and Hall conductivities in the center of the heated spot.

It follows from eq. (33) that the emission becomes more effective when β approaches zero, that is if the speed of the heated spot approaches the speed of wave light. On the contrary for large β the emission strength is reduced. If the ionospheric conductivity is sufficiently large

$$\sigma_H z_o^2 >> ca \tag{34}$$

we obtain the approximate expression for the ELF field when $\beta << 1$ as

$$E_z(0) = \frac{2\sqrt{2\pi a}}{\beta z_o\sigma_H^{1/2}c^{1/2}}E_o(\Delta\Sigma_p^o+\beta\Delta\Sigma_H^o)$$

$$\left[0.6\phi\left(\frac{3}{4};\frac{1}{2};-\right)-0.9\frac{c\Delta t}{a}\phi\left(\frac{9}{4};\frac{3}{2};\frac{\Delta t^2c^2}{a^2}\right)\right] \tag{35}$$

The result (35) is valid if $\sigma_H > \sigma_p$ and σ is not too small

$$1 > \beta \gtrsim \beta_o = \left(\frac{ca}{z_o^2 \sigma_H}\right)^{1/2} \tag{36}$$

For smaller β it is necessary to retain dissipative terms in the denominator in (32) that influence the value of the residue. Estimates show that the maximum of radiation corresponds to $\beta \sim \beta_o$. Comparing eq. (35) with the result obtained in the previous section we find that in both cases similar intensification of the radiation exists for $\beta \to 0$. But in the case of isotropic media such an effect is absent if the vector of the external field $\underline{E}_a$ is orthogonal to the direction of the motion.

5. Concluding Remarks

We presented above an analysis of the excitation of the EIW by using an ionospheric heating source moving at the bottom of the waveguide with speed close to the speed of light. It should be noted that contrary to the amplitude modulation in which a monochromatic wave at the modulated frequency is generated, the Cerenkov emission generates broadband radiation, whose dominant frequency is controlled by the transit time of the heater over its horizontal size. The overall spectrum $E_z(\omega)$ is given by

$$E_z(\omega) = Ae^{j\frac{\omega}{c}(y+\beta x)} \left\{ \left(\frac{a\omega}{2c}\right)^{1/2} e^{-\frac{a^2\omega^2}{4c^2}} \phi\left(0; \frac{5}{4}; \frac{a^2\omega^2}{4c^2}\right) \right.$$

$$\left. -i0.4\frac{c}{a}\frac{\partial}{\partial\omega}\left[\left(\frac{a\omega}{2c}\right)^{7/2} e^{-\frac{a^2\omega^2}{4c^2}} \phi\left(1; \frac{11}{4}; \frac{a^2\omega^2}{4c^2}\right)\right]\right. \tag{37}$$

where $A = 0.6\left(\frac{2\pi a}{\sigma_H c}\right)^{1/2} \frac{\Delta\Sigma_p^o + \Delta\Sigma_H^o}{\beta_H z_o} E_a$, and $\phi(\alpha, \beta, z)$ is confluent hypergeometric function. It

follows from (37) that the intensity of the radiation quickly diminishes at large frequencies (ω > 2c/a). The maximum of the spectrum is at the frequency range

$$f \sim f^* = c/2\pi a \tag{38}$$

At the same time spectral amplitude according to eq. (37) changes slowly for small frequencies. It is interesting to find out how the spectral amplitude at $f = f^*$ depends on the radius a. Let us assume that the full power of the heated is constant. With the increase of a the heating of electrons falls $\sim a^{-2}$, hence the collision frequency of electrons also becomes smaller. So the increase of radius a in the case of weak heating with $\nu_e \sim T_e$ leads to considerable decrease in the amplitude of the spectral maximum.

In our analysis it was implicitly assumed that the source moves uniformly along the y-axis between $y = -\infty$ and $y = +\infty$. It is important to determine the effect on the excitation amplitude resulting from a finite size sweep with length Δy. The analysis of appendix I shows that the requirement for optimum Cerenkov emission is that the sweeping length L satisfy the inequality

$$\frac{L}{a} > \frac{1}{\beta^2} \tag{39}$$

From eqs. (35) and (36) and assuming $\beta \approx \alpha\beta_o$ with $\alpha \approx 2 - 3$, we find

$$\frac{E_z(z=o)}{E_a} \approx \frac{L_z \Delta \sigma_H}{c}$$

Taking $L_z \approx 3 - 5\,\text{km}$, $\Delta\sigma_H \approx 10^4\ \text{sec}^{-1}$ we find $\frac{E_z(z=o)}{E_a} \approx .2 - .5$. Since $E_a \approx 10 - 20\,\text{mV/m}$, fields of a few mV/m can be produced by Cerenkov emission. These far exceed the electric field values measured in experiments using amplitude modulation (Stubbe et al., 1981,). Before closing we should remark on two issues:

First, we should distinguish the Cerenkov process discussed here to the one discussed by Papadopoulos et al. (1994b). In the latter the excited modes were the eigenmodes of the

ionospheric plasma, the whistler and the helicon mode. The required speed of the current source was much smaller than the speed of light $(v/c \approx 0(10^{-2}))$. The excited waves propagate mainly upwards, while the excitation of the EI waveguide modes is indirect. The current Cerenkov scheme, discussed here, leads to direct excitation of the waveguide modes and requires source motion at speeds comparable to the speed of light.

Second, a proper analysis of the problem, requires consideration of the ionopsheric density and collision frequency profile. The equations describing the waveguide modes for exponential conductivity profiles including magnetization have been discussed in Greigfinger and Greigfinger (1979) and Tripathi et al. (19). We are currently using these techniques to examine the role of the inhomogeneities to the Cerenkov excitation. The analysis of the present paper should be considered a good analysis of the problem for sharp ionospheric density profiles, when the height at which the conduction current equals the displacement current, is to within a skin depth of the reflection height.

In concluding we must emphasize that by repetitive excitation of the pulses we can synthesize harmonic or other waveforms with the desirable low frequench characteristics.

Appendix

We examine here the restrictions imposed in our analysis by the fact we assumed a heated spot moving uniformly along the y axis from y = $-\infty$, to y + $+\infty$. To examine the effect of a finite sweepoing distance Δy we consider the following model. We assume that the heated spot is stationary for $t \leq t_1 = -\Delta t/2$ (the coordinate of the spot is $y_1 = \Delta y/2$). During the time interval $-\Delta t/2 < t < \Delta t/2$ the spot moves uniformly and rectilinearly along the y axis with the speed v. From $t = t_2 = -\Delta t/2$ the spot is again stationary and its center is localized $y_2 = \Delta y/2$. We perform the Fourier transformation using the above function y(t). We obtain that in the case of finite distance the δ–function in the equations (12) should be replaced by δ_ω

$$\delta_\omega = \frac{1}{2\pi} \frac{\sin(\omega - k_y v)\,\Delta t/2}{\omega - k_y v} + \mu \qquad \text{(A1)}$$

where the function μ takes into account the contribution of the limits of integration $\pm\Delta t/2$ and does not correspond to the propagating wave. As a result if $\sigma/\omega >> 1$ we obtain

$$E_z(0) = \frac{4\pi i}{c z_o} E_a \int\int\int d\omega dk_x dk_y \,\Delta\,\Sigma\sqrt{-ik_y}$$

$$\frac{\sin(\omega - k_y v)\frac{\Delta t}{2}}{2\pi(\omega - k_y v)} \times \frac{\sqrt{k_o^2\epsilon - k_y^2}}{\epsilon\omega} \frac{k_y}{k_o^2 - k_y^2} e^{-i(k_y y + k_x x - \omega t)} \qquad \text{(A2)}$$

We represent now the frequency of radiation ω in the form ω = kv + Ω where Ω is a small correction due to the finite distance of motion. Substituting this expression in eq. (41) of the main text we find

$$E(0) = \frac{\sqrt{\pi} i E_a}{c^{3/2} z_o \beta \sqrt{\sigma}} \int dk_y e^{-ik_y y} \triangle \Sigma \sqrt{-ik_y}$$

$$\int d\Omega \frac{\sin(\Omega \triangle t)}{\Omega} e^{i(k_y v + \Omega)t - i\sqrt{k_y^2 \beta^2 + \frac{k_y v \Omega}{c^2}}} \tag{A3}$$

The obtained above result (A3) is valid if the corrections to $k_x = \beta k_y$ is small. It implies that the condition

$$\Omega << c k_y \beta^2 \tag{4}$$

should be fulfilled. Taking for the estimations $\Omega << 2c/\triangle L$, $k_y \sim 1/a$, we receive the the condition for the validity of our results is given by

$$\triangle y >> a/\beta^2 \tag{5}$$

If the inequality (4) is fulfilled the phase front in (eq. A2) is approximately a plane wave propagating at the angle $\alpha = \frac{\pi}{2} - \beta$ to the y axis. In the opposite case $\triangle y << a/\beta^2$ it is easy to verify that the Cerenkov radiation is absent. So in the optimum conditions ($\beta << 1$) the distance $\triangle y$ of the heated spot motion, should be considerably larger than the characteristic wavelength $\lambda = \pi a$.

References

Barr, B., P. Stubbe, ELF and VLS radiation from the polar electrojet antenna, Radio Science, 19, 1111–1122, 1984.

Beljaev, P.P., D.S. Kotik, S.N. Mitjakov et al., Generation of the electromagnetic signals of combinative frequencies in the ionosphere, Radiophysics and Quantum Electronics, 30, 248–267, 1987.

Borisov, N.D., A.V. Gurevich and K. Papadopoulos, Generation of radiation by superluminal movement of heated region in the ionosphere, Proceedings of the III SUZDAL URSI symposium on modification of the Ionosphere by Powerful Radio Waves, Moscow, p. 212, 1991.

Getmantsev, G.G., N.A. Zujiov, D.S. Kotik et al., The discovery of combinative frequencies in the interaction of powerful HF radiation with ionospheric plasma, JEATP Lett, 20, 101–105, 1974.

Greigfiner, C. and P. Greigfinger, On the ionopsheric parameters which govern high–latitude ELF propagation in the earth–ionospheric wave guide, Radio Sci., 14, pp. 889–895, 1979

Gurevich, A.V., Nonlinear Phenomena in the Ionosphere, New York, Springe Verlag, 1978.

Kotik, D.C., L.F. Mironenko, S.N. Mitjakov, et al., About the possibility of formation of superluminal source of the Cherenkov radiation by Getmantsev effect. In modification of the ionosphere by powerful radio waves, Proceedings of the International Symposium, Suzdal, p. 79, 1986.

Kotik, D.C., S.V. Poljakovv and V.A. Jashnov, Excitation of the earth — Ionosphere waveguide by low frequency sources in an inhomogeneous ionosphere, Radiophysics and Quantum Electronics, 21, 659–663, 1978.

Landau, L.D. and E.M. Livshits, Electrodynamics of continuous media, Pergamon Press, Oxford, 1960.

Papadopoulos, et al., HAARP - Research and Applications, (in press), 1995

Papadopoulos, K., H. B. Zhou and A. S. Sharma, The role of helicons in magnetospheric and ionospheric physics, Comm. Pl. Phys., 15, 6, 321-337, 1994a.

Papadopoulos, K., H.B. Zhou and C.L. Chang, Cerenkov excitation of whistler/helicon waves by ionospheric HF heating, Geophys. Rev. Lett., 21, 1767-1770, 1994b.

Papadopoulos, K., C.L. Chang, P. Vitello and A. Drobot, On the efficiency of ionospheric ELF generation, Radio Science, 25(6), 1311-1320, 1990.

Papadopoulos, K., R.A. Shanny, L. Susman, M. Machina, and P. Stamboulis, Electrojet modulation ELF communications, AGARD Conference Proceedings No. 485, Editors J.E. Rasmussen, P.A. Kossey and T.B. Jones, 37A-1 — 37A-9, 1990.

Papadopoulos, K., A.S. Sharma and C.L. Chang. On the efficient operation of a plasma ELF antenna driven by modulation of ionospheric currents, Comments Plasma Phys. Controlled Fusion, 13, 1, 1-17, 1989.

Stubbe, P., K. Kopka and R.L. Dowden, Generation of ELF and VLF waves by polar electrojet modulation, experimental results, J. Geophys. Res., 11, 9073-9078, 1981.

Tripathi, V.K., C.L. Chang and K. Papadopoulos, Excitation of the earth-ionosphere waveguide by an ELF source in the ionosphere, Radio Sci., 17, pp. 1321-1326, 1982.

Appendix T

Excitation of ELF/VLF Waves in the Earth-Ionosphere Waveguide by a Moving Current Source

Excitation of ELF/VLF Waves in the Earth-Ionosphere Waveguide by a Moving Current Source

C.L. Chang and K. Papadopoulos
Science Applications International Corporation

ABSTRACT

The excitation of the long range TEM mode in the earth-ionospheric waveguide by a current source moving horizontally in the lower ionosphere has been studied. The moving current source is generated by a novel HF modification scheme that requires the horizontal sweeping of the HF heater beam. Motion of the current source in the ionosphere can excite plasma waves similar to the traditional Cerenkov excitation by a charged particle moving through a media. The excited waves may propagate into the earth-ionosphere waveguide and couple to the TEM waveguide mode at ELF/VLF frequency. The resulting radiation pattern of the TEM wave at the ground level peaks in the direction of the motion. Frequency spectrum, wave amplitudes, and power of injection are derived. Our analysis includes realistic effects such as anisotropic plasma and vertical density profile in the ionosphere.

Introduction

Excitation of ELF/VLF waves by ionospheric modification utilizing high power HF heating facilities have been studied both in theory [Getmantsev et al., 1974; Kotik and Trakhtengerts, 1975; Stubbe and Kopka, 1977; Chang et al., 1981; Tripathi et al., 1982] and in experiments [Ferraro et al., 1982; Stubbe, et al., 1982; Barr et al., 1984]. Successful excitation has been achieved by heating the lower ionosphere periodically with amplitude modulated HF waves. This technique of modification creates a spatially stationary but temporal oscillating current source in the HF modified region in the ionosphere, which then radiates like an antenna in the earth-ionospheric waveguide. The radiation frequency is determined by the combined period of the heating cycle of the HF waves and the cooling cycle of the ambient plasma.

Recent construction of a powerful HF heating facility in Alaska under the High Frequency Active Auroral Reaerch Program (HAARP) has opened the possibility to conduct the ionospheric modification and ELF/VLF excitation more ingeniously. The unique design of the HAARP heater will provide unprecedented HF transmitter power as well as flexibilities in antenna directionality and frequency tunability [Brandt and Kossey, 1993]. Based on the designed capability of the HAARP facility, Papadopoulos et al. [1994] has analyzed alternative schemes for exciting ELF/VLF waves in the lower ionosphere. Specifically, by chirping the frequency of the HF heater continuously during a pulse, the heated region can move vertically in real time, thus creating a current source with upward motion in the ionosphere. Such a moving source in the lower ionosphere can stimulate Cerenkov emissions of whistler/helicon waves at a well defined frequency with high efficiency. The emitted waves propagate upward and downward along the ambient magnetic field line, which results in power injections into the magnetosphere and the earth-ionosphere waveguide, respectively.

Similarly, a current source of horizontal motion may also be induced by sweeping the HF beam horizontally at a fixed HF frequency as being mentioned in the earlier paper [Papadopoulos et al., 1994]. Such maneuver may be readily accomplished by the envisioned heater facility under HAARP, which has superb steering agility in its design. The horizontal moving current source in the ionosphere can generate ELF/VLF waves similar to the source of vertical motion studied previously. A diagram to illustrate the physical situation is given in Figure 1. The problem of ELF/VLF excitation in the earth-ionosphere waveguide has two important aspects. The first is the study of ELF/VLF wave generation by a moving current source in the ionosphere. The second is to understand the propagation of the generated waves into the earth-ionosphere waveguide and

their coupling to the waveguide modes. In this paper, we present the analyses of both aspects in a unified mathematical framework. However, only the most important waveguide mode, namely the TEM wave, is considered. This is because the TEM wave at ELF/VLF frequency has very low spatial attenuation rate. Therefore, it is the single dominant wave at far field zone in the earth-ionospheric waveguide.

It is worth noting that the concept of generating ELF/VLF waves by a moving source in the earth-ionospheric waveguide was explored in the past by Kotik et al. [1986] and Borisov et al. [1991]. These works differ from the current analysis in one important aspect. They considered a two-layer waveguide model that consisted a uniform ionosphere on top of a vacuum region with abrupt transition, while the current study uses a continuous ionosphere profile. The obvious advantage of having a continuous profile is a more accurate description of the waveguide properties and the propagation characteristics of the waves in the ELF/VLF frequency domain. It will be demonstrated in this paper that the efficiency of the excitation scheme varies notably under various ionosphereic conditions (day/night, highly stimulated,etc.). Moreover, by inverting the process, the current analysis opens the possibility of probing the lower ionosphere by studying the propagation characteristics of ELF/VLF waves at different frequencies.

The paper is organized as follows. The first section is the introducation. The second section contains the basic wave equations in the earth-ionosphere waveguide with current source included. The third section contains the form of current source and the calculations of the TEM excitation. The last section contains discussions and interpretations of the results.

Basic Equations

There have been a number of studies to provide simple and approximate solutions for handling ELF/VLF propagation in the earth-ionosphere waveguide. Works by Booker and Lefeuvre [1977], Greifinger and Greifinger [1978,1979], and Booker [1980] all present reasonably simplified procedures for evaluating the effects of the ionosphere on the ELF/VLF propagation. These effects include magnetic field, plasma anisotropy, and plasma density profiles while horizontal stratification is assumed. We will adopt the formalism given in the paper by Greifinger and Greifinger [1979], henceforth referred to as GG, to develop the theoretical framework for the excitation of the TEM waveguide mode by a moving current source in the lower ionosphere.

Choose the geomagnetic field in the vertical (z) direction $\mathbf{B}_o = B_o \hat{\mathbf{e}}_z$ and a time dependence of the form $\exp(-i\omega t)$ for the wave fields. The wave equations governing the electromagnetic propagation

in the earth-ionosphere waveguide are the basic Maxwell's equations

$$\nabla \times \mathbf{E} = \mathrm{i}\frac{\omega}{c}\mathbf{B} \tag{1}$$

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}(\mathbf{J} + \mathbf{J}_s) - \mathrm{i}\frac{\omega}{c}\mathbf{E} \tag{2}$$

where $\mathbf{J}_s$ is the external current source, $\mathbf{J}$ is the wave-induced plasma current defined by the generalized Ohm's law

$$\mathbf{J} = \sigma_P \mathbf{E}_\perp + \sigma_H (\mathbf{E}_\perp \times \hat{\mathbf{e}}_z) + \sigma_Z E_z \hat{\mathbf{e}}_z \tag{3}$$

and σ_P, σ_H, and σ_Z are the Pedersen, Hall, and Parallel conductivities, respectively. In terms of plasma species, these conductivities can be expressed as

$$\sigma_P = \sum_j \left(\frac{nq_j^2}{m_j}\right) \frac{\nu_j}{\Omega_j^2 - (\omega + \mathrm{i}\nu_j)^2}$$

$$\sigma_H = \sum_j \left(\frac{nq_j^2}{m_j}\right) \frac{\Omega_j}{\Omega_j^2 - (\omega + \mathrm{i}\nu_j)^2}$$

$$\sigma_Z = \sum_j \left(\frac{nq_j^2}{m_j}\right) \frac{1}{\nu_j} \tag{4}$$

where n is the density, q is the charge, m is the mass, ν is the effective collision frequency, Ω is the cyclotron frequency including charge sign, j is the index of plasma species, and summation Σ is carried over all the species.

It is convenient at this point to introduce the scalar and vector potentials to replace the fields as

$$\mathbf{E} = -\nabla\phi + \mathrm{i}\frac{\omega}{c}A$$

$$\mathbf{B} = \nabla \times A \ . \tag{5}$$

Combination of Eqs. (1), (2), and (5) yields the governing wave equation of the form

$$\nabla(\nabla\cdot A) - \nabla^2 A - k_o^2 A = \frac{4\pi}{c}(\mathbf{J} + \mathbf{J}_s) + ik_o\nabla\phi \tag{6}$$

where $k_o = \omega / c$.

Generally speaking, the TEM wave is the lowest order of TM mode propagation in a vacuum transmission line. However, since the ionospheric plasma is magnetized and thus anisotropic, TEM propagation in the earth-ionosphere waveguide acquires a finite TE component. An effective representation is adopted in GG to express the TEM vector potential as the combination of a TM component A_z and a TE component u. Thus,

$$A = A_z\hat{e}_z + \frac{i}{k_o}(\nabla_\perp \times u\hat{e}_z) \tag{7}$$

Taking the Fourier transform in the plane perpendicular to z, and replacing the spatial operator by

$$\nabla \to \hat{e}_z\frac{\partial}{\partial z} + \nabla_\perp = \hat{e}_z\frac{\partial}{\partial z} + ik \tag{8}$$

the vector equation (6) can be decomposed into three coupled scalar equations in the forms of

$$\frac{\partial\phi}{\partial z} = ik_o A_z - \frac{k^2}{\frac{4\pi}{c}\sigma_Z - ik_o}A_z + \frac{4\pi}{c\left(\frac{4\pi}{c}\sigma_Z - ik_o\right)}J_{sz} \tag{9}$$

$$\frac{\partial A_z}{\partial z} = ik_o\phi - \frac{4\pi}{c}(\sigma_P\phi - \sigma_H u) - i\frac{4\pi}{k^2 c}k\cdot J_{s\perp} \tag{10}$$

and

$$\frac{\partial^2 u}{\partial z^2} - (k^2 - k_o^2)u = -i\frac{4\pi k_o}{c}(\sigma_H\phi + \sigma_P u) + \frac{4\pi k_o}{k^2 c}(k\times\hat{e}_z)\cdot J_{s\perp} \ . \tag{11}$$

where J_s now is the Fourier component of the external current source in (ω,k) space. Equations (9), (10), and (11) constitute the basic wave equations and k is the eigenvalue corresponding to a appropriate set of boundary conditions on the wave solutions at the ground (z=0) and at large z.

Excitation of the TEM Waveguide Mode

The TEM mode is the least-evanescent mode in the earth-ionosphere waveguide at the ELF/VLF frequency regime. Therefore, it is the only remaining mode at a horizontal distance of several times the height of the waveguide away from the current source. A modal calculation of the TEM excitation is performed in this paper based on a set of ionospheric profiles in a 3D geometry. Specifically, wave amplitudes and power flux at ground level are estimated according to an analytic procedure outlined in GG which produced excellent approximations to the TEM eigenvalue and the eigenfunctions in simple expressions. We note that the advantage of this procedure is its generality to handle a wide range of ionospheric conditions.

The conventional approach of ELF/VLF generation by modulating the ionospheric currents is the use of an amplitude or frequency modulated HF pulse [see references in the Introducation section]. This technique produces a current source that is stationary in space and oscillitory in time. An alternative scheme recently proposed by Papadopoulos et al.[1994] utilizes an unmodulated but moving current source of the form $S(\mathbf{x}-\mathbf{v}_o t)$, where $\mathbf{v}_o$ is the moving velocity and $\mathbf{x}$ is in the direction of motion. Such current source can be created in the lower ionosphere by sweeping the HF transmitter beam horizontally, thus changing the spatial location of the energy deposition continuously in time. A moving source in the ionosphere can induce whistler/helicon radiations by a process similar to the Cerenkov emission by a charged particle moving through a media. The following is an analytic study of exciting the TEM waves from such source. Special emphases have been on the excitation of plasma waves in the ionosphere and the propagation and coupling of these waves to the earth-ionosphere waveguide.

Moving Current Source

The physics of the TEM excitation can be illustrated by considering a three-dimensional finite-size current source of the form

$$\mathbf{J}_s = \frac{I_o L_z}{\pi L_x L_y} e^{-\left(\frac{x-x_o}{L_x}\right)^2} e^{-\left(\frac{y}{L_y}\right)^2} \delta(x_o - v_o t)\, \delta(z - h_s)\, \hat{\mathbf{e}}_x \qquad (12)$$

where h_s is the location of the source in altitude, x_o is the center point of the gaussian profile in x, v_o is the moving speed in the x direction, and L_x ,L_y ,and L_z are the source dimension in x, y, and z, respectively. Assuming that the HF heating occurs at altitudes around 80 km , the current

strength I_o was shown to be related to the change in Pedersen current [Papadopoulos et al., 1989] as

$$I_o = \frac{L_x L_z}{4\pi} \frac{\omega_e^2}{\Omega_e^2} \frac{(T_e - T_o)}{T_o} \nu_e(T_o) E_o \qquad (13)$$

where ω_e is the electron plasma frequency, Ω_e is the electron cyclotron frequency, T_e is the modified electron temperature, T_o is the ambient electron temperature before HF modification, ν_e is the electron collision frequency, and E_o is the ambient electric field induced by the solar wind flow. The current model expressed in equations (12) and (13) is an idealized situation for the modal calculation because the extension in z is compressed to a point representation by delta functions. Therefore, the total current moment I_o is a height-integrated value over the altitude. Such simplification is introduced to make the case analytically tractable. A complete analysis for general current shape in altitude will be conducted numerically and will be reported in a seperate publication.

Taking $\mathbf{k} = k\,\hat{\mathbf{e}}_x$, the Fourier analyzed source expression in the $(\omega,\mathbf{k})$ space can be expressed as

$$J_s = 2\pi I_o L_z \, e^{-\left(\frac{k_x L_x}{2}\right)^2} e^{-\left(\frac{k_y L_y}{2}\right)^2} \delta(\omega - k_x v_o)\, \delta(z - h_s)\, \hat{\mathbf{e}}_x \,. \qquad (14)$$

From this expression, it is clear that phase velocity of the exited TEM wave ω/k_x has to equal the source speed v_o.

Excitation in the Ionosphere

Consider the situation where ionospheric modification occurs in the 70-90 km altitude range. In this region the following condition exists

$$\sigma_Z >> \sigma_H \approx \sigma_P >> \omega \,. \qquad (15)$$

Furthermore, for TEM propagation $k \approx k_o$, equation (9) becomes

$$E_z = - \frac{\partial \phi}{\partial z} + ik_o A_z \equiv 0 \qquad (16)$$

which implies that the vertical electric field is shorted out by the large plasma conductivity along the

magnetic field. Utilizing this relation in Eqs. (10) and (11), and recognizing that $\partial/\partial z >> k$, we have

$$\frac{\partial^2\phi}{\partial z^2} + i\frac{4\pi k_o}{c}\left(\sigma_P\phi - \sigma_H u\right) = \frac{4\pi}{c}\frac{k_o}{k^2}\mathbf{k}\cdot J_{s\perp} \tag{17}$$

$$\frac{\partial^2 u}{\partial z^2} + i\frac{4\pi k_o}{c}\left(\sigma_H\phi + \sigma_P u\right) = \frac{4\pi}{c}\frac{k_o}{k^2}\left(\mathbf{k}\times\hat{\mathbf{e}}_z\right)\cdot J_{s\perp} \ . \tag{18}$$

By introducing the variables representing the O and X propagation

$$\psi_\pm \ = \ -\phi \pm iu \tag{19}$$

we can obtain two uncoupled equations

$$\frac{\partial^2\psi_\pm}{\partial z^2} + \frac{4\pi k_o}{c}\left(i\sigma_P \pm \sigma_H\right)\psi_\pm = \frac{4\pi}{c}\frac{k_o}{k^2}\left[-\mathbf{k} \pm i\left(\mathbf{k}\times\hat{\mathbf{e}}_z\right)\right]\cdot J_{s\perp} \tag{20}$$

Under nominal daytime conditions, the electron Hall conductivity dominates in the upper D and the Lower E region, i.e. $\sigma_H \approx \sigma_{He} > \sigma_P$. In this case the Pedersen conductivity can be neglected from (20). Since the electron Hall conductivity has negative value, the upper (lower) sign of $\pm$ in this equation corresponds to vertical O-ordinary (X-extraordinary) wave propagation. Thus, the O wave is evanecent and only the X wave can propagate in this region. To obtain approximate solutions to (20), an altitude h_1 is introduced in GG where the local wavelength in z becomes equal to the local scale length ζ_1 of the conductivity profile of σ_H. In the vicinity of h_1, the conductivity profile can be approximated by the exponential function

$$\sigma_H = -\left|\sigma_{H1}\right| e^{(z-h_1)/\zeta_1} \ . \tag{21}$$

The requirement that local wavelength and scale length matches at h_1 leads to the condition

$$\frac{16\pi k_o}{c}\left|\sigma_{H1}\right| \zeta_1^2 \ = \ 1 \tag{22}$$

The altitude h_1 can be determined from this condition provided that both ionospheric profile and ELF/VLF frequency are known. Physically, h_1 can be viewed as the reflecting point of the

ELF/VLF wave in the ionosphere. Above h_1, the ionosphere can be considered as a slowly varying medium to the X and O waves propagation. Due to the large value of σ_{He}, the O wave is nearly terminated within a few scale lengths and the phase propagation of the X wave is nearly along the magnetic field.

Substituting Eqs. (14), (21) and (22) into (20), we can obtain a simplified wave equation with the source term

$$p^2 \frac{\partial^2 \psi_\pm}{\partial p^2} + p \frac{\partial \psi_\pm}{\partial p} \mp p^2 \psi_\pm = \chi_{s\pm} \delta(p - p_s) \tag{23}$$

where

$$p = e^{(z-h_1)/2\varsigma_1} \; ; \; p_s = e^{(h_s-h_1)/2\varsigma_1} \tag{24}$$

and

$$\chi_{s\pm} = -\frac{(4\pi)^2}{c} \frac{k_o}{k^2} I_o \zeta_1 \, p_s L_z \left(k_x \mp k_y\right) e^{-\left(\frac{k_x L_x}{2}\right)^2} e^{-\left(\frac{k_y L_y}{2}\right)^2} \delta(\omega - k_x v_o) \; . \tag{25}$$

The homogeneous part of equation (23) is of the Bessel type [Abramowitz and Stegun, 1972]. Therefore, we can express the O wave solutions as

$$\begin{aligned} \psi_+ &= a K_o(p) \; ; && p > p_s \\ &= b I_o(p) + d K_o(p); && p < p_s \end{aligned} \tag{26}$$

where a, b, and d are constants to be determined, and I_o and K_o are the modified Bessel's function of the first and the second kind, respectively. As p increases, I_o increases and K_o decreases exponentially. Therefore, O wave solution above the current source becomes evanescent at large z, which satisfies the boundary condition. However, below the current source, both I_o and K_o are present and O wave can have finite amplitude. This implies that the excited O wave can tunnel into the earth-ionospheric waveguide, thus contributing in part to the coupling with the TEM wave.

Similarly, we can construct the solutions for the X wave excitation. Since the X wave propagates

in the ionosphere ,we can write

$$\begin{aligned} \psi_- &= e\, H_o^{(1)}(p) \;; && p > p_s \\ &= f\, H_o^{(2)}(p) + g\, H_o^{(1)}(p); && p < p_s \end{aligned} \tag{27}$$

where e, f, and g are constants to be determined, and $H_o^{(1)}$ and $H_o^{(2)}$ are the Hankel functions. The $H_o^{(1)}$ function is an upward propagating solution and the $H_o^{(2)}$ function is a downward propagating solution. For $p > p_s$, the X wave propagates only upward. This represents an open boundary at the top of the waveguide, or equivalently, an injection of X wave into the magnetosphere. For $p < p_s$, both upward and downward solutions are present. The amplitude f in (27) is the part of X wave being injected into the earth-ionospheric waveguide, thus directly contributing to the TEM excitation.

Amplitudes of the excited X and O waves can be evaluated from the inhomogeneous solutions of Eq. (25) that includes the source term. Specifically, wave solutions at the source location satisfy the following conditions

$$\psi_\pm(p) \quad \text{is continuous at } p = p_s$$

$$\left.\frac{\partial \psi_\pm}{\partial p}\right|_{p_s-\epsilon}^{p_s+\epsilon} = \mp \frac{\chi_{s\pm}}{p_s^2} \qquad \text{at } p = p_s \tag{28}$$

Applying these conditions to the O and X wave soultions in (26) and (27), we can obtain the wave amplitudes as

<u>O Wave</u>

$$\begin{aligned} b &= \frac{\frac{\chi_{s+}}{p_s^2} K_{os}}{I_{os} K_{os}^{\nabla} - I_{os}^{\nabla} K_{os}} = J_{o1} \\ d &= a - \frac{\frac{\chi_{s+}}{p_s^2} I_{os}}{I_{os} K_{os}^{\nabla} - I_{os}^{\nabla} K_{os}} = a + J_{o2} \end{aligned} \tag{29}$$

<u>X Wave</u>

$$\begin{aligned} f &= \frac{\frac{\chi_{s-}}{p_s^2} H_{os}^{(1)}}{H_{os}^{(2)} H_{os}^{(1)\nabla} - H_{os}^{(1)} H_{os}^{(2)\nabla}} = J_{x1} \\ & \qquad \frac{\chi_{s-}}{p_s^2} H_{os}^{(2)} \end{aligned} \tag{30}$$

where the subscript s on modified Bessel and Hankel functions indicates the functional values at p = p_s, and the superscript symbol ∇ stands for the differentiation (d/dp). It is interesting to note that we can optimize the excitation of the downward propagating X wave by judicious selection of source height h_s so that f is maximized. This is equivalent to maximizing all the Hankel function contributions in (30). The other effect that the source location h_s has on the excitation efficiency is the $(1/p_s)^2$ dependence in Eqs. (28) to (30). This factor suggests that the excitation efficiency decays exponentially if the current source is located above the reflection height h_1. The dependence of h_1 on various ionospheric conditions and ELF/VLF frequencies will be discussed later in the paper.

When σ_P is not negligible in comparing with σ_H, X and O waves can have different reflection height h_1 and scale length ζ_1 that differ only slightly. Under this circumstance, Eqs. (20) may be solved separately for each wave and corrections to the wave solutions retained. This general case was done in the paper by Booker [1980]. For simplicity, we will assume a single reflection height and scale length for both the X and O waves in our calculation.

TEM Waveguide Mode

A simple approximate expression for the TEM wave was obtained in GG extending from ground level up to an altitude h_o, which is the de facto bottom of the ionosphere. This altitude h_o is defined as the point where the parallel current matches the displacement current at the ELF/VLF frequency. Therefore, it is frequecny depedent. Below h_o, the ordering for the TEM wave is

$$\frac{\partial}{\partial z} >> k \approx k_o > \frac{4\pi\sigma_H}{c}, \frac{4\pi\sigma_P}{c}$$

and

$$A_z >> \phi, u . \qquad (30)$$

If the ground is considered as perfectly conducting, the appropriate boundary condition is u(z=0) = ϕ(z=0) = 0. Using the approximations in (30) and excluding the source current terms in Eqs. (9) to (11), the TEM solutions can be approximated as

$$A_z(z) = A_{zo} = \text{constant}$$

$$u(z) = \beta z$$

$$\phi(z) = i k_o A_{zo} \left(z - \int_0^z dz \frac{k^2}{\frac{4\pi}{c} i k_o \sigma_z(z) + k_o^2} \right) \qquad (31)$$

where A_{zo} and β are constants remain to be determined.

In the altitude range $z \geq h_o$, parallel conductivity of the plasma may be approximated by an exponential function in height

$$\sigma_z(z) = \frac{\omega}{4\pi} e^{(z-h_o)/\zeta_o} \qquad (32)$$

where ζ_o is the conductivity scale length at the altitude h_o, and $h_o >> \zeta_o$. Since $\sigma_z = \omega / 4\pi$ at $z = h_o$, parallel plasma current equals the displacement current at this point. At altitudes a few scale heights above h_o, the integrand of (31) becomes negligibly small due to the exponential rising of σ_z, thus a cut-off is in placed at $z = h_o$. Substituting (32) into (31), the approximate potential ϕ becomes

$$\phi(z) \approx i k_o A_{zo} \left[z - \left(\frac{k}{k_o} \right)^2 \left(h_o - i \frac{\pi}{2} \zeta_o + i \zeta_o e^{-(z-h_o)/\zeta_o} \right) \right] \qquad (33)$$

It is worth noting that for TEM wave, the horizontal magnetic field is essentially constant between the ground and h_1. However, as indicated by (16), the vertical electric field falls off rapidly above h_o as it is shorted out by large σ_z. Therefore, the horizontal energy flow associated with the TEM wave is essentially confined in the vaccum region $0 < z < h_o$.

Excitation of the TEM Wave

Below h_1, the excited X and O waves can propagate downward and couple to the TEM wave in the vacuum region. The coupling occurs in the region $h_o < z < h_1$ where the X and O wave solutions in (26) and (27) match the TEM solutions of (31) and (33). Taking the small argument expansion and keeping only the leading terms in Bessel and Hankel functions, we obtain

$$I_o(p) \approx 1$$

$$K_o(p) \approx -\frac{(z-h_1)}{2\zeta_1} - \gamma + \ln 2$$

$$H_o^{(1)}(p) \approx 1 + i\frac{2}{\pi}\left[\frac{(z-h_1)}{2\zeta_1} + \gamma - \ln 2\right]$$

$$H_o^{(2)}(p) \approx 1 - i\frac{2}{\pi}\left[\frac{(z-h_1)}{2\zeta_1} + \gamma - \ln 2\right] \qquad (34)$$

where γ is the Euler's constant. Based on these asymptotic formula and the wave amplitudes in (29) and (30), the approximate expressions of X and O waves in the altitude range $z \le h_1$ can be written as

$$\phi \approx -\frac{1}{2}\left\{J_{o1} + J_{x1} + J_{x2} + e - \left[a + J_{o2} + i\frac{2}{\pi}(J_{x1} - J_{x2} - e)\right]\cdot\left[\frac{(z-h_1)}{2\zeta_1} + \gamma - \ln 2\right]\right\}$$

$$u \approx -\frac{i}{2}\left\{J_{o1} - J_{x1} - J_{x2} - e - \left[a + J_{o2} - i\frac{2}{\pi}(J_{x1} - J_{x2} - e)\right]\cdot\left[\frac{(z-h_1)}{2\zeta_1} + \gamma - \ln 2\right]\right\} . \qquad (35)$$

Term by term comparison between these expressions and the TEM solutions in (31) and (33) yields the TEM wave amplitudes

$$ik_o A_{zo} = \frac{1}{2\zeta_1 \mathbf{D}}\left[-J_{o1} + (4J_{x1} + 2J_{o1})\frac{i}{\pi}\frac{h_1}{2\zeta_1}\right] \qquad (36)$$

$$\beta = \frac{1}{2\zeta_1 \mathbf{D}}\left\{-iJ_{o1} + (4J_{x1} - 2J_{o1})\frac{1}{\pi}\left[\frac{h_1}{2\zeta_1} - \left(\frac{k}{k_o}\right)^2\frac{\left(h_o - i\frac{\pi}{2}\zeta_o\right)}{2\zeta_1}\right]\right\} \qquad (37)$$

and

$$\mathbf{D} = \left(1 - i\frac{2h_1}{\pi\zeta_1}\right)\cdot\left[\frac{h_1}{2\zeta_1} - \left(\frac{k}{k_o}\right)^2\frac{\left(h_o - i\frac{\pi}{2}\zeta_o\right)}{2\zeta_1}\right] + \frac{h_1}{2\zeta_1} . \qquad (38)$$

Note that the factors γ and ln 2 are neglected in comparing with $h_1 / 2\zeta_1$ (>> 1). The quantity **D** in (38) is the dispersion for TEM propagation in the earth-ionosphere waveguide. Assuming $h_1 >> \zeta_1$, $h_o >> \zeta_o$, and letting **D** = 0, we can calculate the TEM wavevector to the first order as

$$\frac{k_{TEM}}{k_o} = \left(\frac{h_1}{h_o}\right)^{1/2} \left[1 + \frac{i\pi}{4}\left(\frac{\zeta_o}{h_o} + \frac{\zeta_1}{h_1}\right)\right] . \qquad (39)$$

This expression is essentially the same as the one given in GG. The real part of (39) corresponds to the phase velocity of the TEM propagation in the waveguide, which is

$$\frac{v_{ph}}{c} = \frac{k_o}{k_{TEM}} = \left(\frac{h_o}{h_1}\right)^{1/2} < 1 . \qquad (40)$$

The imaginary part of (39) corresponds to the attenuation rate of the TEM waves propagating along the earth-ionosphere waveguide

$$\mathrm{Im}\left(\frac{k_{TEM}}{k_o}\right) = \frac{\pi}{4}\left(\frac{h_1}{h_o}\right)^{1/2}\left(\frac{\zeta_o}{h_o} + \frac{\zeta_1}{h_1}\right) . \qquad (41)$$

This expression shows that the attenuation of TEM waves is attributed to two energy dissipation factors on the RHS of (41). The first factor is the plasma heating at the altitude h_o and the second factor is the convective loss due to X wave leakage at altitude h_1. The attenuation rate in decibels per megameter is

$$\alpha \approx 0.143\, f_o \left(\frac{h_1}{h_o}\right)^{1/2}\left(\frac{\zeta_o}{h_o} + \frac{\zeta_1}{h_1}\right) . \qquad (42)$$

where f_o is the ELF/VLF frequency in hertz.

We proceed to calculate the TEM field components at the ground level. At z = 0, the vertical electric field has the form

$$E_z(0) = \left(\frac{k}{k_o}\right)^2 i\, k_o\, A_{zo} \qquad (43)$$

where itallic letters indicate Fourier components. Substituting (36) into (43), and assuming $h_1 / 2\pi\zeta_1 >> 1$, we have

$$E_z(0) \equiv \frac{i}{\pi}\frac{h_1}{2\zeta_1}\left(\frac{k}{k_o}\right)^2 \frac{1}{2\zeta_1 \mathbf{D}}\,\frac{\chi_s}{p_s^2}\,\delta(\omega - k_x v_o)\left[4(k_x + ik_y)\bar{J}_{x1} + 2(k_x - ik_y)\bar{J}_{o1}\right] \qquad (44)$$

where

$$\chi_{s\pm} = \chi_s\left(k_x \mp ik_y\right)\delta(\omega - k_x v_o) \qquad (45)$$

and

$$J_{o1} = \frac{\chi_{s+}}{p_s^2}\bar{J}_{o1}\ ,\quad J_{x1} = \frac{\chi_{s-}}{p_s^2}\bar{J}_{x1}\ . \qquad (46)$$

The electric and magnetic fields in real space and time can be obtained by performing the inverse Fourier transform

$$\begin{bmatrix}\mathbf{E}\\ \mathbf{B}\end{bmatrix} = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} dk_x\ e^{ik_x x}\int_{-\infty}^{\infty} dk_y\ e^{ik_y x}\int_{-\infty}^{\infty} d\omega\ e^{-i\omega t}\begin{bmatrix}E\\ B\end{bmatrix}\ . \qquad (47)$$

Because of the delta function $\delta(\omega - k_x v_o)$ in (44), the ω integration in (47) can be easily carried out by replacing ω with $k_x v_o$. The integration of k_y is equivalent to a contour integration over the complex k_y space. Its value equals to the sum of residues at the TEM poles $\mathbf{D}(k_{yo}) = 0$. From (39), we have

$$k_{yo} = |k_x|(\lambda_R + i\lambda_I) \qquad (48)$$

where

$$\lambda_R = \left[\left(\frac{v_o}{c}\right)^2 \frac{h_1}{h_o} - 1\right]^{1/2} \qquad (49)$$

and

$$\lambda_I = \frac{\pi}{4}\left(\frac{v_o}{c}\right)^2 \frac{h_1}{h_o}\left(\frac{\zeta_o}{h_o} + \frac{\zeta_1}{h_1}\right)\frac{1}{\lambda_R}\ . \qquad (50)$$

After both ω and k_y integrations, the vertical electric field in real space and time can be shown to have the form

$$E_z(0) = -\frac{i}{2}\left(\frac{I_o L_z}{v_o}\right)\left(\frac{v_o}{c}\right)^2 \int_{-\infty}^{\infty} dk_x\ e^{ik_x x}\ e^{ik_{yo}|y|}\left(\frac{\zeta_1}{h_o}\right)\left(\frac{|k_x|}{\lambda_R p_s}\right)\cdot$$

$$\cdot\, e^{-\left(\frac{k_x L_x}{2}\right)^2} e^{-\left(\frac{k_{yo} L_y}{2}\right)^2} \left[4\left(1 + i(\lambda_R + i\lambda_I)\frac{|k_x|}{k_x}\right)\bar{J}_{x1} + 2\left(1 - i(\lambda_R + i\lambda_I)\frac{|k_x|}{k_x}\right)\bar{J}_{o1} \right] \qquad (51)$$

where $\underline{x} = x - v_o t$ is the coordinate in the rest frame of current source. Since $\omega = k_x v_o$, the k_x integration above is equivalent to an integration in frequency ω. As such, the integrand of the k_x integration contains the important information of the frequency spectrum of the ELF/VLF waves excited by the moving current source, to be measured at the ground level. We also note that the quantities h_o, h_1, ζ_o, ζ_1, and k_{yo} are frequency (and ionospheric profile) dependent, therefore, the actual k_x integration of (51) is more complicated than it appears to be.

Under some circumstances, the integration of (51) can be analytically estimated to yield an approximate $E_z(0)$. For instance, if the scale lengths and scale heights are weak functions of k_x (i.e. ω), the integrand of (51) is linearly proportional to k_x at small k_x and decays exponentially as $\exp[-(k_x L_x/2)^2]$ at large k_x. Therefore, the k_x spectrum is peaked at k_{xo} with a width of Δk_x, both can be estimated to be

$$k_{xo} = \frac{\sqrt{2}}{L_x} \;,\quad \Delta k_x = \frac{1}{L_x} \;. \qquad (52)$$

Applying the mean value theory, the k_x integration can be approximated by the width Δk_x times the integrand evaluated at k_{xo}. This gives an estimated electric field amplitude as

$$|E_z(0)| \approx \sqrt{2}\left(\frac{I_o L_z}{c L_x^2}\right)\left(\frac{v_o}{c}\right)\left(\frac{\zeta_1}{h_o}\right)\left(\frac{|G|}{\lambda_R p_s}\right) e^{-\frac{\sqrt{2}}{L_x}\lambda_I |y|}\, e^{-\frac{1}{2}\left[\left(\frac{\lambda_R L_y}{L_x}\right)^2 + 1\right]} \qquad (53)$$

where G is the height gain factor consisting mainly of Bessel's functions specified in (29), (30), and (46)

$$G = 4\left[1 + i(\lambda_R + i\lambda_I)\right]\bar{J}_{x1} + 2\left[1 - i(\lambda_R + i\lambda_I)\right]\bar{J}_{o1} \;. \qquad (54)$$

From (53), we can see that the spatial extend of $E_z(0)$ is confined in the y direction by the attenuation factor $\sqrt{2}\lambda_I / L_x$. Therefore, the excited ELF/VLF waves will peak in a channel along the direction of the motion. More accurate picture of the radiation field can be obtained by evaluating the spectral integration numerically. Such results will be presented in the next section.

Total Power of Injection

It is instructive to calculate the total power of TEM radiation injected by the moving current into the earth-ionospheric waveguide. Recognizing that the power injection is mainly in the x direction, the Poynting flux in that direction can be expressed as

$$S_x = - \frac{c}{4\pi} E_z^* B_y . \tag{55}$$

where * denotes compex conjugate quantity. The magnetic component B_y is related to the vertical electric field through the Maxwell's equation

$$\frac{\partial}{\partial z} E_x - \frac{\partial}{\partial x} E_z = - \frac{1}{c} \frac{\partial}{\partial t} B_y . \tag{56}$$

From (51), we know that the excited fields depend only on the moving coordinate $\underline{x} = x - v_o t$. This implies that the time differentiation and the space differentiation are interchangeable, i.e. $\partial/\partial t = - v_o \partial/\partial x$. Using this relation in (56), the B_y field can be expressed as

$$B_y = - \frac{c}{v_o} \left(E_z - \int dx \frac{\partial}{\partial z} E_x \right) , \tag{57}$$

which can be substituted directly into (55) and yields

$$S_x = \frac{c^2}{4\pi v_o} \left(|E_z|^2 - E_z^* \int dx \frac{\partial}{\partial z} E_x \right) . \tag{58}$$

Knowing the Poynting flux S_x, total power of injection in x can be estimated by integrating over the waveguide cross-section in y and z, which is

$$P_x = \int_{-\infty}^{\infty} dy \int_0^{\infty} dz \, S_x = \frac{c^2}{4\pi v_o} \int_{-\infty}^{\infty} dy \int_0^{\infty} dz \, |E_z|^2 . \tag{59}$$

We note that in deriving (59), the conditions $E_x(0)=0$, $E_z(\infty)=0$, and $\partial E_z/\partial z \approx 0$ for $z < h_o$ are used. Since the E_z field of the TEM wave is nearly constant at $z < h_o$, and is cut off rapidly by the ionosphere at h_o, the z integration in (59) equals $h_o |E_z(0)|^2$. Substituting (53) into (59) and

carrying out the y integration, we can express the total power of injection in x as

$$P_x \approx \frac{\sqrt{2}}{4\pi\lambda_I}\frac{I_o^2}{c}\left(\frac{h_o}{L_x}\right)\left(\frac{L_z}{L_x}\right)^2\left(\frac{v_o}{c}\right)\left(\frac{\zeta_1}{h_o}\right)^2\left(\frac{|G|}{\lambda_R}\right)^2 e^{-\frac{(h_s-h_1)}{\zeta_1}} e^{-\left[\left(\frac{\lambda_R L_y}{L_x}\right)^2+1\right]} \qquad (60)$$

where λ_R, λ_I, and G are defined in (49), (50), and (54), respectively. We remark that the current strength I_o, as given by (13), contains all the physics of plasma heating by the HF beam.

Profile Studies and Discussions

In this section, comprehensive studies of TEM waveguide characteristics and its excitation by a moving current source under various ionospheric conditions are presented. The changing ionosphere is exemplified in a set of background electron density profiles taken from the paper of Barr et al. [1984] as shown in Figure 2. Profiles 1-3 correspond to nightime conditions while 4-6 to daytime conditions. Ascending numbers represent low, moderate, and high auroral activity, respectively. Electron density of the daytime profiles is generally higher than that of the nightime. The notable exception is profile 4, which is comparable to profile 3 at or below 70 km but is markedly lower above 70 km. Sharp density increase of profile 4 at 80 km causes sudden shift in the waveguide parameters as will be shown below.

TEM Waveguide Characteristics

For a given ELF/VLF frequency, scale lengths ζ_1, ζ_o and scale heights h_1, h_o are calculated based on Eqs. (22) and (32), respectively. The scale height h_o is computed directly by matching the propagation frequency with the parallel conductivity. It signifies the bottom of the ionosphere, or more specifically, the point where ionospheric effect becomes important to the TEM propagation. The scale height h_1, on the other hand, is computed numerically by an iteration procedure that converges upon Eq. (22). It can be understood as the reflection point where the local TEM wavelength matches the scale length of the conductivity profile. Since most of the TEM energy is confined between h_1 and the ground, this scale height can be alternatively viewed as the top of the TEM waveguide.

In figure 3, the scale length ζ_o at h_o for various profiles is shown versus ELF/VLF frequency. We can see from this figure that the typical size of ζ_o is about 2 to 4 km. Since it is much shorter than the vertical TEM wavelength, the transition from vacuum region to the ionosphere is very sharp.

The frequency dependence of scale length ζ_1 at h_1 is shown in figure 4. In contrast to figure 3, ζ_1 varies considerably with frequency and ionospheric condition. While this scale length depends only weakly with frequency in profiles 2 and 3, it decreases in general with the increasing frequency in profiles 1, 5, and 6. In profile 4, the sudden jump of ζ_1 around 600 -700 Hz is due to the sharp increase of electron density at 80 km as mentioned before.

The scale heights h_0 and h_1 that define the TEM waveguide are plotted in figure 5 as functions of frequency. The lower set of curves corresponds to h_0 and the higher one corresponds to h_1. For a fixed frequency, both scale heights become lower with the increasing level of electron density. Therefore, TEM waveguide tends to contract during daytime/high auroral activity and expand during nightime/low auroral activity. Moreover, figure 5 shows that for a given ionospheric profile, h_0 moves upward and h_1 moves downward as the ELF/VLF frequency increases. Both scale heights coalesce at about 1 kHz, beyond which the TEM waveguide can be viewed as consisting mainly two distinct layers, the plasma layer sitting on top of a vacuum layer, seperating by a sharp interface. Therefore, only at high frequencies the simplified two-layer waveguide model applies [Borisov et al., 1991].

From earlier analysis, we can see that the current source is preferably situated in an altitude window between h_0 nad h_1. This is based on two considerations. First, the location of the current source should be below h_1 to ensure good coupling to the TEM waveguide. And secondly, from Eq. (13), appreciable current strength can only be achieved in a region of appreciable ambient plasma density, thus at a location above h_0. However, this altitude window narrows down quickly with the increasing ELF/VLF frequency according to figure 5. At or above 1 kHz, the optimal choice of the current source location is practically at the boundary of the lower ionosphere. This places an additional constraint to the HF modification procedure in order to achieve high efficiency of the ELF/VLF excitation.

In Figure 6, we plot the spatial attenuation rate of the TEM wave as defined in Eq. (42) versus the wave frequency for different ionospheric profiles. Typically, the attenuation rate ranges from a few dB per mega-meter at low ELF frequencies to ~ 100 dB/Mm at low VLF frequencies. Between 400 Hz and 1 kHz, the attenuation rate varies considerably with the ionospheric profile, especially during the nightime (curves 1-3). It shows a consistent increasing trend with the increasing auroral activity, or equivalently, the electron density in the lower D region. However, such trend is diminished once the electron density at 70 km exceeds 10^3 / c.c., as shown by the coalescence of daytime curves 5 and 6. This interesting feature opens the possibility of using the excitation scheme as a diagnostic tool to monitor the variation of electron population at the lower

ionosphere when the population level is low. For instance, a viable approach may be to scan the TEM frequency by changing the moving speed of the current source, and to measure the attenuation-frequency curve at various receiving stations. The empirical results may be further benchmarked with a set of well caliberated attenuation rate curves to construct the electron density profile regressively. This potential application could be an interesting alternative to the in-situ rocket measurements because of its ability to provide continuous coverage at all times.

Threshold on Sweeping Speed

In the previous section we have shown that X and O waves can be excited by a moving current source in the ionosphere. Frequency of the excited waves is related to the source speed and the extend of the source in the direction of motion by the delta function $\delta(\omega - k_x v_o)$ in Eq. (44), that is, $\omega = k_x v_o$. Physically, this can be understood as the transit time effect, i.e. a finite size source creates a pulse-like response as it moves past a fixed reference point in space. The time pulse can be made sharper, or its bandwidth shifted to higher frequencies, by increasing the speed v_o and/or reducing the scale length ~ $1/k_x$ of the current source. Therefore, the frequency content of the X and O waves can be readily tuned by controlling the sweeping speed of the HF heater and/or the size of the heated area.

In order to convert the excited waves into TEM propagation, their frequency and wavenumber must satisfy the TEM dispersion given by Eq. (40). By letting $\omega = k_x v_o$ and $\mathbf{k}_{TEM} = k_x \hat{\mathbf{e}}_x + k_y \hat{\mathbf{e}}_y$, Eq. (40) can be written as

$$\left(\frac{k_y}{k_x}\right)^2 = \left(\frac{v_o}{c}\right)^2 \frac{h_1}{h_o} - 1 \ . \qquad (61)$$

It is easy to see that if $v_o < c\sqrt{(h_o/h_1)}$, the above condition cannot be satisfied unless k_y/k_x is imaginary, which implies that the TEM wave is below the waveguide cutoff and can not propagate. Therefore, to excite the TEM waveguide mode efficiently the HF heater must sweep at a speed exceeding the threshold value, defined by

$$v_o = c\sqrt{\frac{h_o}{h_1}} \ . \qquad (62)$$

This threshold speed varies with the ELF/VLF frequency because both h_o and h are frequency

dependent. From Figure 5, $h_1 > h_o$ at low frequencies, the threshold v_o is below the speed of light. However, as frequency increases h_o approaches h_1, the threshold speed approaches the speed of light.

Amplitude of the TEM Waves and Power Injection

In Eq. (51), the vertical electric field at z=0 is transformed back into real space and time by integrating over the k_x spectrum. Since k_x and ω are interchangeable, the frequency spectrum of $E_z(0)$ can be readily obtanied from the integrand of Eq. (51) by replacing k_x with ω/v_o, provided that v_o is given. By the same token, the approximate peak and width of the frequency spectrum can be calculated from k_{xo} and Δk_x specified in Eq. (52). This is done in Figure 7, where the peak (horizontal bar) and the width (vertical bar) of the TEM frequency spectrum are evaluated versus the size of the current source assuming $v_o = c$. As an illustration to the transit time effect, this figure shows that a narrower source in motion generates a shorter time pulse of higher frequency content. Moreover, it shows that major portion of the ELF/VLF frequency spectra can be covered by varying the source size from a few km to hundreds of km. Therefore, frequency selection can be readily achieved by adjusting the antenna gain (i.e. the source size) of the HF heater, which is a relatively simple task for the HAARP facility.

The vertical electric field E_z at z=0 can be calculated according to Eq. (51). However, instead of performing the k_x integration directly, an approximate $|E_z(0)|$ is obtained in Eq. (53) assuming that the scale lengths and heights depend only weakly on k_x and $\underline{x} = 0$. We plot in Figure 8 the approximated E_z amplitude, as given by Eq. (53) with y = 0, versus the source size L_x for various ionospheric profiles. The current source in this case, as shown by Eqs. (12) and (13), has the following parameters: $L_y = L_x = L_z = 10$ km, $v_o = c$, $E_o = 10$ mV/m, $B_o = 0.35$ gausses, $T_e = 2T_o$ and $h_s = (h_o + h_1)/2$. We note that these numbers represent a very modest set of heating specs in the conventional practice of HF modification. Therefore, the $|E_z(0)|$ estimate presented in Figure 8 is a rather conservative one. From this figure, we can see that the electric field amplitude increases as the electron density (and the profile number) increases in the D region for all L_x. Moreover, $|E_z(0)|$ decreases with the increasing L_x, or with the decreasing ELF/VLF frequency since $\omega \sim 1/L_x$. In a highly activated ionosphere, such as the one represented by daytime profile 6, the amplitude of ground electric field can reach a high level of 10 mV/m at frequencies close to 1 kHz. Under more subdued ionospheric conditions, such as those of profiles 3 and 4, the electric field can be a fraction of 1 mV/m or so over the entire frequency range. Since the projected HAARP facility allows HF heating at a much higher level than what is considered in this study, it will be relatively easy to achieve the estimated results in a proof-of-principle experiment.

Analytic expression for the total power injected into the earth-ionospheric waveguide along x is given in Eq. (60). Using the same heating parameters, we plot in Figure 9 the total power injection as a function of L_x under various ionospheric conditions. These power curves exhibit a similar profile dependence as the ground amplitude in the previous figure. For instance, the total injection power increases with the increasing ionospheric electron density and, for a fixed profile, it decreases with the increasing L_x, or equivalently, the decreasing ELF/VLF frequency. The power level ranges from hundreds of Watts under highly activated situation to a few Watts under nominal conditions. Excitation efficiency is typically the lowest for profile 1, corresponding to nighttime low level of auroral activity. In this case, the power level remains below 0.1 Watts for the entire frequency spectrum.

Radiation Pattern

An important question to ask is how the excited TEM fields are distributed in the earth-ionospheric waveguide, especially at the ground level where measurements are made. To address this question, it is necessary to construct the $E_z(z=0)$ profile on a 2D x-y plane by performing the spectral integration in Eq. (51). One of such example is presented in Figure 10 based on ionospheric profile 5, same set of parameters used in Figure 8, and a fixed source height at $h_s = 65$ km. The k_x integration in Eq. (51) is carried out in the frequency range between 0 to 10^4 Hz. In Figure 10, the E_z amplitude is normalized to its maximun value, and both $\underline{x}$ (= $x-v_0t$) and y axis are normalized to L_x. From this figure, we can see that the TEM field is peaked directly below the current source at $\underline{x} = 0$. From this peak, the E_z amplitude falls off both in the $\underline{x}$ and the y directions. The amplitude decreases monotonically in the y direction because of the attenuation factor $\exp(-\sqrt{2}\lambda_I y/L_x)$ in Eq. (53). However, the amplitude decays in an oscillatory fashion in the x direction and remains finite along the $\underline{x}$ axis to a long distance. Therefore, the radiation pattern suggests a highly directional TEM propagation. We can view the TEM waves as being injected by the moving current source mainly along the direction of motion, i.e. x direction in this case.

Since the radiation pattern is plotted in a reference frame moving together with the current source, for a ground based instrument situated at $x > 0$ and $y = 0$, the field pattern appears moving toward it initially at $t = 0$, reaching its peak value when $x-v_0t = 0$, and falling behind the peak at later time when $x-v_0t < 0$. Therefore, the stationary instrument will register a transient E_z pulse due to the transit time effect. Field strength will be maximal if the instrument is located along the line of sight defined by the source motion.

It is intersting to point out that at the threshold v_o as defined by Eq. (62), $\lambda_R = 0$ and λ_I is singular.

For v_o close to the threshold, $\lambda_I >> \lambda_R$, and the Ez field can attain very large value due to the factor $\exp[-(k_{yo}L_y/2)^2]$ in Eq. (51). Since λ_I is singular, such high field can only exist on the x axis. Moreover, for $v_o \leq c$, it is always possible to find a frequency (or k_x) that λ_I becomes singular. However, it remains to be seen that this singularity can contribute significantly to the field because in reality the singular k_x can be quite large, so its importance is diminished by the $\exp[-(k_xL_x/2)^2]$ dependence in the frequency spectrum.

Summary

The novel concept of ELF/VLF generation in the earth-ionospheric waveguide by a moving current source produced by sweeping the HF beam horizontally across the lower ionosphere has been studied analytically using a three dimensional model. To facilitate the design of a proof-of-principle experiment, this concept has been tested with a set of realistic plasma profiles encompassing a whole range of ionospheres varying from undisturbed nightime to moderately disturbed daytime conditions. Realistic features such as plasma anisotropy and density variation in height are included in the analysis. Computations have been made first to specify the waveguide properties of the TEM propagation under different ionospheres; second to estimate the field strength and the excitation power in the waveguide based on a modal source current in the ionosphere; and finally to construct the two dimensional TEM field pattern on the ground.

In general it has been demonstrated that the excitation by a moving current source can be as efficient as the conventional method of current modulation, depending on the ionosphere profile. Plots of waveguide excitation with frequency over the range 100 Hz to 10 kHz have shown that such excitation produces a broad spectrum of ELF/VLF waves in the waveguide. Frequency bandwidth is shifted upward for a faster moving source or for a smaller source moving at a fixed speed, which is readily understood as the transit time phenomenon. Generally speaking, field strengths produced in the waveguide have been found to increase with increasing electron density. The excitation has also been shown to be more efficient at higher frequencies. A more elaborate estimate of the field pattern has shown that the radiation field is highly directed toward the direction of source motion. This is markedly different from the conventional excitation scheme using a stationary but oscillating current source, which produces an isotropic radiation pattern on the ground. As such, this method of excitation can be utilized to inject ELF/VLF waves into the earth-ionospheric waveguide and selectively boost the radiation power in desired directions.

References

Abramowitz, M., and I. A. Stegun, Bessel functions of integer order, Chapter 9 in Handbook of Mathematical Functions, Dover Publications, New York, 1972.

Barr, R., and P. Stubbe, ELF and VLF radiation from the "polar electrojet antenna", Radio Sci., 19, pp. 1111-1122, 1984.

Booker, H. G., and F. Lefeuvre, The relation between ionospheric profiles and ELF propagation in the earth-ionosphere transmission line, J. Atmos. Terr. Phys., 39, pp. 1277-1292, 1977.

Booker, H. G., A simplified theory of ELF propagation in the earth-ionosphere transmission line, J. Atmos. Terr. Phys., 42, pp. 929-941, 1980.

Borisov, N. D., A. V. Gurevich, K. Papadopoulos, Generation of Radiation by Superluminal Movement of Heated Region of the Ionosphere, Proc. of the III Suzdal URSI Symposium on Modification of the Ionosphere by Powerful Radio Waves, USSR Academy of Science, p. 212, 1991.

Brandt, R. G., and P. A. Kossey, The HAARP high-latitude research facility, EOS Transactions, AGU, 74, SA11A-2, invited paper, Fall meeting, 1993.

Chang, C.L., V. K. Tripathi, K. Ko, K. Papadopoulos, J. Fedder, P. Palmadesso, and S. Ossakow, Wireless generation of ELF/VLF radiation in the ionosphere, in Effects of the Ionosphere on Radiowave Systems, edited by J. M. Goodman, pp. 91-99, U.S. Government Printing Office, Washington, D.C., 1981.

Ferraro, A. J., H. S. Lee, R. Allshouse, K. Carroll, A. A. Tomko, F. J. Kelly, and R. G. Joiner, VLF/ELF radiation from the ionospheric dynamo current system modulated by powerful HF signals, J. Atmos. Terr. Phys., 44, pp. 1113-1122, 1982.

Getmantsev, G. G., N. A. Zuikov, D. S. Kotik, L. F. Mironenko, N. A. Mityakov, V. O. Rapaport, Yu. A. Sazonnov, V. Yu. Trakhtengerts, and V. Ya. Eidman, Combination Frequencies in the interaction between high power short-wave radiation and ionospheric plasma, JETP Lett., Engl. Transl., 20, pp. 101, 1974.

Greifinger, C., and P. Greifinger, On the ionospheric parameters which govern high-latitude ELF propagation in the earth-ionospheric waveguide, Radio Sci., 14, pp. 889-895, 1979.

Greifinger, C., and P. Greifinger, Approximate method for determining ELF eigenvalues in the earth-ionosphere waveguide, Radio Sci., 13, pp. 831-837, 1978.

Kotik, D. S., and V. Yu. Trakhtengerts, Mechanism of excitation of combination frequencies in ionospheric plasma, JETP Lett., Engl. Transl., 21, pp. 51, 1975.

Kotik, D. S., L. F. Mironenko, S. N. Mityakov, Modification of the Ionosphere by Powerful Radiowave, Trans. of the International Symposium, p. 91, 1986.

Papadopoulos, K., H. B. Zhou, and C. L. Chang, Cerenkov excitation of whistler/helicon waves by ionospheric HF heating, Geophys. Rev. Lett., 21, pp. 1767-1770, 1994.

Stubbe, P., and H. Kopka, Modulation of the polar electrojet by powerful HF waves, J. Geophys. Res., 82, pp. 2319-2325, 1977.

Stubbe, P., H. Kopka, M. T. Reitveld, and R. L. Dowden, ELF and VLF wave generation by modulated HF heating of the current carrying lower ionosphere, J. Atmos. Terr. Phys., 44, pp. 1123-1136, 1982.

Tripathi, V. K., C. L. Chang, and K. Papadopoulos, Excitation of the earth-ionosphere waveguide by an ELF source in the ionosphere, Radio Sci., 17, pp. 1321-1326, 1982.

Figure Captions

Figure 1. A diagram to illustrate the excitation of ELF waves by sweeping the heater beam, thus creating a current source moving horizontally in the lower ionosphere. The downward propagating ELF waves can couple to the long range TEM waves of the earth-ionospheric waveguide.

Figure 2. Electron density profiles in the ionosphere as given by Barr and Stubbe [1984]. The solid curves as marked by 1-3 are nightime profiles and the dashed curves are daytime profiles. The increasing number represents the increasing level of auroral activities.

Figure 3. Scale lengths at the bottom of the ionosphere h_o are plotted as functions of ELF/VLF frequency for various ionosphere profiles.

Figure 4. Scale lengths at the top of the TEM waveguide h_1 are plotted as functions of ELF/VLF frequency for various ionosphere profiles.

Figure 5. Characteristic heights of the TEM waveguide versus the ELF/VLF frequency for various ionosphere profiles.

Figure 6. Attenuation rates (in dB/Mega-meter) of TEM propagation in the earth-ionospheric waveguide as functions of ELF/VLF frequency and ionosphere profiles.

Figure 7. Frequency spectrum of the ELF/VLF waves excited by a moving current source in the ionosphere as a function of the scale length (L_x) of the source in the direction of motion.

Figure 8. Amplitude of the vertical electric field at z=0 and y=0 as excited by a moving current source at $v_o = c$, $h_s = (h_o + h_1)/2$, and of size $L_y = L_x$.

Figure 9. Total power injected by the moving current source into the waveguide as a function of L_x and ionosphere profiles.

Figure 10.Field pattern of TEM radiation on the ground is plotted in the ($\underline{x}$, y) plane. The field amplitude is calculated based on ionosphere profile 5. Field amplitude is normalized to its maximum value at $\underline{x}$=0 and y=0.

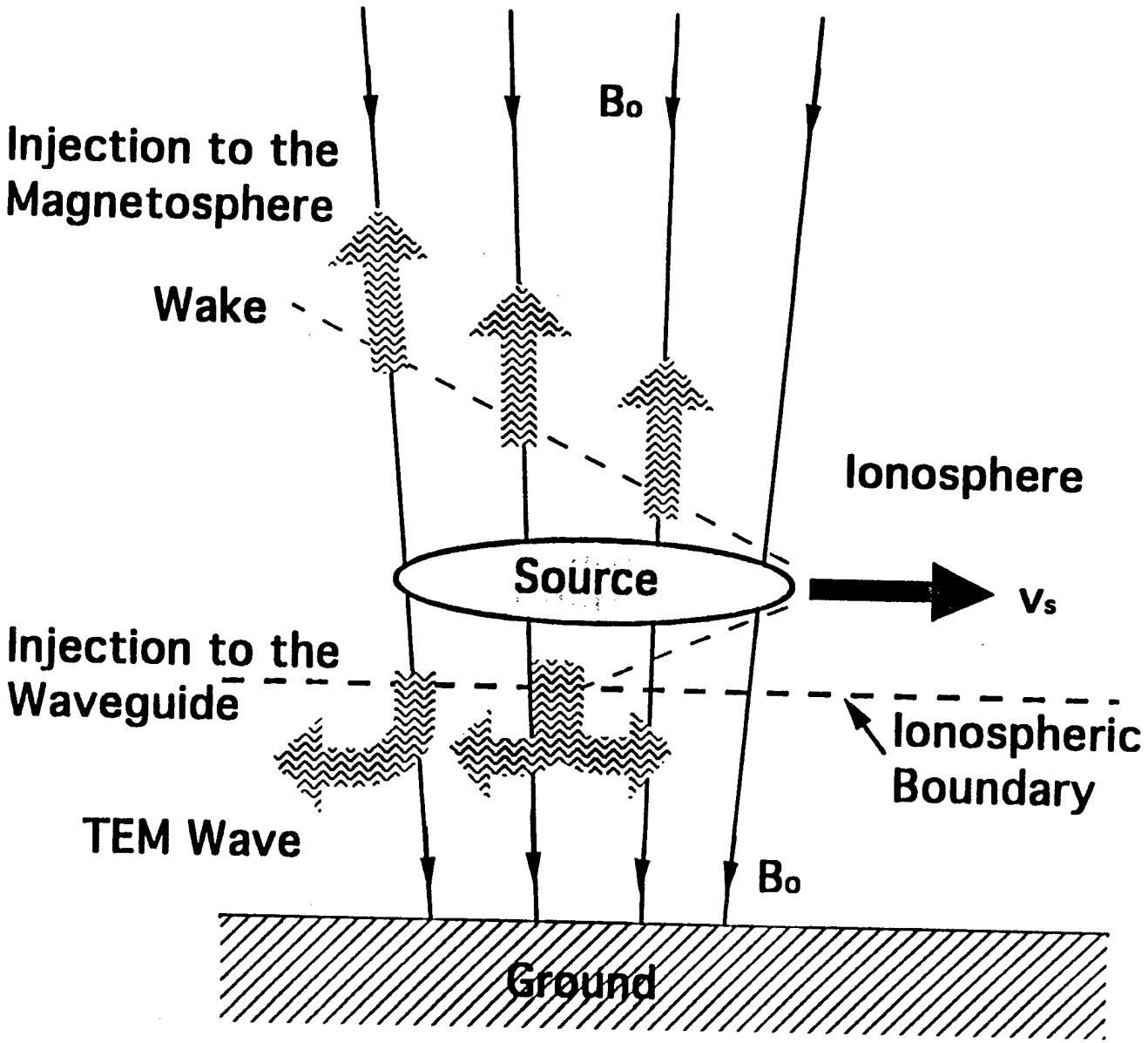

Injection to the
Magnetosphere
Wake
Bo
Ionosphere
Source
Vs
Injection to the
Waveguide
Ionospheric
Boundary
TEM Wave
Bo
Ground

FIGURE 1

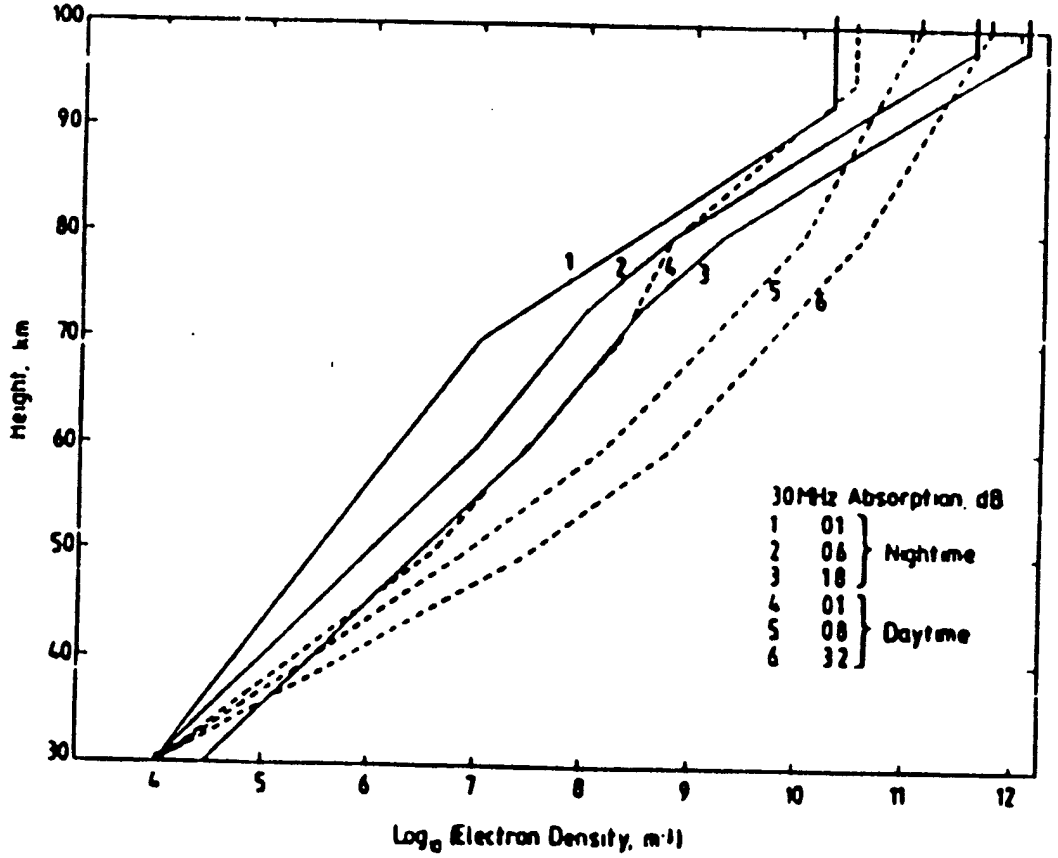


Fig. 2. Electron density profiles.

FIGURE 2

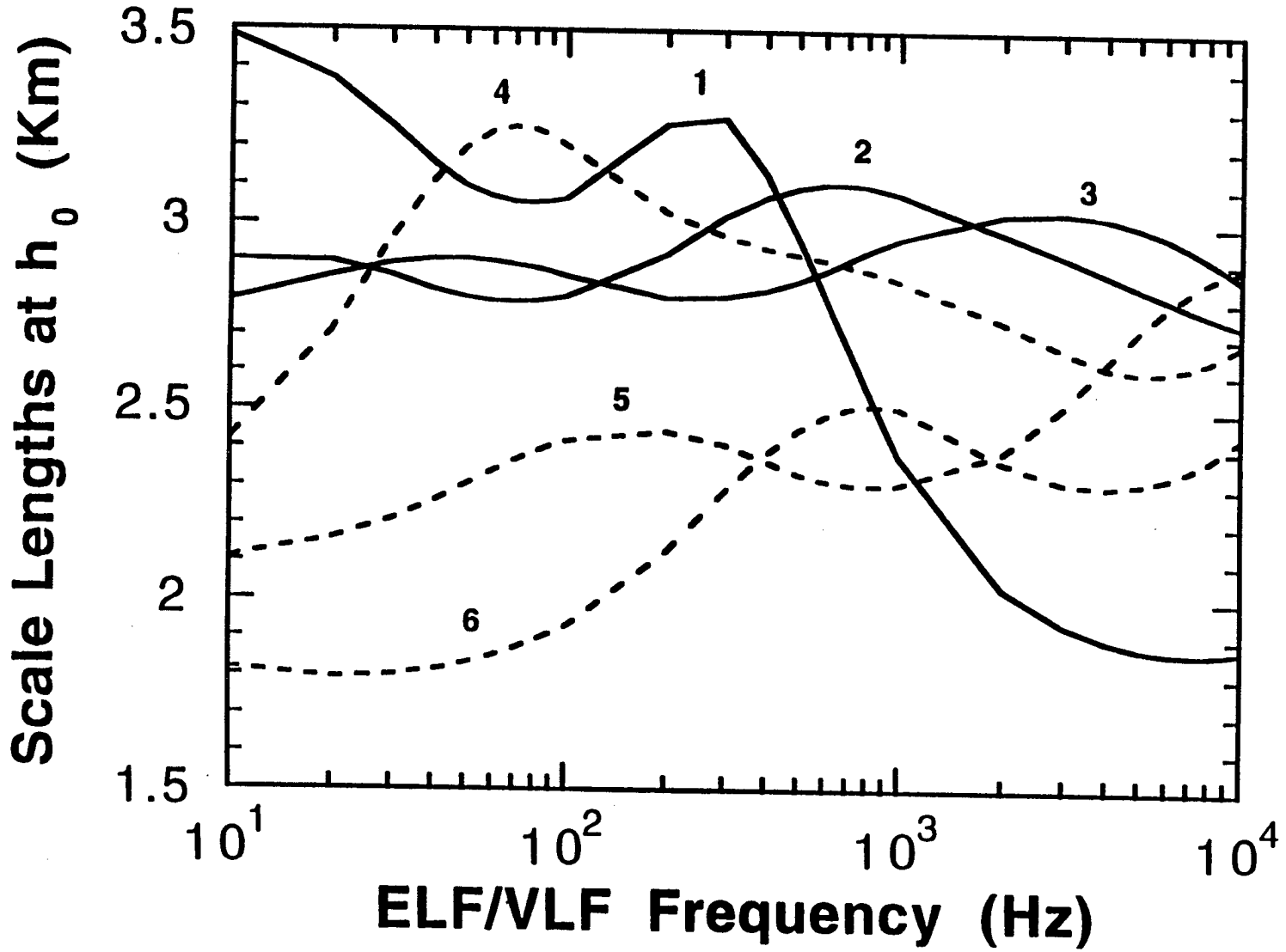

Scale Lengths at h_o (Km)
3.5
3
2.5
2
1.5
10^1
10^2
10^3
10^4
ELF/VLF Frequency (Hz)
1
2
3
4
5
6

FIGURE 3

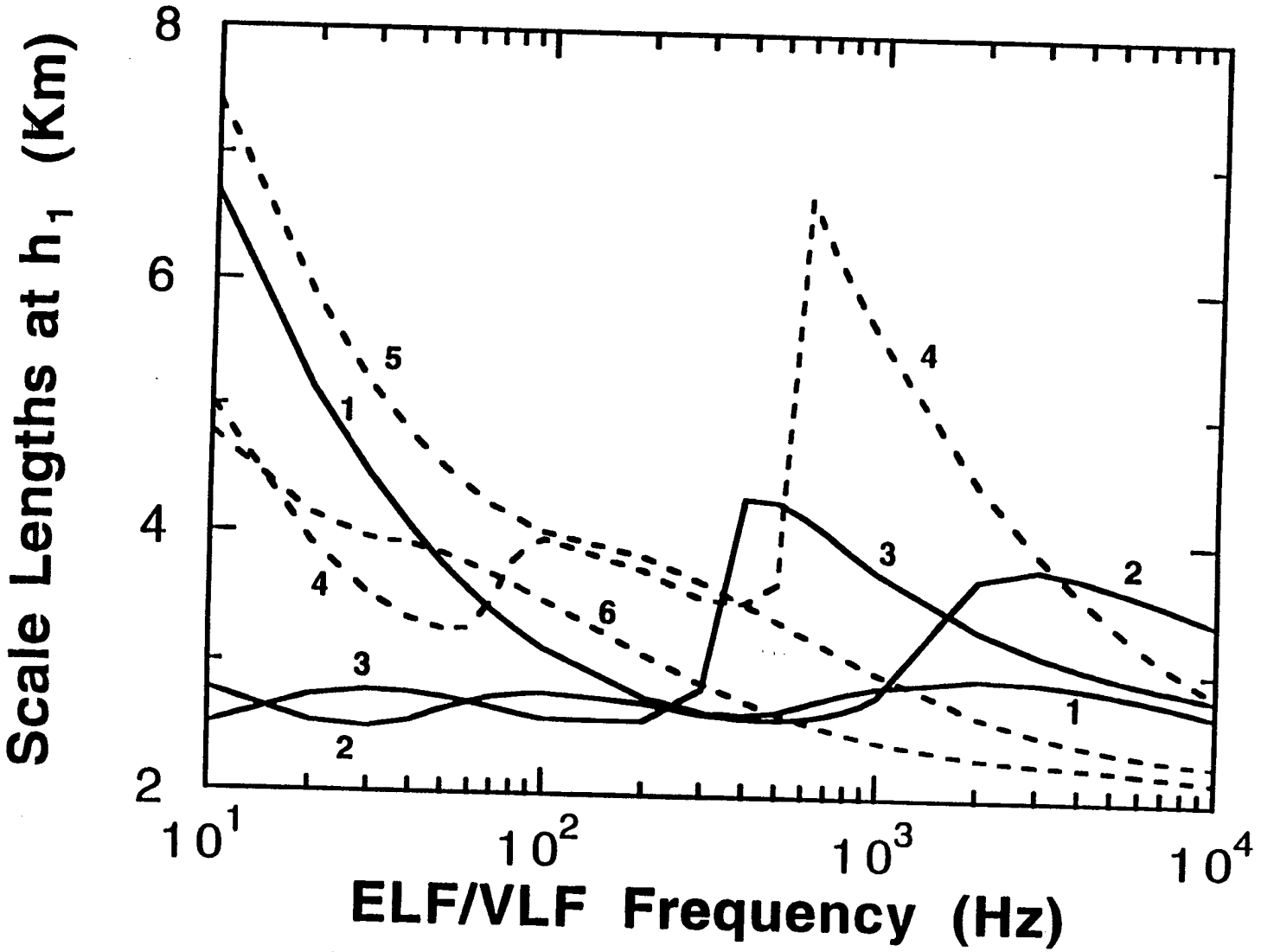

Scale Lengths at h_1 (Km)
8
6
4
2
10^1
10^2
10^3
10^4
ELF/VLF Frequency (Hz)
1
2
3
4
5
6

FIGURE 4

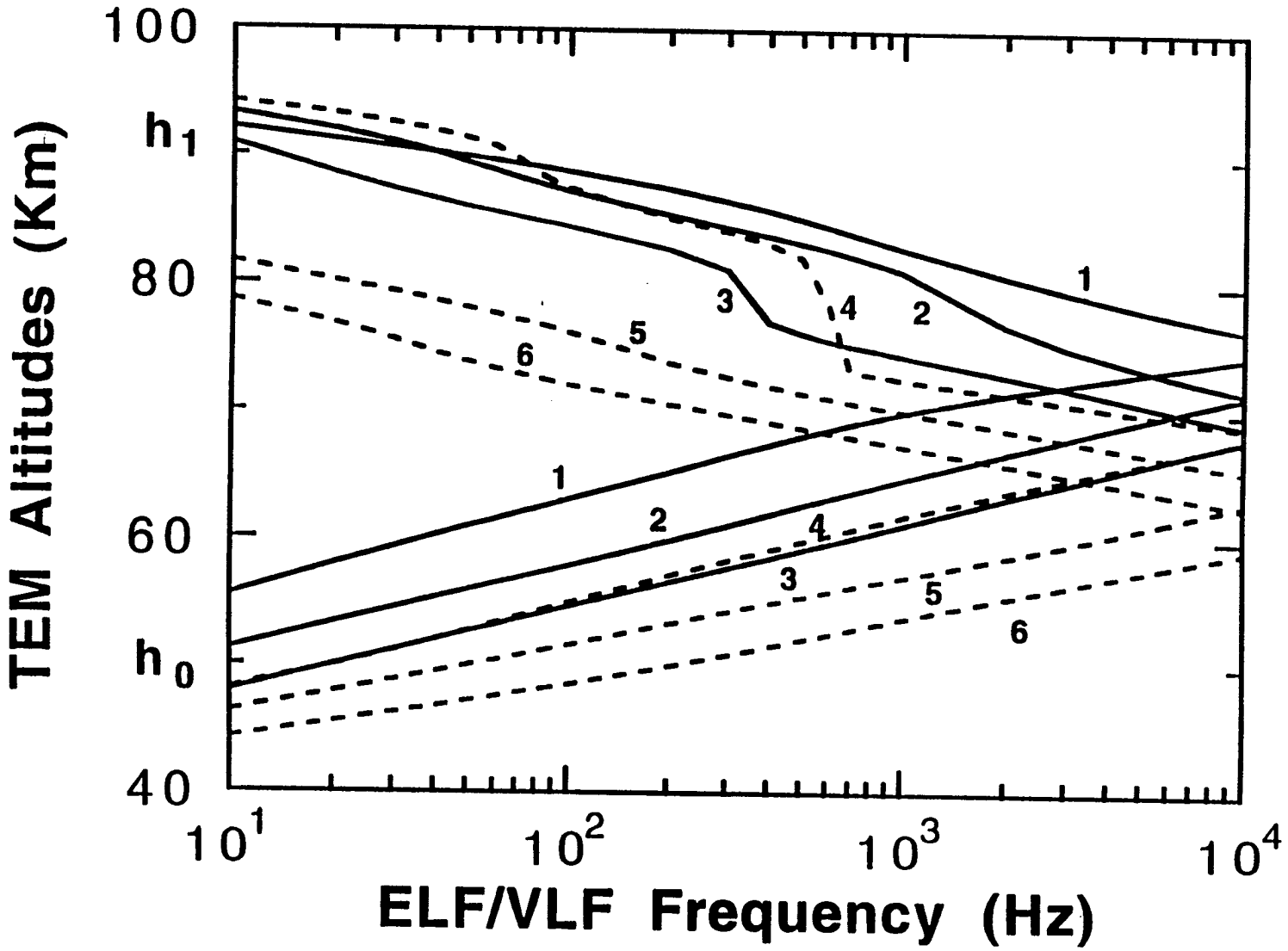

100
h1
80
60
h0
40
TEM Altitudes (Km)
10^1
10^2
10^3
10^4
ELF/VLF Frequency (Hz)
1
2
3
4
5
6
1
2
4
3
5
6

FIGURE 5

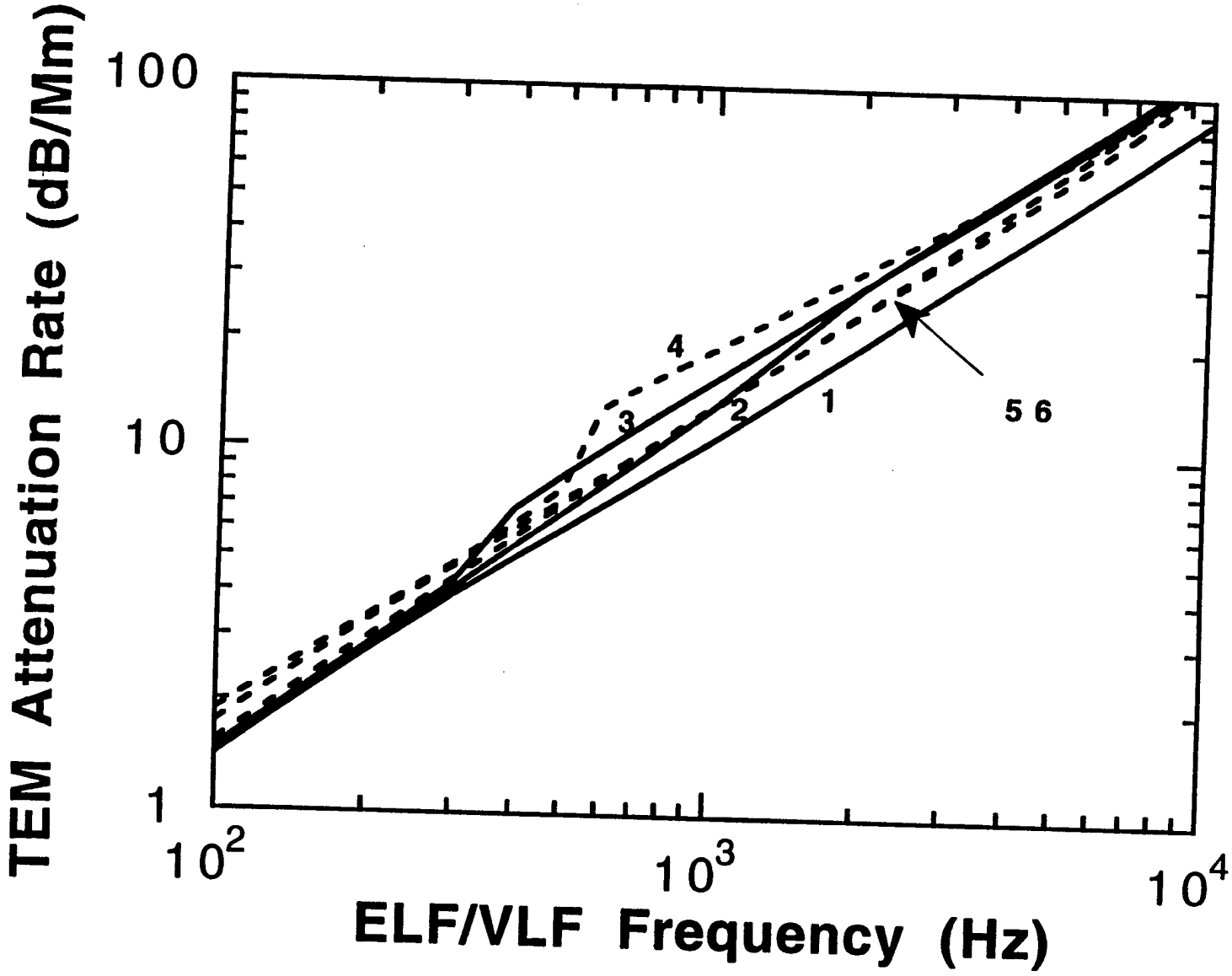

TEM Attenuation Rate (dB/Mm)
100
10
1
10^2
10^3
10^4
ELF/VLF Frequency (Hz)
1
2
3
4
5 6

FIGURE 6

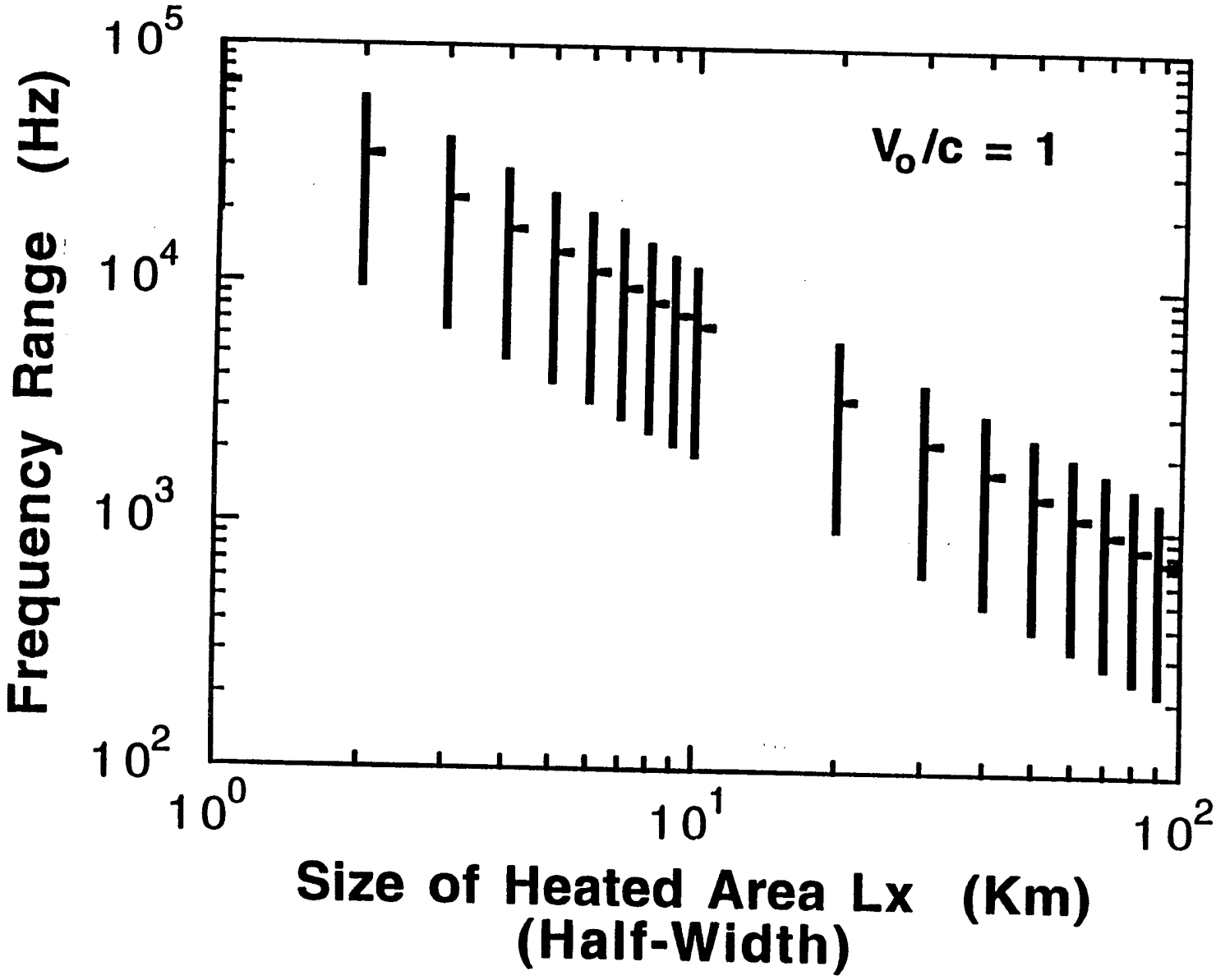

Frequency Range (Hz)
10^5
10^4
10^3
10^2
$V_o/c = 1$
10^0
10^1
10^2
Size of Heated Area Lx (Km)
(Half-Width)

FIGURE 7

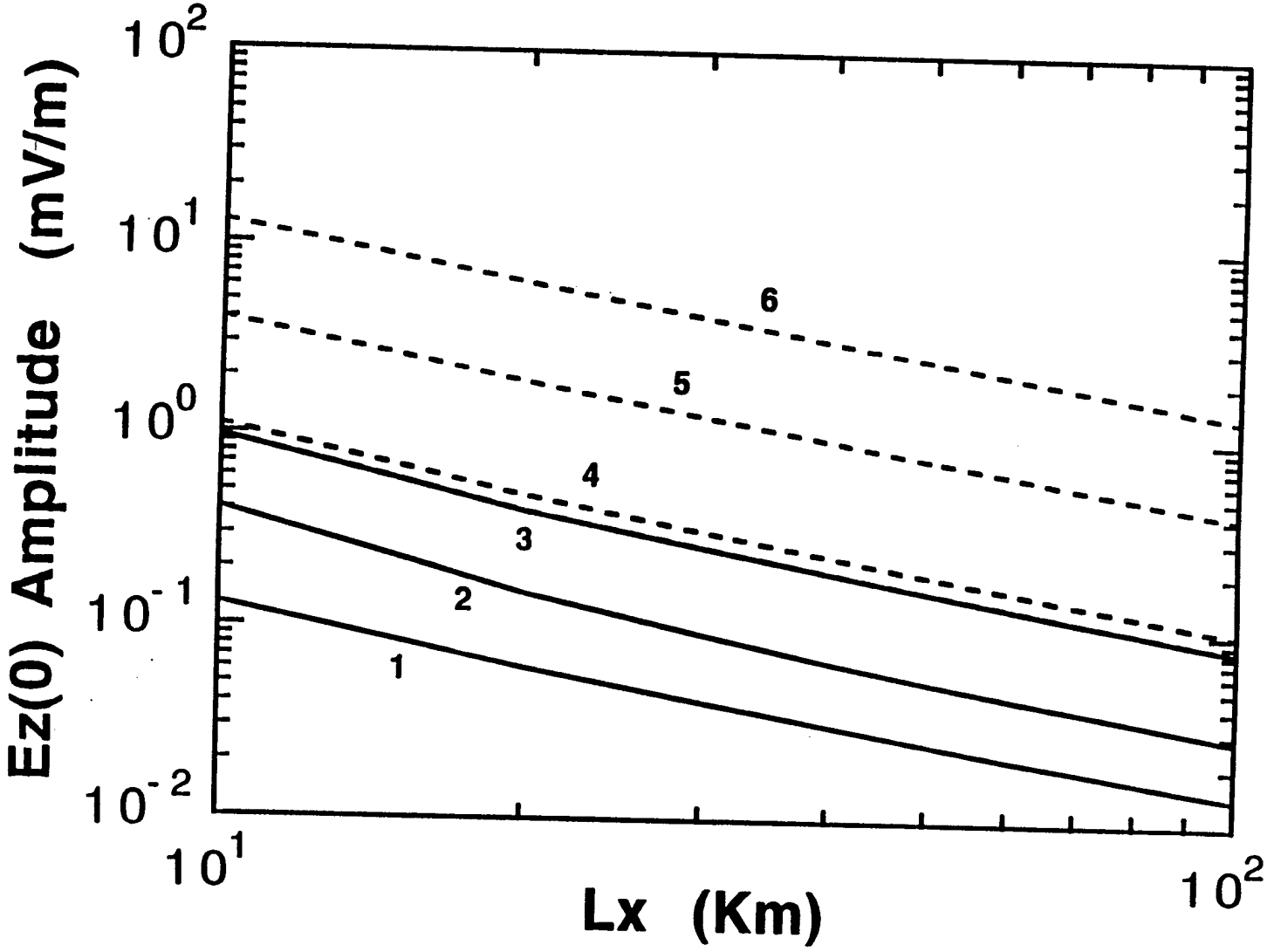


FIGURE 8

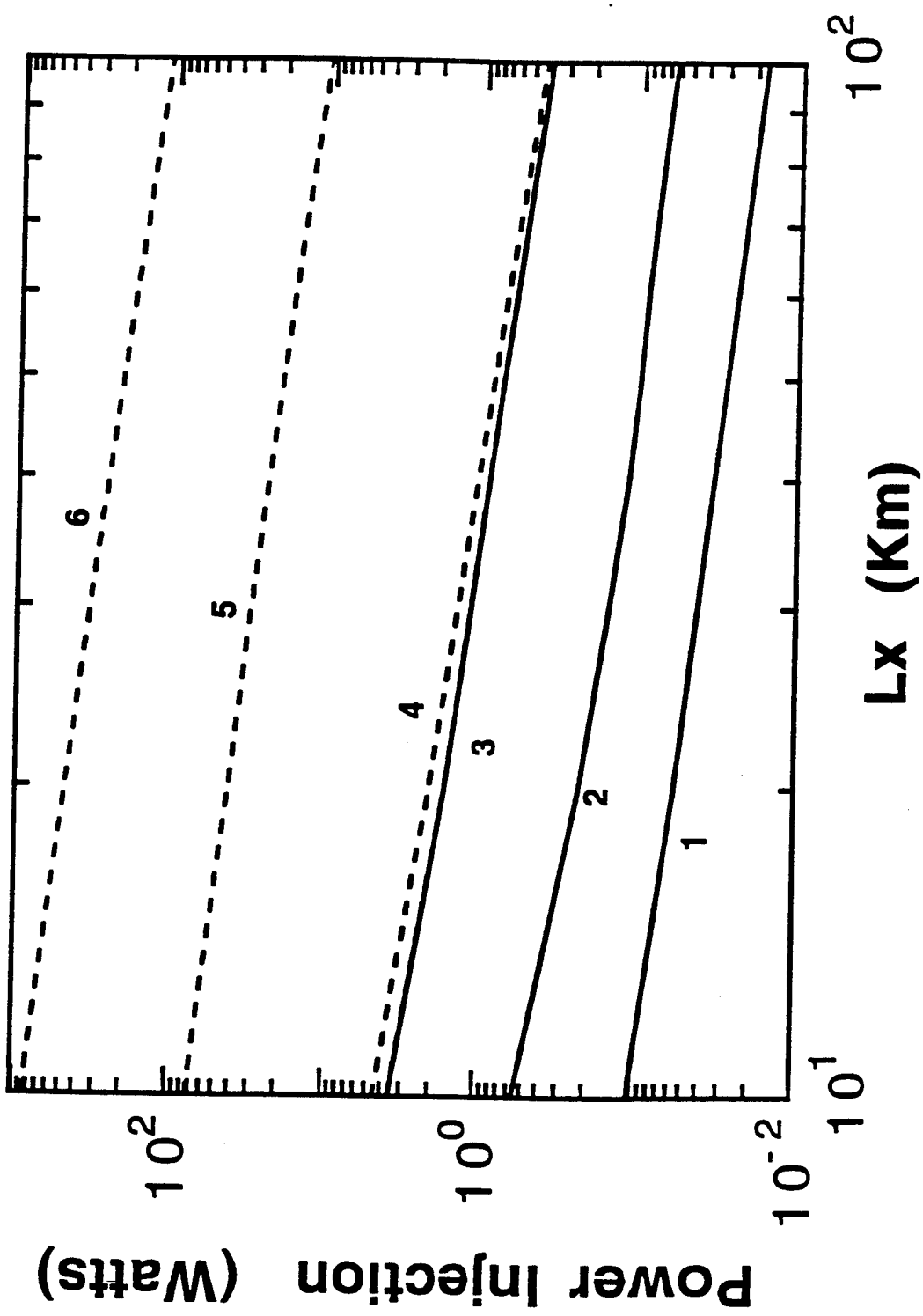


FIGURE 9

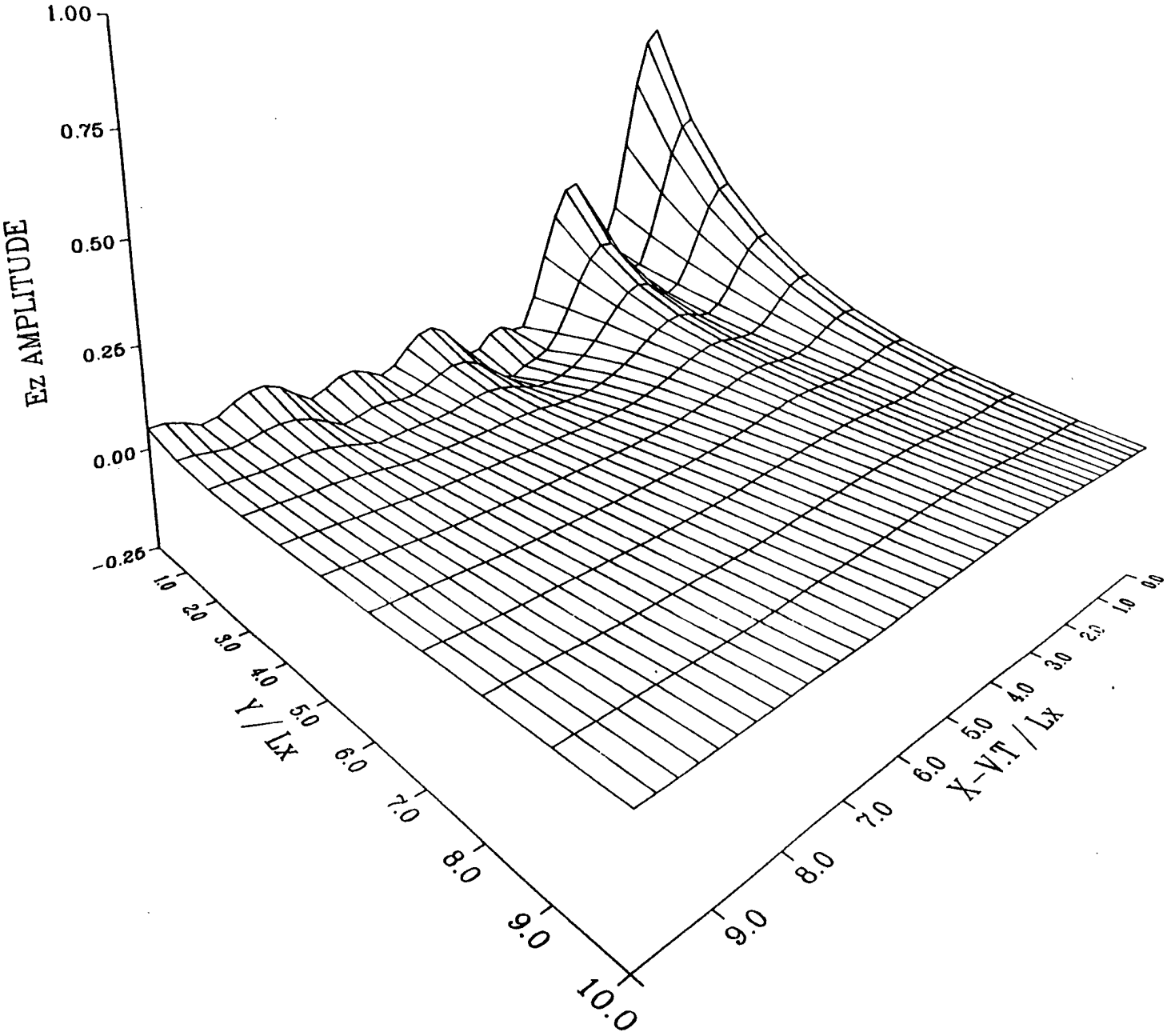

1.00
0.75
0.50
0.25
0.00
-0.25
Ez AMPLITUDE
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
Y / Lx
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
X-VT / Lx

FIGURE 10

Appendix U

Excitation of ULF Waves in the Ionospheric Heating Experiments

EXCITATION OF ULF WAVES IN THE IONOSPHERIC HEATING EXPERIMENTS

by

Chia-Lie Chang

Science Applications International Corporation
McLean, Viginia

at

High Power RF Ionospheric Modification Workshop
Santa Fe, New Mexico
April, 1996

HIGHLIGHTS

- Large amplitude ULF waves (Pc5) were generated during electrojet modulation experiments at Tromso (Stubbe & Kopka, JGR, p. 1606, 1981)
- Large amplitude ULF waves are desirable in subterranean Imaging
- Electron-ion chemistry, NOT electron collision frequency, is responsible for the excitation of ULF waves
- HF heating of electrons reduces their recombination rate to NO^+ and O_2^+ ; electron density increases as a result, so is the Hall conductivity at lower E region
- Density modulation occurs at ULF timescale, much slower than the heating time
- Strong Pc1- Pc2 signals (~ 400pT) is expected based on HIPAS parameters

OBSERVATIONS OF ULF EXCITATION

- ULF amplitude is 2 orders of magnitude larger than those of ELF/VLF waves

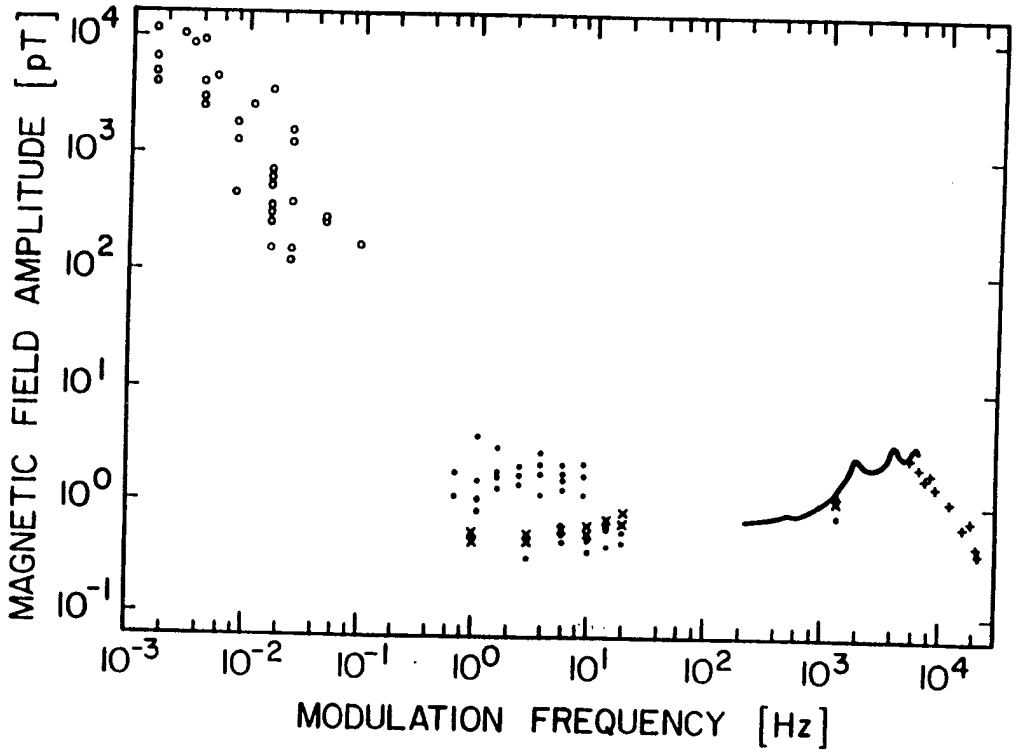


FIGURE 1 A typical magnetic field amplitude vs. frequency measured on the ground at the Max Planck Tromso "Heater" facility (Stubbe, Ref. 21).

OBSERVATIONS OF ULF EXCITATION

- Only H component (N-S) of the magnetic field is affected, corresponding to an east-west (E-W) oriented modulation current (Hall current)

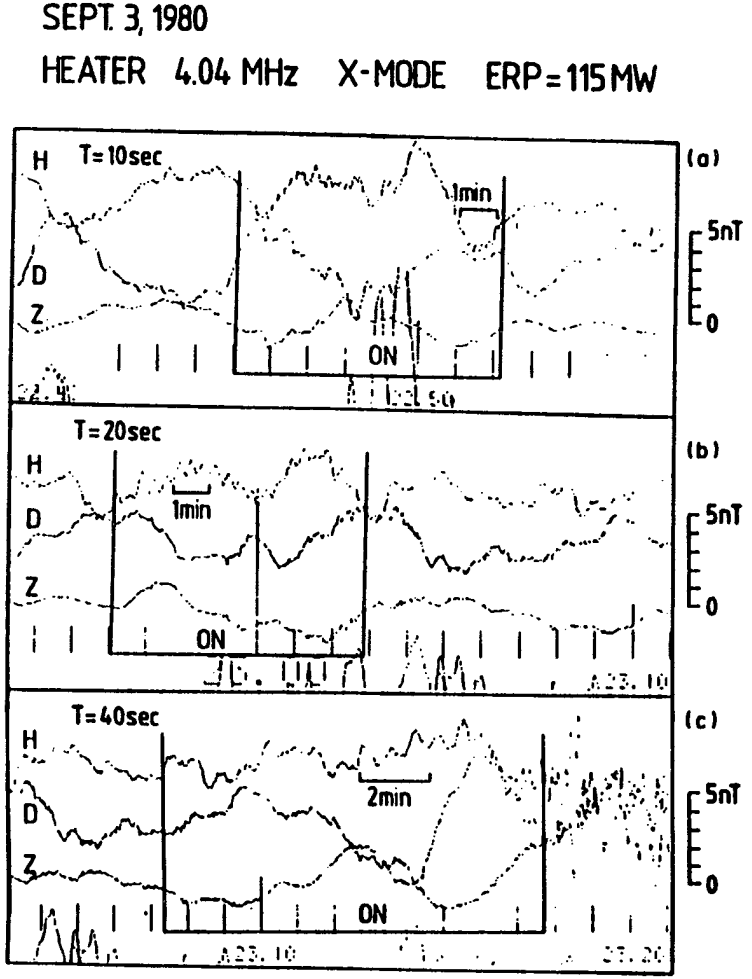


Fig. 4. Magnetometer recording of micropulsations excited by modulated HF heating. The heater-on times are indicated by bars. During these times, the HF wave was 100% modulated with the modulation periods (a) 10 s, (b) 20 s, and (c) 40 s.

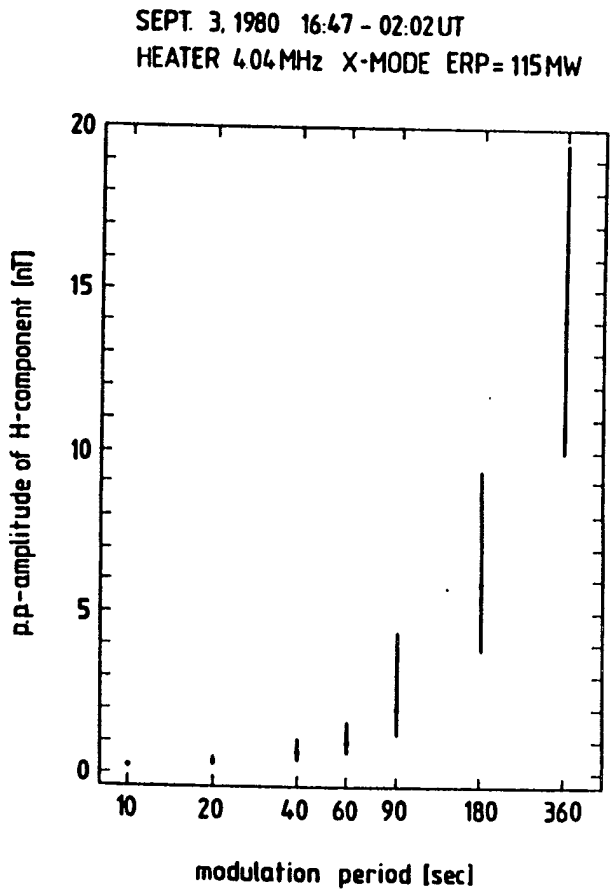


Fig. 6. Peak–peak amplitude of *H* component vs modulation period, based on data obtained on 3/4 September 1980, from 16:47 to 02:02 UT.

OBSERVATIONS OF ULF EXCITATION

- Stubbe & Kopka, "Generation of Pc5 Pulsations by Polar Electrojet Modulation: First Experimental Evidence", JGR p.1606 (1981)

- Stubbe et. al., "Ionospheric Modification Experiments in Northern Scandinavia" J. Atmos. Terr. Phys. p. 1025 (1982)

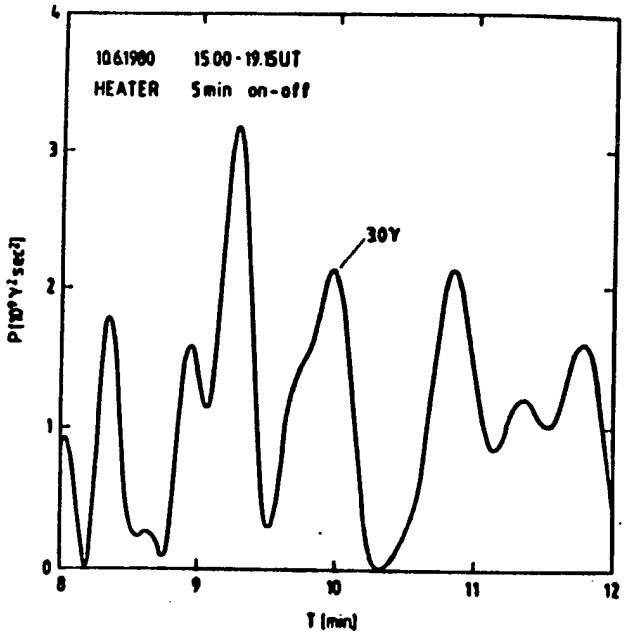


Fig. 2. Power spectrum of $H(t)$ for the period 1500–1915 UT. During this time the heater was 5 min on/5 min off.

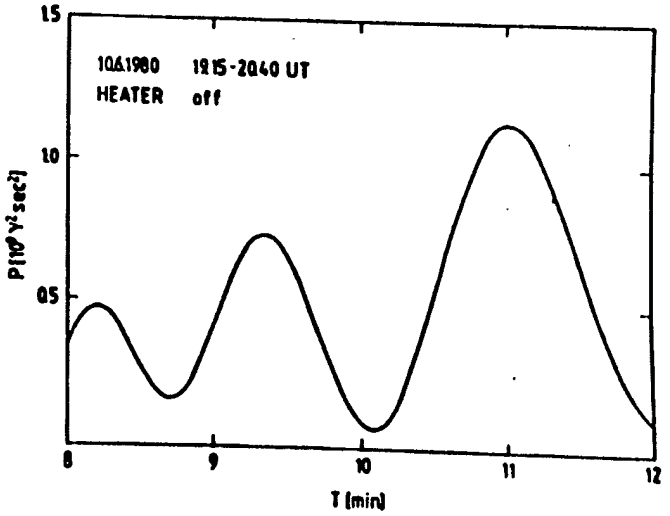


Fig. 4. Power spectrum of $H(t)$ for the period 1915–2040 UT. During this time the heater was off.

HF HEATING IN THE D AND THE LOWER E REGIONS

- The heating equation

$$\frac{\partial T}{\partial t} = \frac{P_{\mathbf{ERP}}\,\kappa(z)}{6\pi z^2 N(z)}\, exp\left(-\int_{z_o}^{z} \kappa(z')\,dz'\right) - \sum_j \delta_j \nu_{ej}(T - T_o)$$

where

$$\kappa(z) = \frac{\omega_{pe}^2}{2c}\frac{\nu_e}{\left[\nu_e^2 + (\omega \mp \Omega_e)^2\right]}$$

- Both $\kappa(z)$ and ν_{ei} depend on the time-varying density $N(z)$
- The ν_{ej} include collisions with O_2, N_2, O, H, He

ELECTRON DENSITY ENHANCED BY HF HEATING

- As electron temperature T increases, the **recombination rate** α decreases, so the electron density N increases with time
- Dissociative recombination processes in the D and lower E region (Gurevich 1978)

$$O_2^+ + e^- \rightarrow O + O \qquad \text{and} \qquad NO^+ + e^- \rightarrow N + O$$

- Electron density specified by a rate equation

$$\frac{\partial N}{\partial t} = \alpha_o N_o^2 - \alpha N^2$$

where

$$\alpha = \alpha_1\left(N_{NO^+} / N^+\right) + \alpha_2\left(N_{O_2^+} / N^+\right) \quad \text{in} \quad \left(cm^3 / sec\right)$$

and

$$\alpha_1 \approx 5 \cdot 10^{-7}\left(300^\circ / T\right)^{1.2} \quad ; \quad \alpha_2 \approx 2.2 \cdot 10^{-7}\left(300^\circ / T\right)^{0.7} \text{ in } \quad \left(cm^3 / sec\right)$$

TIME SCALES FOR DENSITY MODULATION

- Electron **temperature** rises and falls at a **fast timescale** (less than msec) in the D and lower E region, while **density** varies at a **much slower timescale**

- Maximun density increase : $N_\infty = N_o \sqrt{\alpha_o / \alpha}$

- Density rise-time : $\tau_r = 1/2N_o\sqrt{\alpha_o\alpha}$ for $N << N_\infty$

$$N(t) = N_o\left\{(\alpha_o / \alpha)\left[1 - exp(t / \tau_r)\right] + exp(t / \tau_r)\right\}$$

- Density decay-time : $\tau_d = 1/2N_o\alpha_o$

$$N(t) = N_o \frac{1 + \beta exp(-t / \tau_d)}{1 - \beta exp(-t / \tau_d)} \quad \text{and} \quad \beta = \frac{N(t=0) - N_o}{N(t=0) + N_o}$$

- Both **density timescales** depend on the **ambient electron density** N_O

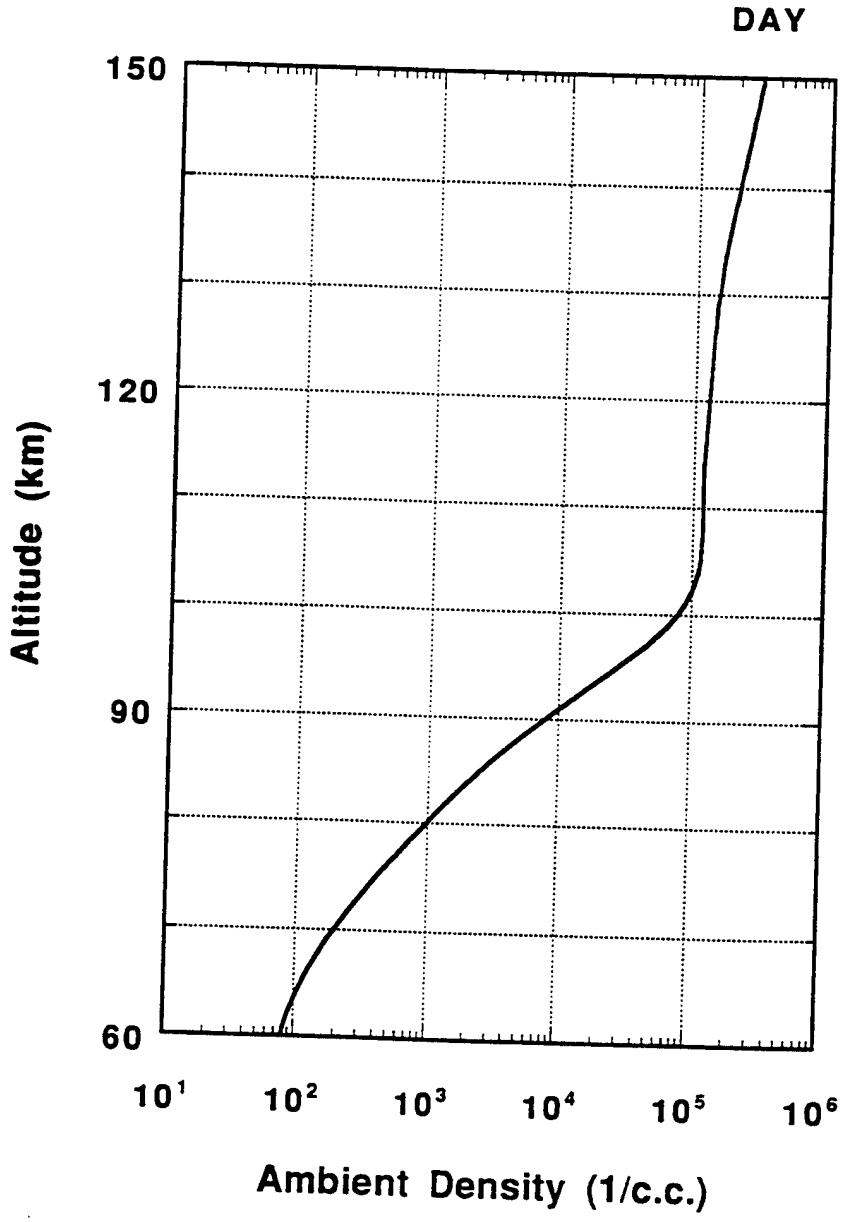
COMPUTATIONAL RESULTS

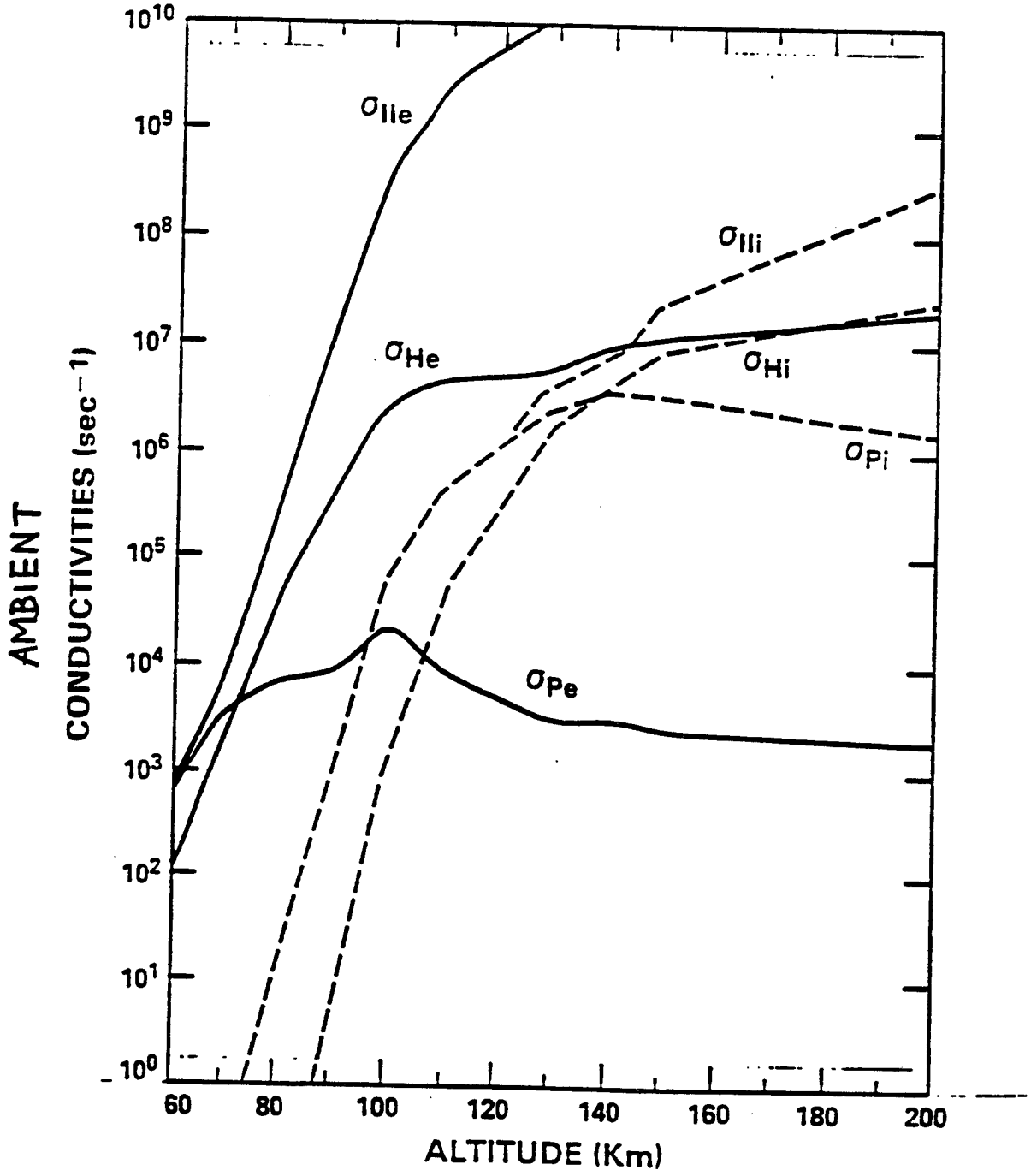
- The prominent **density enhancement** occurs at **lower E region** (above 95 km) while the major **temperature increase** occurs at **D region** (60 km to 75 km)
- The density modulation occurs at **higher altitudes** because the recombination rate is proportional to the **square of ambient density**
- The **density modulation** affects **Hall** conductivity
  the **temperature** modulation affects **Pederson** conductivity
- The **density enhancement spreads widely** from 90 km to 150 km and beyond while the **temperature increase** is very **localized** around 70 km
- Density increase of **a few percent** in the E region is typical for nominal heating
- Since σ_H at lower E region is **2 to 3 orders of magnitude higher** than σ_P at lower D region, **density change** leads to **much larger** overall conductivity modulation at **ULF** heating timescale

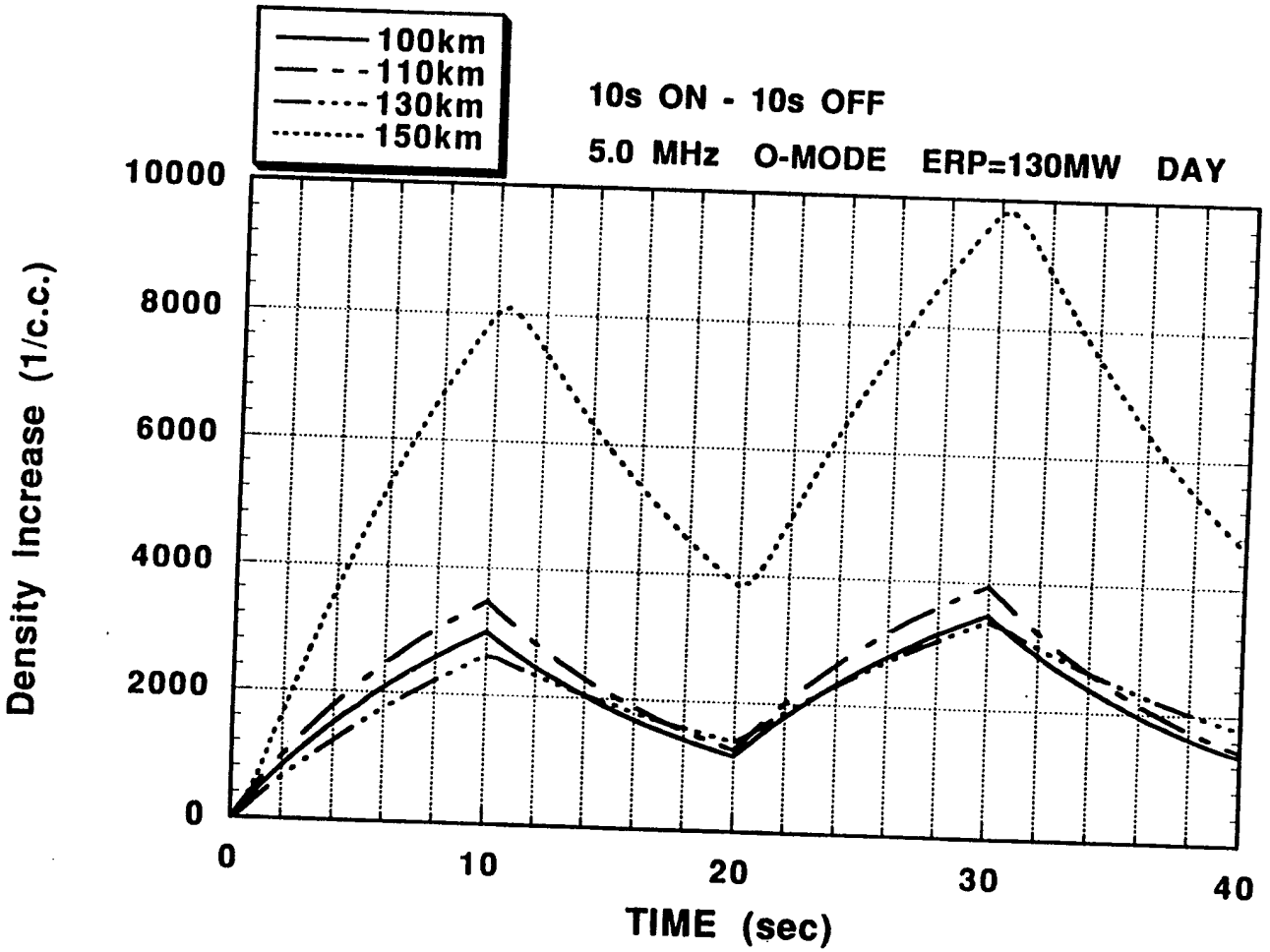
COMPUTATIONAL RESULTS

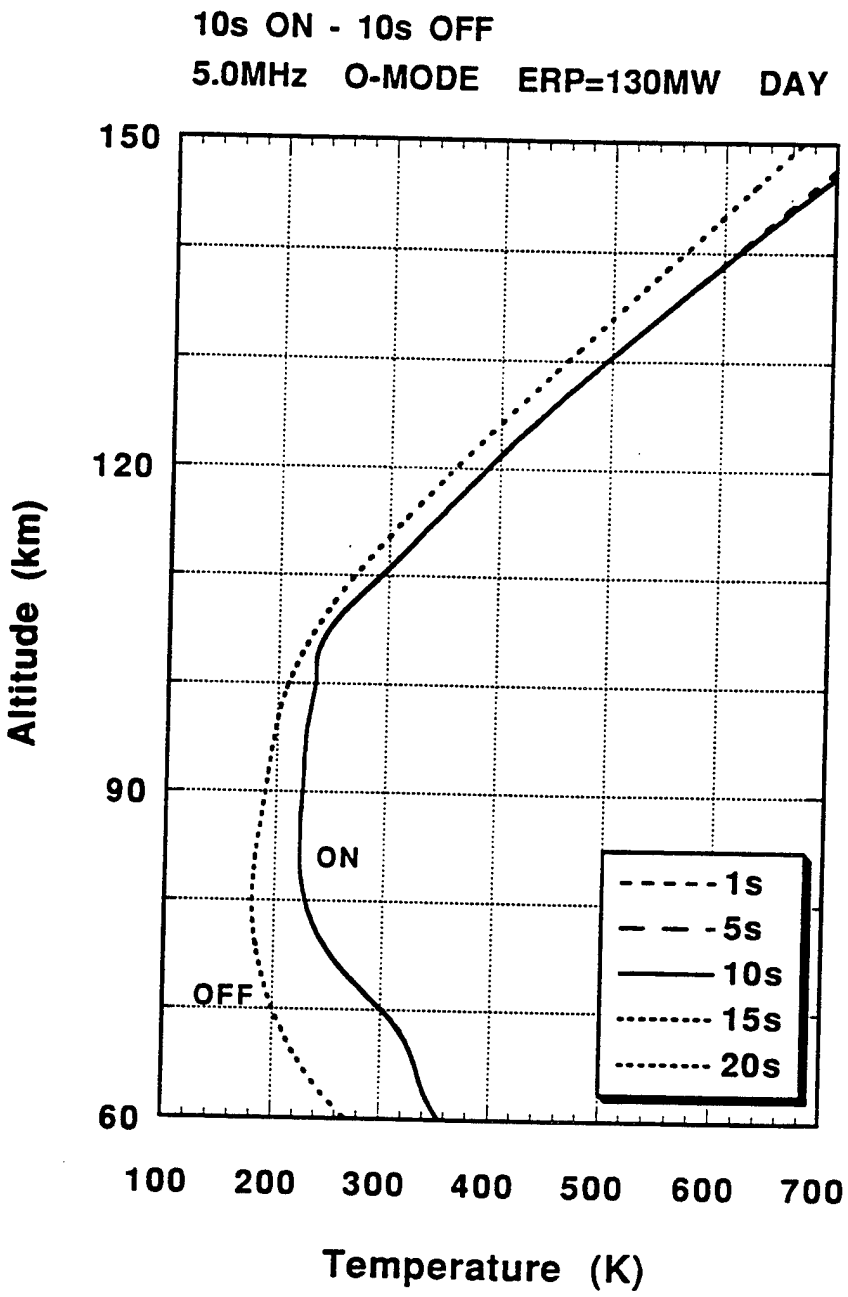
- Example 1: Daytime, X Mode, ERP = 128 MW, 5.0 MHz
  (straight heating of 1s, 10s, 20s, 30s ; Pc1 - Pc2 pulsations)

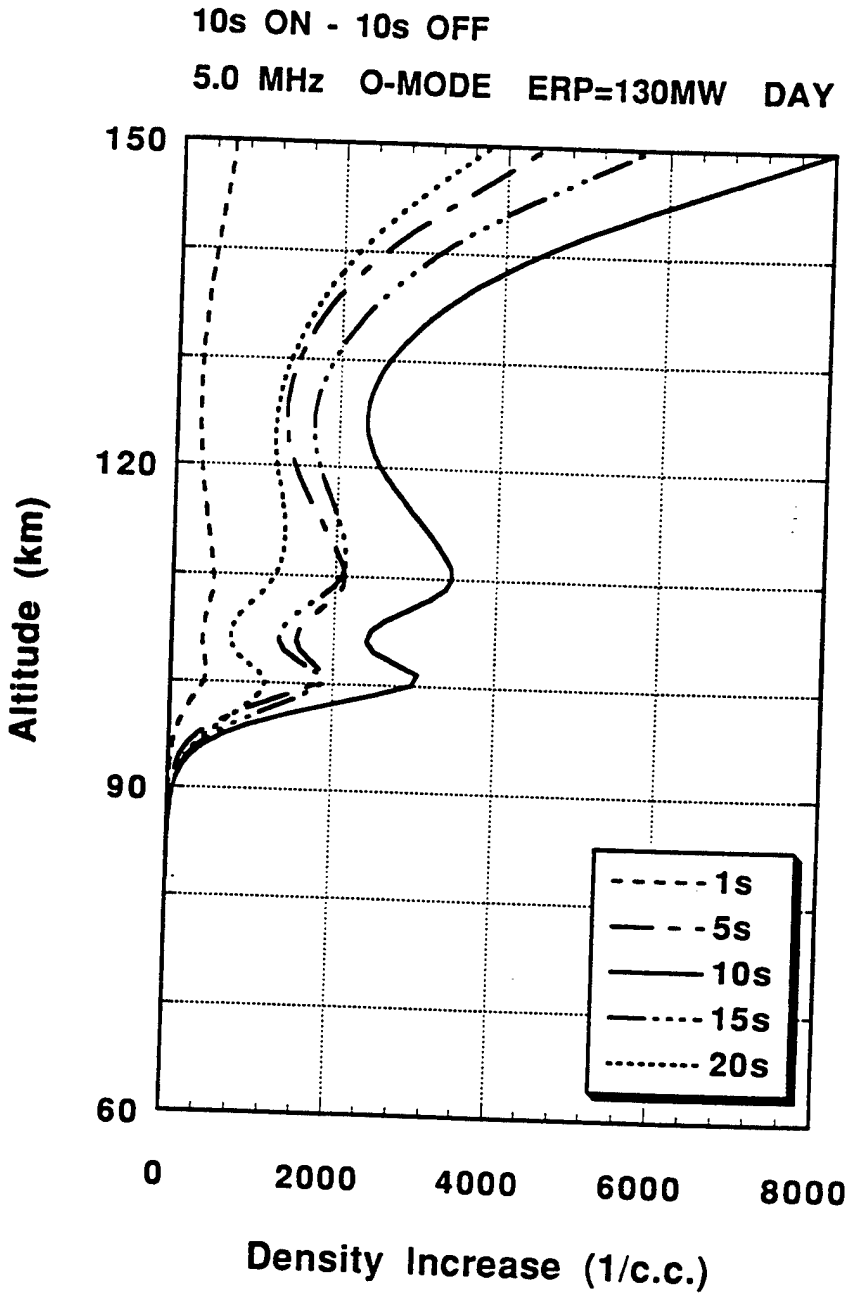
- Example 2: Daytime, O Mode, ERP = 128 MW, 5.0 MHz, density modulation
  (heating cycle: 10s ON - 10s OFF ; Pc2 pulsation)

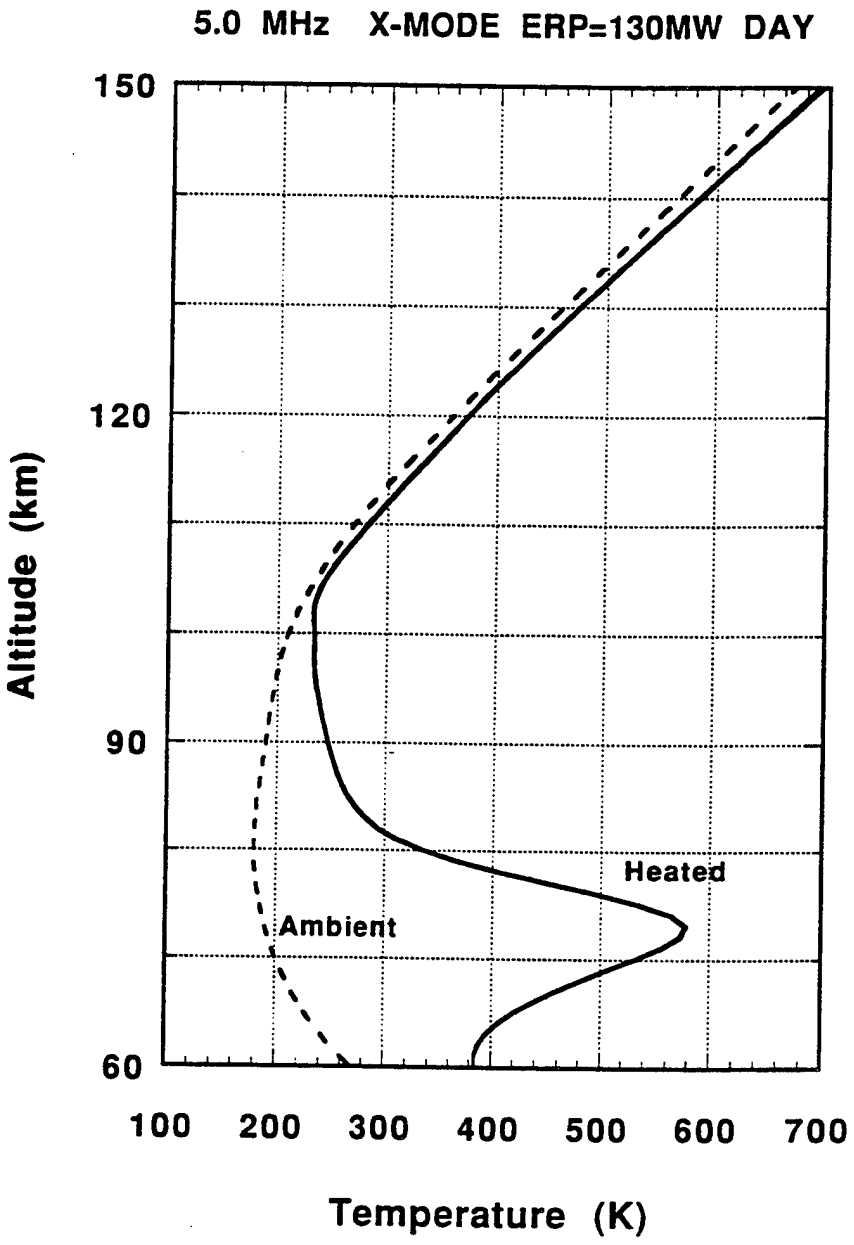

DAY
150
120
90
60
Altitude (km)
10^1
10^2
10^3
10^4
10^5
10^6
Ambient Density (1/c.c.)

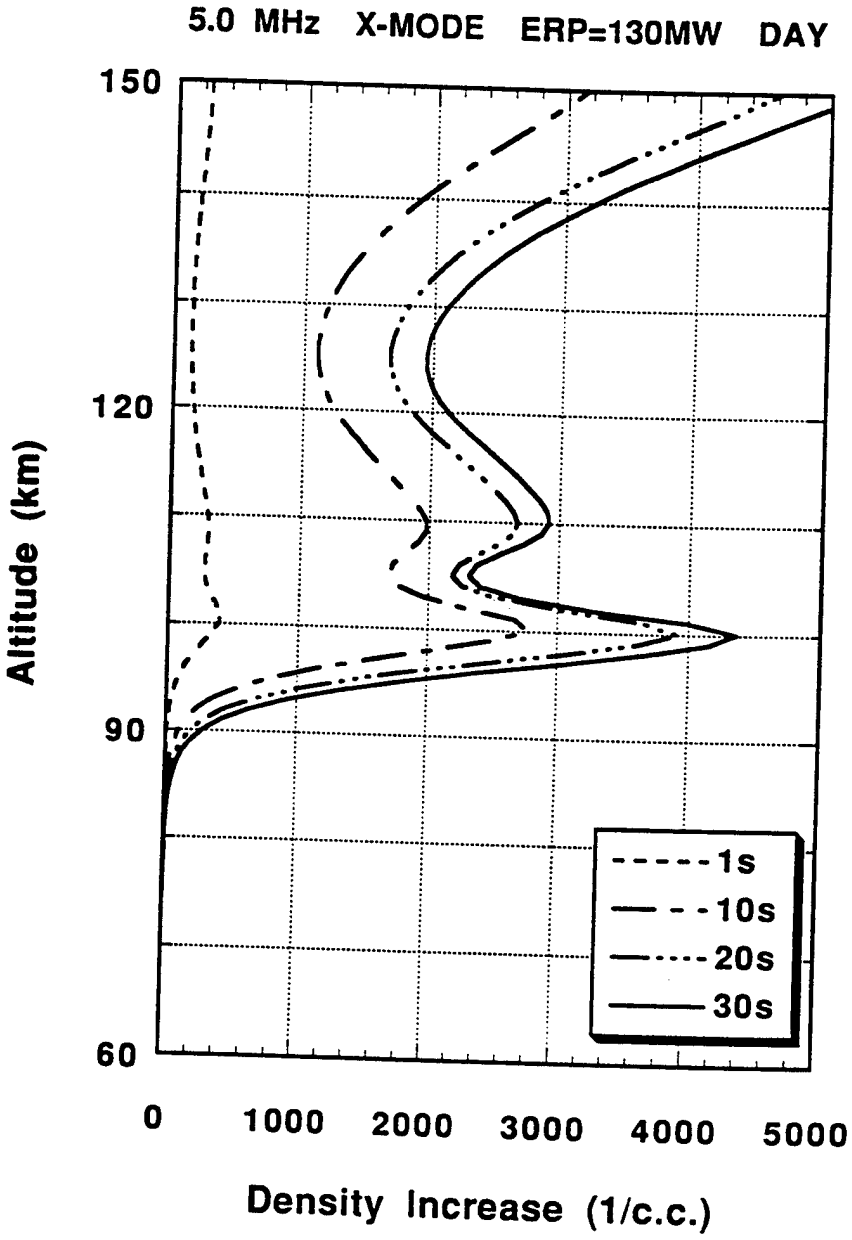

AMBIENT CONDUCTIVITIES (sec⁻¹)
10¹⁰
10⁹
10⁸
10⁷
10⁶
10⁵
10⁴
10³
10²
10¹
10⁰
σIIe
σIIi
σHe
σHi
σPi
σPe
60
80
100
120
140
160
180
200
ALTITUDE (Km)


100km
110km
130km
150km
10s ON - 10s OFF
5.0 MHz O-MODE ERP=130MW DAY
Density Increase (1/c.c.)
10000
8000
6000
4000
2000
0
0
10
20
30
40
TIME (sec)


10s ON - 10s OFF
5.0MHz O-MODE ERP=130MW DAY
150
120
90
60
Altitude (km)
ON
OFF
1s
5s
10s
15s
20s
100
200
300
400
500
600
700
Temperature (K)


10s ON - 10s OFF
5.0 MHz O-MODE ERP=130MW DAY
150
120
90
60
Altitude (km)
0
2000
4000
6000
8000
Density Increase (1/c.c.)
1s
5s
10s
15s
20s


5.0 MHz X-MODE ERP=130MW DAY
150
120
90
60
Altitude (km)
Heated
Ambient
100
200
300
400
500
600
700
Temperature (K)


5.0 MHz X-MODE ERP=130MW DAY
150
120
90
60
Altitude (km)
0
1000
2000
3000
4000
5000
Density Increase (1/c.c.)
1s
10s
20s
30s

CONDUCTIVITY MODULATION AND CURRENT MOMENT

- The **Hall conductivity** is **enhanced** by the **density increase** in the heated region

$$\Delta\sigma_{He} = \frac{\Delta N(z)e^2}{m_e}\frac{1}{\Omega_e} \qquad \text{and} \qquad \Delta\vec{J} = \Delta\sigma_{He}\,\vec{E}_o \times \hat{b}_o \;; \qquad \text{also valid for ions}$$

- Since magnetic field is vertical, and the ambient electric field is in the north-south (NS) direction, the modulated current ΔJ is mainly in the **east-west** (EW) direction; as indicated by the experimental observations

- Based on Example 3 (Pc2 of 0.05 Hz), density increase $\Delta N(z) \approx 4 \cdot 10^3$ /c.c. , and $E_o \approx 10\ mV/m$, the magnitude of **current density** modulation is

$$\Delta J \approx 6.6 \cdot 10^{-9}\ Amp/m^2$$

- Integrate from 100 km to 150 km based on an antenna gain of 80, the **current moment** is

$$IL \approx 1.6 \cdot 10^7\ Amp - m$$

AMPLITUDE OF ULF WAVES AT GROUND LEVEL

- At Pc2 frequency (0.05 Hz), the lower E layer may be transparent, thus

$$\nabla \times B = \mu_o \Delta J \qquad \text{and} \qquad B \approx \mu_o \Delta J \cdot \Delta z$$

the **magnetic amplitude** is roughly $B \approx 400\,pT$, consistent with the observations

- Density change ΔN **increases** with the modulation period, so the ULF amplitude **increases** as the frequency **decreases**

- More work is need to improve the ULF propagation model, which should include

  Ground conductivity

  Earth-Ionospheric waveguide - curvature effects, etc.

  Alfvenic duct in the F region

POTENTIAL IMPACT OF THE ULF EXCITATION RESEARCH

- Create large magnetic amplitude at the ground level for **subterranean mapping** and exploration, with **controlled penetration depths** for depth profiling

- Extend the **ULF modulation** into the **F region**, excite **ducted** propagation of large amplitude **Alfvenic waves** in the F layer

    - ULF injection into the magnetosphere, geomagnetic survey

- Explore a ULF wave-generation mechanism for **ULF band communication** system

- Excite ULF waves as a diagnostic tool for **ionospheric chemistry** in the E region

Appendix V

The Ionospheric Focused Heating Experiment

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. A9, PAGES 17,331–17,345, SEPTEMBER 1, 1995

The Ionospheric Focused Heating Experiment

P. A. Bernhardt,[1] C. L. Siefring,[1] P. Rodriguez,[1] D. G. Haas,[1] M. M. Baumback,[1] H. A. Romero,[1] D. A. Solin,[1] F. T. Djuth,[2] L. M. Duncan,[3] D. E. Hunton,[4] C. J. Pollock,[5] M. P. Sulzer,[6] C. A. Tepley,[6] L. S. Wagner,[7] and J. A. Goldstein[7]

Abstract. The Ionospheric Focused Heating rocket was launched on May 30, 1992. The sounding rocket carried an instrument and chemical payload along a trajectory that crossed the intersection of the beams from the 430-MHz incoherent scatter radar and the 5.1-MHz high-power radio wave facility near Arecibo. The release of 30 kg of CF_3Br into the *F* region at 285 km altitude produced an ionospheric hole that acted like a convergent lens to focus the HF transmissions. The power density inside the radio beam was raised by 12 dB immediately after the release. A wide range of new processes were recorded by in situ and ground-based instruments. Measurements by instruments flying through the modified ionosphere show small-scale microcavities (<1 m) and downshifted electron plasma (Langmuir) waves inside the artificial cavity, electron density spikes at the edge of the cavity, and Langmuir waves coincident with ion gyroradius (4 m) cavities near the radio wave reflection altitude. The Arecibo incoherent scatter radar showed 20 dB or greater enhancements in ion acoustic and Langmuir wave turbulence after the 5.1-MHz radio beam was focused by the artificial lens. Enhancements in airglow from chemical reactions and, possibly, electron acceleration were recorded with optical instruments. The Ionospheric Focused Heating experiment verified some of the preflight predictions and demonstrated the value of active experiments that combine high-power radio waves with chemical releases.

1. Introduction

The Ionospheric Focused Heating (IFH) experiment was designed to demonstrate the use of chemical releases to enhance the effective radiated power of high-power HF facilities and to measure in situ the turbulence in the heated and chemically modified plasma. The rocket payload for the IFH experiment consisted of a chemical canister and a set of diagnostic instruments. The IFH payload was launched on May 30, 1992, during the Combined-Release and Radiation Effects Satellite (CRRES) program. All of the CRRES Caribbean rockets took off from the north coast of Puerto Rico. The CRRES rocket campaign, called *El Coqui*, was organized to take advantage of the ionospheric heating facility, incoherent scatter radar, and other ground support diagnostics located near the Arecibo Observatory in Puerto Rico [*Djuth*, 1993; *Kelley et al.*, this issue; *Djuth et al.*, this issue].

The term "ionospheric heating" refers to ionospheric modification by high-power radio waves. The transmitted high-frequency (HF) electromagnetic wave is called the pump. The effective radiated power (ERP) of the pump is calculated as the product of the transmitter power and the antenna gain for the facility. The ERP of the Arecibo HF facility was about 50 MW at the 5.1 MHz frequency used during the *El Coqui* campaign. The ERP was about 10 MW lower than usual because of reduced transmitter antenna gain resulting from tilting the normally vertical beam. The HF beam was tilted 12° to the north to allow penetration by the rockets without endangering populated areas.

The Arecibo incoherent scatter radar (ISR) is a primary diagnostic instrument for the heated plasma. The transmitted wave (f_T = 430 MHz) from the radar is scattered from plasma waves f_p, yielding scattered electromagnetic waves f_s that are received with the 305-m dish antenna located at Arecibo. The scattering conditions require that the scattering wavelength is equal to one half the transmitted wavelength ($\lambda = c/f_T = 0.698$ m) of the radar.

When the powerful HF transmitter is turned on, enhanced ion lines may be produced by scatter from HF-induced ion acoustic waves [*Showen and Kim*, 1978]. All ion line spectra reside within a 20-kHz band centered on the radar frequency. Plasma line spectra are produced by electron plasma waves scattering the radar signal. The frequencies of these lines fall near $f_T \pm f_{HF}$ where f_{HF} is the pump frequency.

The electromagnetic pump may be scattered by plasma waves to yield stimulated electromagnetic emissions (SEE) [*Thidé et al.*, 1989]. SEE is often characterized by a downshifted maximum (DM) located about 15 kHz below the pump frequency. During the IFH experiment, SEE observations were conducted with a radio frequency spectrum analyzer connected to a broadband receiving antenna.

[1]Plasma Physics Division, Naval Research Laboratory, Washington, D. C.
[2]Geospace Research, Incorporated, El Segundo, California.
[3]School of Applied Science and Engineering, University of Tulsa, Tulsa, Oklahoma.
[4]Phillips Laboratory, GPID, Hanscom Air Force Base, Massachusetts.
[5]NASA Marshall Space Flight Center, Huntsville, Alabama.
[6]Arecibo Observatory, Arecibo, Puerto Rico.
[7]Information Technology Division, Naval Research Laboratory, Washington, D. C.

Paper number 94JA01887.
0148-0227/95/94JA-01887$05.00

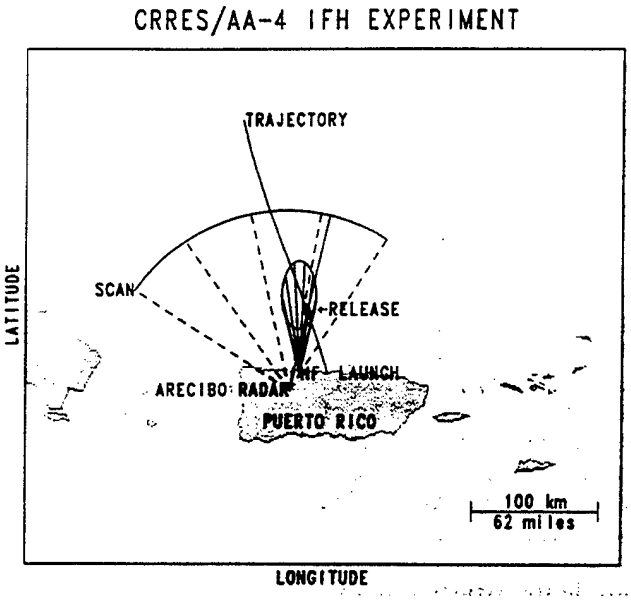


Figure 1. Ground projection of the Ionospheric Focused Heating (IFH) rocket trajectory relative to the island of Puerto Rico. The location of the 5.1-MHz HF heater and the 430-MHz UHF radar beams are also indicated. The CF_3Br release occurred along the trajectory slightly to the east of the center of the HF beam.

The process for conversion of high-power radio waves into plasma waves has been explained in terms of parametric instabilities [*Fejer*, 1979] or Langmuir turbulence [*Dubois et al.*, 1990]. The pump electromagnetic wave can decay into two plasma waves, an electron plasma wave and an ion acoustic wave propagating in opposite directions. The frequency of the electron plasma waves is offset from the pump frequency by the frequency of the ion acoustic wave.

High-power radio waves simultaneously produce large-scale (>10 km) cavities, ion gyroradius structures (~10 m), medium-scale irregularities (~1 km), and small-scale Langmuir turbulence (~1 m). Often, these irregularities are located in the same region of heated plasma. At Arecibo, during special conditions found at winter and solar minimum, large-scale cavities produced by thermal expansion of the plasma have been observed near the reflection point of the HF beam [*Duncan et al.*, 1988; *Hansen et al.*, 1990]. Inside these cavities the ISR often detects enhanced ion lines from ion acoustic waves [*Duncan et al.*, 1988; *Bernhardt et al.*, 1989].

The effects of electron acceleration are seen as airglow clouds attached by magnetic field lines to the HF wave reflection region [*Bernhardt et al.*, 1988, 1989, 1991b]. Suprathermal electrons produced by electron plasma waves acting on the thermal population in the *F* layer stream down magnetic field lines to collide with and excite atomic oxygen in the neutral atmosphere. Fluxes of energetic electrons can excite the red-line (630.0 nm) and green-line (557.7 nm) emissions from atomic oxygen in the upper atmosphere. The electron energies required to excite measurable intensities for the two lines are 3.5 and 6.0 eV, respectively. The intensities of 630.0-nm and 557.7-nm emissions have been analyzed to yield estimates of the suprathermal electron energy spectra [*Bernhardt et al.*, 1989].

All of the phenomena mention above (i.e., enhanced ion lines, SEE, parametric instabilities, electron acceleration, enhanced airglow) occur if the power density from the HF pump wave exceeds a threshold. One way of exceeding existing thresholds is to focus the beam with an artificial ionospheric lens. The purpose of the IFH experiment was to (1) use a chemical release to form an artificial lens that focuses the radio beam of the Arecibo HF facility and (2) measure the effects of the resulting increase in ERP on the ionospheric heating processes. This paper presents an overview of the observations that were affected by the focusing. The next section describes the rocket payload that formed the artificial lens and measured the effects of the chemical release in situ. Section 3 outlines the observations made by the Langmuir probe, plasma wave receiver, and electric field instruments on the rocket. The data from the Arecibo incoherent scatter radar are described in section 4. Airglow enhancements following the chemical release are discussed in section 5. Conclusions regarding the IFH experiment are given in section 6.

2. Experiment Concept: Chemically Produced Cavity

At Arecibo, large-scale cavities and enhanced ion acoustic waves have been found to be coincident with electron acceleration that produces airglow [*Bernhardt et al.*, 1989]. Self-action of the HF wave generates the large-scale density structures that "self-focus" the HF wave [*Hansen et al.*, 1990, 1992]. The processes that couple the HF-induced cavities to enhanced ion acoustic waves are not understood, so an experiment was devised to control the cavity formation and then see the response of the enhanced ion lines and the optical emissions. The active experiment used a chemical to form a large electron scale cavity in the HF beam. The effects of the artificial cavity on the heated ionosphere were recorded with in situ and ground-based instruments. This experiment was called "Ionospheric Focused Heating" (IFH) because it was thought that the chemically produced cavity would focus the HF, yielding substantially larger field strengths for the HF waves [*Bernhardt and Duncan*, 1987].

The ionospheric hole produced by the release of CF_3Br removed electrons by dissociative attachment. The effectiveness of this chemical was demonstrated in the *F* region during the NASA-sponsored first Negative Ion Cation Release (NICARE 1) experiment [*Bernhardt et al.*, 1991a]. The dissociative attachment reaction is

$$CF_3Br + e^- \rightarrow Br^- + CF_3 \qquad (1)$$

where rate coefficient $k_1 \sim 10^{-7}$ cm^3/s. With a 30-kg release of this material the electron population was depleted within 10 km of the release point [*Bernhardt et al.*, 1991a]. The negative and positive ions are mutually neutralized by the reaction

$$Br^- + O^+ \rightarrow Br^* + O^* \qquad (2)$$

where rate coefficient $k_2 \sim 10^{-8}$ cm^3/s and asterisk denotes excited states. Care must be taken to distinguish between airglow from O* excited by surprathermal electron fluxes and by chemical reactions in the negative ion plasma.

The plasma waves measured in situ during the IFH experiment were expected to be strongly influenced by the presence of the negative ions. During the NICARE 1 and 2

Table 1. Instrumentation for the IFH Rocket

Experiment	Group	Sensor	Specifications	Sample Rate
		*Langmuir Probes**		
LP1	NRL	20-cm probe	26 to 4 × 10^6 cm^{-3}	5.556 ksample/s
LP2	NRL	20-cm probe	3 to 4 × 10^5 cm^{-3}	5.556 ksample/s
$(\delta n/n)1$	NRL	20-cm probe	0 to 3	11 ksample/s
$(\delta n/n)2$	NRL	20-cm probe	0 to 3	11 ksample/s
DP1	PL	3.0-inch disk	10 to 10^6 cm^{-3}	1.4 ksample/s
DP2	PL	1.5-inch disk	10 to 10^6 cm^{-3}	11.2 ksample/s
		Electric Fields†		
dc, LF, ac 14	NRL	4.3-m dipole	0 to 5.6 kHz	11.2 ksample/s
dc, LF, ac 12	NRL	0.3-m dipole	0 to 5.6 kHz	11.2 ksample/s
dc, LF, ac 34	NRL	0.3-m dipole	0 to 5.6 kHz	11.2 ksample/s
dc, LF, ac 56	NRL	0.3-m dipole	0 to 5.6 kHz	11.2 ksample/s
DC1	NRL	1.5-m RF	2.5 to 12 MHz	one 1600 point
DC2	NRL	1.5-m RF	2.5 to 12 MHz	spectrum/s
DC3	NRL	5.5-cm RF	5.050 MHz ± 75 kHz	4 Mbit/s
DC4	NRL	11.0-cm RF	5.075 MHz ± 75 kHz	4 Mbit/s
		Energy Spectra‡		
TECHS	MSFC	cylinder	0.5 to 50 eV (32)	320-ms scan
TECHS	MSFC	cylinder	45° × 15° (8)	32 × 8 spectra

Abbreviations are (column 1) LP, Langmuir probe; DP, disk probe; LF, low frequency: DC, downconverter; and TECHS, thermal-electron capped hemispherical spectrometer and (column 2) NRL, Naval Research Laboratory; PL, Phillips Laboratory; and MSFC, Marshall Space Flight Center; ksample is kilosample.

*Parameters are as follows: for LP1 and LP2, electron density; for $(\delta n/n)1$ and $(\delta n/n)2$, relative density; and for DP1 and DP2, ion density.

†Parameters for all dc, LF, and ac experiments are dc and LF E fields; for DC1–DC4 they are Langmuir and pump waves.

‡Parameters are thermal electrons. Numbers in parentheses indicate number of channels.

experiments, new wave modes were excited inside the negative-ion cloud even without the presence of a high-power radio wave [*Ganguli et al.*, 1992]. Enhanced electric field fluctuations come from lower hybrid waves, ion plasma waves, and ion cyclotron waves excited by the presence of Br^- and O^+ in the chemically modified plasma [*Scales et al.*, 1993]. With the IFH experiment the addition of an electromagnetic pump field added to the complexity of the plasma wave structures in the negative-ion cloud.

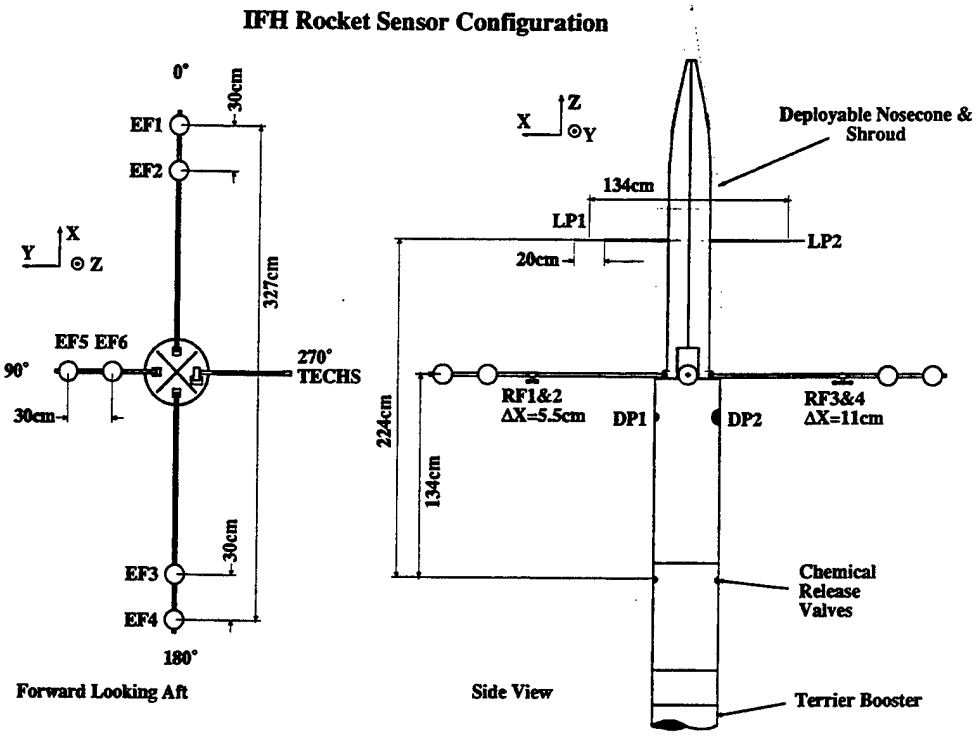


Figure 2. Sensors and instruments on the IFH rocket. The chemical canister was located between the instrument section and the rocket motor.

The magnitude of the focusing by the chemically produced cavity was estimated to be 20 dB in the theoretical paper by *Bernhardt and Duncan* [1987]. The heating at the focus yielded channeling of the HF beam.

The IFH experiment was designated as the CRRES AA 4 rocket in March 1988. In situ instruments were delivered by Naval Research Laboratory (NRL), Phillips Laboratory, and Marshall Space Flight Center for integration into the payload in 1990. The launch of AA 4 occurred on May 30, 1992, at 0411 Atlantic standard time (0811 universal time) from the launcher located near Vega Baja in Puerto Rico. The chemical payload consisted of 30 kg of CF_3Br inside a pressurized tank heated to 340 K. The release occurred at 0813:49.3 UT within 5 km of the center of the Arecibo HF beam. The radar track of the IFH rocket trajectory established the release location was 18.97°N latitude, 66.60°W longitude, and 283 km altitude. Using real-time trajectory information, the chemical tank valve was opened with a command signal so that the release was deposited 20 km below the reflection altitude of the 5.1-MHz HF transmissions. The gas expansion and inertial motion caused the electron depletion to pass through the reflection altitude.

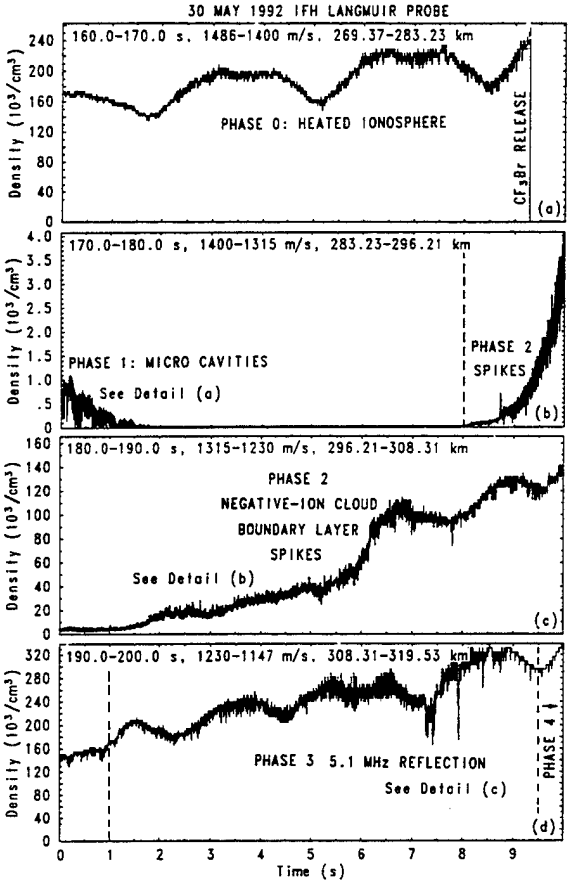


Figure 3. Electron density measured by the IFH rocket in four 10-s periods. The values are derived assuming that electron density is proportional to the measured current from the Naval Research Laboratory Langmuir probe LP2. In regions of very low density, such as phase 1 and the early part of phase 2, the negative ions may provide a substantial portion of the Langmuir probe current. Density fluctuations are evidence of (a) rocket spin, (b) microcavities, (c) spikes, and (d) heater-induced cavities. The rocket speed and altitude at the beginning and end of each period is indicated, along with the range of times after launch.

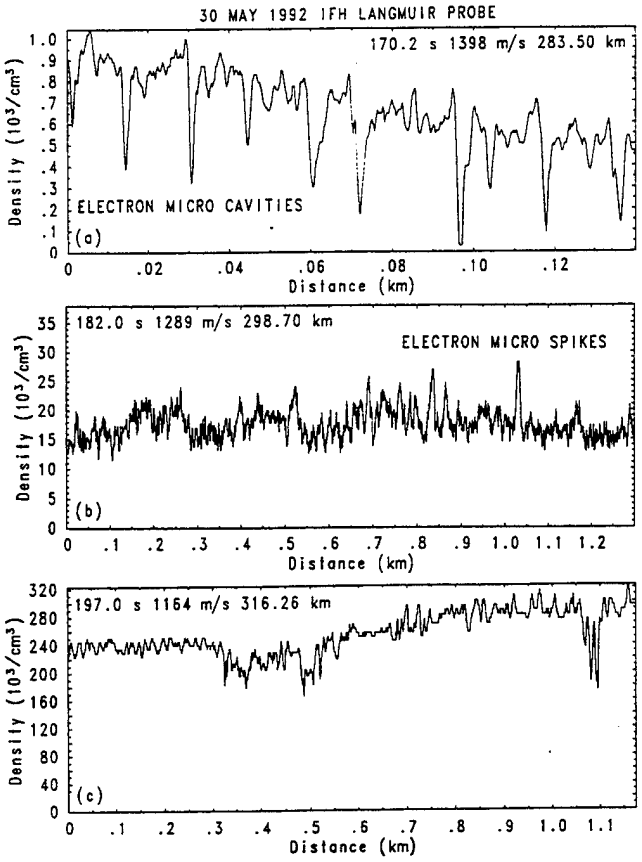


Figure 4. Electron density details along several portions of the IFH rocket trajectory at (a) 170.2 s, (b) 182.0 s, and (c) 197.0 s. The time, speed, and altitude for the start each segment is given along the top of each data sample. The discrete density steps in the data are due to the finite length for the digital word from the logarithmic amplifiers of the Langmuir probe.

Figure 1 illustrates the locations of the flight trajectory, heater beam, and ISR scan projected onto the ground plane containing the island of Puerto Rico. The CF_3Br release occurred slightly to the east of the HF beam center. Data at 630.0 nm from the Arecibo Fabry-Perot interferometer were analyzed to yield a thermospheric wind velocity of 40 m/s westward. The ionospheric hole created by the release drifted with this velocity into the center of the heater beam. By using postlaunch extrapolations for the rocket trajectory and commanded release, the ISR beam was positioned at the release point. After the ionospheric hole was well developed, the ISR was scanned in azimuth to keep track of the westward convection of the ionospheric hole.

The IFH rocket was fully instrumented to measure the plasma densities, electric fields, plasma waves, and suprathermal electrons expected to be detected in the modified plasma (Table 1). The placement of the sensors is shown in Figure 2. All of the booms were located 1 m or more ahead of the two chemical release valves. Upon release the instruments were surrounded by the CF_3Br cloud. If T is initial temperature and m the molecular mass, the cloud initially expanded with a mean thermal speed $(2\,KT/m)^{1/2} = 200$ m/s, keeping the rocket payload as its center of mass. After

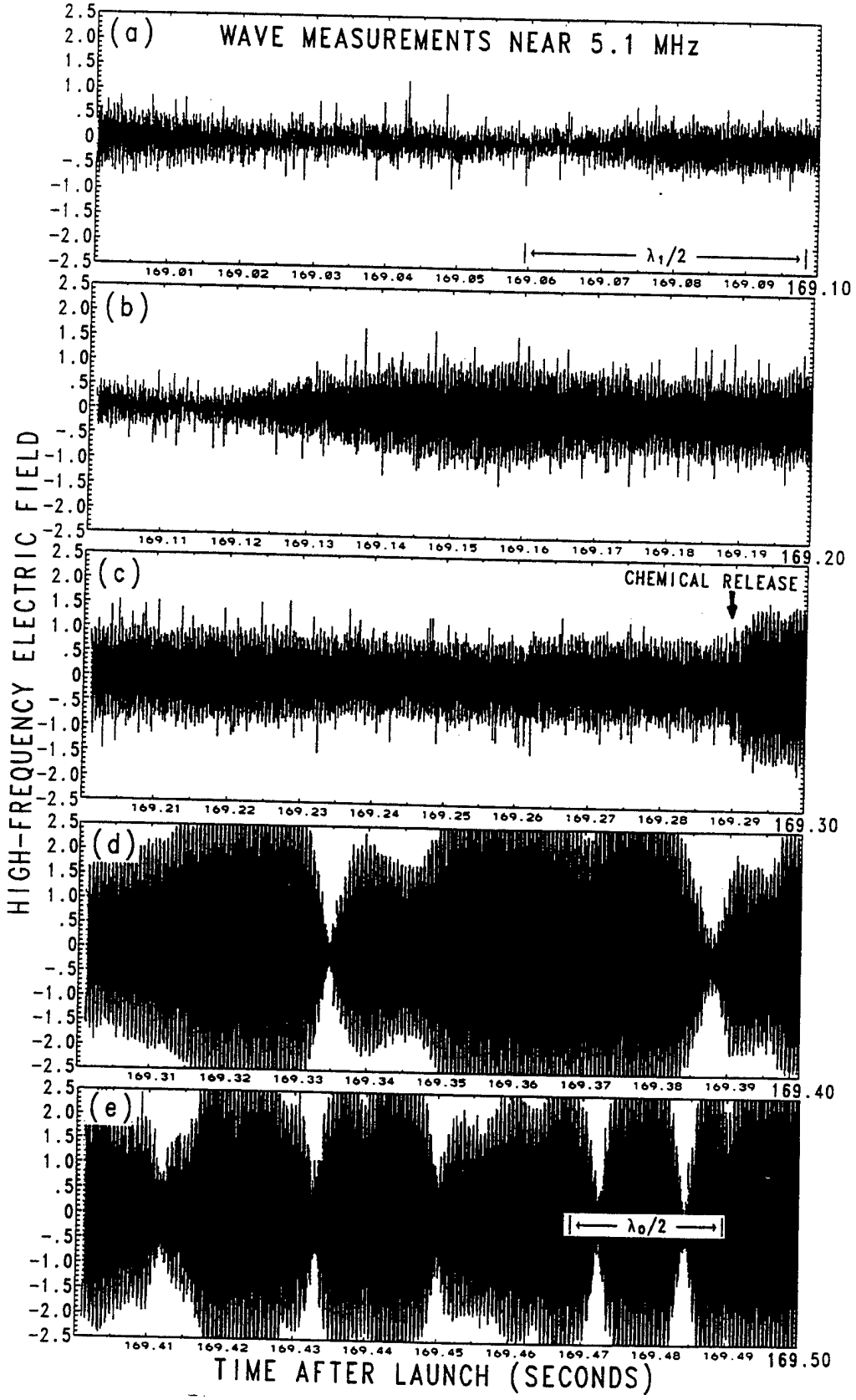


Figure 5. In situ measurements of the electric fields (volts per meter) near the 5.1-MHz pump wave from (a) 169.0 s to (b) 169.1 s to (c) 169.2 s to (d) 169.3 s to (e) 169.4–169.5 s after launch. A large increase in wave amplitude follows the CF_3Br release.

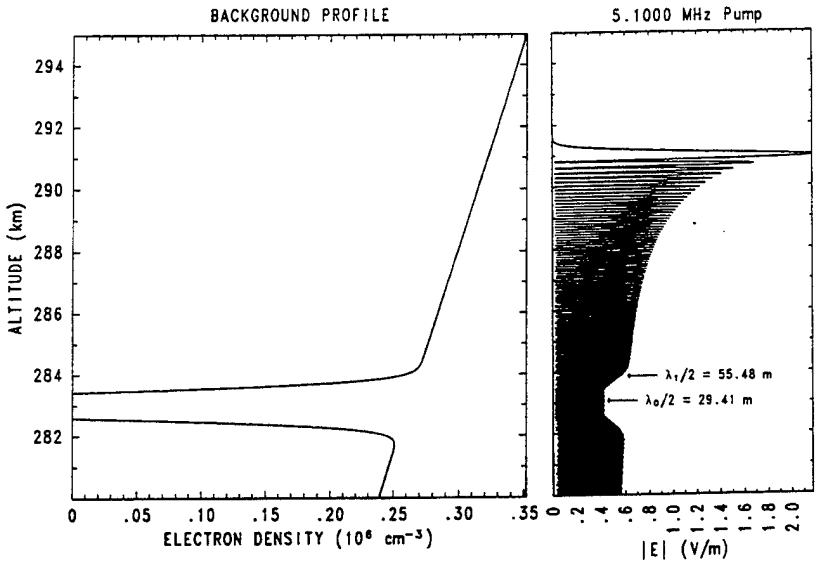


Figure 6. Computed standing wave in a horizontally stratified ionosphere with a localized electron depletion.

electron attachment the negative ions became tied by gyro orbits to the magnetic field lines and the rocket passed through the top edge of the electron depletion.

Both electron and ion densities were measured using probes on the IFH rocket. The Naval Research Laboratory provided a double Langmuir probe that was biased at a fixed voltage (+6 V) to collect electrons. The separation between the two probes was 1.26 m. Two disk probes biased to collect positive ions were mounted on the payload skin just aft of the E field booms. The disk probes, one small (1.5 inches) and one large (3 inches), were provided by the Phillips Laboratory.

Multiple electric field and plasma wave receivers were provided by the Naval Research Laboratory covering the 0- to 12-MHz frequency range. The electric fields were detected with booms deployed from the payload after launch. The spacings between sensors on the booms varied from 0.3 to 3.4 m. Plasma waves near the HF pump frequency of 5.1 MHz were detected with two radio frequency sensors with spacings of 5.5 and 11 cm. The radio frequencies are downconverted to a 150-kHz band and digitized with 12-bit words. Details of the downconverter are given by *Haas et al.* [1995].

The energy spectrum of suprathermal electrons between 0.5 and 50 eV was to be detected with the thermal-electron capped hemispherical spectrometer (TECHS) instrument provided by Marshall Space Flight Center. The TECHS was mounted on the opposite side of the EF5 and EF6 booms. Unfortunately, after launch the voltage sweep was lost on the instrument and the energies of the detected electrons have not been determined.

3. In Situ Measurements

The data from the flight instruments can be divided into five phases (Figure 3). In phase 0, before the release, the electron density measurements showed both 3.4-s spin modulation and fine scale irregularities (Figure 3a). At this time, the 5.1-MHz reflection altitude was 305 km. The irregularities were probably generated by the heater because their amplitude grew as the rocket moved closer to the HF reflection height. The normalized fluctuations ($\Delta n/n$) increased from 3% at 270 km altitude to 7% at 283 km just before release. The low-frequency electric field data also show enhanced fluctuations during this period. These fluctuations may provide evidence for field-aligned irregularities extending down from the reflection altitude. These irregularities could be the result of the constant ionospheric heating for more than 1 hour prior to launch. The $\Delta n/n$ fluctuation level was less than 1% on the downleg of trajectory in the same altitude range when the rocket was well away from the heated volume. During another launch from Puerto Rico, *Kelley et al.* [this issue] reported in situ measurements of small-scale irregularities down to 18 km below the 5.1-MHz reflection altitude.

Before the rocket could cross the ambient reflection level for the 5.1-MHz HF waves, the electron attachment chemical was expelled from a heated tank. For the first 9 s (phase 1) after the CF_3Br release, the depressed electron plasma is pitted with narrow microcavities with sizes of <1 m (see Figure 3b), where size is assumed to be the rocket velocity times the measurement time. Since the rocket was moving at an oblique angle (123°) with respect to the magnetic field, the sizes of stationary, field-aligned structures would be smaller by a factor of $\sin(123°) = 0.839$ than the scales given in Figures 3 and 4. A high-resolution sample of this data is plotted versus distance along the trajectory in Figure 4a. The microcavities were recorded by both Langmuir probes. Usually, the features smaller than the distance between the two probes were detected only at one of the probes. The microcavities larger than the 2-m probe separation were usually measured with both probes simultaneously. Strong electric fields saturated the E field instrument during this early time.

During phase 2 the IFH payload entered the negative-ion cloud boundary layer where the electron density was 1 to 10% of ambient (Figures 3b and c). The density fluctuations ($\Delta n/n \sim 17\%$) were dominated by spikes with widths of about 7 m. A detailed spatial sample of these microspikes is

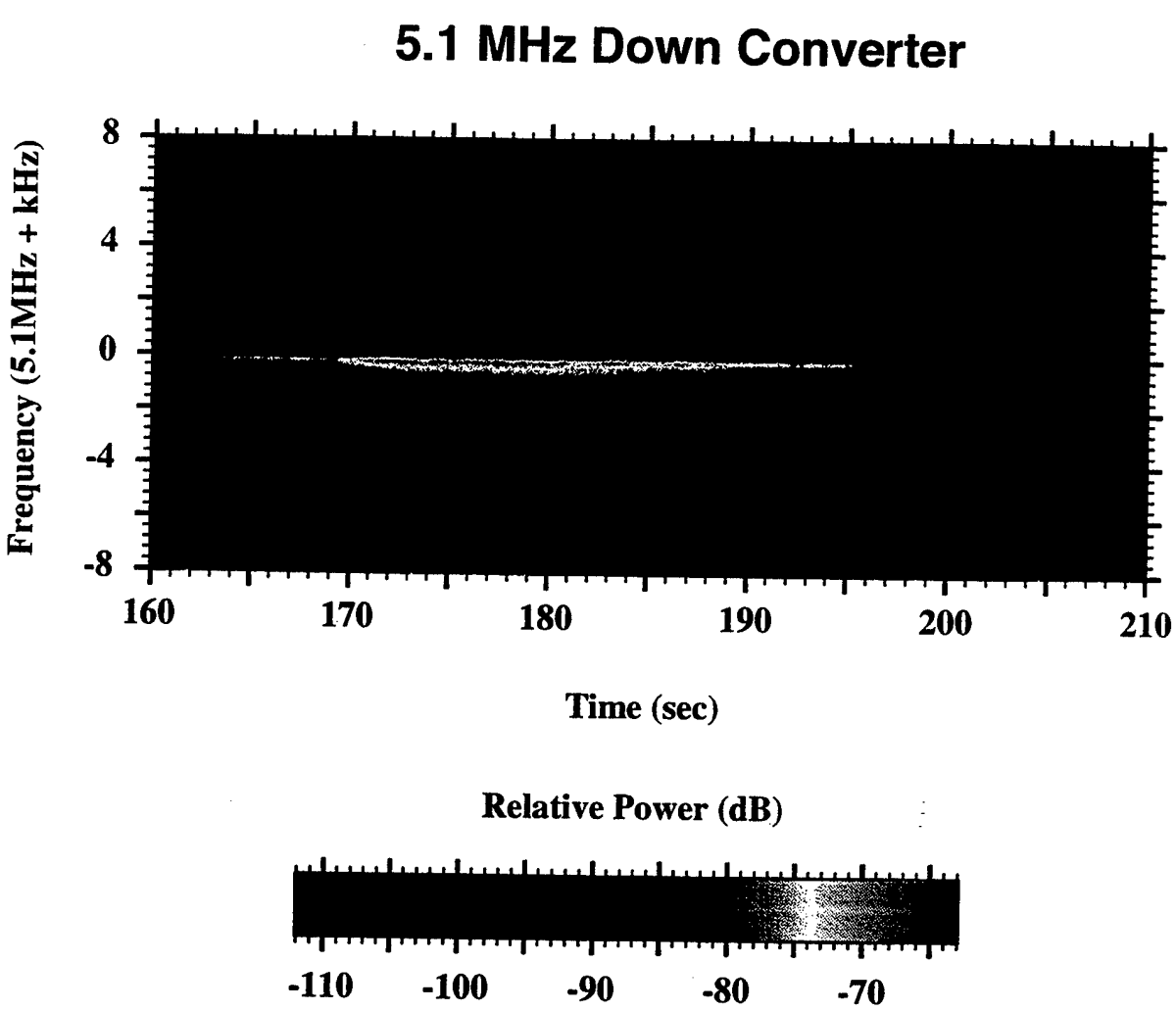


Plate 1. Spectrogram of the pump fields and associated Langmuir waves. The spectra is naturally divided into (1) a prerelease narrow line, (2) downshifted broadening after release, (3) a burst when passing through the HF reflection level 196 to 199 s after launch, and (4) a disappearance of the pump thereafter.

illustrated in Figure 4b. In situ measurements showed electric field reversals simultaneous with the microspikes and cavities.

The transition between electron microcavities in phase 1 to electron microspikes in phase 2 has been explained by *Scales et al.* [1994, 1995] using electrostatic simulations of the three-component plasma. A shear in the electron velocity at the negative-ion/electron boundary layer seems to drive the instability that produces the irregularities. At early times, in the center of the electron depletion, the negative ions form density spikes that are neutralized by electron density cavities. The background positive-ion plasma remains relatively uniform. At later times, in the boundary layer, the electron spikes form as the negative-ion and positive-ion plasma evolves into density cavities. Unfortunately, the frequency response of the positive-ion disk probe measurements was too low to verify the computed variations in positive-ion density.

The amplitude of density fluctuations was at a minimum ($\Delta n/n \leq 7\%$) when the payload was between the negative-ion boundary layer and the HF wave reflection altitude. During phase 3 the rocket payload approached the new 5.1-MHz reflection region (between 311 and 316 km altitude at times 21 through 30 s after release). As with phase 0, the irregularity fluctuation level increased as the rocket approached the critical density of 3.23×10^5 cm^{-3}. High-resolution spatial plots of selected data are illustrated in Figure 4. The time, rocket speed, and altitude for the start of each data sample is given. The deepest cavities between 197 and 198 s in Figure 3d show internal structures in greater detail in Figure 4c, where the distance along the trajectory is used as the horizontal axis. The two 40% cavities near 1.1-km distance (317.4-km altitude) at the right side of Figure 4c have radii of about 4 m.

The chemical release at 283 km expanded over the altitude (290 km) where the HF wave reflected before the release. Consequently, any irregularities at that level are overwhelmed by the dissociative attachment of electrons. When the rocket passed through the new reflection level at the top of the electron density hole, the observed irregularities must have been newly formed because the HF wave did not penetrate to this level before the release, 10 s earlier. The average radius of the cavities in this region were on the order of an O^+ ion cyclotron radius (4 m). Irregularities of this size have been detected with radio backscatter by *Minkoff et al.* [1974], *Belenov et al.* [1977], and *Coster et al.* [1985]. In situ

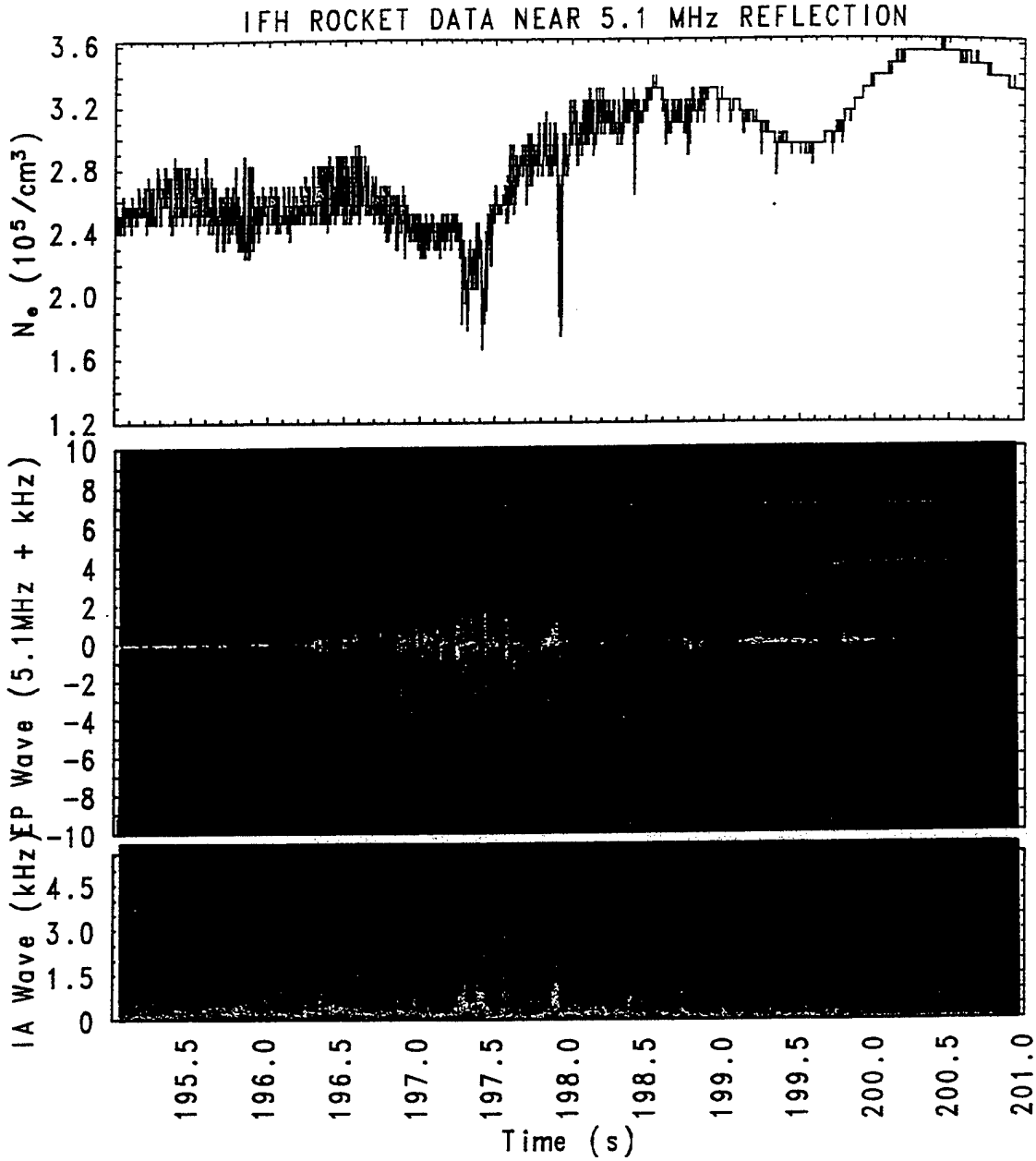


Plate 2. Detail of the electron density, Langmuir waves around 5.1 MHz and low-frequency ion acoustic waves near the HF reflection level. The Langmuir waves and ion acoustic waves seem to be trapped or guided by the density cavities. Spectra of low-frequency electric fields are measured between sensors EF1 and EF4 of Figure 2.

measurements of similar size structures have been reported by *Kelley et al.* [this issue] during a rocket flight through the Arecibo heater beam on June 9, 1992.

After passing through the reflection level, the rocket entered the unheated plasma labeled phase 4 in Figure 3d. Here the electron density fluctuations decreased to values of $\Delta n/n < 1\%$.

The chemical release and subsequent electron depletion modified the HF wave electric field. Figure 5 shows the changes in the 5.1-MHz electric field measured in situ between 169.0 and 169.5 s after launch. At this time the rocket velocity was 1411 m/s, with an angle of 17° with the vertical. The electric field measurements shows successive minima as the rocket passes through the nulls in the standing wave (or Airy) pattern of the reflected 5.1-MHz transmission. The minimum distance between the nulls is expected to be $\lambda_1/2$, where $\lambda_1 = c/(n_1 f_1)$; c is the speed of light, n_1 is the refractive index in the magnetoplasma, and $f_1 = 5.1$ MHz is the transmitter frequency. The plasma density just below release is $n_e = 2.6 \times 10^{11}$ m^{-3}, the gyro frequency in the ionosphere over Arecibo is $f_{ce} = 1.07$ MHz, and the propagation angle between a vertical wave vector and the magnetic field is $\Theta = 40°$. With these parameters the refractive index for the 5.1-MHz wave is $n_1 = 0.530$ and the wavelength in the plasma is $\lambda_1 = 111$ m. Dividing $\lambda_1/2$ by the rocket velocity gives the scale length shown in Figure 5a. During the upleg of the flight the nulls in the 5.1-MHz wave measurements are greater than $\lambda_1/2$ (Figures 5a, 5b, and 5c).

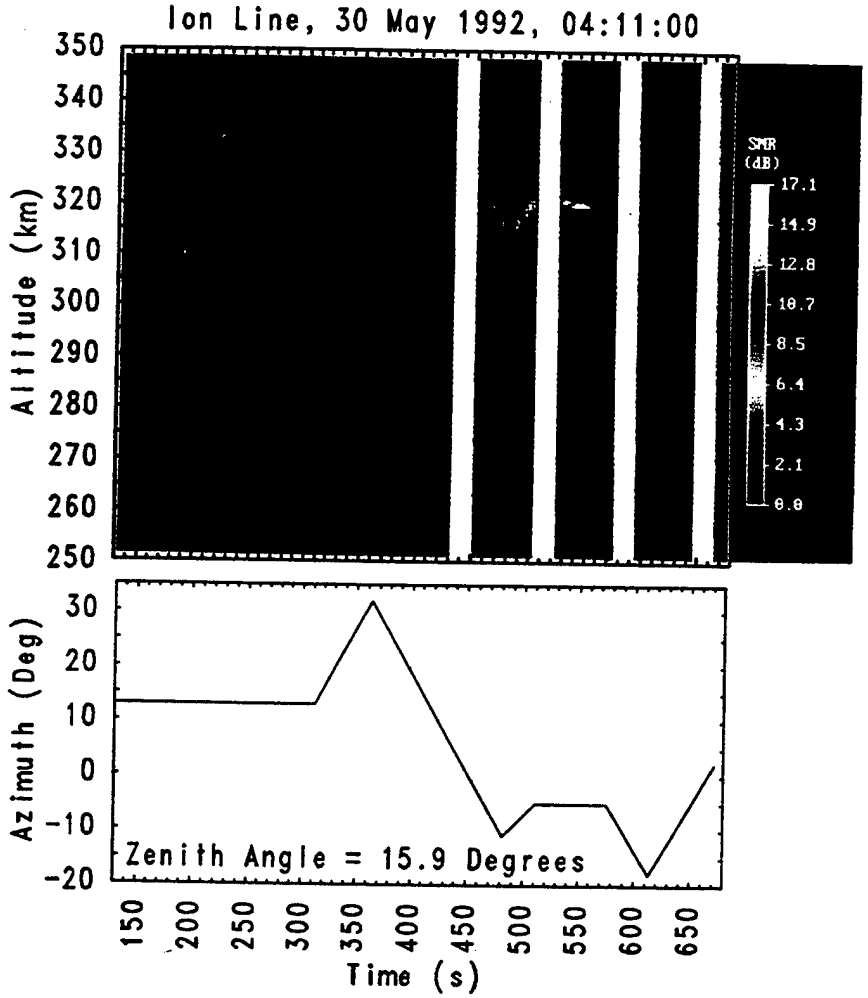


Plate 3. Incoherent scatter ion line obtained by the Arecibo radar during the IFH experiment. The radar backscatter is affected by (1) reduction of electron density in the *F* layer, (2) sidelobe scatter from the rocket body, and (3) enhanced ion acoustic waves at the top of the *F* layer hole. The radar was operated in a spectral mode during the times of data gaps.

After the release at 169.29 s the free electrons vanish and a horizontally stratified standing wave would have nulls spaced by $\lambda_0/2$, where $\lambda_0 = c/f_1 = 58.82$ m in free space. Also, the electric field amplitude should drop by a factor of $(n_1)^{1/2} = 0.73$ to maintain that same power density as in the unmodified plasma. This effect is illustrated by the one-dimensional computation of the standing wave at 5.1 MHz, reflecting above an electron density depletion in a horizontally stratified layer (Figure 6). The ionospheric profile represents a vertical cut through the ionosphere 3 s after the CF_3Br release. The one-dimensional calculations for these fields do not account for the limited horizontal extent of the ionospheric hole and the bending of the HF wave fronts.

The in situ measurements of the electric fields for the 5.1-MHz wave (Figure 5) differ from the calculated fields (Figure 6) in several ways. The electric field amplitude increases by a factor of 3 after the release, and the nulls have spacings both greater and less than $\lambda_0/2$. The power density immediately after release increases by a factor of 16 or 12 dB. This enhancement is attributed to focusing by the ionospheric hole. The erratic variations in the wave nulls result from the rocket passing through a time-varying interference pattern established in the electron density cavity. A multidimensional model of these fields is being constructed to simulate this pattern.

The frequency spectrum from the 5.1-MHz downconverter on the IFH rocket is shown in Plate 1. While the rocket payload was inside the ionospheric hole, the 5.1-MHz pump was broadened into a downshifted spectrum with a width of 1.3 kHz. This broadening started at the time of the release (169.3 s after launch). The intensities of the downshifted waves were >10 dB above the prerelease level. The source of the broadening may have been scatter of electromagnetic waves by the small-scale irregularities at the edge of the negative-ion cloud [*Scales et al.*, 1994, 1995].

The width and intensity of the waves returned to the prerelease level 26 s after release when the rocket had passed through the phase 2 irregularities. A localized burst of broadband HF noise ($\Delta f \leq 8$ kHz) was next recorded when the payload transited the critical region of the ionosphere during phase 3. Accurate determination of the HF reflection point is not possible from the electron density data because of modulation from the spin and precession of the payload. On the basis of the disappearance of the 5.1-MHz

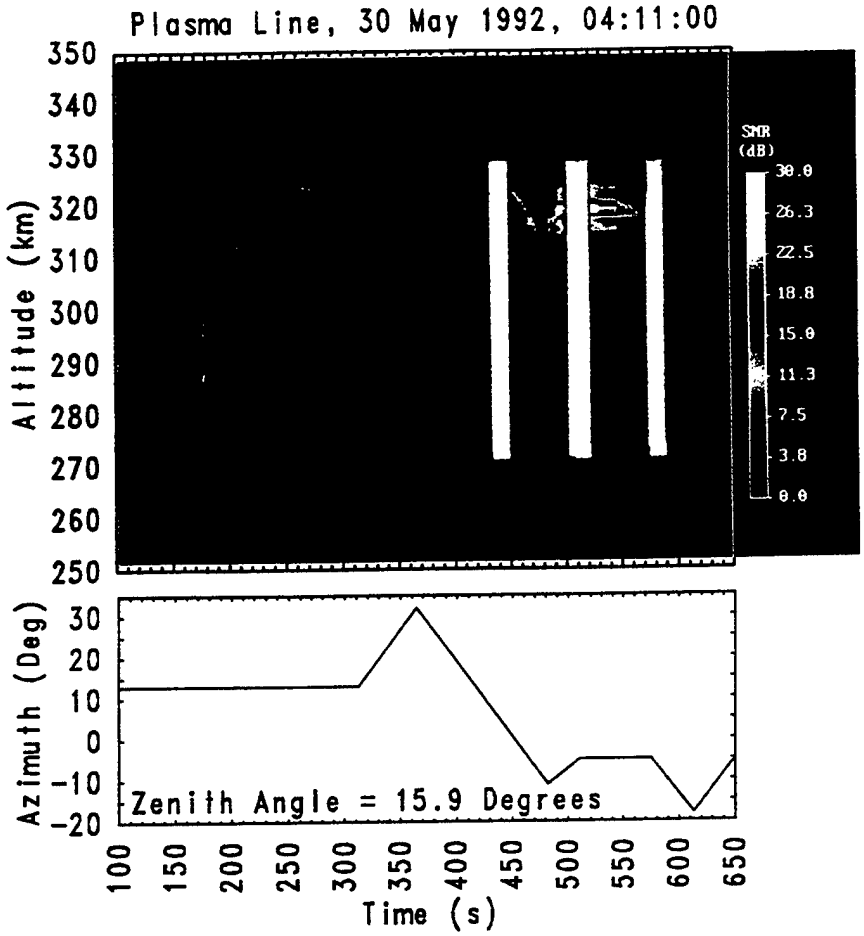


Plate 4. Incoherent scatter measurement of the downshifted plasma line at 430 − 5.1 MHz. The plasma line shows (1) the HF reflection layer before the release, (2) a brief focusing after release, (3) a new reflection height at the top of the ionospheric hole, and (4) a 20-dB increase in the plasma line after focusing occurs.

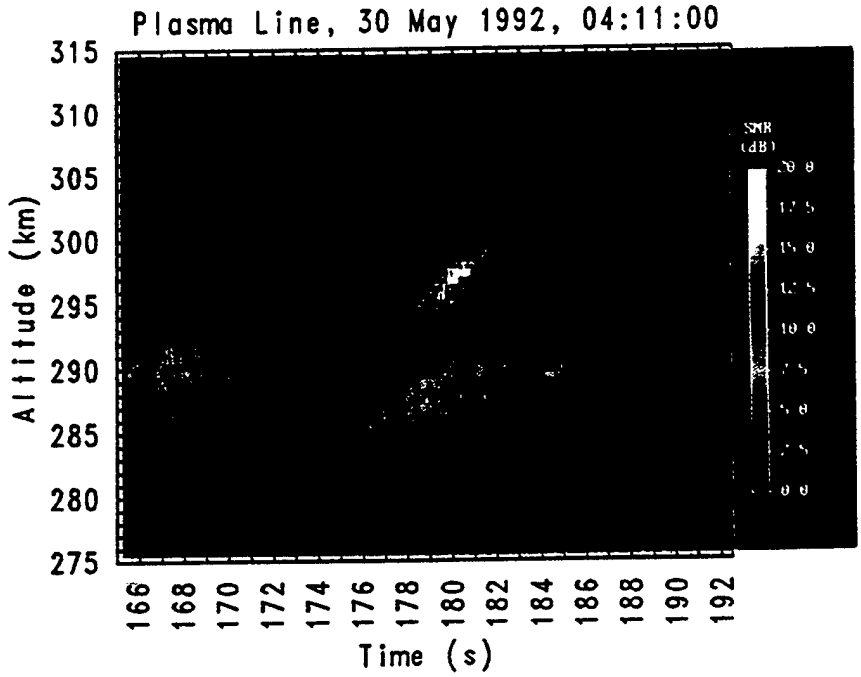


Plate 5. Plasma line splitting during the 10 s after the release of CF_3Br. The strongest backscatter from Langmuir waves is 10 dB stronger than the prerelease plasma line. The downshifted Langmuir waves illustrated by the in situ observations of Figure 6 may be the source of the scattered radar signal.

signals from the plasma wave receiver (Plate 2), the critical level was penetrated between 199 and 200 s after launch.

A plot of the electron density irregularities, the spectral deviations from 5.1 MHz, and low-frequency electric fields shows evidence of Langmuir and ion acoustic waves coincident with the density cavities near the reflection level (Plate 2). These data are consistent with formation of field-aligned irregularities that guide Langmuir generated by parametric decay of the pump electromagnetic wave. The measured irregularities seem to be field-aligned ducts described by *Muldrew* [1978, 1988] that guide Langmuir waves.

The width of the high-frequency spectral broadening near the reflection level is about 2 kHz. As will be shown later in Figure 7, the broadening in the Langmuir wave scatter of the 430-MHz radar is on the order of 10 or 20 kHz. The apparently narrow in situ spectral measurements may be due to the presence of negative ions at the top of the ionospheric hole. The narrowness of the spectrum is not due to instrument instrumental which had a ±75-kHz bandwidth capability (Table 1). A more detailed discussion of the in situ wave measurements near the HF reflection region are given by P. Rodriguez et al. (Evidence of parametric wave interactions in the Ionospheric Focused Heating Experiment, submitted to *Geophysical Research Letters*, 1995).

4. Incoherent Scatter Radar Measurements

Even though the 5.1-MHz HF transmitter was in operation continuously, no enhanced ion lines were recorded up to 450 s after the launch of the IFH rocket. The trajectory of the rocket was measured by the sidelobes of the Arecibo radar. Backscatter by the thermal ion line showed the reduction in electron density following the chemical release. Both of these effects are shown in Plate 3 (left side). The azimuth of the radar was fixed at 13°, 300 s after launch. During this period, no enhanced scatter from ion acoustic waves was measured even though strong low-frequency electric fields were measured in situ (Plate 2). This may have been because the Arecibo radar beam was not aligned with the turbulent area inside the rapidly evolving ionospheric hole. After this time the azimuth was scanned from one edge of the ionospheric hole to the other.

Plate 3 (right side) shows the effects of focusing through the artificial hole. Because of a 40 m/s westward drift, the hole center moved to an azimuth of −5°, 500 s after release. Around this time a strongly enhanced ion line (EIL) was observed where the HF wave reflected at the critical density region. This region occurred where the pump frequency equaled the electron plasma frequency across the top of the artificial hole. The intensity of this EIL was the largest recorded during the May–July period of the *El Coqui* rocket campaign. This unusual intensification of the EIL was most likely the result of focusing by the chemically produced hole in the *F* region.

The radar-downshifted plasma line (PL) went through a number a changes during the IFH experiment (Plate 4). Before the chemical release a weak PL was recorded near the HF reflection height of 290 km in the *F* region. During the first 10 s following the release the plasma line was recorded at multiple altitudes between 285 and 300 km. The intensity increased by 10 dB in a path aligned with the rocket trajectory between 295 and 300 km. After this initial enhancement the PL vanished, only to reappear at the top of the chemically produced cavity with an intensity comparable to the prerelease value. The absence of plasma lines when the instruments on the rocket were recording strong Langmuir waves (Plate 2) is again attributed to the misalignment of the radar beam inside the active region of the ionospheric hole.

The strongest enhancements in the PL occurred between 280 and 430 s after release. The 10-km range spread at this time (Plate 4) may be the result of sidelobe smearing associated with the coded radar pulse. The actual range in altitude is probably similar to the 2-km spread shown from the strongest enhanced ion line in Plate 3. At 480 s after launch (310 s after release) the azimuth scan reversed directions. Both the plasma line and enhanced ion line show spatial symmetry around this time. This indicates that turbulence was excited at the HF reflection layer by focusing through a long-lived ionospheric lens.

The plasma line 10 s after release was enhanced by 10 dB (Plate 5). After release at 169.3 s the HF reflection layer splits into a region along the rocket trajectory near 300 km altitude and a region near the original 290-km level. The large plasma line enhancement at 300 km altitude is coincident with the strong downshifted Langmuir waves recorded on the rocket (Plate 1). The transient disturbance displayed in Plate 5 may be associated with turbulence in the boundary layer of the negative ion cloud [*Scales et al.*, 1994].

The spectra of the plasma lines and enhanced ion lines were obtained during the data gaps shown as white vertical bars in Plate 4. A sample of the spectra during the period of strong enhancements 335 s after release (504 s after launch) is illustrated in Figure 7. The plasma line spectrum in Figure 7a shows all of the characteristic features (i.e., decay line, cascade bump) of an unfocused plasma line at Arecibo. The enhanced ion line is easily distinguishable from the background (thermal) ion line (Figure 7b). A typical feature of enhanced ion lines is that the upshifted and downshifted components are nearly symmetrical. This is the case for the focused enhanced ion line (Figure 7b). The only unusual features of the observed plasma or ion lines were their large intensities.

Besides electromagnetic wave focusing, processes such as linear mode conversion [*Mjølhus*, 1990] could be responsible for enhanced coupling to Langmuir waves. If linear mode coupling were the source of the enhanced ion lines, one would expect asymmetries in the plasma line and ion line spectra. This is because the upgoing electromagnetic wave would be expected to mode convert into an upgoing *Z* mode which reflects and become a downgoing Langmuir wave [*Mjølhus*, 1990]. With a downgoing Langmuir wave, the downshifted plasma line and the enhanced ion lines should be absent. We have also ruled out direct mode coupling because the wave normals of the rays do not have the critical angle required for linear conversion.

On the basis of ray tracing through a model ionosphere, the source of the enhanced turbulence has been attributed to focusing by the chemically produced cavity. Without the chemical release the pump beam reflects near 300 km, independent of ray launch angle (Figure 8a). In the early phase of the release the strong electron density gradients form a focal point inside the ionospheric hole. Using an electron density model that matches the ISR electron density profile 10 s after the CF_3Br release, the focus is 8 km from the center of the hole, whereas the upper reflection boundary

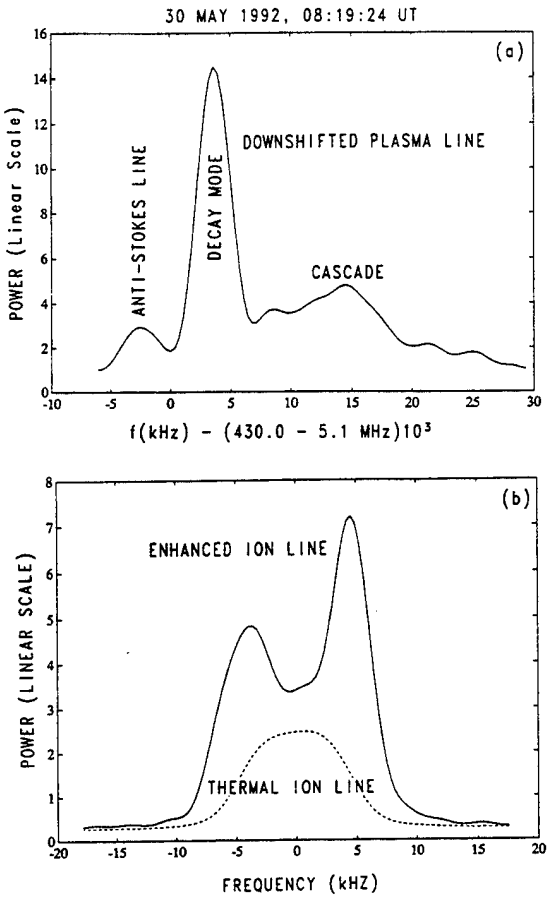


Figure 7. Spectrum of the (a) downshifted plasma line and (b) enhanced ion line during the time of intense HF focusing.

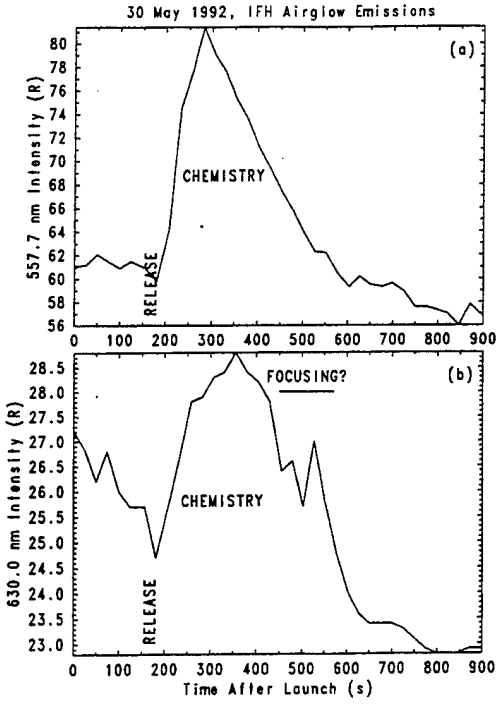


Figure 9. (a) Green-line (557.7 nm) and (b) red-line (630.0 nm) enhancements following the CF_3Br release during the IFH experiment. The bulk of these enhancements can be explained by excitation of atomic oxygen after mutual neutralization of O^+ and Br^- in the chemically modified plasma. A small red-line increase is observed during the "focusing" period of strong enhanced ion and plasma lines shown in Plates 3 and 4.

of the hole is 15 km from the center (Figure 8b). The hole evolves by the action of chemistry and plasma transport. At some point in the evolution, the HF focal point and the reflection boundary coalesce, yielding a region of large electric fields. This condition occurs at about 300 s after release. The large electric fields responsible for the enhanced radar backscatter occur because (1) the HF waves are focused onto a reflection contour of the electron density and (2) the group velocity goes to zero at the reflection level, causing an accumulation of the wave energy. The focusing processes are limited by diffraction [*Bernhardt and Duncan*,

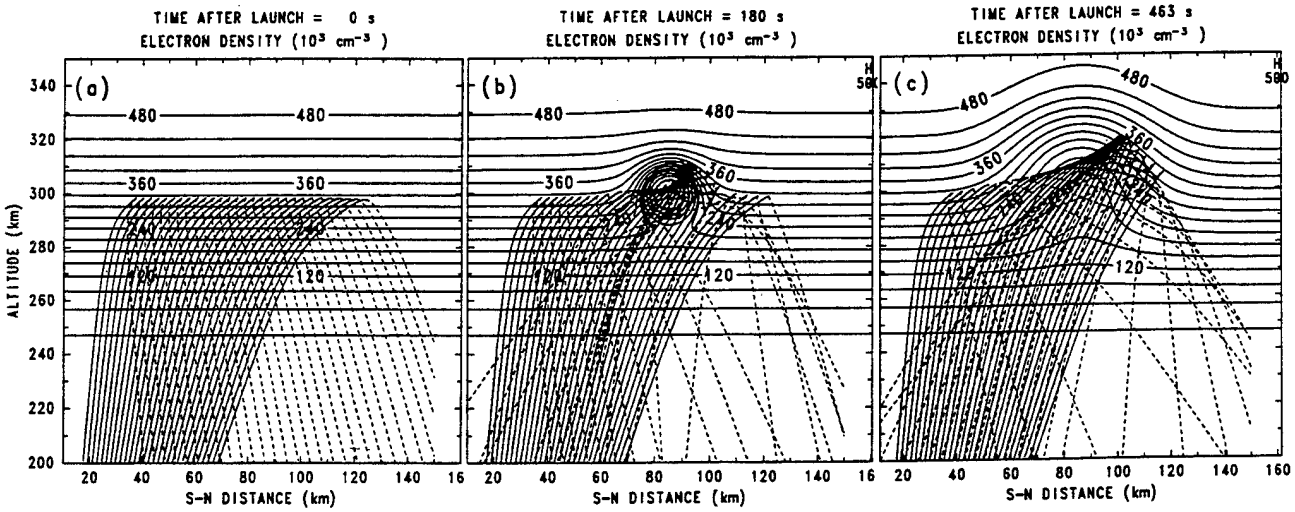


Figure 8. Rays from the 5.1-MHz HF transmitters propagating through (a) the unmodified *F* layer, (b) the ionospheric hole 10 s after release, and (c) the hole at a later time when the focal point is at the HF reflection level. The electron density contours are derived from a spherical release model adjusted to fit density measurements obtained from the incoherent scatter radar.

1987] and by irregularities that may scatter the electromagnetic waves. Figure 8c shows a tube of rays focused on the upper boundary of the ionospheric hole 293 s after release. This is the time period when the unusual enhancements of ion acoustic and electron plasma waves were observed in the ISR data.

5. Ground-Based Optical and Radio Observations

A number of optical instruments were operated at the Arecibo Observatory during the IFH experiment. Photometers with 5° fields of view were pointed at the release location and recorded the 557.7- and 630.0-nm emissions. A low-light-level charge-coupled device (CCD) camera with a 60° field of view recorded 630-nm emissions from the release and HF focusing regions. A description of these instruments is given by *Bernhardt et al.* [1988].

An enhancement of airglow during the IFH experiment occurred immediately after the CF_3Br release. Figure 9 illustrates the 20 Rayleigh increase in 557.7-nm (green line) and the 4 Rayleigh increase in the 630.0 nm (red line) of atomic oxygen. These enhancements probably come from the excited atomic oxygen that was a product of mutual neutralization reaction (2). *Bernhardt* [1987] predicted that $O(^1D)$ should be the primary excited oxygen state from Br^- and O^+ mutual neutralization. On the basis of the observations that the 557.7-nm intensities were 5 times larger than the 630.0-nm intensities, we conclude that $O(^1S)$ was the primary product of the reaction

$$Br^- + O^+ \rightarrow Br(^2P^0) + O(^1S) + 6.065\ \text{eV} \quad (3)$$

and that green-line and red-line emissions came from

$$O(^1S) \rightarrow O(^1D) + h\nu(557.7\ \text{nm}) \qquad \tau_1 = 0.71\ \text{s} \quad (4)$$

$$O(^1D) \rightarrow O(^3P) + h\nu(630.0,\ 636.4\ \text{nm}) \qquad \tau_2 = 147\ \text{s} \quad (5)$$

where τ_1 and τ_2 are the radiative lifetimes of the states.

The process described by (3), (4), and (5) is consistent with the observations. The red-line emission peaked 80 s after the maximum green-line emission. Also, the red-line intensity was 20% of the green-line intensity. Both of these effects can be attributed to the relatively long lifetime of the $O(^1D)$ state which is formed by electronic transition from the shorter-lived $O(^1S)$ state. The integrated volume emission rate was reduced by diffusion of the $O(^1D)$ atoms and by collisional quenching before radiation [*Bernhardt et al.*, 1989].

At the time of the strongly enhanced ion line, between 450 and 570 s after launch, a small 1 Rayleigh enhancement is observed in the 630.0-nm (red line) channel of the photometer. This peak is coincident with the enhanced ion lines produced by focusing of the HF radio waves. Its magnitude is surprisingly small. Since HF-induced perturbations in the green line are typically 5% of the red-line values [*Bernhardt et al.*, 1989], a corresponding peak would be too weak to be detectable in the 557.7-nm channel (Figure 9a). The low-light-level CCD camera did not show any enhancements in red-line airglow.

Previous experiments at Arecibo have indicated that extremely strong enhanced ion lines are accompanied by larger (~50 Rayleigh) increases in red-line airglow [*Bernhardt et al.*, 1988, 1989]. This was not the case for the IFH experiment. We are currently investigating a number of explanations for the apparent weakness of heater-induced airglow. These include the effects of residual negative ions in the plasma and tilting of the HF beam from vertical.

Using ground-based spectrum analyzers attached to wide-band HF antennas, observations of stimulated electromagnetic emissions (SEE) were attempted at Arecibo, Puerto Rico, and Providenciales, Caicos. No SEE was detected during the entire *El Coqui* campaign. This is not totally unexpected because previous experiments have demonstrated that SEE is more often observed at higher-latitude heating facilities [*Thidé et al.*, 1989].

Los Alamos and NRL set up bistatic HF propagation links between Dominica and Providenciales, Caicos, during the IFH release and other CRRES rocket experiments. Propagation frequencies were chosen to reflect near the release points for chemicals. The CF_3Br release produced a large Doppler shift in the HF wave passing through the disturbed regions. A manuscript describing the results of these experiments is in preparation (P. E. Argo, private communication, 1994).

6. Conclusions

The IFH chemical release produced a large (>30-km diameter) hole in the *F* layer centered within 6 km of the 5.1-MHz HF beam. The reactions between CF_3Br and the ambient electrons yielded a turbulent negative-ion plasma. The action of the high-power radio wave on the negative ions, positive ions, and electrons produced microcavities with diameters of <1 m and factor of 10 or greater reductions in the already depressed plasma density. In the boundary layer, between the negative ion and electron dominated regions, electron density spikes were recorded with the in situ Langmuir probes. These spikes may result from cavities in the negative ions that are filled by electrons to maintain charge neutrality. The negative-ion, positive-ion, and electron simulation models of *Scales et al.* [1993, 1994, 1995] are being used to study this process.

The in situ measurements of the electron density near the HF reflection region at the top of the ionospheric hole show electron plasma waves trapped or guided by cavities the size of the positive-ion gyroradius. The frequency spread of these Langmuir waves is 8 kHz or less. The structure observed at the reflection region may be associated with field-aligned cavities guiding Langmuir waves generated by the parametric decay instability (PDI) as described by *Muldrew* [1978, 1988]. Processes that can generate field-aligned irregularities by the thermal parametric instability have been summarized by *Mjølhus* [1993]. Horizontal structures may be formed by the ponderomotive force in the peaks in the standing wave of the HF pump [*Leyser and Thidé*, 1988].

From the in situ and ground-based observations we deduce that the artificial ionospheric hole yielded 10- to 20-dB enhancements in the pump wave intensity by focusing. This degree of focusing is consistent with the predictions by *Bernhardt and Duncan* [1987]. The strongest focusing lasted for 150 s after the ionospheric hole had evolved so that the focal point was located on the reflection level at the top of the hole. The amplitudes of the ion acoustic waves and electron plasma oscillations were increased by more than 20 dB during the period of the strongest focusing. The weak-

ness of the red-line airglow enhancements during this time is difficult to explain. Airglow induced by RF heating usually comes from surprathermal electrons colliding with ambient atomic oxygen. The negative ions trapped on the magnetic field lines below the HF reflection level may have scattered the suprathermal electrons before they could excite the oxygen.

An experiment complementary to IFH used a barium release from a sounding rocket. This experiment (CRRES AA 2) produced an electron density enhancement in the powerful radio beam from the Arecibo HF facility [*Djuth et al.*, this issue]. It is instructive to compare the results of the IFH and AA 2 radar measurements. Following the barium release of the AA 2 experiment, the intensity of the upshifted plasma line increased by about 12 dB and the downshifted plasma line vanished below the noise floor of the radar receiver. The downshifted and upshifted enhanced ion lines showed corresponding asymmetries. The measured asymmetric plasma lines may have been the result of ion cloud expansion or steep density gradients. Because an electron density enhancement defocuses an HF beam, focusing does not explain the AA 2 results. The source of the asymmetric radar spectra is not understood and should be a subject for future research [*Djuth et al.*, this issue].

Both the IFH and AA 2 experiments yielded large enhancements in HF-induced Langmuir turbulence. The IFH experiment produced the largest increase in plasma line strength (20 dB), but AA 2 yielded the largest absolute plasma line intensity. A quantitative comparison of the two techniques is difficult because the intensity of the plasma line before the IFH release was much less than the intensity before the AA-2 release. It is an open question whether ionospheric holes or electron enhancements produce the largest effect on ionospheric heating.

The IFH experiment has demonstrated that the release of an electron attachment chemical will form an artificial lens in the *F* region that may be used to increase the power density of any HF facility. The effective radiated power of these facilities can be increased by about 20 dB for periods of time longer than 2 min. Similar IFH experiments should be attempted in the future to give a temporary boost in the power levels available for ionospheric heating. This boost may yield phenomena that are not produced with current heating facilities.

Acknowledgments. The authors gratefully acknowledge the support of the staff of Arecibo Observatory and graduate students from the Department of Physics and Astronomy, Clemson University, and the Department of Electrical Engineering, Cornell University. The Arecibo Observatory is part of the National Astronomy and Ionosphere Center, operated by Cornell University under contract with the National Science Foundation. The research at NRL was supported by the National Aeronautics and Space Administration and the Office of Naval Research. One of the authors (F.T.D) acknowledges support from NASA under contract NAS8-39075 and from the Phillips Laboratory, Hanscom Air Force Base, under contract F19628-92-C-0168.

The Editor thanks T. L. Arce and J. D. Hansen for their assistance in evaluating this paper.

References

Belenov, A. F., et al., Parameters of artificial small-scale ionospheric irregularities, *Radiophys. Quantum Electron.*, Engl. Transl., *20*, 1240–1245, 1977.

Bernhardt, P. A., A critical comparison of ionospheric depletion chemicals, *J. Geophys. Res.*, *92*, 4617–4620, 1987.

Bernhardt, P. A. Field-aligned dynamics of chemically induced perturbations to the ionosphere, *J. Geophys. Res.*, *96*, 13,887–13,900, 1991.

Bernhardt, P. A., and L. M. Duncan, Ionospheric focused heating–A hybrid modification experiment, *J. Atmos. Terr. Phys.*, *49*, 1107–1117, 1987.

Bernhardt, P. A., L. M. Duncan, and C. A. Tepley, Artificial airglow excited by high-power radio waves, *Science*, *242*, 1022–1027, 1988.

Bernhardt, P. A., C. A. Tepley, and L. M. Duncan, Airglow enhancements associated with plasma cavities formed during ionospheric heating experiments, *J. Geophys. Res.*, *94*, 9071–9092, 1989.

Bernhardt, P. A., P. Rodriguez, C. L. Siefring, and C. S. Lin, Field-aligned dynamics of chemically induced perturbations to the ionosphere, *J. Geophys. Res.*, *96*, 13,887–13,900, 1991a.

Bernhardt, P. A., W. A. Scales, S. M. Grach, A. N. Karashtin, D. S. Kotik, and S. V. Polyakov, Excitation of artificial airglow by high power radio waves from the "SURA" ionospheric heating facility, *Geophys. Res. Lett.*, *18*, 1477–1480, 1991b.

Coster, A. J., F. T. Djuth, R. J. Jost, and W. E. Gorden, The temporal evolution of 3-m striations in the modified ionosphere, *J. Geophys. Res.*, *90*, 2807–2818, 1985.

Djuth, F. T., *El Coqui* active experiments explore the ionosphere, *Soundings*, *2*, 1–3, 1993.

Djuth, F. T., M. P. Sulzer, J. H. Elder, and K. M. Groves, The CRRES AA-2 release: HF wave-plasma interactions in a dense Ba^+ cloud, *J. Geophys. Res.*, this issue.

Dubois, D. F., H. A. Rose, and D. Russell, Excitation of strong Langmuir turbulence in plasmas near critical density: Application to HF heating of the ionosphere, *J. Geophys. Res.*, *5*, 21,221–21,272, 1990.

Duncan, L. M., S. P. Sheerin, and R. A. Behnke, Observations of ionospheric cavities generated by high-power radio waves, *Phys. Rev. Lett.*, *61*, 239–242, 1988.

Fejer, J. A., Ionospheric modification and parametric instability, *Rev. Geophys.*, *17*, 135–157, 1979.

Ganguli, G., P. A. Bernhardt, W. A. Scales, P. Rodriguez, C. L. Siefring, and H. A. Romero, Physics of negative ion plasmas created by chemical releases in space, in *Physics of Space Plasmas*, *SPI Conf. Proc. and Reprint Ser.*, vol. 12, edited by T. Chang, pp. 161–183, Scientific, Cambridge, Mass., 1993.

Haas, D. G., C. L. Siefring, M. M. Baumback, P. Rodriguez, and D. P. McNutt, Rocket-borne downconverter for measuring space plasma turbulence, *Rev. Sci. Instrum.*, *66*, 1056–1060, 1995.

Hansen, J. D., G. J. Morales, L. M. Duncan, J. E. Maggs, and G. Dimonte, Large-scale ionospheric modifications produced by nonlinear refraction of an HF wave, *Phys. Rev. Lett.*, *65*, 3285–3288, 1990.

Hansen, J. D., G. J. Morales, and J. E. Maggs, Large-scale HF-induced ionospheric modifications: Theory and modeling, *J. Geophys. Res.*, *97*, 17,019–17,032, 1992.

Kelley, M. C., T. L. Arce, J. Saloway, M. P. Sulzer, T. Armstrong, M. Carter, and L. M. Duncan, Density depletions at the 10-m scale induced by the Arecibo heater, *J. Geophys. Res.*, this issue.

Leyser, T. B., and B. Thidé, Effects of pump-induced density depletions on the spectrum of stimulated electromagnetic emissions, *J. Geophys. Res.*, *93*, 8681–8688, 1988.

Minkoff, J., P. Kugelmenn, and I. Wiessman, Radio frequency scattering from a heated ionospheric volume, 1, VHF/UHF field-aligned and plasma-line backscatter measurements, *Radio Sci.*, *9*, 941–955, 1974.

Mjølhus, E., On linear conversion in a magnetized plasma, *Radio Sci.*, *25*, 1312–1339, 1990.

Mjølhus, E., On the small scale striation effect in ionospheric radio modification experiments near harmonics of the electron gyro frequency, *J. Atmos. Terr. Phys.*, *55*, 907–918, 1993.

Muldrew, D. B., The role of field-aligned ionization irregularities in the generation of HF-induced plasma line at Arecibo, *J. Geophys. Res.*, *83*, 2552–2560, 1978.

Muldrew, D. B., Duct model explanation of the plasma line overshoot observed at Arecibo, *J. Geophys. Res.*, *93*, 7598–7604, 1988.

Scales, W. A., P. A. Bernhardt, and G. Ganguli, Simulations of

negative ion clouds and electron density depletions produced during electron attachment chemical release experiments, in *Physics of Space Plasmas, SPI Conf. Proc. and Reprint Ser.*, vol. 12, edited by T. Chang, pp. 597–608, Scientific, Cambridge, Mass., 1993.

Scales, W. A., P. A. Bernhardt, G. Ganguli, C. L. Siefring, and P. Rodriguez, Small-scale plasma irregularities produced during electron attachment chemical releases, *Geophys. Res. Lett.*, *21*, 605–608, 1994.

Scales, W. A., P. A. Bernhardt, and G. Ganguli, Early time evolution of a chemically produced electron depletion, *J. Geophys. Res.*, *100*, 269–280, 1995.

Showen, R. L., and D. M. Kim, Time variations of HF-induced plasma waves, *J. Geophys. Res.*, *83*, 623–628, 1978.

Thidé, B., A. Hedberg, J. A. Fejer, and M. P. Sulzer, First observations of stimulated electromagnetic emission at Arecibo, *Geophys. Res. Lett.*, *16*, 369–372, 1989.

M. M. Baumback, P. A. Bernhardt, D. G. Haas, P. Rodriguez, H. A. Romero, and C. L. Siefring, and D. A. Solin, Plasma Physics Division, Naval Research Laboratory, Code 6794, 4555 Overlook Road, SW, Washington, DC 20375. (e-mail: bern@ppd. nrl.navy.mil).

F. T. Djuth, Geospace Research, Inc., 550 N. Continental Boulevard, Suite 110, El Segundo, CA 90245.

L. M. Duncan, School of Applied Science and Engineering, University of Tulsa, Tulsa, OK 74104.

J. A. Goldstein and L. S. Wagner, Information Technology Division, Naval Research Laboratory, Code 5554, 4555 Overlook Road, SW, Washington, DC 20375.

D. E. Hunton, Phillips Laboratory, GPID, Hanscom Air Force Base, MA 01731.

C. J. Pollock, NASA Marshall Space Flight Center, ES53 Space Science Laboratory, Huntsville, AL 35812.

M. P. Sulzer and C. A. Tepley, Arecibo Observatory, P.O. Box 995, Arecibo, PR 00613.

(Received September 16, 1993; revised May 20, 1994; accepted June 23, 1994.)

Appendix W

Relaxation of the Stressed Plasma Sheet Boundary Layer

GEOPHYSICAL RESEARCH LETTERS, VOL. 21, NO. 8, PAGES 645-648, APRIL 15, 1994

Relaxation of the stressed plasma sheet boundary layer

H. Romero

Science Applications International Corporation

G. Ganguli

Beam Physics Branch, Plasma Physics Division, Naval Research Laboratory

Abstract. Observations indicate that the plasma sheet boundary layer (PSBL) becomes stressed during active times resulting in enhanced plasma flows, intense wave activity, and its minimum spatial width becoming on the order of the ion Larmor radius. We present simulations showing that the Electron-Ion-Hybrid (EIH) instability plays a key role during of the nonlinear relaxation of such a stressed PSBL. The EIH is a long wavelength mode ($k_y L_E \sim 1$, where k_y is the perpendicular wavelength and L_E the gradient scale size), its spectrum is broadband (extending from below the lower hybrid frequency to the electron plasma frequency), and its onset can generate substantial perpendicular ion acceleration and anomalous viscosity leading to particle mixing across the PSBL.

Introduction

The plasma sheet boundary layer (PSBL) has long been recognized as one of the primary transport regions of the magnetosphere [*Eastman, et al.*, 1984]. The PSBL is a low beta plasma which sustains magnetic field-aligned currents, high speed ion flows, large cross-field electric fields, and a significant component of ionospheric plasma [*Frank et al.*, 1981; *Eastman, et al.*, 1984; *Parks et al.*, 1984; *Cattell et al.*, 1986]. Using the two ISEE spacecraft, the typical spatial extent of the PSBL (during quiet times) has been determined to lie in the range of 10,000 to 30,000 km, or 100 to 300 local ion Larmor radii, ρ_i [*Parks et al.*, 1984]. More recently, a low-energy particle layer has been detected on the outer edge of the PSBL whose typical width is on the order of 500 km [*Parks et al.*, 1992]. During active times, however, the PSBL becomes stressed and its minimum width can become smaller than 1 ρ_i [*Parks et al.*, 1984]. Commonly occurring phenomena during which the PSBL can attain this width are substorms in their late growth phase during which the entire extent of the central plasma sheet, including the PSBL, becomes as narrow as 1 ρ_i [*Mitchell et al.*, 1990].

Most previous theoretical studies of the PSBL have explored the role played by infinite, uniform, field-aligned beams on the generation of broadband electrostatic noise (BEN). Here, we concentrate on a different aspect of PSBL dynamics, i.e., on how its transverse structure evolves during active times. Our principal conclusion is that as the PSBL's scale size approaches ρ_i, a class of instabilities can be excited by the strong velocity shear that is self-consistently generated during periods of high magnetospheric activity. These modes can: (i) provide dissipation, (ii) give rise to strong anomalous transport, and (iii) help to relax the stress build-up. This study complements efforts in global magnetohydrodynamic (MHD) modeling which have been successful in elucidating the mechanisms responsible for the increase of stress in the central plasma sheet but which fail to address the dominant physics governing the magnetotail's relaxation insofar as these processes are dominated by microturbulence which involve spatial (on the order of ρ_i or less) and temporal scale sizes not accessible to the MHD formalism.

Paper number 93GL03385
0094-8534/94/93GL-03385$03.00

Equilibrium Model

In a previous study [*Romero et al.*, 1990], a boundary layer equilibrium was constructed by assuming: (i) that the variation in pressure at the layer [*Parks et al.*, 1992] is caused by a density gradient, and (ii) that the magnetic field across this structure is uniform. We have recently presented an equivalent boundary layer equilibrium formalism that is simpler to implement in numerical simulations [*Romero et al.*, 1992b; *Romero and Ganguli*, 1993]. Here, we generalize the equilibrium model described in *Romero and Ganguli*, 1993, to apply to general PSBL conditions, i.e., we include self-consistently the spatial variation of the equilibrium magnetic field in this structure. This is motivated by observations [*Frank et al.*, 1981; *Parks et al.*, 1984] showing that the lobe magnetic field, typically 40 nT, is reduced in the central plasma sheet to about 20 nT.

Since the present work deals with physical processes occurring at a boundary layer (i.e., in a region where plasma parameters vary significantly), the following convention is adopted: any reference to values of the plasma frequency, cyclotron frequency, or Larmor radius will be made with respect to the region of higher plasma density.

A description of the equilibrium configuration and simulation setup has been given, for the case of a uniform magnetic field, in *Romero and Ganguli*, 1993. Here, we adopt GSM coordinates and present the modifications required to incorporate effects due to a nonuniform magnetic field, $\mathbf{B}(z)$. At the boundary layer: (i) the ion density, $n_i(z)$, varies strongly in the z (northward) direction, and (ii) the dominant component of $\mathbf{B}(z)$ is assumed to lie in the x (earthward) direction, i.e., $\mathbf{B}(z) = B_0 b(z) \mathbf{a}_x$, where B_0 denotes the strength of the magnetic field, $b(z)$ is a dimensionless function that describes its spatial variation, and $\mathbf{a}_x$ is a unit vector along the x axis. There develops an electron east-west $\mathbf{E} \times \mathbf{B}$ flow, $V_E(z)$, in the region of strong variation of the ion density. The shear frequency of the system can be naturally defined as follows: $\omega_s = V_0 / L_E$, where V_0 gives the peak value of $V_E(z)$, and L_E is the scale size of variation of the flow (here, L_E

also equals the density gradient scale size). Since during active times the PSBL's minimum L_E is on the order of ρ_i, we find that $V_0 \sim (0.1 - 0.2)V_e$, where V_e denotes the electron thermal velocity [*Romero et al.*, 1990]. Then, $\omega_s \gg \Omega_i(z)$ which implies that the unmagnetized limit can be taken in obtaining the ion equilibrium distribution function. The PSBL's temperature gradient (ignored here) will further increase the pressure gradient leading to larger values of ω_s thus implying that our present estimates of velocity shear in the PSBL are conservative. In conformity with observations [*Parks et al.*, 1984], we assume that no mean ion flow takes place across the PSBL-lobe interface so that ion force balance is achieved via a north-south electric field $\mathbf{E}(z) = E(z)\mathbf{a}_z$, where $\mathbf{a}_z$ is a unit vector directed along the z axis. We choose the spatial variation of $\mathbf{E}(z)$ as follows:

$$E(z) = \frac{E_0}{N_0}\{F(\xi - z_0) - F(\xi + z_0)\} \ , \quad (1)$$

where $\xi = z - L_z/2$, $F(\zeta) = \mathrm{Sech}^2(\zeta/L_E)$, $N_0 = 1 - \mathrm{Sech}^2(2z_0/L_E)$, L_z gives the simulation's length in the z direction, and z_0 is the distance separating the two regions in which the externally imposed electric field is either positive or negative. We choose $z_0 \gg L_E$ so that $N_0 \approx 1$. Then, the peak value of $E(z)$ is $\pm E_0$, and occurs at $L_z/2 \pm z_0$, respectively. The ion temperature is T_i and the ion density is given by $n_i(z) = N_i \exp((C_i/N_0)D(\xi))$. Here, $C_i = (2E_0L_E)/(B_0\rho_iV_i)$, V_i is the ion thermal velocity, ρ_i is the ion Larmor radius, N_i is a constant, $D(\xi) = G(\xi - z_0) - G(\xi + z_0)$, and $G(\zeta) = \mathrm{Tanh}(\zeta/L_E)$.

Since the ions are unmagnetized, the electric field given in Eq.(1) causes an $\mathbf{E} \times \mathbf{B}$ electron flow in the y (east-west) direction. The equilibrium distribution function deviates from being a simple drifting Maxwellian [*Ganguli et al.*, 1988a; *Romero and Ganguli*, 1993]. Expansion in terms of the small parameter ρ_e/L_E yields:

$$F_e = \frac{n_e(Z_g)}{\sqrt{\eta(z)}(\pi V_e)^{3/2}} \exp\left\{-\frac{w_\perp^2 + v_z^2}{V_e^2}\right\} ,$$

where Z_g is the electron guiding center (a constant of the motion), $\Omega_e(z)$ is the spatially dependent electron cyclotron frequency, $w_\perp^2 = v_x^2 + w_y^2/\eta(z)$, $w_y = v_y - V_E(z)$, $V_E(z) = E(z)/(B_0 b(z))$ is the sheared electron $\mathbf{E} \times \mathbf{B}$ flow, V_e is the electron thermal velocity, and $\eta(z) = 1 + V_E'/\Omega_e(z)$. We use a prime to denote taking the derivative with respect to z. It has been shown that $w_\perp^2$ represents the particle kinetic energy perpendicular to the ambient magnetic field and is conserved to better than second order in ρ_e/L_E [*Romero and Ganguli*, 1993]. Note that the presence of shear in the electron flow causes the distribution function to be anisotropic within the plane perpendicular to the equilibrium magnetic field. Finally, to initialize the system, the electron density is determined from the quasineutrality condition: $n_e(z) = n_i(z)$, i.e., it is assumed that the electric field given in Eq. (1) is maintained by an externally-imposed time-independent charge density.

The net electron cross-field flow can be found by taking the first moment of F_e. The result is: $U_y = V_E(z)(1 - Z_iT_e/T_i)$, where Z_i is the ion charge and T_e is the electron temperature. Since the ions have been assumed to be stationary and the electric field to be as shown in Eq. (1), the use of Ampere's law yields the following self-consistent expression for the spatial variation of the equilibrium magnetic field:

$$\frac{1}{2}\frac{db^2(z)}{dz} = \frac{\omega_{pe}^2(z)}{c^2\Omega_{e0}}\left\{\frac{E(z)}{B_0}\right\}\left(1 + \frac{Z_iT_e}{T_i}\right) \ , \quad (2)$$

where $\omega_{pe}(z)$ is the electron plasma frequency, and $\Omega_{e0} = -eB_0/m_e$.

Note that $\mathbf{E}(z)$ and $\mathbf{B}(z)$ have been chosen so that the equilibrium ion density is symmetric about $L_z/2$, the mid point of the simulation box. This implies that the present simulation contains two back-to-back boundary layers, each centered about the points $L_z/2 \pm z_0$, respectively. So long as z_0 is chosen large enough (its exact length is to be determined a posteriori, but in practice tens of electron Larmor radii or Debye lengths will suffice), no correlations arise between the relevant physics at the two layers thus justifying the use of periodic boundary conditions in the present simulation. Their use implies that the pattern in the region $0 < z < L_z$ is repeated indefinitely, while the lack of correlation between the layers implies that the results obtained for any one of the two layers is also applicable to its companion (mirror image).

Nonuniform Magnetic Field Effects on the EIH Mode

We consider in this section the linear stability properties of the EIH mode in the presence of nonuniform electric and magnetic fields. To this end, we employ the formalism presented in *Romero et al.*, 1992a and consider the linear stability of the boundary layer described in the previous section. The equilibrium ion density, magnetic field, and electron cross-field flow profiles are characterized as follows: (i) $n_i(z) = N_{i0} - (\Delta N_i/2)(1 + \mathrm{Tanh}(z/L_E))$, (ii) $b(z) = 1 + (\Delta b/2)(1 + \mathrm{Tanh}(z/L_E))$, and (iii) $V_E(z) = V_0\,\mathrm{Sech}^2(z/L_E)$. Figure 1 shows the dependence of the growth rate of the EIH mode, γ/ω_{LH}, on the normalized perpendicular wave number k_yL_E. Two cases are considered: One with $\Delta b = 0$ and the other with $\Delta b = 0.5$ (the magnetic field variation in this latter case is often encountered in PSBL-lobe crossings [*Frank et al.*, 1981; *Parks et al.*, 1984]). The parameters used to obtain these results

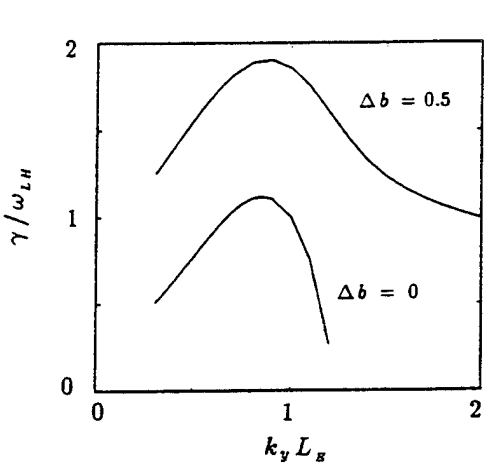


Fig. 1 Linear growth rates of the EIH mode corresponding to a magnetic field variation across the boundary layer of 0% ($\Delta b = 0$) and 50% ($\Delta b = 0.5$).

are: $\delta = \omega_{pe}/\Omega_e = 1.0$, $\alpha_e = V_0/(\Omega_e L_E) = 0.25$, $\Delta N_i/N_{i0} = 0.8$, and $M_i/m_e = 400$ (M_i and m_e denote the ion and electron masses, respectively). All of the quoted values are attained in the limit $z \to -\infty$.

The results presented in Fig. 1 imply that the inclusion of the variation of the equilibrium magnetic field for conditions representative of the PSBL-lobe interface results in: (i) an increase in the maximum value of the growth rate of the EIH mode, and (ii) no appreciable modification in the wavelength of the most unstable EIH mode.

Nonlinear Simulations

The nonlinear evolution of the system is investigated using an electrostatic, 2-1/2 D particle-in-cell (PIC) code first described in *Romero et al.*, 1992b. The physical parameters used here are: $L_z/\rho_i = L_y/\rho_i = 8$, $\rho_i/L_E = z_0/L_E = 6.4$, and $\rho_e/\lambda_D = \omega_{pe}/\Omega_{e0} = 1.2$. At $t = 0$, the system is in equilibrium, its length in the y direction is L_y, ω_{pe} is the electron plasma frequency in the region of higher plasma density, and $\lambda_D = V_e/\omega_{pe}$ is the Debye length. The ion to electron mass ratio is 400, $T_i = 4T_e$, and the total number of particles in the simulation is 786 432. Since $\mathbf{B}(z)$ is directed along the x axis, the waves reported in this study are flute-like, i.e., $k_\parallel = 0$. We choose E_0 such that $\omega_s = E_0/(B_0 L_E) = 10.8\omega_{LH}$ (i.e., $\kappa = \omega_s/\omega_{LH} = 10.8$). Finally, the electric field is decomposed into two components: the first is doubly periodic in the two spatial dimensions, and the second is time-independent and given by Eq.(1). The time-independent field component represents the stress imposed on the system by the solar wind flow, while the time-dependent doubly periodic field component models the stress relief caused by the onset of the various microinstabilities excited in the system. This latter field component is obtained by solving Poisson's equation at each time step in a mesh containing 64 nodes in the z direction and 128 nodes in the y direction.

The nonlinear evolution of the EIH mode has been described in considerable detail for the case of a uniform magnetic field [*Romero et al.*, 1992b; *Romero and Ganguli*, 1993]. In this regime, the EIH mode causes: (i) the onset of long wavelength ($k_y L_E \sim 1$, $L_E \gg \rho_e$) closed contours in the electrostatic potential (indicative of the formation of vortices in the electron flow) whenever $\kappa = \omega_s/\omega_{LH} > 1$, (ii) considerable ion acceleration perpendicular to both the equilibrium electric and magnetic fields, (iii) a broadband spectrum extending in frequency from below the lower hybrid frequency to near the electron plasma frequency, (iv) the development of significant anomalous viscosity which decreases the magnitude of the cross-field electron flow and eventually limits the further growth of the instability, and (v) a smooth transition, as κ is varied, from a regime dominated by the EIH mode ($\kappa > 1$) to one dominated by the lower hybrid drift instability ($\kappa < 1$). These general features are preserved in character by including in the simulation the self-consistent variation of the equilibrium magnetic field (as described by Eq. (2) of this work) across the boundary layer. In what follows, we describe the nonlinear evolution of the transverse structure of the PSBL under stressed conditions and use parameters corresponding to PSBL plasmas: $\omega_{pe}/\Omega_e > 1$, a magnetic field variation (between the PSBL and lobe) of the order of 50%, and $T_i/T_e = 4$ (this number can readily be further increased without affecting the principal physical results).

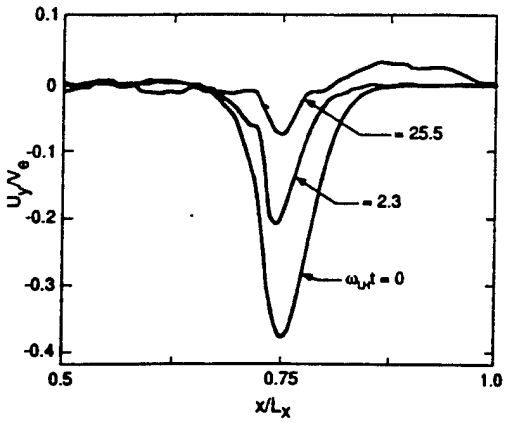


Fig. 2 Time evolution of the sheared electron $\mathbf{E} \times \mathbf{B}$ flow showing that the EIH mode leads to substantial anomalous viscosity.

First, we find that substantial anomalous viscosity develops due to the excitation of the EIH waves. This is shown in Fig. 2 where the $\mathbf{E} \times \mathbf{B}$ flow is shown at the three indicated times in the vicinity of the boundary layer. It is seen that the magnitude of the flow decays to less than half its original value in a time interval shorter than 5 lower hybrid periods. In addition, the flow spreads in the region of low magnetic field to a distance of nearly 2 ρ_i in a time interval of about 20 lower hybrid periods. This decreases the value of ω_s thereby relaxing the stress build-up in the system. Details as to how one can quantify the EIH-induced anomalous viscosity can be found in *Romero and Ganguli*, 1993. We also note that the final flow profile shown in Fig. 2 extends between 1 and 2 ρ_i and that its magnitude is larger than the ion thermal velocity. This implies that the state of the system is now susceptible to the Inhomogeneous Energy Density Driven Instability (IEDDI) [*Ganguli, et al.*, 1988a]. Further depletion of shear due to IEDDI waves may seed the system for the Kelvin-Helmholtz ($\omega < \Omega_i$) instability unless a

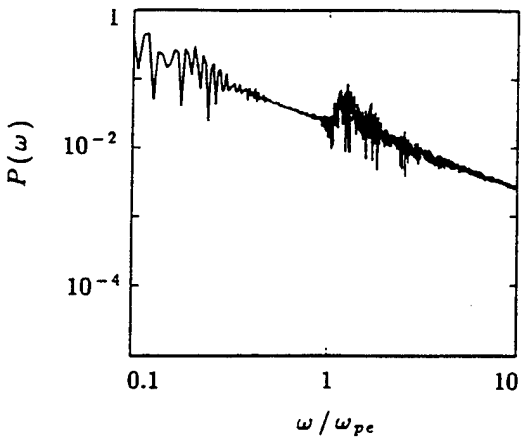


Fig. 3 Fourier transform of the time sequence corresponding to the dominant mode's energy density. It is seen that the frequency spectrum of the EIH mode is broadband, extending from below the lower hybrid frequency up to the electron plasma frequency.

balance is reached between the external stress input and the local dissipation processes.

It is also found that neither a spatially varying magnetic field nor the condition $\omega_{pe}/\Omega_e > 1$ can change the broadband character of both the z and y components of the electric field of the EIH mode (as first reported in [*Romero et al.*, 1992b]). This is illustrated in Figure 3, which shows the Fourier transform of the time sequence corresponding to the dominant mode's energy density. It is seen that this frequency spectrum is broadband, extending from below the lower hybrid frequency, through the electron plasma frequency and even higher. We note that the peak in the power spectrum occurs in the vicinity of the lower hybrid frequency and that a relative power enhancement occurs near the plasma frequency.

Another effect found in this study is the occurrence of a slight decrease in the wavelength of the EIH mode as the value of ω_{pe}/Ω_e is increased. This effect is in conformity with linear theory predictions [*Ganguli, et al.*, 1988b] which also indicate that the wavelength of the fastest growing mode ceases to become smaller as $\omega_{pe}/\Omega_e \gg 1$. We find that for the present simulation the maximum fluctuating wave energy occurs at $k_y L_E \sim 0.8$, whereas for the simulations presented in *Romero et al.*, 1992b (corresponding to $\omega_{pe}/\Omega_e = 0.52$), it occurs at $k_y L_E \sim 0.4$. Nonetheless, the present results indicate that the mode retains its long wavelength character: $k_y L_E \sim 0.8 < 1$.

Discussion and Conclusions

We have shown the EIH mode to be broadband in the absence of any cold plasma component and that the inclusion of the self-consistent variation of a PSBL-like magnetic field does not affect this result. This implies that the broadband character of the waves observed in the PSBL may be ascribed to the the local nonuniformities of this system in conjunction with effects due to field-aligned beams. As described in the present work (see also *Romero and Ganguli* 1993), significant anomalous viscosity and resistivity develop due to the EIH mode. The anomalous viscosity leads to a rapid broadening of the steepened layer and leaves the system in a state such that an inverse frequency cascade is possible where the IEDDI [*Ganguli, et al.*, 1988a] and Kelvin-Helmholtz modes are successively excited. The excitation of the EIH mode leads to substantial particle acceleration and mixing across the boundary layer[*Romero et al.*, 1992b]. This implies that unique distribution functions should be observed within this structure. Since this distribution of particles is accelerated at lower altitudes where it triggers auroral activity, the mechanism described here is of interest to ionospheric studies as well. The EIH-induced anomalous resistivity may be of interest during fast reconnection processes in magnetic substorms. Finally, the results of this study point to the following conclusion: that the width of the PSBL results from a balance between the stress caused by the solar wind flow and relaxation processes caused by microinstabilities such as the one described in this work. The observed width of this layer is then clearly related to the level of magnetospheric activity.

Acknowledgements. This work was supported by ONR.

References

Cattell, C. A., et al., ISEE observations of the plasma sheet boundary, plasma sheet, and neutral sheet. 1. Electric field, magnetic field, and ion composition, *J. Geophys. Res.*, 91, 5663, 1986.

Eastman, T. E., et al., The plasma sheet boundary layer, *J. Geophys. Res., 89*, 1553, 1984.

Frank, L. A., et al., Field-Aligned currents in the earth's magnetotail, *J. Geophys. Res., 86*, 687, 1981.

Ganguli, G., Y. C. Lee, and P. J. Palmadesso, Kinetic theory for electrostatic waves due to transverse velocity shears, *Phys. Fluids, 31*, 823, 1988a.

Ganguli, G., Y. C. Lee, and P. J. Palmadesso, Electron-Ion hybrid mode due to transverse velocity shear, *Phys. Fluids, 31*, 2753, 1988b.

Mitchell, D. G., et al., Current carriers in the near-earth cross-tail current sheet during substorm growth phase, *Geophys. Res. Lett., 17*, 583, 1990.

Parks, G. K., et al., Particle and field characteristics of the high-latitude plasma sheet boundary layer, *J. Geophys. Res.*, 89, 8885, 1984.

G. K. Parks, et al., Low-Energy particle layer outside of the plasma sheet boundary, *J. Geophys. Res.*, 97, 2943, 1992.

Romero, H., et al., Equilibrium structure of the plasma sheet boundary layer-lobe interface, *Geophys. Res. Lett., 17*, 2313, 1990.

Romero, H., et al., Electron-Ion hybrid instabilities driven by velocity shear in a magnetized plasma, *Phys. Fluids B, 4*, 1708, 1992a.

Romero, H., G. Ganguli, and Y. C. Lee, Ion acceleration and coherent structures generated by lower hybrid shear-driven instabilities, *Phys. Rev. Lett., 69*, 3503, 1992b.

Romero, H., and G. Ganguli, Nonlinear evolution of a strongly sheared cross-field plasma flow, *Phys. Fluids B, 9*, 3163, 1993.

G. Ganguli, Beam Physics Branch, Plasma Physics Division, NRL, Washington, D. C., 20375.

H. Romero, Applied Optics Branch, Optical Sciences Division, NRL, Washington, D. C., 20375.

(Received July 15, 1993; revised: October 21, 1993; accepted: November 16,1993.)

Appendix X

Radiation Generation by Photo-Switched, Periodically-Biased Semiconductors

Radiation Generation by Photo–Switched, Periodically–Biased Semiconductors

E. Esarey, P. Sprangle, B. Hafizi,[†] and P. Serafim[‡]

Beam Physics Branch, Plasma Physics Division

Naval Research Laboratory, Washington DC 20375-5346

Abstract

A laser pulse, propagating nearly parallel to the surface of a planar semiconductor wafer, will generate electron-hole pairs. If the semiconductor is spatially biased with a static electric field of period λ_0, the laser pulse acts as a fast switch and generates a periodic current. The rapid switching of the current generates radiation, which propagates along the surface and can be confined by a conducting wall placed parallel to the wafer. The wavelength of the radiation can be tuned by adjusting λ_0, the wafer-wall separation, and/or the carrier density. In the absence of collisional damping, N_0 periods of the static bias electric field will generate N_0 periods of radiation. Under idealized conditions, the maximum electric field of the radiation is equal to the applied static field and the maximum efficiency of converting the static electric field energy to electromagnetic energy is 30%. In practice for typical parameters, tunable electromagnetic radiation can be generated with wavelengths in the 50 μm to 500 μm range, pulse durations in the pico- or subpicosecond range, and peak powers on the order of 100 W.

[†]Icarus Research, P.O. Box 30580, Bethesda, MD 20824-0780

[‡]Northeastern University, Department of Electrical Engineering, Boston, MA 02115

I. Introduction

Over the past few years, several methods for the generation of pico- and subpicosecond far-infrared radiation by optical switching of biased semiconductors have been investigated [1-10]. Most of these investigations have involved a planar photoconductor (typically, GaAs or InP) which is illuminated by ultrashort laser pulses. When the semiconductor is biased with a uniform static electric field, the photo-switching of the semiconductor produces a current transient which results in the generation of a single cycle of coherent radiation. For example, D. You et al. [7] report the generation of high power ($\sim$ MW) half-cycle pulses with durations < 0.5 ps and energies < 0.8 μJ. This was achieved by illuminating a 3.5 cm $\times$ 3.5 cm GaAs wafer, biased with an external field of 11 kV/cm, with a 120 fs pulse from a Ti:sapphire laser at normal incidence.

Other photoconductor-based schemes have been investigated for generating many cycles of far-infrared radiation. For example, N.M. Froberg et al. [5] used a photoconducting antenna array to produce terahertz radiation pulses. This array was formed by depositing 64 electrodes, of width 20 μm and spaced 150 μm apart, on a GaAs substrate. The electrodes were then used to bias the photoconductor with a periodic electric field of wavelength λ_0. When illuminated with laser pulses incident on the substrate at an angle of ϕ from the normal, radiation was produced over a wide range of angles and the wavelength of the radiation at an angle θ from normal was found to be $\lambda \simeq (\sin\theta + \sin\phi)\lambda_0$. The frequency of the radiation could be adjusted by changing the angles ϕ and θ, and/or the biasing period λ_0. This method was used to produce low power ($\sim$ nW) 500 GHz radiation in 20 ps pulses.

Recently, an alternative concept for generating far-infrared radiation has been proposed by W.B. Mori et al. [11], which relies on the conversion of a periodic static electric field into radiation by a relativistic ionization front. In this device, a parallel plate capacitor array is used to generate a periodic electric field within a gas. An intense, short laser pulse is injected in the gas, propagating parallel to the capacitor plates. As the laser pulse propagates, it ionizes the gas and produces free plasma electrons (photocarriers). A current transient is generated along the relativistic ionization front, resulting in the gener-

ation of electromagnetic radiation which propagates in either the direction of or opposite to the laser pulse. In the one-dimensional (1-D) limit and in the limit where the ionization front is moving at the speed of light in vacuum c, radiation is generated with frequency ω given by $\omega/c = (k_0^2+k_p^2)/2k_0$, where $k_0 = 2\pi/\lambda_0$, λ_0 is the period of the static electric field, $k_p = \omega_p/c$, $\omega_p = (4\pi n_e e^2/m_e)^{1/2}$ is the electron plasma frequency and n_e is the density of the plasma electrons. Furthermore, the number of cycles of output radiation is roughly equal to the number of periods of the static electric field. Hence, the frequency of the radiation can be tuned by adjusting the static field wavelength, λ_0, and/or the electron plasma density, n_e. Furthermore, the output pulse duration can be controlled by adjusting the number of periods of the static electric field. In principle, high peak powers (MW) can be obtained over a wide wavelength range ($10-100$ μm) with varying pulse durations (1-10 ps).

In the following, we propose and analyze a device which combines some of the features of photo-switched semiconductors [1-10] and of radiation conversion using a periodic capacitor array [11]. This device will be referred to as a Photo-switched Periodically-biased Semiconductor (PPS). Radiation is generated in the PPS by applying a static, periodic electric field across the surface of a planar semiconductor (e.g., GaAs or InP). A laser pulse is injected into the device such that it propagates along and nearly parallel to the semiconductor surface and perpendicular to the applied static field. As the pulse propagates it generates photocarriers within the semiconductor and, hence, a transient current develops in response to the applied periodic electric field. These current transients generate electromagnetic radiation. A planar conductor, placed parallel to the photoconductor, can be used to confine and enhance the generated radiation (see Fig. 1).

In the one-dimensional (1-D) limit, it will be shown that radiation is generated with frequency ω given by $\omega/c = (k_0^2+\bar{k}_p^2)/2k_0$, where $k_0 = 2\pi/\lambda_0$, λ_0 is the period of the static electric field, $\bar{k}_p = \bar{\omega}_p/c$, $\bar{\omega}_p = (4\pi\bar{n}_n e^2/m^*)^{1/2}$ is the effective 1-D plasma frequency, $\bar{n}_n$ is the effective 1-D photocarrier density, and m^* is the effective mass of the photocarriers. In the absence of collisions, N_0 periods of applied static electric field generate N_0 periods of radiation. Furthermore, it will be shown that the maximum electric field of the radiation generated within the device is less than or equal to the amplitude of the applied static

electric field. In the 1-D limit, the maximum efficiency of converting the energy in the static electric field into electromagnetic radiation is 30%. Such a PPS device may be capable of generating hundreds of watts of coherent radiation in the $50 - 500$ μm range, which can be tuned by adjusting the period of the static electric field, the carrier density and/or the device dimensions. The duration of the electromagnetic pulse can be controlled by adjusting the number of periods of the static electric field. In principle, one can choose how many cycles of radiation will be generated. Such a source of tunable, ultrashort pulse radiation would have various applications, including absorption spectroscopy, time-resolved studies in physics and chemistry, remote sensing and radar, and high-speed multiplexing.

One can envision many possible configurations for a PPS device. In this paper, we will analyze the following simplified and idealized two-dimensional (2-D) configuration (see Fig. 1). Consider a single planar semiconductor, the surface of which lies in the y, z plane and is located at $x = b$ along the x-axis. A static, periodic, bias electric field of the form $\mathbf{E}_B = E_0 \cos k_0 z \ \mathbf{e}_y$ is applied in the y-direction, where $k_0 = 2\pi/\lambda_0$ and λ_0 is the bias period. The device is of length L_0 is the z-direction, biased with $N_0 = L_0/\lambda_0$ periods of static electric field, and the photoconductor is assumed to be uniform and infinite (i.e., some length L_y large compared to other characteristic dimensions) in the y-direction. To confine the generated radiation, a conducting wall is positioned at a distance $x = a$ above and parallel to the semiconductor surface. A short laser pulse is injected between the conduction wall and the semiconductor, such that it propagates at the speed of light c in the positive z direction nearly parallel to the semiconductor and skims along its surface. Photocarriers are assumed to be generated within a penetration depth b (between $x = 0$ and $x = b$) inside the semiconductor. For simplicity, the device is assumed to be symmetric about $x = 0$, i.e., conducting walls are located at $x = \pm a$, the semiconductor is a thin slab of thickness $2b$ centered about $x = 0$ and laser pulses propagate along both surfaces of the semiconductor slab. In principle, one can envision using more than one semiconductor slab, placed parallel to one another, with laser pulses injected between their surfaces. The 1-D limit corresponds to many closely spaced semiconductor slabs. In the following, the effects of collisional damping within the semiconductor are included; however, other nonideal effects, such as fringe electric fields or nonuniform carrier densities, are neglected.

II. Wave Equation and Current Response

The generation of electromagnetic radiation by the photo-induced current is described by the wave equation

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{E} = \frac{4\pi}{c^2}\frac{\partial \mathbf{J}}{\partial t}, \tag{1}$$

where $\mathbf{E}$ is the electric field generated by transients in the current density $\mathbf{J}$. Prior to the injection of the laser pulse, a static electric field of the form $\mathbf{E}_B = E_0 F(z)\cos k_0 z \mathbf{e}_y$ is applied to the semiconductor sheet, where $\lambda_0 = 2\pi/k_0$ is the wavelength of the static field, E_0 is the amplitude, $F(z) = u(z) - u(z - L_0)$, $u(z)$ is the heaviside unit step function and L_0 is the axial length of the device, i.e., $F(z) = 1$ for $0 < z < L_0$ and $F(z) = 0$ otherwise. As the laser pulse propagates along the surface of the semiconductor, it generates carriers and a photo-induced current $\mathbf{J}$ within the semiconductor.

The following simplified model for the photo-induced current will be assumed. The photo-induced current is given by $\mathbf{J} = -en_n v\mathbf{e}_y$, where n_n is the effective carrier density and v is the effective carrier velocity. The carrier density is determined approximately by $\partial n_n/\partial t = q\alpha I/\hbar\omega_L$, where I is the laser intensity inside the semiconductor, α is the absorption coefficient, $\hbar\omega_L$ is the energy of the laser photons, and q is the quantum efficiency for converting photons into charge carriers, and the effects of diffusion and recombination have been neglected. The laser pulse duration τ_L is assumed to be less than the characteristic response time of the device. Hence, the carriers are assumed to be generated instantaneously by the laser pulse, $n_n = n_0 u(ct - z)$, where $n_0 = q\alpha I\tau_L/\hbar\omega_L$ is the peak density, $ct - z = 0$ gives the location of the laser pulse front, and the carriers exist in the region $ct - z \geq 0$ (behind the laser pulse). Furthermore, the carrier density is assumed to be uniform within the penetration depth $0 < |x| < b$, i.e., $n_n = n_0 u(ct - z)u(b - |x|)$.

Carriers which are produced in the presence of the electric fields $\mathbf{E}$ and $\mathbf{E}_B$ obey the equation

$$\frac{\partial}{\partial t}(n_n \mathbf{v}) = -\frac{en_n}{m^*}(\mathbf{E} + \mathbf{E}_B) - c\nu n_n \mathbf{v}, \tag{2}$$

where m^* is the effective mass of the carrier and $c\nu$ is the effective collision frequency. Hence, the current response is given by

$$\left(\frac{\partial}{\partial t} + c\nu\right)\mathbf{J} = \frac{\omega_p^2}{4\pi}(\mathbf{E} + \mathbf{E}_B), \tag{3}$$

where $\omega_p = (4\pi n_n e^2/m^*)^{1/2}$ is the effective plasma frequency. On the right-side of Eq. (3), the electric field of the laser is neglected, i.e., it is assumed that the only role of the laser field is to generate the carrier population, $\omega_p^2 = \omega_{p0}^2 u(ct-z)u(b-|x|)$, where $\omega_{p0} = (4\pi n_0 e^2/m^*)^{1/2}$. Note that the current induced by a static, uniform electric field E_0 is given by $J = J_0[1-\exp(-\zeta\nu)]$, where $J_0 = -en_n\mu E_0$ is the steady-state current, $\mu = e/(m^* c\nu)$ is the carrier mobility, $\zeta = ct - z$ and $1/c\nu$ is the relaxation time.

It is convenient to introduce the independent variables $\zeta = ct - z$, z and x. Using these variables, the 2-D wave equation is given by

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} - 2\frac{\partial^2}{\partial \zeta \partial z}\right)\mathbf{E} = \frac{4\pi}{c}\frac{\partial \mathbf{J}}{\partial \zeta} \tag{4}$$

and the induced current is given by

$$\left(\frac{\partial}{\partial \zeta} + \nu\right)\mathbf{J} = \frac{\omega_p^2}{4\pi c}(\mathbf{E} + \mathbf{E}_B), \tag{5}$$

where $\omega_p^2 = \omega_{p0}^2 u(\zeta)u(b-|x|)$. It will also prove convenient to introduce the Laplace transform $Q_s(s)$ of the quantity $Q(\zeta)$ with respect to ζ, i.e., $Q_s = \int_0^\infty d\zeta Q \exp(-s\zeta)$. Taking the Laplace transform of Eqs. (4) and (5) yields the wave equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} - 2s\frac{\partial}{\partial z} - \frac{sk_p^2}{s+\nu}\right)E_s = \frac{k_p^2(x)}{s+\nu}E_0 F(z)\cos k_0 z, \tag{6}$$

where $E(\zeta = 0) = 0$ has been assumed, $k_p^2(x) = \omega_{p0}^2/c^2$ for $|x| \le b$ and $k_p^2(x) = 0$ for $|x| > b$, i.e., the charge carriers only exist within the semiconductor of thickness $2b$. Within the region of the semiconductor, $0 < z < L_0$, the device is periodic in z and the induced electric field can be written as the real part of $E_s = \hat{E}_s \exp(ik_0 z)$. Hence, the wave equation becomes

$$\left(\frac{\partial^2}{\partial x^2} - k_0^2 - 2isk_0 - \frac{sk_p^2}{s+\nu}\right)\hat{E}_s = \frac{k_p^2(x)}{s+\nu}E_0. \tag{7}$$

III. One-Dimensional Limit

To gain understanding of the basic mechanism of radiation generation and the characteristics of the radiation (e.g., the wavelength, group velocity, pulse duration, etc.), it is insightful to solve Eq. (7) in the one-dimensional (1-D) limit. The 1-D limit is obtained by letting $\partial/\partial x \to 0$ and $k_p(x) \to \bar{k}_p$ in Eq. (7), where $\bar{k}_p$ is the effective 1-D plasma wavenumber and is constant. The 1-D limit corresponds to a PPS device consisting of many thin semiconductor wafers placed closely together. In this case, the effective 1-D plasma wavenumber is given by $\bar{k}_p^2 \simeq \delta_1 k_p^2/\Delta$, where δ_1 is the wafer thickness and Δ is the separation between wafers. In effect, the laser pulse is propagating through a uniform, continuous medium producing photoconductors characterized by a plasma wavenumber $\bar{k}_p^2$. The 1-D limit is also directly applicable to the case of a photoionized gas [11], i.e., the laser pulse propagates into a gas with a periodic static electric field and ionizes the gas to produce a plasma with a plasma frequency $c\bar{k}_p$. Since the ionizing laser is assumed to propagate with velocity c, it is implicitly assumed that $\omega_L \gg c\bar{k}_p$, where ω_L is the frequency of the ionizing laser.

In the 1-D limit, the wave equation, Eq. (7), becomes

$$\left[(2isk_0 + k_0^2)(s+\nu) + s\bar{k}_p^2\right]\hat{E}_s = -\bar{k}_p^2 E_0. \tag{8}$$

Hence, $\hat{E}_s = -\bar{k}_p^2 E_0/D(s)$, where the 1-D dispersion relation is given by

$$D(s) = (2isk_0 + k_0^2)(s+\nu) + s\bar{k}_p^2. \tag{9}$$

The inverse Laplace transform is governed by the zeros of $D(s)$. Equation (9) can be written as $D = 2ik_0(s-s_1)(s-s_2)$, where $s_{1,2}$ are the zeros of $D(s)$ given by

$$s_{1,2} = -\frac{\nu}{2} + \frac{i}{4k_0}(k_0^2 + \bar{k}_p^2) \pm \frac{1}{2}\left\{\left[\nu - \frac{i}{2k_0}(k_0^2 + \bar{k}_p^2)\right]^2 + 2ik_0\nu\right\}^{1/2}. \tag{10}$$

The inverse transform of $\hat{E}_s$ is given by

$$\hat{E}(\zeta) = \frac{i\bar{k}_p^2 E_0}{2k_0(s_1 - s_2)}\left[\exp(s_1\zeta) - \exp(s_2\zeta)\right]. \tag{11}$$

Equations (10) and (11) can be analyzed in various limits.

In the absence of collisions, $\nu = 0$, the zeros are $s_1 = i(k_0^2 + \bar{k}_p^2)/2k_0$ and $s_2 = 0$. Hence, $E = \hat{E}\exp(ik_0 z)$ is given by

$$E = \frac{\bar{k}_p^2 E_0}{(\bar{k}_p^2 + k_0^2)} \left\{ \exp\left[\frac{i}{2k_0}(\bar{k}_p^2 + k_0^2)\zeta + ik_0 z \right] - \exp(ik_0 z) \right\}. \tag{12}$$

The first term on the right of Eq. (12) represents the induced electromagnetic wave and the second term represents an induced static electric field. The electric field of the electromagnetic wave can be written in terms of z and t as

$$E = \frac{\bar{k}_p^2 E_0}{(\bar{k}_p^2 + k_0^2)} \exp\left[i(\omega t - k_z z) \right], \tag{13}$$

where the axial wavenumber and frequency are given by

$$k_z = (\bar{k}_p^2 - k_0^2)/2k_0, \tag{14}$$

$$\omega/c = (\bar{k}_p^2 + k_0^2)/2k_0, \tag{15}$$

respectively. Notice that $\omega^2/c^2 = k_z^2 + \bar{k}_p^2$, which is the dispersion relation for electromagnetic radiation propagating in a plasma with a plasma frequency $c\bar{k}_p$. Since the frequency of the radiation remains constant as the radiation exits the device, the wavelength of the generated radiation is $\lambda = 2\pi c/\omega$. Also notice that the electric field amplitude of the radiation is maximum when $\bar{k}_p^2 \gg k_0^2$ and equal to the bias electric field amplitude E_0.

Within the device, the axial group velocity v_g and axial phase velocity v_p of the radiation are related by $v_p/c = \omega/ck_z = c/v_g$. Hence, the axial group velocity is given by

$$v_g/c = (\bar{k}_p^2 - k_0^2)/(\bar{k}_p^2 + k_0^2). \tag{16}$$

In the limit $\bar{k}_p^2 \ll k_0^2$, the radiation wavelength is $\lambda = 2\pi c/\omega \simeq 4\pi/k_0 = 2\lambda_0$ and the group velocity is $v_g \simeq -c$, i.e., the wavelength is twice that of the static electric field and is moving in the negative z-direction (the direction opposite of the incident laser pulse) at approximately the speed of light c. In the limit $\bar{k}_p^2 \gg k_0^2$, the radiation wavelength is $\lambda \simeq 2\lambda_0 k_0^2/\bar{k}_p^2 \ll 2\lambda_0$ and the group velocity is $v_g \simeq c(1 - 2k_0^2/\bar{k}_p^2)$, i.e., the wave is moving in the positive z-direction (the direction of the incident laser pulse) at approximately the

speed of light c. Notice that when $\bar{k}_p = k_0$, $k_z \to 0$ and $v_g \to 0$, i.e., the wave has a only time dependence with frequency $\omega = ck_0$ and no energy leaves the device.

The length L of the electromagnetic pulse generated by the device can be determined as follows. The incident laser pulse encounters the periodic static electric field at $z = 0$ and $t = 0$, at which point the radiation begins to be generated. The time it takes the laser pulse to reach the end of the periodic structure ($z = L_0$) is given by $T = L_0/c$. During this time the "front" of the radiation pulse, which was generated at $z = 0$, has traveled a distance $v_g T = L_0 v_g/c$ and is now located at $z = L_0 v_g/c$. Since the "back" of the radiation pulse at time $t = T$ is located at $z = L_0$, the total length of the radiation pulse is given by $L = (1 - v_g/c)L_0$. Hence,

$$L = \frac{2k_0^2 L_0}{(\bar{k}_p^2 + k_0^2)}, \tag{17}$$

or $L = ck_0 L_0/\omega$. Since the wavelength of the radiation existing the device is $\lambda = 2\pi c/\omega$, $L = \lambda L_0/\lambda_0$. In terms of the number of periods of the radiation pulse, $N = L/\lambda$, and the number of periods of the static electric field, $N_0 = L_0/\lambda_0$, $N = N_0$. In other words, a device consisting of N_0 periods of a static, sinusoidal electric field of wavelength λ_0 will generate N_0 periods of radiation of wavelength $\lambda = 2\pi c/\omega$, where ω is given by Eq. (14). Within the device, the pulse is traveling with a group velocity given by Eq. (16). For example, in the limit $\bar{k}_p^2 \ll k_0^2$, $v_g \simeq -c$ (the radiation is traveling opposite to the incident laser pulse), $\lambda \simeq 2\lambda_0$ and $L \simeq 2L_0$. In the limit $\bar{k}_p^2 \gg k_0^2$, $v_g \simeq c(1 - 2k_0^2/\bar{k}_p^2)$ (the radiation is traveling in the direction of the incident laser pulse), $\lambda \simeq 2\lambda_0 k_0^2/\bar{k}_p^2$ and $L \simeq 2L_0 k_0^2/\bar{k}_p^2$.

The average intensity, or power flux, of the radiation within the device is given by $I = \langle|\mathbf{S}|\rangle$, where $\mathbf{S} = c(\mathbf{E} \times \mathbf{B})/4\pi$ is the Poynting flux and the angular brackets signify a time average. In the 1-D limit, $\mathbf{S} = (v_g E^2/4\pi)\mathbf{e}_z$. Hence, $I = |v_g||E|^2/8\pi$, or

$$I = \frac{c}{8\pi} \frac{\bar{k}_p^4 |\bar{k}_p^2 - k_0^2|}{(\bar{k}_p^2 + k_0^2)^3} E_0^2. \tag{18}$$

The power exiting the device is $P = IL_x L_y$, where L_x and L_y are the x and y device dimensions. Notice that the intensity (power) is maximum when $\bar{k}_p^2 \gg k_0^2$ and is given by $I \simeq cE_0^2/8\pi$. The total energy W in the radiation pulse is found by multiplying P by the pulse length L/c, Eq. (17), i.e., $W = IL_x L_y L/c$. In terms of the total energy in the

static electric field initially stored within the device, $W_0 = (E_0^2/16\pi)L_xL_yL_0$, the total pulse energy is given by

$$W = \frac{4k_0^2\bar{k}_p^4|\bar{k}_p^2 - k_0^2|}{(\bar{k}_p^2 + k_0^2)^4}W_0. \tag{19}$$

The energy conversion efficiency, $\eta = W/W_0$, is maximum when $\bar{k}_p^2 = (5+\sqrt{17})k_0^2/2$, which implies $\lambda \simeq (0.6)^2\lambda_0$ and $W/W_0 \simeq 31\%$.

The above results have assumed $\nu = 0$. In the limit of weak collisional damping, $\nu \ll (\bar{k}_p^2 + k_0^2)/2k_0$, the two roots are given by

$$s_1 = \frac{i}{2k_0}(\bar{k}_p^2 + k_0^2) - \frac{\nu\bar{k}_p^2}{(\bar{k}_p^2 + k_0^2)}, \tag{20}$$

and $s_2 = -\nu k_0^2/(\bar{k}_p^2 + k_0^2)$. The above results, Eqs. (12)-(19), still apply, only now the mode is damped in ζ, i.e., the electric field, Eq. (13), is multiplied by the damping factor $\exp(-\zeta/L_d)$, where the damping distance L_d is given by $L_d^{-1} = \nu k_0^2/(\bar{k}_p^2 + k_0^2)$. In effect, this limits the length of the radiation pulse, $L \lesssim L_d$.

In the limit of strong collisional damping, $\nu \gg (\bar{k}_p^2 + k_0^2)/2k_0$, the two roots are given by $s_1 \simeq ik_0/2 - \bar{k}_p^2/4\nu$ and $s_2 \simeq -\nu$. The electric field of the radiation is given by

$$E \simeq \frac{i\bar{k}_p^2 E_0}{2\nu k_0} \exp\left[\left(i\frac{k_0}{2} - \frac{\bar{k}_p^2}{4\nu}\right)\zeta + ik_0 z\right]. \tag{21}$$

The frequency, axial wavenumber, and group velocity of the mode are $\omega/c = k_0/2$, $k_z = -k_0/2$, and $v_g \simeq -c$. Notice that in the strong damping limit, the amplitude of the electric field is reduced by the factor $[(\bar{k}_p^2 + k_0^2)/(2k_0\nu)]\exp(-\zeta/L_d)$ compared to Eq. (13), where the damping distance is $L_d = 4\nu\bar{k}_p^2$. Hence the intensity I and the total pulse energy W are reduced by $(\bar{k}_p^2 + k_0^2)^2/(2k_0\nu)^2$, and the length of the radiation pulse is limited to $L \lesssim L_d$.

IV. Two-Dimensional Solutions

The generation of radiation in 2-D can be described by solving Eq. (7) inside and outside the semiconductor. Within the semiconductor, $|x| < b$, the Laplace transform of the electric field is given by $\hat{E}_s = \hat{E}_{s1} = F + A_1 \cos k_1 x$, and outside, $b < |x| < a$, $\hat{E}_s = \hat{E}_{s2} = A_2 \sin k_2(a-x)$, where

$$k_1^2 = -2ik_0 s - k_0^2 - sk_p^2/(s+\nu), \tag{22}$$

$$k_2^2 = -2ik_0 s - k_0^2, \tag{23}$$

$F = (k_p/k_1)^2 E_0/(s+\nu)$, and $A_{1,2}$ are constants (independent of x). The functional forms of $E_{s1,2}$ have assumed the boundary conditions $\partial E_s/\partial x = 0$ at $x = 0$ and $E_s = 0$ at $x = a$, i.e., the symmetric solution. The coefficients $A_{1,2}$ can be found by requiring E_s and $\partial E_s/\partial x$ to be continuous at $x = b$. This gives

$$\hat{E}_{s1} = F\left[1 + D^{-1} k_2 \cos k_2(a-b) \cos k_1 x\right], \tag{24}$$

$$\hat{E}_{s2} = FD^{-1} k_1 \sin k_1 b \sin k_2(a-x), \tag{25}$$

where the dispersion relation $D(s)$ is given by

$$D(s) = k_1 \sin k_2(a-b) \sin k_1 b - k_2 \cos k_2(a-b) \cos k_1 b. \tag{26}$$

The inverse Laplace transform of Eqs. (24) and (25) is determined largely by the zeros of $D(s)$, i.e., $s = s_n$ where $D(s_n) = 0$. Near a zero, $D \simeq (s - s_n)\partial D/\partial s$. Using standard theory of residues and assuming simple poles, the asymptotic behavior of the inverse transform of the electric field is given by

$$E_1 = \sum_n \frac{k_p^2 b E_0 \sin k_2(a-b)}{(s_n+\nu)\partial D/\partial s} \cos k_1 x \exp(s_n \zeta + ik_0 z), \tag{27}$$

$$E_2 = \sum_n \frac{k_p^2 b E_0}{(s_n+\nu)\partial D/\partial s} \sin k_2(a-x) \exp(s_n \zeta + ik_0 z), \tag{28}$$

where the right-side is evaluated at $s = s_n$ and the sum is over all zeros s_n. For the modes of interest, s_n can be written as $s_n = ik_n - \gamma_n$, where γ_n is the damping rate, $\omega_n = ck_n$ is

the frequency, $k_z = k_n - k_0$ is the axial wavenumber, and $v_g = c^2 k_z/\omega_n$ is the axial group velocity.

In the limit $b \ll a$, the total power is dominated by the field E_2 in the region $b < |x| < a$. The intensity of the radiation $I = c\langle|\mathbf{E}\times\mathbf{B}|\rangle/4\pi$ is given by $I \simeq c^2|v_g||E_2|^2/8\pi$ and the power is $P \simeq L_y \int dx I$. Hence, the power in the n^{th} mode can by written as

$$P_n \simeq \frac{caL_y|k_n - k_0|}{8\pi k_n}\left|\frac{bk_p^2 E_0}{(s_n+\nu)\partial D/\partial s}\right|^2 \exp(-2\gamma_n\zeta). \tag{29}$$

In the absence of collisions ($\nu = 0$, $\gamma_n = 0$), the pulse length of the radiation is given by $L_n \simeq (1 - v_g/c)L_0$, where $v_g = c^2 k_z/\omega_n = c(k_n - k_0)/k_n$ is the axial group velocity. Hence, $L_n \simeq k_0 L_0/k_n$, i.e., a device consisting of $N_0 = L_0/\lambda_0$ periods of bias field will produce a radiation pulse consisting of $N = N_0$ periods of radiation, where $N = L_n/\lambda$, $\lambda = 2\pi c/\omega_n$ and $\omega_n = ck_n$. The total energy in the n^{th} mode radiation pulse is given by $W_n \simeq P_n L_n/c \simeq (k_0/k_n)P_n L_0/c$, where P_n is given by Eq. (29) with $\gamma_n = 0$.

To further evaluate the inverse Laplace transforms, Eqs. (27)-(28), it is necessary to find the zeros of $D(s)$. Insight can be gained by noting that in either the limit $k_p \to 0$ or the limit $b \to 0$, $D = -k_2 \cos k_2 a$. In these limits, $D = 0$ implies $k_2 = (2\ell+1)\pi/2a$, where ℓ is an integer. Hence, the zeros of $D(s)$ are given by $s = s_n$, where $s_n = i(k_0^2 + k_x^2)/2k_0$, $k_x = n\pi/2a$, and $n = 2\ell + 1$ is odd. This implies that radiation is generated in discrete modes, where the frequency of n^{th} mode is $\omega_n = -is_n$.

A more relevant limit can be analyzed in the limit of a thin current sheet $b/a \ll 1$ by assuming $|k_{1,2}b| \ll 1$. Then,

$$D \simeq -\hat{k}_p^2 b \sin k_2 a - k_2 \cos k_2 a, \tag{30}$$

where $\hat{k}_p^2 = sk_p^2/(s+\nu)$. Notice that $|\hat{k}_p^2 b^2| \ll 1$ is implied by $|k_{1,2}b| \ll 1$ and that $\hat{k}_p = k_p$ in the absence of collisions ($\nu = 0$). The zeros of $D(k_2) = 0$ can be approximated by letting $k_2 = k_x + \delta k_x$, and assuming $|\delta k_x a| \ll 1$. Analytical expressions for the quantities k_x and δk_x can be found in various limits. Specifically, when (A) $|\hat{k}_p^2 ab| \ll 1$ and (B) $|\hat{k}_p^2 ab| \gg 1$.

Consider the limit (A) $|\hat{k}_p^2 ab| \ll 1$. Notice that this along with $|\hat{k}_p^2 b^2| \ll 1$ imply $|\hat{k}_p^2 b^2| \ll b/a \ll 1$. In this limit, the zeros of $D(k_2)$, Eq. (30), are given by $k_2 = k_{x1} + \delta k_{x1}$,

where $k_{x1} = n\pi/2a$, $n = 2\ell + 1$ is odd, and $\delta k_{x1} \simeq b\hat{k}_p^2/k_{x1}a$. Using the definition of k_2, the zeros of $D(s)$ are given by $s = s_n$, where

$$2s_n = -\nu + ik_t \pm \left[(\nu - ik_t)^2 + \frac{2i}{k_0}\nu(k_0^2 + k_{x1}^2)\right]^{1/2}, \tag{31}$$

and $k_t = (k_0^2 + k_{x1}^2 + 2bk_p^2/a)/2k_0$. Notice that this reduces to the 1-D limit when $k_{x1} = 0$ and $2bk_p^2/a = \bar{k}_p^2$. Furthermore, $\partial D/\partial s \simeq i(-1)^\ell ak_0$. Simplified expressions for the zeros, Eq. (31), can be found in the limits of weak, $\nu \ll k_t$, or strong, $\nu \gg k_t$, collisions. The zeros can be written as $s_n = ik_n - \gamma_n$, where the mode frequency $\omega_n = ck_n$ and damping rate γ_n are given by

$$k_n = \begin{cases} (k_0^2 + k_{x1}^2 + 2bk_p^2/a)/2k_0, & \text{for } \nu \ll k_t, \\ (k_0^2 + k_{x1}^2)/2k_0, & \text{for } \nu \gg k_t, \end{cases} \tag{32}$$

and

$$\gamma_n = \begin{cases} \nu bk_p^2/ak_0k_n, & \text{for } \nu \ll k_t, \\ bk_nk_p^2/a\nu k_0, & \text{for } \nu \gg k_t. \end{cases} \tag{33}$$

The power in the n^{th} mode can be written, in the limit $|\hat{k}_p^2 ab| \ll 1$, as

$$P_n \simeq \frac{caL_y|k_n - k_0|}{8\pi k_n(k_n^2 + \nu^2)} \left(\frac{bk_p^2E_0}{k_0a}\right)^2 \exp(-2\gamma_n\zeta), \tag{34}$$

where k_n and γ_n are given by Eqs. (32)-(33) in the appropriate limits. Equation (34) indicates that the power in the various modes increases with increasing k_p. In fact, when $\nu = 0$, the expression for the power becomes independent of k_p for sufficiently large k_p. However, Eq. (34) assumed that $|\hat{k}_p^2 ab| \ll 1$, i.e., it is not valid in the large k_p limit.

Consider the limit (B) $|\hat{k}_p^2 ab| \gg 1$. Notice that this along with $|\hat{k}_p^2 b^2| \ll 1$ imply $b/a \ll |\hat{k}_p^2 b^2| \ll 1$. In this limit, the zeros of $D(k_2)$, Eq. (30), are given by $k_2 = k_{x2} + \delta k_{x2}$, where $k_{x2} = \ell\pi/a$ (ℓ is an integer) and $\delta k_{x2} \simeq -k_{x2}/\hat{k}_p^2 ab$. Hence, $s_n = ik_n - \gamma_n$, where the frequency $\omega_n = ck_n$ and damping rate γ_n are given by

$$k_n = (k_0^2 + k_{x2}^2)/2k_0, \tag{35}$$

$$\gamma_n = \nu k_{x2}^2/abk_0k_nk_p^2. \tag{36}$$

Furthermore, $\partial D/\partial s \simeq (-1)^{\ell} iabk_0\hat{k}_p^2/k_{x_2}$. The power in the n^{th} mode can be written as

$$P_n \simeq \frac{caL_y k_{x2}^4 |k_{x2}^2 - k_0^2| E_0^2}{2\pi^3 \ell^2 (k_{x2}^2 + k_0^2)^3} \exp(-2\gamma_n \zeta). \tag{37}$$

Since the axial wavenumber of the mode is $k_z = k_n - k_0$, the axial group velocity is $v_g/c = (k_{x2}^2 - k_0^2)/(k_{x2}^2 + k_0^2)$. The length of the radiation pulse is approximately $L_n \simeq (1 - v_g/c)L_0$ which gives $L_n \simeq 2k_0^2 L_0/(k_{x2}^2 + k_0^2)$, assuming $L_n \lesssim 1/2\gamma_n$. Neglecting collisions, $\nu \to 0$, the total energy in the n^{th} mode is $W_n \simeq P_n L_n/c$, which gives

$$W_n \simeq \frac{8k_0^2 k_{x2}^4 |k_{x2}^2 - k_0^2|}{\pi^2 \ell^2 (k_{x2}^2 + k_0^2)^4} W_0, \tag{38}$$

where W_0 is defined as $W_0 = aL_y L_0 E_0^2/8\pi$, i.e., the energy density of the electrostatic field multiplied by the device volume. As with the 1-D limit , W_n/W_0 is maximum when $k_{x2}^2 = (5 + \sqrt{17})k_0^2/2$, i.e., when $a/\ell \simeq 0.16\lambda_0$. This corresponds to a wavelength $\lambda_1 = 2\pi c/\omega_1 \simeq 0.36\lambda_0$ and a maximum efficiency of $W_n/W_0 \simeq (6.3/\ell^2)\%$.

V. Numerical Examples

The analytic expressions presented in Sec. IV for the mode frequency $\omega_n = ck_n$, the damping rate γ_n, and the power P_n were based on solving the simplified dispersion relation, Eq. (30), in the limit $|\hat{k}_p b|^2 \ll 1$. To verify the analytic theory, the zeros of the full dispersion relation, Eq. (26), were found numerically as a function of $k_p^2 \propto n_0$, where n_0 is the density of charge carriers. The numerical solutions assumed a current layer thickness of $b = 1$ μm, a conducting wall located at $a = 100$ μm, a bias field period of $\lambda_0 = 400$ μm, and a collision frequency $c\nu$ given by $\nu^{-1} = 100$ μm. The effective plasma frequency $\omega_p = ck_p$ was varied over the range $1\ \text{cm}^{-1} < k_p < 10^5\ \text{cm}^{-1}$.

The lowest order zero to Eq. (26), $s = s_1$, is plotted in Fig. 2, where $s_1 = ik_1 - \gamma_1$. The mode wavenumber k_1 is given by the imaginary part of s_1, which is plotted in Fig. 2(a). Notice that in the limit $k_p \to 0$, Eq. (32) predicts that $k_1 \simeq (k_0^2 + k_{x1}^2)/2k_0 \simeq 157$ cm^{-1}, which is in agreement with Fig. 2(a). In the limit $abk_p^2 \gg 1$ ($k_p \gg 10^3\ \text{cm}^{-1}$), Eq. (35) predicts that $k_1 \simeq (k_0^2 + k_{x2}^2)/2k_0 \simeq 393\ \text{cm}^{-1}$, again in agreement with Fig. 2(a). The damping rate is given by the real part of s_1, which is plotted in Fig. 2(b). A maximum value of $\gamma_1 = 24\ \text{cm}^{-1}$ ($\gamma_1^{-1} = 420$ μm) is obtained at $k_p \simeq 1.5 \times 10^3\ \text{cm}^{-1}$. Away from this value, i.e., $k_p \ll 10^3\ \text{cm}^{-1}$ or $k_p \gg 10^3\ \text{cm}^{-1}$, γ_1 rapidly diminishes as predicted by Eqs. (33) and (36).

The real and imaginary parts of $\partial D/\partial s$, evaluated at $s = s_1$, are shown in Fig. 3(a) and (b), respectively. These quantities are important for determining the radiation power, as indicated by Eq. (29). It is useful to define a normalized power

$$\hat{P}_n = \left| \frac{(k_n - k_0)bk_p^2}{k_n(s_n + \nu)\partial D/\partial s} \right|^2, \tag{39}$$

where the maximum power in the n^{th} mode is given by $P_n = (caL_y E_0^2/8\pi)\hat{P}_n$. The quantity $\hat{P}_1$, evaluated at $s_n = ik_1 - \gamma_1$, is plotted in Fig. 4 as a function of k_p. Figure 4 shows that the power achieves a maximum of $\hat{P}_1 = 0.135$ at a value of $k_p \simeq 5 \times 10^3$ cm^{-1}. At $k_p \simeq 5 \times 10^3\ \text{cm}^{-1}$, where the power is maximum, the mode frequency ck_1 is given by $k_1 = 375\ \text{cm}^{-1}$ (a wavelength of $\lambda_1 = 2\pi/k_1 = 168$ μm) and the damping rate is

$\gamma_1 \simeq 6$ cm^{-1} ($\gamma_1^{-1} = 0.17$ cm). In the limit $k_p^2 ab \gg 1$ ($k_p > 10^3$ cm^{-1}), theory predicts

$$\hat{P}_1 = \frac{4k_{x2}^2 |k_{x2}^2 - k_0^2|}{a^2 (k_{x2}^2 + k_0^2)^3}, \tag{40}$$

as indicated by Eq. (37). For the values used in Fig. 4 ($a = 100$ μm, $k_{x1} = 314$ cm^{-1}, $k_0 = 157$ cm^{-1}), $\hat{P}_1 \simeq 0.156$, somewhat greater (15%) than the numerical maximum.

As an example, consider a device with a conducting wall at $a = 100$ μm, a width of $L_y = 1$ cm, and a GaAs semiconductor with a photo-induced current thickness of $b = 1$ μm. The applied static electric field has an amplitude of $E_0 = 10$ kV/cm and a period of $\lambda_0 = 400$ μm. The carrier mobility for GaAs is $\mu_n = 8500$ cm^2/V-s, which gives an effective collision frequency $c\nu = e/m^* \mu_n$ of $1/c\nu = 0.32$ ps or $1/\nu = 96$ μm, where m^* is the effective mass of the carriers ($m^*/m_e = 0.067$ for GaAs). These values are nearly identical to those used in Figs. 2-4. A value of $abk_p^2 = 25$ is assumed, i.e., $k_p \simeq 5 \times 10^3$ cm^{-1}, which is the value at which the power is maximum in Fig. 4. This corresponds to a carrier density of $n_n = 4.7 \times 10^{17}$ cm^{-3} (recall, $k_p^2 = 4\pi n_n e^2/m^* c^2$). From Figs. 2-4, $k_1 \simeq 375$ cm^{-1}, $\gamma_1 \simeq 5.8$ cm^{-1}, and $\hat{P}_1 \simeq 0.135$. This corresponds to a radiation wavelength of $\lambda_1 = 2\pi/k_1 \simeq 170$ μm, a damping length of $\gamma_1^{-1} \simeq 0.17$ cm, and peak radiation power of $P_1 \simeq 180$ W. In the absence of collisions, a device consisting of N_0 periods of static field produces N_0 periods of radiation. The radiation pulse length would be $L = N_0 \lambda$. However, for this example, collisions limit the radiation pulse length to $L \lesssim \gamma_1^{-1}$. This limits the number of radiation periods to $N \lesssim (\gamma_1 \lambda)^{-1} \simeq 10$, i.e., because of collisions, the device needs to consist of $N_0 \leq 10$ static field periods. This gives a radiation pulse duration of $L/c \lesssim 5.7$ ps. The laser power absorbed in producing a carrier population of $n_n = 4.7 \times 10^{17}$ cm^{-3} is one the order of 500 kW. Hence, a laser pulse energy on the order of a few μJ should be sufficient.

VI. Discussion

A PPS (photo-switched, periodically-biased, semiconductor) device has been proposed and analyzed as a compact source of short-pulse, tunable, far-infrared radiation. The analysis assumed a simplified geometry, shown in Fig. 1, which consists of a planar semiconductor (in the y, z-plane) biased with a periodic electric field of the form $\mathbf{E}_B = E_0 \cos k_0 z \mathbf{e}_y$. A laser pulse is injected into the device, such that it propagates in the z-direction, skimming along the surface of the wafer. The laser pulse produces photocarriers and a transient current response, leading to the generation of far-infrared radiation. A conducting wall, placed parallel to the wafer at a separation distance of $x = a$, is used to confine the radiation.

The PPS concept combines some of the features of radiation generation by optical switching of a biased semiconductor [1-10], and of radiation generation using an ionization front in a periodically-biased gas [11]. Notice that, in contrast to the present work, the work of Froberg et al. [5] used a biased field of the form $\mathbf{E}_B = E_0 \cos k_0 z \mathbf{e}_z$, no conducting wall was present, and the semiconductor was illuminated at angles near normal to the wafer. Hence, the details of the mechanism for radiation generation are different. The work of Mori et al. [11] on ionization fronts in gases assumes a periodic bias field of the form $\mathbf{E}_B = E_0 \cos k_0 z \mathbf{e}_y$, similar to the present work. The results of Sec. III, describing the 1-D PPS limit, can be applied directly to ionization of a periodically biased gas. Included for the first time are the effects of collisions.

In the 1-D limit, radiation is generated with frequency $\omega = c(\bar{k}_p^2 + k_0^2)/2k_0$, wavenumber $k_z = (\bar{k}_p^2 - k_0^2)/2k_0$, and group velocity $v_g = c(\bar{k}_p^2 - k_0^2)/(\bar{k}_p^2 + k_0^2)$. Hence, the wavelength of the radiation $\lambda = 2\pi c/\omega$ can be tuned by adjusting the bias period $\lambda_0 = 2\pi/k_0$ or the effective plasma wavenumber $\bar{k}_p$ (i.e., via the carrier density $n_n \propto \bar{k}_p^2$). In the absence of collisions, N_0 periods of of bias field generates N_0 periods of radiation. For $\bar{k}_p^2 \gg k_0^2$, the electric field of the radiation is maximum and equal to the bias field E_0. The efficiency of converting the electrostatic field energy into radiation energy, $\eta = W/W_0$, is maximum when $\bar{k}_p^2 = (5+\sqrt{17})k_0^2/2$, which implies $\lambda \simeq (0.6)^2 \lambda_0$ and $W/W_0 \simeq 31\%$. A finite collision frequency $c\nu$ damps the radiation pulse and limits the pulse length L to a length shorter than the damping length L_d, $L \lesssim L_d$, where $L_d \propto \nu^{-1}$.

The 2-D theory of a PPS device is presented in Sec. IV. Analytic solutions describing the characteristics of the radiation pulse were obtained in the limit $|\hat{k}_p^2 b^2| \ll 1$. In 2-D, radiation is generated in various modes characterized by the zeros of the dispersion relation, Eq. (30), $s_n = ik_n - \gamma_n$, where n is the mode number, $\omega_n = ck_n$ is the mode frequency, $v_g = c(k_n - k_0)/k_n$ is the axial group velocity, and $\gamma_n \simeq L_d^{-1}$ is the damping rate. In the limit $|\hat{k}_p^2 b^2| \ll b/a \ll 1$, k_n and γ_n are given by Eqs. (32) and (33), respectively, and the power in the n^{th} mode is given by Eq. (34). As k_p^2 (proportional to the carrier density n_n) increases, the mode frequency and power increases. In the limit $b/a \ll |\hat{k}_p^2 b^2| \ll 1$, k_n and γ_n are given by Eqs. (35) and (36), respectively, and the power in the n^{th} mode is given by Eq. (37). Equation (38) implies the the energy conversion efficiency, W_n/W_0, is maximum when $k_{x2}^2 = (5 + \sqrt{17})k_0^2/2$, i.e., when $a/\ell \simeq 0.16\lambda_0$. This corresponds to a wavelength $\lambda_1 = 2\pi c/\omega_1 \simeq 0.36\lambda_0$ and a maximum efficiency of $W_n/W_0 \simeq (6.3/\ell^2)\%$.

To verify the analytic solutions, the full 2-D dispersion relation, Eq. (26), was solved numerically in Sec. V. In particular, the mode frequency ck_n, damping rate γ_n, and normalized power, $\hat{P}_n$, Eq. (39), were plotted as a function of k_p (proportional to the square root of the carrier density n_n) in Figs. 2(a), 2(b), and 4, respectively. For the parameters chosen (which are relevant to GaAs), $a = 100$ μm, $b = 1$ μm, $L_y = 1$ cm, $E_0 = 10$ kV/cm, $\lambda_0 = 400$ μm, and $1/\nu \simeq 100$ μm, the normalized power for the fundamental ($n = 1$) was found to maximize at a value of $k_p \simeq 5 \times 10^3$ cm^{-1} (corresponding to a carrier density of $n_n = 4.7 \times 10^{17}$ cm^{-3}). This gave a radiation wavelength of $\lambda_1 = 170$ μm, a damping length of $\gamma_1^{-1} = 0.17$ cm, and a peak power of $P_1 = 180$ W. The length of the radiation pulse is limited by collisions $L \leq \nu_1^{-1}$, which corresponds to a pulse duration $L/c \lesssim 5.7$ ps. The length of radiation pulse can be controlled by adjusting the number of bias periods N_0 and/or the damping rate γ_n. The wavelength of the radiation can be adjusted by adjusting the bias period λ_0, the carrier density, and/or the device dimensions.

This analysis indicates that a PPS device should be capable of generating peak radiation powers in the hundreds of watts and radiation wavelengths in the $50 - 500$ μm range. The pulse durations can be ultrashort, $\lesssim 10$ ps, and, in principle, the device can provide control over the number of cycles of radiation generated. Such a source of radiation might have applications in the areas of far-infrared spectroscopy [6], the study of atomic (e.g.,

Rydberg) states [12], the characterization of materials (dielectrics, semiconductors, and superconductors) [13,14], and remote sensing.

Acknowledgments

The authors acknowledge useful conversations with D. Papadopoulos, W.B. Mori and T. Katsouleas, and the numerical assistance of J. Krall. This work supported by the Office of Naval Research.

References

1. G. Mourou, C.V. Stancampiano, and D. Blumenthal, Appl. Phys. Lett. **38**, 470 (1981).
2. D.H. Auston, K.P. Cheung, and P.R. Smith, Appl. Phys. Lett. **45**, 284 (1984).
3. B.B. Hu, J.T. Darrow, X.C. Zhang, and D.H. Auston, Appl. Phys. Lett. **56**, 886 (1990).
4. B.I. Greene, J.F. Federici, D.R. Dyakaar, R.R. Jones, and P.H. Bucksbaum, Appl. Phys. Lett. **59**, 893 (1991).
5. N.M. Froberg, B.B. Hu, X.C. Zhang, and D.H. Auston, IEEE J. Quantum Electron. **28**, 2291 (1992).
6. B.I. Greene, P.N. Saeta, D.R. Dykaar, S. Schmitt-Rink, and S.L. Chuang, IEEE J. Quantum Electron. **QE-28**, 2303 (1992).
7. D. You, R.R. Jones, P.H. Bucksbaum, and D.R. Dykaar, Opt. Lett. **18**, 290 (1993).
8. R.R. Jones, D. You, and P.H. Bucksbaum, Phys. Rev. Lett. **70**, 1236 (1993).
9. A.S. Weling, B.B. Hu, N.M. Froberg, and D.H. Auston, Appl. Phys. Lett. **64**, 137 (1994).
10. D. You, R.R. Jones, P. H. Bucksbaum, and D. R. Dykaar, J. Opt. Soc. Am. B **11**, 486 (1994).
11. W.B. Mori, T. Katsouleas, J.M. Dawson, and C.H. Lai, Phys. Rev. Lett. **74**, 542 (1995).
12. R.R. Jones, N.E. Tielking, D. You, C. Raman, and P.H. Bucksbaum, Phys. Rev. A **51**, R2687 (1995).
13. D. Grischkowsky, S. Keiding, M. van Exter, and Ch. Fattinger, J. Opt. Soc. Amer. B **7**, 2006 (1990).
14. M.C. Nuss, K.W. Goossen, J.P. Gordon, P.M. Mankiewich, and M.L. O'Malley, J. Appl. Phys. **70**, 2238 (1991).

Figure Captions

Fig. 1. Schematic of a 2-D PPS device. A semiconductor wafer of thickness $2b$ lies in the y, z-plane. A conducting wall is placed parallel to the wafer at $x = a$. A periodic bias electric field of the form $\mathbf{E}_b = E_0 \cos k_0 z \mathbf{e}_y$ is applied to the wafer, where $\lambda_o = 2\pi/k_0$ is the bias period, and z is the distance along the axis of laser propagation.

Fig. 2. Numerical solution to the 2-D dispersion relation, Eq. (26), for the parameters $b = 1$ μm, $a = 100$ μm, $\lambda_0 = 400$ μm, and $\nu^{-1} = 100$ μm. The lowest order zero $s_1 = ik_1 - \gamma_1$ is plotted as a function of k_p. (a) shows the imaginary part of s_1, i.e., the mode frequency $\omega_1 = ck_1$, and (b) shows the real part of s_1, i.e., the damping rate $\gamma_1 = L_d^{-1}$.

Fig. 3. The (a) real and (b) imaginary parts of $\partial D/\partial s$, evaluated at $s = s_1$, versus k_p, for the parameters of Fig. 2.

Fig. 4. The normalized power, $\hat{P}_1$, Eq. (39), evaluated at $s = s_1$, versus k_p, for the parameters of Fig. 2.

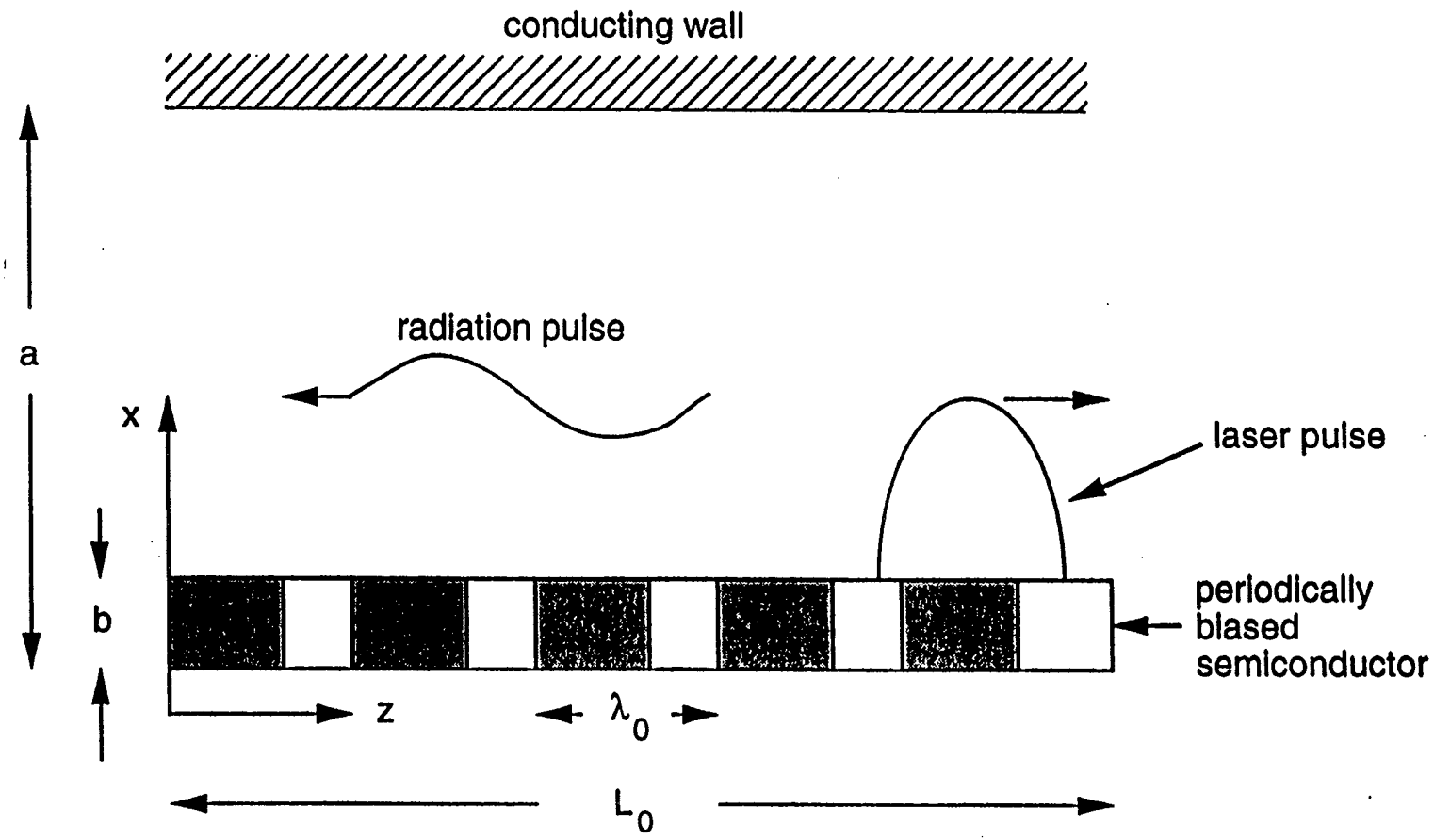

conducting wall
a
radiation pulse
x
laser pulse
b
periodically
biased
semiconductor
z
λ_0
L_0

Fig. 1

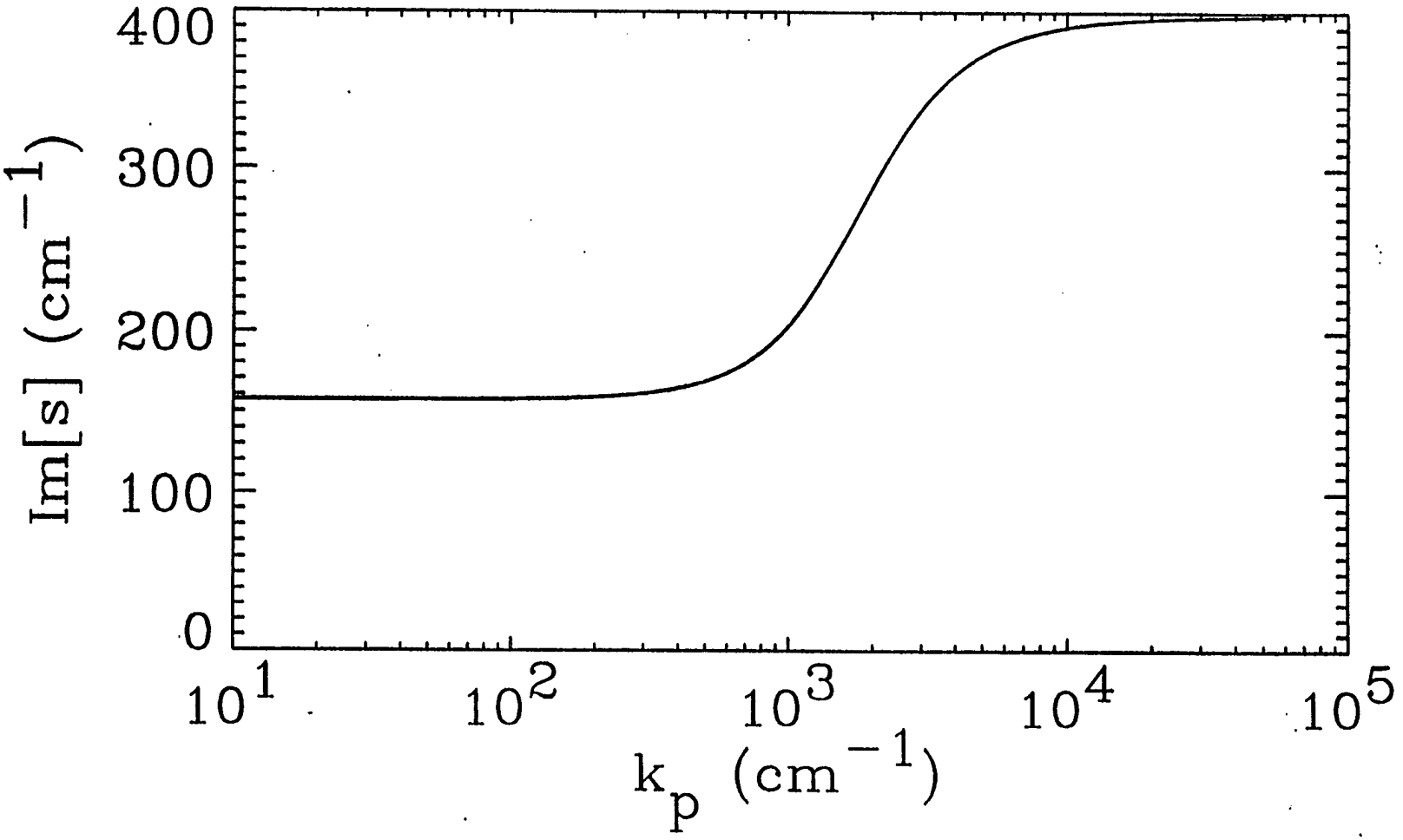


Fig. 2(a)

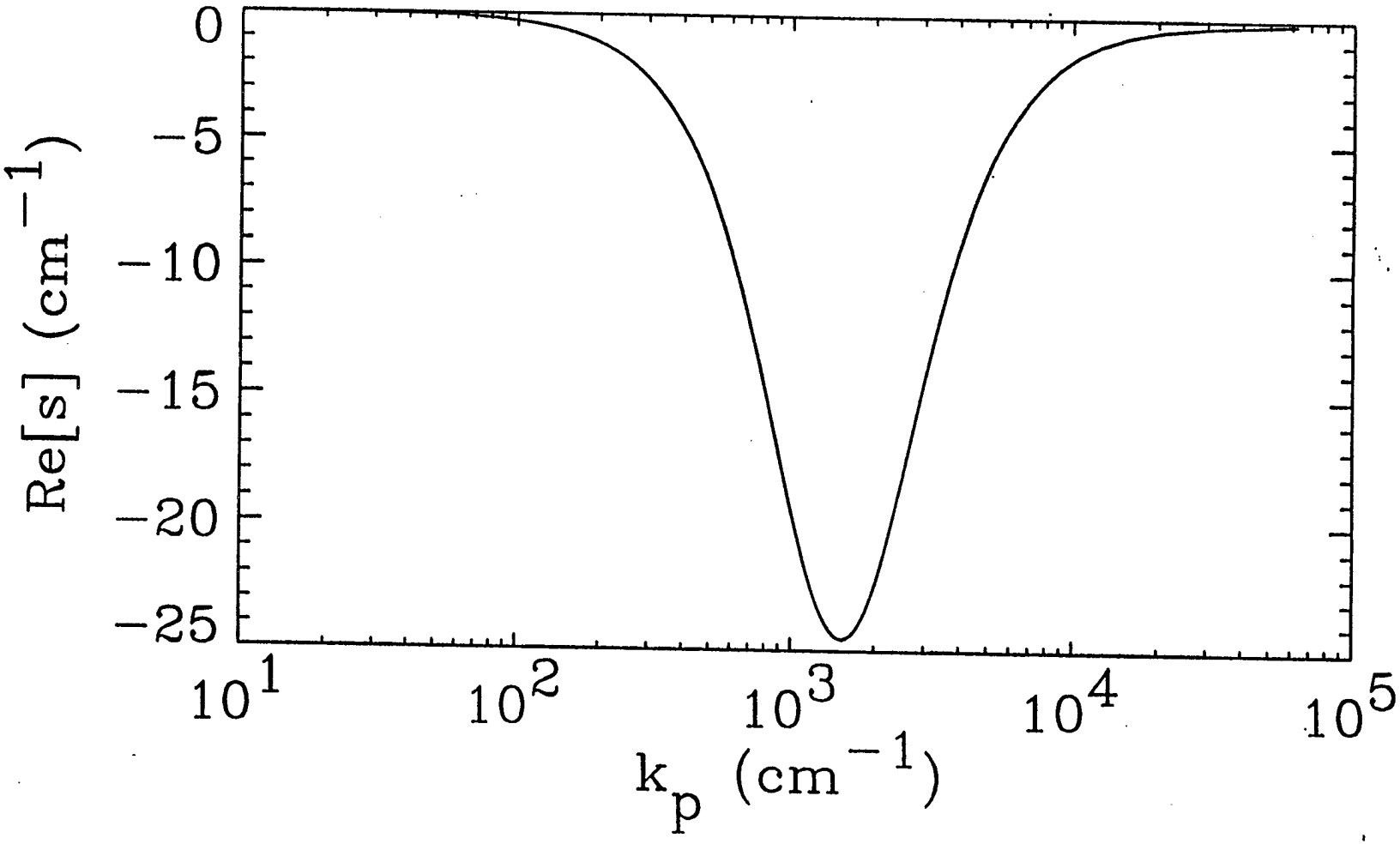


Fig. 2(b)

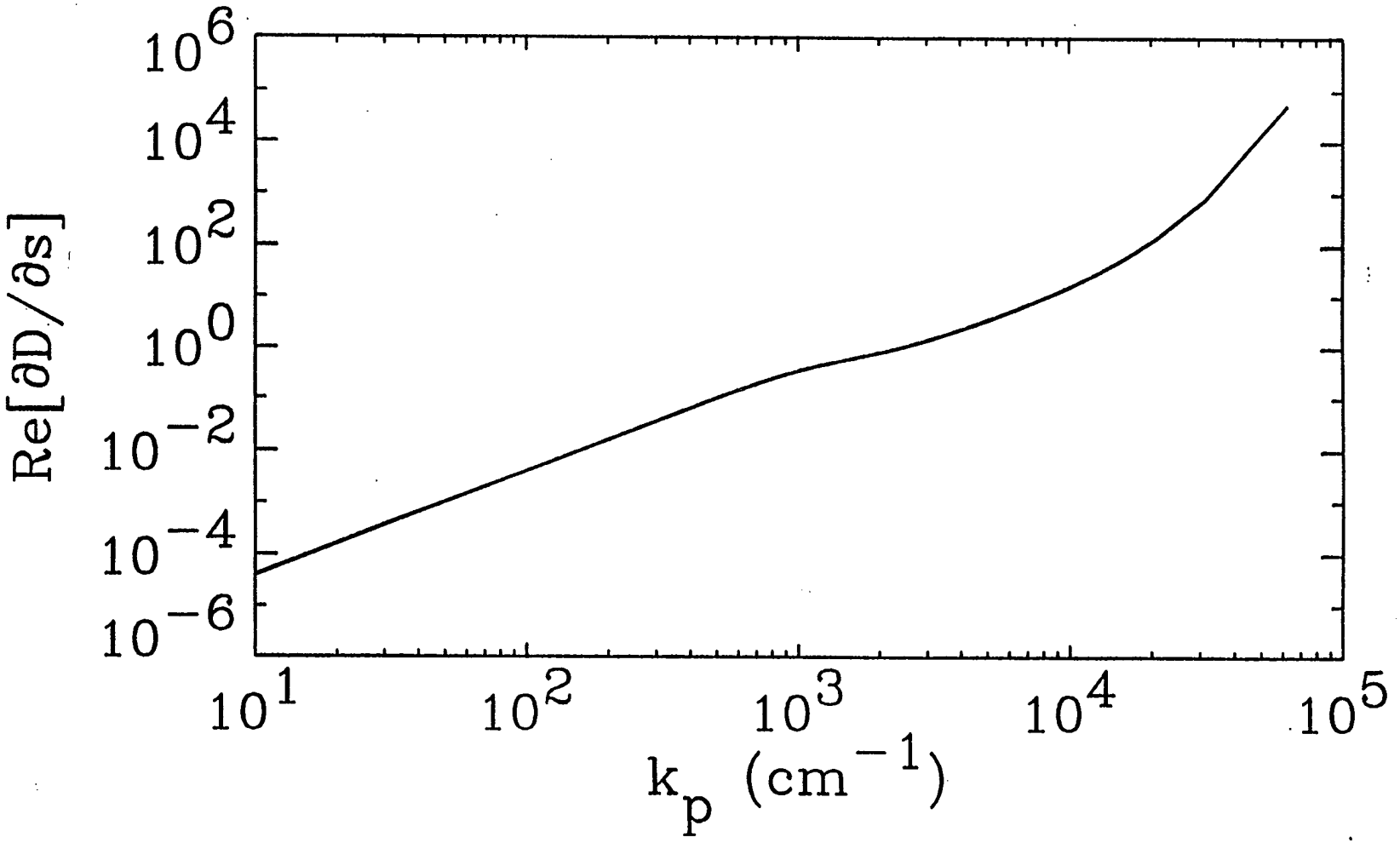


Fig. 3(a)

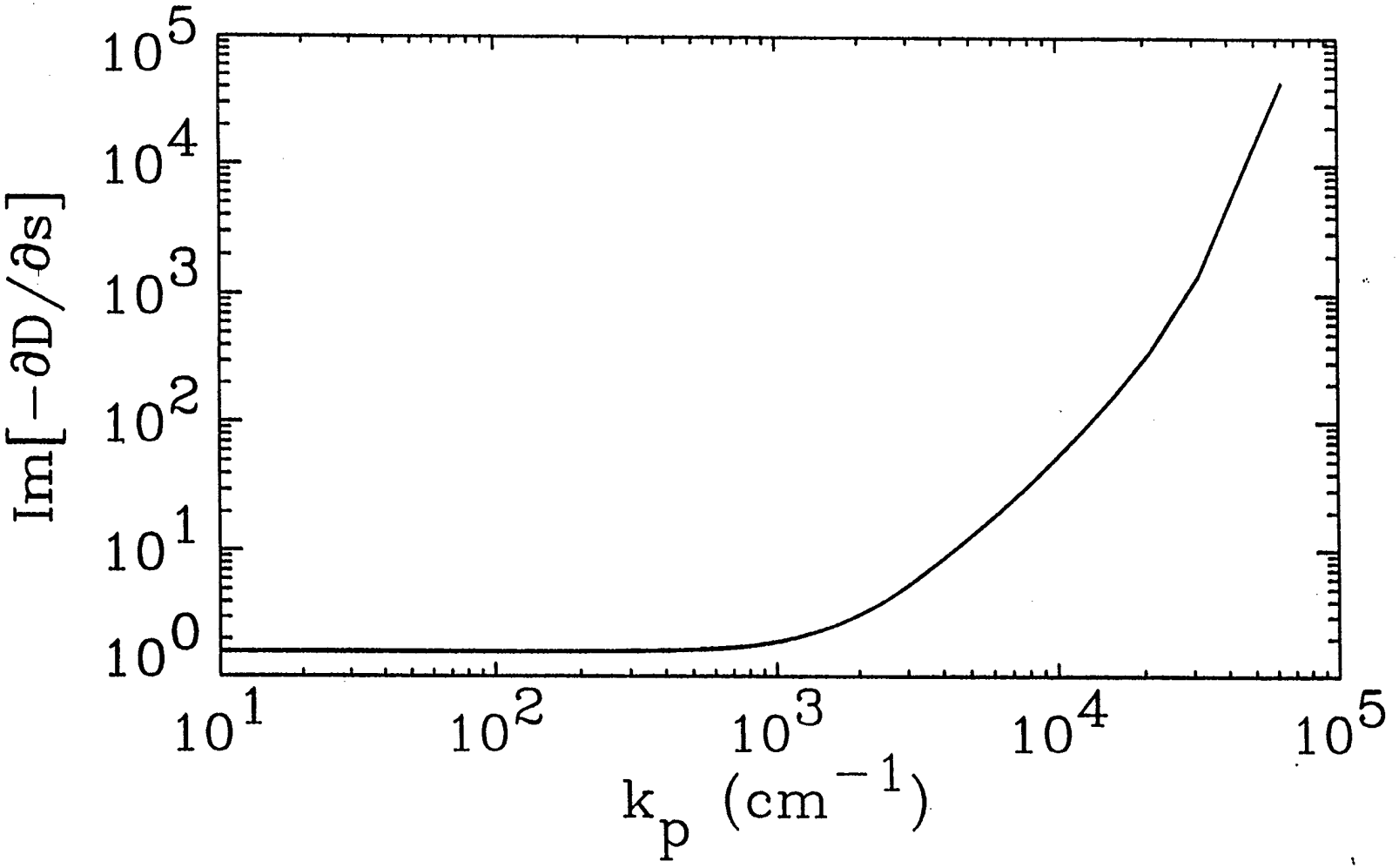

10^5
10^4
10^3
10^2
10^1
10^0
Im[$-\partial D/\partial s$]
10^1
10^2
10^3
10^4
10^5
k_p (cm^{-1})

Fig. 3(b)

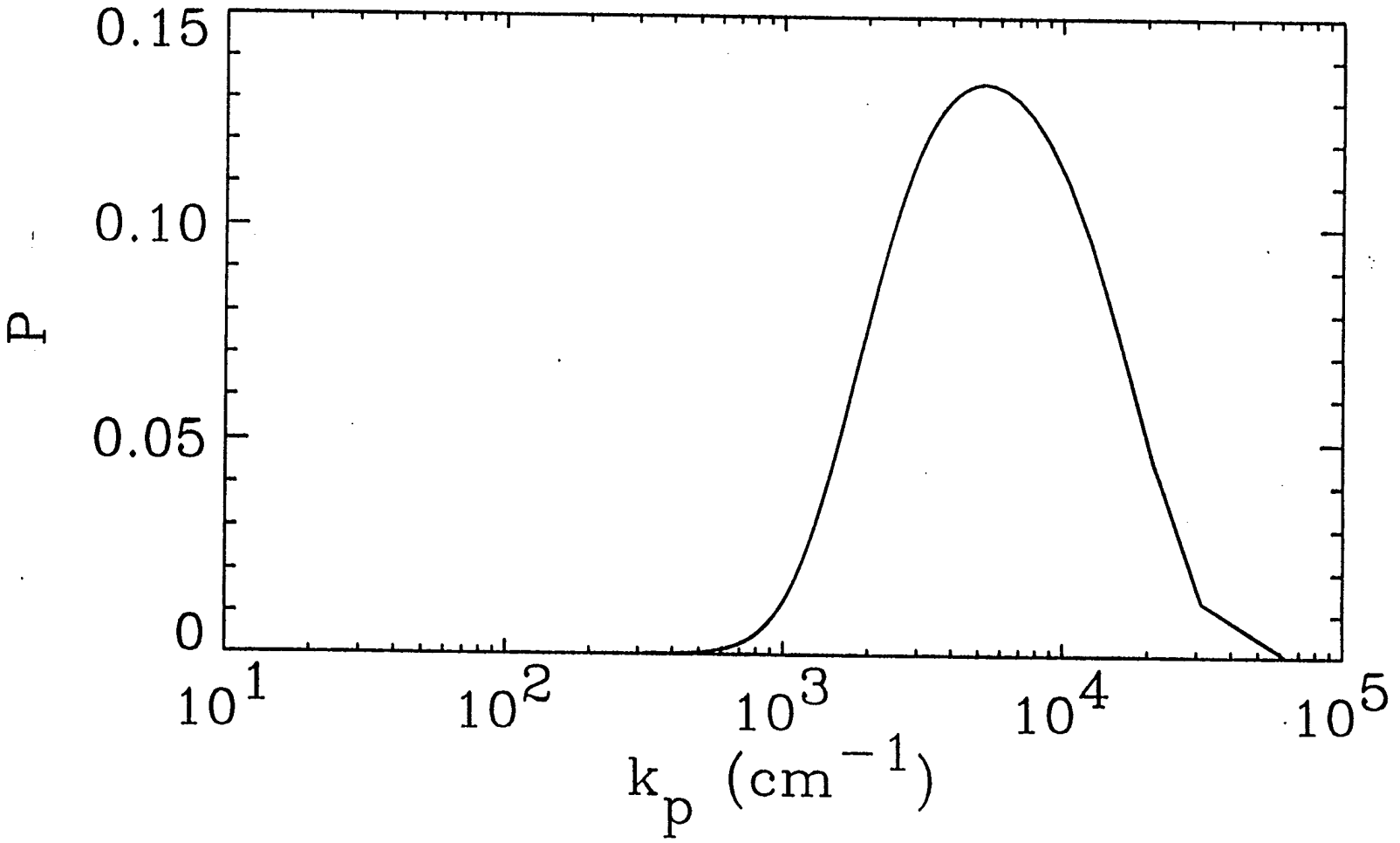


Fig. 4

Appendix Y

One-Dimensional Intense Laser Pulse Solitons in a Plasma

One-Dimensional Intense Laser Pulse Solitons in a Plasma

R.N. Sudan, Y.S. Dimant, and O.B. Shiryaev

Laboratory of Plasma Studies,

Cornell University, Ithaca, NY 14853 USA

Abstract

We develop an analytical framework for the nonlinear dispersion relation between the group velocity and laser amplitude of a class of one-dimensional isolated envelope solitons for modulated light pulse coupled to electron plasma waves, numerically ivestigated by Kozlov et al. [1] and later by Kaw et al. [2]. The principal feature of these large amplitude solitons, viz. that the group velocity has a discrete spectrum, is derived analytically.

In this letter we develop the analytical basis for the existence of a class of one dimensional, isolated soliton solutions for short intense laser pulse coupled to electron plasma waves, previously obtained numerically first by Kozlov et al. [1] and confirmed by Kaw et al. [2]. The nonlinear effects due to high laser intensity i.e., the reduction of electron density through the ponderomotive force and the relativistic increase in the electron mass [3] overcome the natural dispersion of an electromagnetic wave in a plasma. The group velocity of these solitons is related to the laser amplitude and frequency. At low intensity the well-known envelope soliton solution [4] is recovered. The principal features of these solitons are as follows:

(1) The solitons are classfied by the integer N which is the number of nodes of the envelope of the laser vector potential. An eigenvalue, $0 < s^2 < 1$ related to its nonlinear group velocity, is associated with each soliton.

(2) The $N = 0$ soliton [4] is the continuum limit of this class of solitons. In this limit

$\Delta \equiv 1 - s^2 \ll \epsilon^2 \equiv \omega_e^2/\omega^2 \ll 1$ where ω_e is the plasma frequency and ω is the laser frequency; $\Delta \approx \frac{1}{2}e\tilde{\phi}_0/mc^2 \approx \frac{1}{8}(e\tilde{A}_0/mc^2)^2$ where $\tilde{\phi}_0$ and $\tilde{A}_0$ are the maximum electrostatic and vector potential, respectively.

(3) As the laser vector potential amplitude is increased, the continuum of eigenvalues Δ corresponding to $N = 0$ is restricted by a certain value $D_0\epsilon^2$ where $D_0 = 0.012$. The first axisymmetric soliton with $N = 1$ occurs when $\Delta = D_1\epsilon^2$ where $D_1 = 0.495$. The amplitude of the soliton also becomes quantised. Unlike the $N = 0$ soliton which can have arbitrary amplitude limited only by $e\phi_0/mc^2 \ll \epsilon^2 \ll 1$, the $N = 1$ soliton exists only with a particular amplitude.

(4) In the large N limit, $\Delta \approx 0.3(\epsilon)^{8/5}(N\pi/2)^{4/5}$ and $e\tilde{\phi}_0/mc^2 \approx \epsilon^2(N\pi/2)^2$. The shift in the nonlinear group velocity from the linear group velocity is given by $\Delta/2$. The amplitude and group velocity of the solitons with $N \geq 1$ are discrete. This provides the analytical basis for the conclusion arrived at by Kozlov et al. [1] and Kaw et al. [2].

(5) Unlike the well-known $N = 0$ soliton the width $\bar{q}$ of the $N \gg 1$ solitons (in units of c/ω_p) increases with soliton amplitude i.e. $\bar{q} \propto (e\tilde{\phi}_0/mc^2)^{1/2}$.

The analysis is based on Maxwell's equations and the relativistic pressureless cold fluid equations for the electrons while assuming the ions to be completely immobile because of the short duration of the laser pulse. In previous work the nonlinear equations for laser propagation have been developed by many investigators [5]. We begin our treatment by taking the one-dimensional limit of the three dimensional equations formulated by Chen and Sudan [6], which are identical to those obtained by Bulanov et al. [7], viz.,

$$i\mathbf{A}_\tau + \mathbf{A}_{\zeta\zeta} + \mathbf{A}_{\zeta\tau} + \left(1 - \frac{1}{\Phi}\right)\mathbf{A} = 0, \tag{1}$$

$$\Phi_{\zeta\zeta} = \frac{\epsilon^2}{2}\left(\frac{1 + |\mathbf{A}|^2}{\Phi^2} - 1\right). \tag{2}$$

Here $\zeta = z - vt, \tau = t$, where v is the linear group velocity at the laser frequency $\omega/2\pi$; $\mathbf{A} = e\tilde{\mathbf{A}}/\sqrt{2}mc^2$ is the normalised transverse vector potential of the laser, $\Phi = 1 + e\tilde{\phi}/mc^2$ where $\tilde{\phi}$ is the actual electrostatic potential and $\epsilon = \omega_e/\omega \ll 1$. The time variations in the coordinate ζ frame are much slower than when the laser pulse is described in the laboratory coordinate z. The axial coordinates z and ζ are normalized to the laser vacuum wavelength $\lambda_0/2\pi$, the time t by ω. Note that the term $\mathbf{A}_{\zeta\tau}$ neglected in Eqn. (35) of Ref. [6] is included here because it is crucial in the present context and furthermore because it allows the system of equations (1) and (2) to be represented in the Hamiltonian formalism.

We now restrict our treatment to a circularly polarised wave of amplitude A and take a stationary solution of the form

$$A(\zeta, \tau) = a(\zeta) \exp[i(s-1)(\zeta - 2\tau)], \tag{3}$$

$$\Phi(\zeta) = \phi(\zeta), \tag{4}$$

where a is real and s is a real valued parameter; it will turn out that $1 - s \approx \Delta/2$ is the shift in the nonlinear group velocity from the linear value. Substituting (3) and (4) in (1) and (2) and allowing for a coordinate rescaling $q = \epsilon\zeta$, we obtain

$$\epsilon^2 a_{qq} = -p^2 a \equiv \frac{1}{2}\frac{\partial V}{\partial a}, \tag{5}$$

$$\phi_{qq} = \frac{1}{2}\left(\frac{1+a^2}{\phi^2} - 1\right) \equiv -\frac{1}{2}\frac{\partial V}{\partial \phi}, \tag{6}$$

with $p^2 \equiv s^2 - 1/\phi$ and $V \equiv \phi + 1/\phi - a^2(s^2 - 1/\phi)$. These equations agree precisely with those of Kaw et al. [2] in the limit $\epsilon^2 \ll 1$; note that $s^2 = \lambda^2/(1-\beta^2)$ in the notation of Kaw et al. [2].

Equations (5) and (6) are equivalent to those that describe the dynamics of a particle in a two-dimensional potential well $V(a, \phi)$ with the peculiarity that it has mass 2 in the ϕ-coordinate and $-2\epsilon^2$ in the a-coordinate. The energy E of such a particle is a constant of motion that is given by

$$E = \phi_q^2 - \epsilon^2 a_q^2 + V. \tag{7}$$

The kinetic energy component $-\epsilon^2 a_q^2$ is negative because of the negative mass in the a-coordinate. We seek a bounded solution of Eqns. (5) and (6) such that

$$\phi = 1\,, \qquad \phi_q = a_q = a = 0 \qquad \text{at} \qquad q = \pm\,\infty\,. \tag{8}$$

These boundary conditions determine $E = 2$.

We may solve the set (5) and (6) by taking advantage of the wide discrepancy in the equivalent masses $2, -2\epsilon^2$. It is clear that we may employ a quasi-static ϕ in (5). Once (5) is solved for a implicitly (because ϕ is still unknown) we substitute for a in (6) and then solve (6) for ϕ to complete the solution. We recognize that (5) is isomorphic to Schrödinger's equation for a particle of momentum p, energy s^2, in the potential well $1/\phi$. Let ϕ have a maximum at $q = 0$, i.e.,

$$\phi = \phi_0\,, \qquad \phi_q = 0 \qquad \text{at} \qquad q = 0\,, \tag{9}$$

and let ϕ fall off monotonically to unity at $|q| \to \infty$. We assume symmetry about $q = 0$ for ϕ, i.e.

$$\phi(-q) = \phi(q), \tag{10}$$

with the turning points determined by setting $p = 0$ at $q = \pm q_0$, i.e.

$$s^2 = 1/\phi(\pm q_0) \leq 1\,, \tag{11}$$

since $\phi(\pm q_0)$ exceeds unity.

Furthermore, because $\epsilon^2 \ll 1$ the WKB solution of (5) takes the form:

$$a(q) = \frac{\alpha}{\sqrt{p}} \sin\left(\frac{1}{\epsilon}\int_0^q p dq + \sigma\right) + 0(\epsilon^2); \qquad 0 < q < q_0\,, \tag{12}$$

where α is a real constant and $\sigma = \pi/2$ or 0 for symmetric and antisymmetric solutions respectively. Because $\phi(-q) = \phi(q)$ we only consider the range $0 < q < \infty$. Substituting (12) in (7) and setting $E = 2$ we obtain to lowest order in ϵ^2,

$$2 = \phi_q^2 + \phi + \frac{1}{\phi} - \alpha^2 \sqrt{s^2 - \frac{1}{\phi}}, \qquad \text{for} \qquad -q_0 < q < q_0 . \tag{13}$$

Employing (9) in (13), we obtain to lowest order in ϵ^2,

$$\alpha^2 = \left(\phi_0 + \frac{1}{\phi_0} - 2\right) / \sqrt{s^2 - \frac{1}{\phi_0}}, \tag{14}$$

and

$$\phi_q^2 = F(\phi) \equiv 2 - \phi - \frac{1}{\phi} + \left(\phi_0 + \frac{1}{\phi_0} - 2\right) \frac{\sqrt{s^2 - \frac{1}{\phi}}}{\sqrt{s^2 - \frac{1}{\phi_0}}}. \tag{15}$$

It can be established that the WKB solution within $| \pm q_0|$ is valid for

$$\epsilon^2 \ll \phi_0 - 1 \ll \epsilon^{-2}. \tag{16}$$

However, we notice that $F(\phi)$ defined by (15) has two roots ϕ_0 and ϕ_1 with $1 < s^2\phi_1 < s^2\phi_0$. The turning points $\pm q_0$ are just outside $\pm q_1$, where $\phi_1 = \phi(\pm q_1)$. Thus ϕ_q^2 becomes negative in the domain of integration in (12). A more careful analysis is therefore needed for the region $q \lesssim q_1$ to ∞, where the WKB solution (12) is no more valid. Now Eqns. (5), (6), and (7) with $E = 2$ may be combined into one equation in terms of $Z(\psi) = \phi_q^2 + (\phi + 1/\phi - 2) \equiv \epsilon^2 a_q^2 + (s^2 - 1/\phi)\, a^2$, where $\psi = 1/\phi$,

$$(Z_{\psi\psi})^2 = \frac{4}{\psi^4} \frac{a^2 a_q^2}{\phi_q^2} = \frac{4}{\epsilon^2 \psi^4} \frac{Z_\psi [Z + (s^2 - \psi) Z_\psi]}{\psi + 1/\psi - 2 - Z}, \tag{17}$$

and $a^2 = -\, Z_\psi$. Equation (17) must be solved with the boundary conditions

$$Z = Z_\psi = 0, \qquad \text{at} \qquad \psi = 1, \tag{18}$$

and $Z = Z_0 \equiv (1 - \psi_0)^2/\psi_0$ and either $Z_\psi = 0$ or $Z_\psi = Z_0/(\psi_0 - s^2)$ at $\psi = \psi_0$. The former condition $Z_\psi = 0$ leads to antisymmetric solution for $a(q)$ and the latter condition is for symmetric solutions. These boundary conditions lead to the requirement that two relations between ϕ_0, s^2 and N (the integer which determines the number of zero's of a) must be satisfied for $N \geq 1$.

The WKB solution (12) satisfies (17). It can be shown that the parameter s is close to unity, so that in the vicinity of the turning point q_0, as well as in the whole "outer" range

$q_0 < q < \infty$, ϕ is also close to unity. In the limit $\phi = 1 + \delta$, $\delta \ll 1$, $s^2 = 1 - \Delta$, $\Delta \ll 1$, (17) reduces to

$$Z_\delta[Z - Z_\delta(\delta - \Delta)] = \frac{\epsilon^2}{4}(Z - \delta^2)Z_{\delta\delta}^2\,, \tag{19}$$

and we have $a^2 \approx Z_\delta$, $Z \approx \delta^2 + \delta_q^2$. In the region $\delta \ll \Delta$ the asymptotic solution of (19) can be shown to be

$$Z \approx P\delta^2\,, \qquad P = 1 + \frac{4\Delta}{\epsilon^2}\,, \qquad \phi - 1 \equiv \delta \approx \frac{a^2(q)}{2P}\,,$$

$$a(q) \approx \left(1 - \frac{a_0^2(q)}{16P\Delta}\right) a_0(q), \qquad a_0(q) = C\exp\left[-\frac{\sqrt{\Delta}}{\epsilon}(q - q_0)\right]. \tag{20}$$

The conventional WKB solution of (5) for the "outer" range $\delta < 1/s^2 - 1$ would formally be $a_0(q)$ with a preexponential factor $\propto |p|^{-1/2}$ similar to (12). However, the actual solution (20) contains another factor in front of $a_0(q)$ which is $\propto |p|^{1/4}$. This factor is of crucial importance to satisfy the identity $Z = \epsilon^2 a_q^2 + (s^2 - 1/\phi)a^2$. In the "inner" region $\delta - \Delta \gtrsim (\epsilon\alpha)^{4/5}$ the asymptotic solution of (19) that matches the WKB solution (12) is

$$Z \approx \alpha^2\left[\sqrt{\delta - \Delta} - \frac{\epsilon}{4}\frac{\delta_q}{(\delta - \Delta)}\sin 2\theta\right],$$

and

$$a \approx \alpha \sin\theta/(\delta - \Delta)^{1/4}\,, \tag{21}$$

with $\delta_q = -\sqrt{Z - \delta^2}$ and $\theta = \theta_0 + \frac{1}{\epsilon}\int_{\delta_1}^{\delta}\sqrt{\delta - \Delta}d\delta/\delta_q$; $\delta_1 = \phi_1 - 1$; ϕ_1 is the second zero in the expression $F(\phi)$ discussed earlier.

From a careful matching [8] of the solutions (20) and (21) in the intermediate region $\delta - \Delta \sim (\epsilon\alpha)^{4/5}$ which includes the WKB turning point $q = q_0$, we can show that Δ is an unambiguous function of ϕ_0, given by

$$\Delta = 1 - s^2 \approx 0.3(\epsilon\alpha)^{4/5}\,, \tag{22}$$

where $\alpha(\phi_0)$ is given by (14) and may be approximated by $\alpha = [(\phi_0 - 1)^3/\phi_0]^{1/4}$ by setting $s^2 \approx 1$. By matching the phase from (12) with θ from (21) we evaluate the WKB phase integral $\oint p dq$,

$$\int_{\phi_1}^{\phi_0} \sqrt{s^2 - \frac{1}{\phi}} \frac{d\phi}{\phi_q} = \epsilon\,(N+\kappa)\,\frac{\pi}{2}, \tag{23}$$

where s^2 is given by (22), ϕ_q is given by (15), and integer N is the number of zeroes of a. The relation (23) differs from the standard WKB phase integral by the value of the phase shift $\kappa\pi/2 \approx 0.5$. The difference is due to the fact that the "energy" level of the Schrödinger Eqn. (5) determined by the parameter s is rather close to the separatrix $\phi = 1$.

The integral in (23) is very insensitive to contributions from the region $\phi \approx \phi_1$. The principal contribution comes from the region $\phi \simeq \phi_0$ and evaluating this contribution we obtain,

$$\sqrt{\phi_0 - 1}\left[1 - \frac{0.3(\epsilon\alpha)^{4/5}}{(\phi_0 - 1)}\right] \approx \epsilon\,(N+\kappa)\,\frac{\pi}{2}, \tag{24}$$

with $\kappa \approx 0.3$ obtained numerically. We arrive at the important conclusion that in the range of ϕ_0 given by (16), the mutually coupled quantities $s^2 = 1 - \Delta$ and $\phi_0 = 1 + \delta_0$ become quantised. The soliton width is given by $\bar{q} \approx \int_{-q_1}^{q_1} dq \approx 2\int_1^{\phi_0} d\phi/\phi_q \approx 3(\phi_0 - 1)^{1/2} + 4\pi/3$ which increases with amplitude unlike the $N = 0$ case.

To summarize, from (14), (22), and (24), we obtain (a) $\Delta = 1 - s^2 \simeq 0.3\,\epsilon^{8/5}(N\pi/2)^{4/5}$ and (b) $\phi_0 \approx 1 + \epsilon^2(N\pi/2)^2$. The relative spacing $\hat{\phi}_0$ between these discrete levels of ϕ_0 is $\hat{\phi}_0 \sim \epsilon\sqrt{\phi_0 - 1}$. The relations (a) and (b) for Δ and ϕ_0 provide the *quantitative* basis for the discrete spectrum of $N \gg 1$ solitons. The existence of such discrete solitons was first demonstrated by Kozlov et al. [1] through numerical simulations and confirmed by Kaw et. al [2] also by numerical means.

The above solution is valid under condition (16) and $\epsilon^2 \ll \Delta \ll 1$, when we automatically have $N \gg 1$. On the other hand, the standard envelope solution for a low amplitude soliton is obtained in the opposite limit $\Delta \ll \epsilon^2$, when $P \approx 1$ in (20) and we have

$$Z \approx \delta^2 \qquad \text{in the entire range } -\infty < q < \infty. \tag{25}$$

This leads to the following approximate equation from (5) and (6),

$$\epsilon^2 a_q^2 + a^2\left(\frac{a^2}{4} - \Delta\right) = 0, \tag{26}$$

with the solution:

$$a = 2\sqrt{\Delta}\,\mathrm{sech}\left(\sqrt{\frac{\Delta}{\epsilon^2}}\,q\right),$$

$$\delta = \frac{a^2}{2} \qquad \text{and} \qquad \Delta \approx \frac{1}{2}\,\delta|_{q=0} \equiv \frac{1}{2}\,\delta_0\,. \tag{27}$$

This shallow well limit is not covered by the WKB solution (12) because $N = 0$, i.e. a vanishes only at $q = \pm\infty$ and nowhere else. In this limit the eigenvalue $\Delta \equiv 1 - s^2 \ll \epsilon^2$ forms a continuum which depends on the soliton amplitude δ_0.

As the amplitude δ_0 of the soliton is increased from unity, the continuum of eigenvalues Δ corresponding to a nonoscillatory level ($N = 0$) for a is restricted by a certain value [8] of $\Delta = D_0\epsilon^2$ where $D_0 = 0.012$. The first antisymmetric soliton with $N = 1$ occurs [8] when $\Delta = D_1\epsilon^2$ where $D_1 = 0.495$. The amplitude of the soliton becomes quantised. At $N \gg 1$, with $\delta_0 = \phi_0 - 1 \gg \epsilon^2$ and Δ given by (22), we have the WKB solution (12).

To conclude, we have obtained one-dimensional soliton solutions for arbitrary amplitude for intense laser pulses propagating in a uniform plasma by asymptotic analysis. In the large N limit we observe that the relation between the normalised nonlinear shift in the group velocity $1 - s$, as a function of laser frequency ϵ^{-1} is determined by (22) where $\alpha \approx \phi_0^{1/2}$ for $\phi_0 \gg 1$; therefore $1 - s \approx \Delta/2 \approx 0.15\,\epsilon^{4/5}\phi_0^{2/5}$. Thus the shift decreases with higher laser frequency. For fixed N, $\phi_0 \propto \epsilon^2$ from (24) i.e. the amplitude decreases with the increase in laser frequency. Both these conclusions verify the Kaw et al. [2] simulations.

ACKNOWLEDGMENTS

We are indebted to D. Pfirsch for insisting on the precise formulation of the boundary conditions of (17). Many discussions on the topic of this paper with S. Cowley and P. Sprangle are also acknowledged. The work was supported in part by the U.S. Naval Research Laboratory.

REFERENCES

[1] V.A.Kozlov, A.G. Litvak and E.V. Suvorov, *Zh. Eksp. Teor. Fiz.* **76**, 148 (1979) [*Sov. Phys. JETP* **49**, 75 (1979)].

[2] P.K. Kaw, A. Sen and T. Katsouleas, *Phys. Rev. Lett.* **68**, 3172 (1992).

[3] A.I. Akhiezer and R.V. Polovin, *Zh. Eksp. Teor. Fiz.* **30**, 915 (1956) [*Sov. Phys. JETP* **3**, 696 (1956)].

[4] N.L. Tsintsadze and D.D. Tshakaya, *Zh. Eksp. Teor. Fiz.* **72**, 480 (1977) [*Sov. Phys. JETP* **45**, 252 (1977)].

[5] See for example: P. Sprangle, E. Esarey and A. Ting, *Phys. Rev. Lett.* **64**, 2011 (1990); G.Z. Sun, E. Ott, Y.C. Lee and P. Guzdar, *Phys. Fluids* **30**, 526 (1987); P. Sprangle, E. Esarey and A. Ting, *Phys. Rev. A* **41**, 4463 (1990); A.B. Borisov, A.V. Borovsky, O.B. Shiryaev, V.V. Korobkin, A.M. Prokhorov, J.C. Salem, T.S. Luk, R. Boyer and C.K. Rhodes, *Phys. Rev. Lett.* **69**, 2200 (1992).

[6] X.L. Chen and R.N. Sudan, *Phys. Fluids B* **5**, 1336 (1993).

[7] S.V. Bulanov, I.N. Inovenkov, V.I. Kirsanov, N.M.Naumova and A.S. Sakharov, *Phys. Fluids B* **4**, 1935 (1992).

[8] Y.S. Dimant and R.N. Sudan, (in preparation).

Appendix Z

Power Flow Between a Plasma-Opening Switch and a Load Separated by a High-Inductance Magnetically Insulated Transmission Line

Power flow between a plasma-opening switch and a load separated by a high-inductance magnetically insulated transmission line

S. B. Swanekamp,[a] J. M. Grossmann, P. F. Ottinger, and R. J. Commisso
Plasma Physics Division, Naval Research Laboratory, Washington, DC 20375

J. R. Goyer
Physics International Company, San Leandro, California 94577

(Received 10 March 1994; accepted for publication 17 May 1994)

Results are presented from particle-in-cell simulations of the electron flow launched from a plasma opening switch (POS) into a magnetically insulated transmission line (MITL) as the POS opens. The opening process of the POS is treated by removing plasma from a fixed anode-cathode gap with an opening time of τ_{rise}. To be similar to opening switch experiments at Physics International, the simulations were performed with the same inductance L_{MITL} between the POS and load. When L_{MITL}/τ_{rise} is large compared to the POS flow impedance, this inductance effectively isolates the POS from the load during the opening process and the POS voltage is insensitive to changes in the load impedance. Analysis and simulations show that the peak load power is maximized when the load impedance is equal to the POS flow impedance. In contrast to previous theories and simulations of magnetically insulated flows, a large amount of electron flow in the MITL is concentrated near the anode. This is a result of the high effective impedance imposed on the POS by the inductive load which causes a significant electron current loss in the POS. As a result, many electrons lose insulation on the load side of the POS gap and those that do flow into the MITL have been accelerated to nearly the full POS potential. Electrons then $\mathbf{E}\times\mathbf{B}$ drift on equipotential lines close to the anode as they enter the MITL and flow toward the load. Current losses in the MITL are observed due to the proximity of the electron flow to the anode. Some electrons flow from the MITL directly into the load and are registered as load current while others $\mathbf{E}\times\mathbf{B}$ drift back toward the POS along the cathode surface. This is possible because the electron flow launched into the MITL from the POS is large enough to cause sufficient positive image charges on the cathode so that the electric field points out of the cathode surface.

I. INTRODUCTION

The plasma opening switch (POS) is the key component for compact terawatt inductive-energy-store pulsed-power systems.[1] In the closed state, the POS must conduct megampere currents for times approaching 1 μs. During this time electrical energy is converted into magnetic energy and stored in the primary storage inductor. To achieve high output powers, the POS must open on a time scale short compared the conduction time and deliver the magnetic energy to a particle-beam-diode or imploding-plasma load. Considerable advances have been made in understanding the conduction and opening phases of the μs-conduction-time POS;[2] however, much less is understood about the power flow out of the POS as the switch opens into the magnetically insulated transmission line (MITL) that connects the POS and load.

There are currently two classes of opening switches being investigated for pulsed-power applications. One class relies on applied magnetic fields in the POS to control the conduction and opening processes.[3] A second class of opening switches uses no external fields in the POS and relies on $J\times B$ forces in the plasma and gap formation processes to control the conduction and opening processes.[2] Since no applied magnetic fields were used in the simulations, the results of this article apply primarily to opening switches with no applied magnetic fields; however, the results of this article may also apply to opening switches in which the self-magnetic field is much larger than the applied fields at the time the switch opens.

A simplified circuit for an inductive-energy-store (IES) pulsed-power system is shown in Fig. 1. In referring to Fig. 1, I_G and I_L are the generator and load currents, C_{TC} is the power supply capacitance, L_S is the storage inductance, L_{MITL} is the inductance between the POS and load, and Z_L is the load impedance. Switch opening can be accompanied by a large inductive voltage as magnetic energy flows past the POS and into the load circuit. This produces large electric fields in the POS that can result in a large fraction of the current switched into the load circuit to exist as vacuum electron flow. To fully understand the coupling of energy between the POS and the load, it is important to understand the distribution and nature of this electron flow. This article examines the flow of magnetic energy between a current charged inductor into an uncharged MITL as the POS opens. Simulations were performed with the particle-in-cell (PIC) code, MAGIC,[4] to analyze POS experiments on the Decade Prototype Module 1 (DPM1)[5] at Physics International. The general electron flow pattern observed in the simulations is presented along with predictions of the division of current between the vacuum flow and the conductor boundaries.

[a]Science Applications International Corporation, McLean, VA 22102.

 0021-8979/94/76(5)/2648/9/$6.00

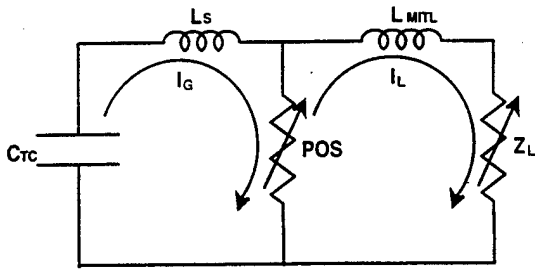


FIG. 1. Simplified circuit of an IES system.

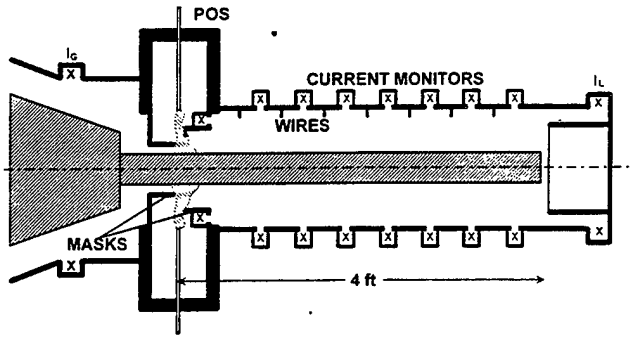


FIG. 2. Schematic of the DPM1 POS and load geometry.

It should be pointed out that both the experiments and simulations were performed with the same large inductance L_{MITL} between the POS and load. This large MITL inductance effectively decouples the POS from the load during the opening process. In this case the POS voltage is limited by the generator current I_G and the gap size that can be achieved in the POS and is insensitive to changes in the load impedance. For this reason this regime of POS operation has been termed switch limited.[6] All the results presented in this article apply to POS operation in the switch-limited regime. This type of POS operation should be contrasted to the regime where the POS and load are closely coupled so that, for small values of load impedance, the POS voltage is limited by the load impedance. This regime of POS operation has been called load limited.[6] The general characteristics of the electron flow in the load-limited regime will be the focus of a future article.

Other authors have considered vacuum electron flow in MITLs introduced by large vacuum impedance mismatches.[7,8] Reference 7 used a fixed gap with both the current and voltage initially zero. Reference 8 uses a static model to predict the electron flow pattern launched by the POS into the MITL. The simulations presented here differs from this work in that the generator current ($I_G \cong 1.2$ MA) flows in the primary circuit at $t=0$. In addition, a dynamic POS model is used where the POS starts in the closed state and evolves into the open state as plasma is removed from a fixed anode-cathode gap.

The remainder of the article is organized as follows. A short description of the DPM1 experiments is presented in Sec. II. A brief overview of the POS model used in the PIC analysis is presented in Sec. III, and the nature of the electron flow out of the POS is described in Sec. IV. The effect of varying the load impedance and MITL inductance on the power transfer between the switch and the load is discussed in Sec. V. The main conclusions of the article are summarized in Sec. VI.

II. DPM1 EXPERIMENTS

The primary energy source for the DPM1 experiment charges a 0.75 μF water capacitor C_{TC} to an initial voltage of approximately 700 kV. During the conduction phase, the capacitor discharges through the POS as current and magnetic energy build up in the 250 nH primary storage inductor L_S. The quarter cycle of the $L_S C_{TC}$ discharge is 680 ns and the peak current is approximately 1.2 MA. A schematic of the DPM1 POS and load circuit geometry is shown in Fig. 2. Plasma is introduced into the POS prior to the main pulse by a set of eight cable guns spaced around the azimuth of a coaxial transmission line.[9] The cable guns are powered separately from the main power pulse and the initial plasma density in the switch is controlled by varying the delay between the plasma guns and the main pulse. The location of plasma in the switch region is controlled by two masks which limit the axial extent of the plasma introduced into the MITL. Plasma from the cable guns flows into the POS through a 2.5 cm opening between the masks as shown in the figure. The mask on the generator side (upstream) of the POS had a radius of $r=7.5$ cm while the mask on the load side (downstream) was placed at $r=10$ cm. The load circuit in the DPM1 experiment consists of either a 60- or 120-cm-long MITL terminated by an electron-beam-diode load. The cathode radius in the downstream MITL was $r_c=4.5$ cm and the anode radius was $r_a=17.2$ cm giving a vacuum impedance of $Z_0=80\ \Omega$.

Current and voltage monitors upstream of the POS were used to measure the generator current and the switch voltage. On most shots only the switch voltage, the generator current, and load current was measured; however, on a few shots a series of anode current monitors were used to measure the current in the MITL at several axial locations. The first of these current monitors was placed on the mask just downstream of the POS with additional anode current monitors spaced every 15 cm along the MITL. Another diagnostic used in these experiments consisted of a series of wires that protruded 2.5 cm through the anode of the MITL. The first of these wires protruded from the downstream mask with additional wires placed 15 cm apart and halfway between the anode current monitors. The damage patterns on these wires were used to detect the presence of high-energy electrons near the anode in the MITL.

Typical wave forms for the generator current, switch voltage, and load currents are shown in Fig. 3 for a 60-cm-long MITL terminated by a pinched electron-beam-diode load with an operating impedance of approximately 5 Ω. During the 600 ns that the POS is closed, no current flows in the load circuit and magnetic energy is stored in the primary storage inductance. As the switch opens, current flows into the load and an inductive voltage of about 2.7 MV is generated. The difference between the generator and load currents I_L after the switch opens represents current loss from either

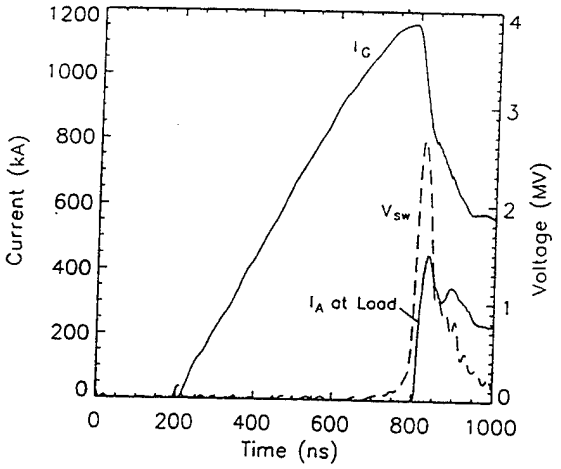


FIG. 3. The generator current, POS voltage, and load current wave forms for DPM1 with approximately 1.2 MA of conducted current and opening into a 60 cm, 80 Ω transmission line terminated by a pinched-beam diode with a 5 Ω operating impedance.

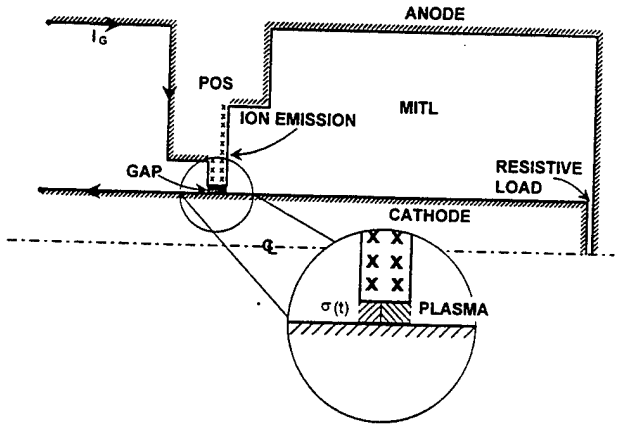


FIG. 4. A schematic of the POS, MITL, and load regions used to simulate DPM1. The anode surfaces marked with × signify space-charge-limited emission of C^{++} ions and the entire cathode surface is a space-charge-limited electron emitter.

current across the POS or from vacuum electron flow to the anode in the MITL. For this shot the current loss at peak load power is approximately 350 kA out of approximately 800 kA. Since this current loss is nearly as large as the net current delivered to the load, substantial gains in system efficiency can be achieved by reducing these current losses. In subsequent sections of the article the loss mechanisms predicted by simulations are presented.

III. SWITCH MODEL

Because of the relatively large densities expected in the DPM1 experiments ($n>10^{15}$ cm^{-3})[10] and fairly long conduction times ($\tau_c>300$ ns) used in the DPM1 experiment, it is not practical to model the entire conduction and opening phases of the POS with a PIC code. Data taken on the HAWK accelerator at NRL suggests that the POS opens by the formation of a vacuum gap in the plasma that allows magnetic energy to flow past the POS and into the load circuit.[4] Gap formation in the low density ($n<10^{13}$ cm^{-3}), short-conduction-time ($\tau_c<50$ ns) POS has been studied extensively with both PIC[11–13] and fluid[14,15] codes. These simulations show the formation of a localized two-dimensional sheath and associated current channel that penetrates the plasma along the cathode. Behind the current front, the magnetic field is high and electrons are magnetically insulated in a vacuum gap which forms in the wake of the sheath. Opening occurs in these simulations when the current channel reaches the load end of the plasma and the gap extends along the entire length of the plasma. Recent PIC simulations[16] of the end of the conduction phase in a high-density ($n\geqslant 10^{15}$ cm^{-3}) POS show similar behavior to the low-density simulations and suggest that the POS opens by erosion even in regimes where MHD distortion of the plasma may control the conduction process.[2,17–19]

Since the purpose of this article is to study the vacuum electron flow that accompanies the opening process and not in the details of the conduction process, a dynamic gap opening model was used to approximate the final stages of POS opening observed in PIC simulations of the gap formation process. A schematic of the POS, MITL, and load models used in the simulations is shown in Fig. 4. The model assumes that the plasma distortion and redistribution that takes place during the conduction phase has already occurred and that the POS plasma carries the entire generator current at $t=0$. In the POS model, a 2.5-cm-long, 0.35-cm-high switch gap is assumed to exist between the anode and cathode. The last 1.5 cm length of the switch gap is bridged with plasma so that the POS gap is initially very conductive. The plasma density of the prefill plasma was chosen to be 10^{14} cm^{-3} and the ions in the plasma prefill were taken to be 10^6 times more massive than the proton mass with an initial velocity sufficient to clear the switch gap in 30 ns. Because of the artificially high ion mass, the prefill ions are not affected by electromagnetic forces during the opening process. This allows for control over the switch opening process in the simulations. As the prefill ions clear the gap, magnetic energy flows past the switch and into the downstream MITL.

The portion of the anode that would be high-density plasma in the experiment is modeled as a conducting boundary and treated as a space-charge-limited C^{++} ion emitter. The entire cathode surface in the MITL is modeled as a space-charge-limited source of electrons. To achieve quiet start conditions in the simulation, the POS must carry the generator current at $t=0$. One way to accomplish this is to initially distribute the current in the POS plasma to carry the generator current. Since it is difficult to self-consistently determine this distribution *a priori*, an alternate method was adopted in the simulations. An enlarged view of the POS gap is shown in Fig. 4. The first 1 cm of the gap is treated as region of time varying conductivity so that the generator current initially flows here. The conductivity of this region drops in time to produce a resistance rise of 2 Ω/ns. As the conductivity region becomes more and more resistive, more and more current is transferred to the prefill plasma. It is found that nearly all the generator current flows in the prefill plasma within the first few ns so that the POS impedance is

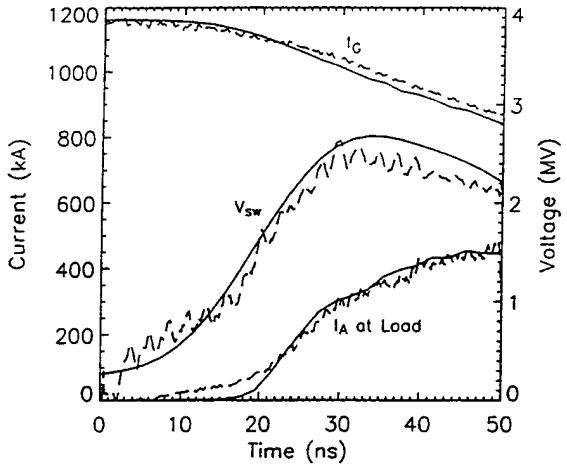


FIG. 5. A comparison of the generator current, POS voltage, and load current predicted by simulation (dashed line) with that measured on DPM1 (solid line).

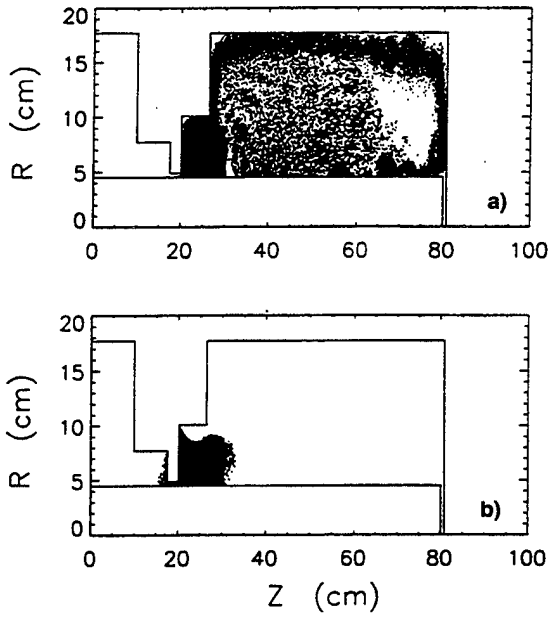


FIG. 6. (a) Electron positions predicted by the simulation at $t=30$ ns. (b) C^{++} positions predicted by the simulation at $t=30$ ns.

quickly dominated by the charged particle dynamics in the switch gap.

The water capacitor and the primary storage inductor (not shown in the figure) are approximated by lumped circuit elements using short one-dimensional transmission line sections and matched to the 2D simulation region across the left-hand-side boundary. The currents and voltages in the transmission line along with the electric and magnetic fields in the 2D region prior to the POS are preloaded at $t=0$ to be consistent with $I_G=1.2$ MA. The switch is assumed to be closed at $t=0$ so that the electric and magnetic fields downstream of the POS are initially zero. A static conductivity model in the load region was used to simulate a resistive load.

IV. ELECTRON FLOW AND CURRENT TRANSFER

The simulation results described in this section were performed with a 60-cm-long MITL with a vacuum impedance of $Z_0=80$ Ω. For comparison with experimental data, the load impedance for the results presented in this section was 5 Ω; however, because the large MITL inductance causes the POS to operate in the switch-limited regime, the general electron flow pattern for load impedances between $0\leq Z_L\leq 5$ Ω and MITL lengths of 30, 60, and 120 cm (downstream inductances 80, 160, and 320 nH) is very similar to that presented in this section.

In this geometry, the switch voltage along with the generator and load currents were measured in the experiment. A comparison of the measured generator current, switch voltage, and load current with that predicted by the PIC simulation are shown in Fig. 5. This figure shows good agreement between the measured electrical parameters and the predicted voltages and currents from the simulation. The small differences between the measured and simulated load currents early in time occur because of the initial high impedance phase of the pinched beam diode used in the experiments which was instead modeled by a constant 5 Ω load in the simulations.

The electron and emitted C^{++} ion positions at $t=30$ ns are shown in Figs. 6(a) and 6(b). Figure 6(a) shows a significant fraction of the electron flow launched in the MITL is concentrated near the anode where the electrons $\mathbf{E}\times\mathbf{B}$ drift parallel to the anode conductor. As the electrons approach the vertical anode plate the electron orbits bend 90° and continue to $\mathbf{E}\times\mathbf{B}$ drift toward the load region. Damage produced on the wires protruding from the anode over several shots is shown in Fig. 7. Notice that the first few wires are relatively undamaged. This may be the result of plasma in the MITL which can shield the wires from the high energy electrons. The remaining wires show a large amount of damage indicating the presence of high-energy electrons near the anode. This agrees qualitatively with the electron flow pattern observed in the simulations.

POS opening is also accompanied by the introduction of a large amount of negative charge into the MITL. This causes positive image charges to appear on both the anode

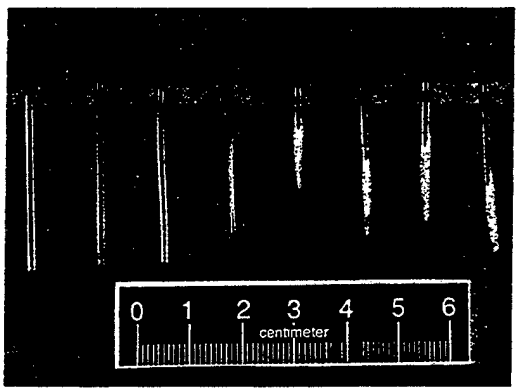


FIG. 7. Damage on the series of wires protruding through the anode along the length of the MITL on DPM1.

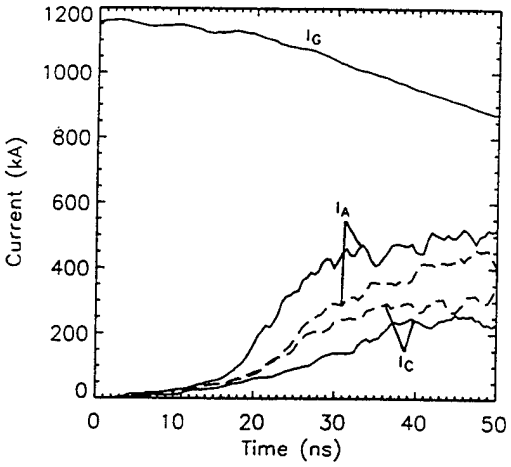


FIG. 8. Time histories of the anode I_A and cathode I_C currents predicted by simulation just downstream of the POS region (solid line) and near the load (dashed line). For comparison the generator current I_G is also shown.

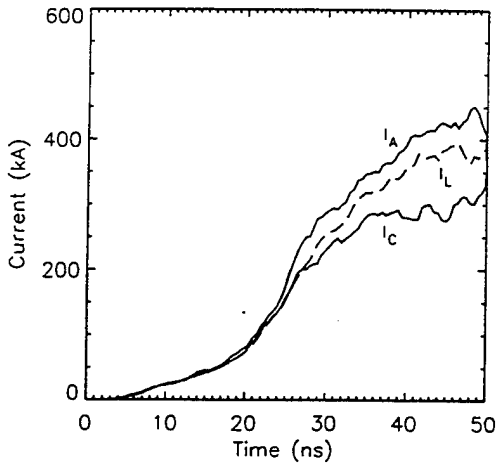


FIG. 9. Time histories of the anode I_A and cathode currents I_C near the load (solid line) and the net current flowing to the anode within the load region I_L (dashed line).

and cathode surfaces. As a result, the electric field on the cathode surface is reversed (i.e., points out of the surface) and electron emission from the cathode is suppressed. If plasmas were formed on the electrode surfaces, the fields on both electrodes would favor ion emission. In this work no ion emission was allowed from either the anode or cathode in the MITL and hence the effect of ion flow in the MITL was not considered.

Field reversal on the cathode also allows some of the electrons that approach the cathode in the load region to **E×B** drift back along the cathode surface toward the POS. This electron flow pattern is very different than that presented by either parapotential[20] or quasilaminar[21] theories of MITL flow. Since these theories assume that the electric field is zero on the cathode surface, they do not treat the case of field reversal on the cathode. In addition these theories are time independent and apply only to situations where variations in the MITL occur gradually so that the electrons remain in a quasistatic equilibrium. Both of these conditions are violated in these experiments. Consequently many previous MITL theories have not been very useful for understanding the type of flow observed in the simulations.

Electrons which encounter the anode boundary along the length of the MITL are removed from the simulation and are registered as anode current loss. Electrons which encounter the cathode boundary along the length of the MITL are retrapped and removed from the simulation. These electrons add to the cathode boundary current and are registered as an increase in the cathode boundary current. If electrons followed the **E×B** guiding center trajectory, no electrons could encounter the anode or cathode boundaries since the electric field must be perpendicular to the conductor. Because of their finite gyroradius and other drift motions, the actual electron orbit is more complicated and it is possible for electrons to encounter the conductor boundaries.

Figure 8 shows the time history of the generator current along with the anode and cathode boundary currents just downstream of the POS ($z=30$ cm) and near the load ($z=80$ cm). The difference between the generator current and the anode current just downstream of the POS indicates current loss in the POS. For this simulation it is seen that nearly 350 kA out of 850 kA is lost in the POS at $t=50$ ns. The amount of vacuum electron flow launched by the POS into the downstream MITL is given by the difference between the anode and cathode currents just downstream of the POS. The vacuum electron flow in the MITL near the POS at $t=50$ ns is approximately 250 kA. Therefore, 250 kA of the 500 kA transferred into the downstream MITL is in the form of vacuum electron flow; however, as indicated by the difference between the anode and cathode currents at the load, the amount of vacuum electron flow at the load is substantially smaller than the vacuum electron flow near the POS. This reduction in vacuum electron flow occurs as a result of losses at the anode and retrapping at the cathode. In Fig. 8 the amount of current loss in the MITL is given by the difference between anode currents near the POS and near the load. The amount of retrapped current is given by the difference between the cathode current near the load and the cathode current near the POS. It is important to note that, in general, the current lost in the POS and the amount of vacuum electron flow in the MITL depends on the load impedance.

In addition to the current flowing in the resistive load, a significant fraction of the vacuum electrons that approach the load region flow directly into the load and are registered as load current. Because of field reversal on the cathode, some of the electrons that approach the load region can also **E×B** drift back along the cathode surface towards the POS. Figure 9 shows the time histories of the anode and the cathode current at $z=80$ cm (0.7 mm upstream from the vertical anode plate). The dashed curve in Fig. 9 is the total current in the load determined by a current monitor on the vertical anode plate at the cathode radius. The difference between the load current and the cathode current is the net vacuum electron current flowing into the load from the MITL. The difference between the anode and load currents represents the net electron flow to the anode plate outside the cathode radius. Because electron flow from the MITL into the load requires an axial field for the electrons to radially **E×B** drift, the amount

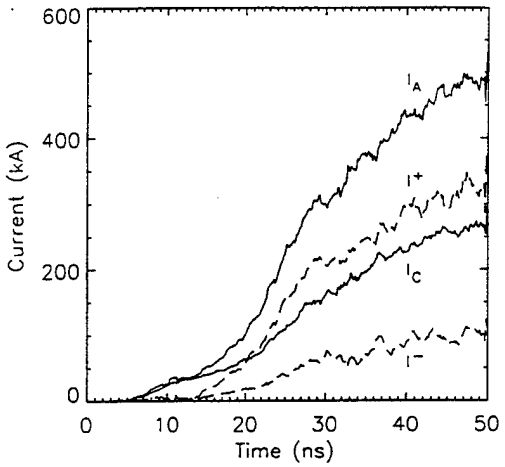


FIG. 10. Time histories of the magnitude of the vacuum electron current flowing in the positive (I^+) and negative (I^-) z directions at $z=50$ cm (dashed line). For comparison the anode (I_A) and cathode (I_C) boundary currents at $z=50$ cm (solid line) are also shown.

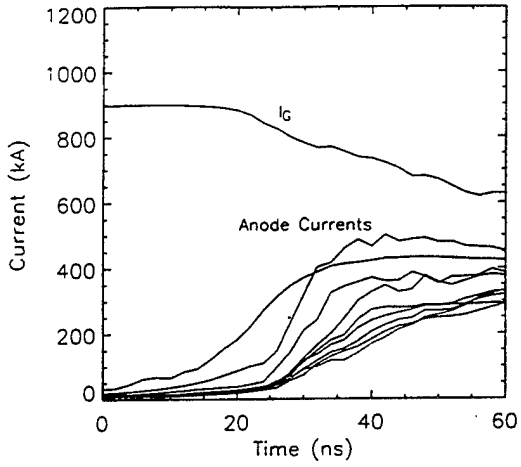


FIG. 11. DPM1 data from the series of anode current monitors shown in Fig. 2.

of vacuum electron flow into the load from the MITL depends on the load impedance.

Figure 10 shows the time histories of the magnitude of vacuum electron current flowing in the positive and negative z directions (the dashed curves) at $z=50$ cm along with the anode and cathode boundary currents (the solid curves). The positive going flow is produced by those electrons $\mathbf{E}\times\mathbf{B}$ drifting along the anode surface and the negative going flow is produced by the electrons $\mathbf{E}\times\mathbf{B}$ drifting back along the cathode surface toward the POS region. Notice that the negative going flow is a substantial fraction of the net current and cannot be neglected. Furthermore, notice that net electron flow predicted by the difference of the anode and cathode boundary currents is actually the difference between the positive and negative going flows.

From the ion positions shown in Fig. 6(b), we see that an ion space-charge cloud evolves off of the anode emission surfaces. This positive space charge is quickly neutralized by electron flow from the POS region forming a quasineutral plasma that moves down the MITL at about 0.3 cm/ns. To explore the effect of this plasma on the electron flow pattern, simulations were done with ion emission turned off everywhere except in the POS gap. These simulations showed that, even without the plasma in the MITL, electron flow patterns were similar to that presented here; however, in the experiment it is possible for plasma to be pushed into the MITL during the conduction and opening processes with a distribution that is different than what is shown here. This plasma could be responsible for shielding the first few wires from electron damage (see Fig. 7).

Experimental data from the series of anode current monitors along the length of the DPM1 MITL are shown in Fig. 11. This data was taken with a 120-cm-long MITL terminated by a short-circuit load. A similar plot for a 120-cm-long MITL simulation terminated by a 2.5 Ω load is shown in Fig. 12. In general, the highest of the anode currents in Fig. 12 was measured closest to the POS. Each successive anode current monitor in Fig. 12 shows a decrease in anode current due to vacuum electron current losses to the anode. Although the conduction currents and the load impedances are somewhat different, Figs. 11 and 12 show similar current losses in the POS and MITL.

V. DISCUSSION AND ANALYSIS

In this section a plausible explanation for the electron flow pattern observed in the simulations is presented and the role of different load impedances and MITL inductances is explored. The maximum radial extent of the electron flow in the POS gap can be estimated from the critical current formula and expressed as[20,22]

$$D_{\text{eff}}(t)=\kappa\ 8500(\gamma_{\text{sw}}^2(t)-1)^{1/2}r_c/I_G(t), \tag{1}$$

where $\gamma_{\text{sw}}(t)=1+eV_{\text{sw}}(t)/mc^2$, V_{sw} is the switch voltage, e is the magnitude of the electron charge, m is the electron rest mass, c is the speed of light, and κ is an empirical factor. It has been shown that, for voltages in the 1–3 MV range, $\kappa=1.6$ gives good agreement with experiment and will be used here.[23] Note that if $\kappa=1$ then D_{eff} is an expression for

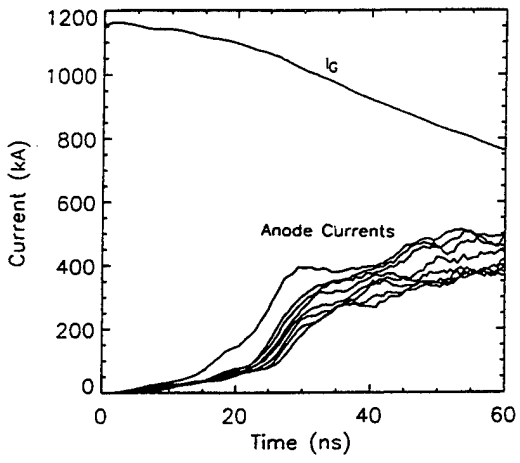


FIG. 12. Anode current predictions from simulation at axial locations similar to the DPM1 data shown in Fig. 11.

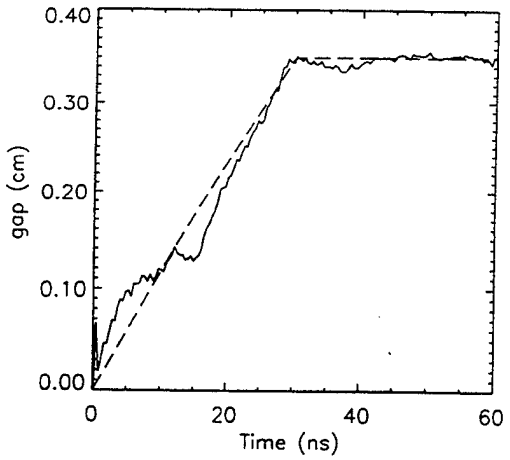


FIG. 13. A comparison of the POS gap size predicted from the critical current formula (solid line) with the dynamic gap formed between the receding plasma and the anode (dashed line).

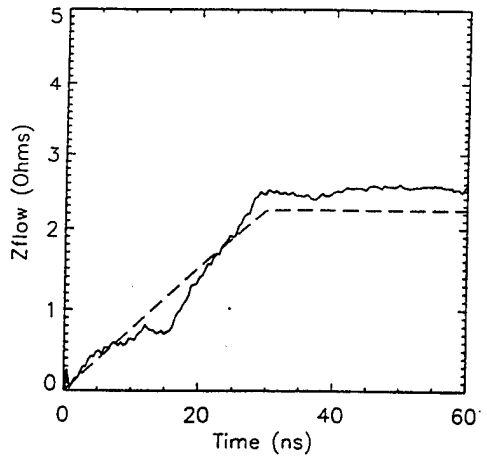


FIG. 14. A comparison of the POS flow impedance predicted from the simulation (solid line) with half the vacuum impedance of the dynamic gap (dashed line).

the Larmor radius of an electron with a speed consistent with the full anode potential. A plot of D_{eff} using V_{sw} and I_G from the simulation is shown in Fig. 13. For comparison the dashed curve shows the time-dependent gap $D(t)$, opened up between the plasma and the anode as the heavy ions are removed from the simulation. This figure suggests that, when the POS operates in the switch-limited regime, electrons are critically insulated during the entire opening process. Furthermore, Eq. (1) shows that, in the switch-limited regime, the POS voltage is limited by the primary storage current and the size of the POS gap. However, in the load-limited regime it is possible for electrons to be better insulated in the POS gap since the POS voltage is not limited by POS gap but limited by the load impedance.

Another concept that has been used to describe vacuum electron flow in a POS or MITL is the flow impedance.[3] For the POS, flow impedance is given by

$$Z_{\text{flow}}^{\text{POS}}=\frac{V_{\text{sw}}}{(I_G^2-I_{CD}^2)^{1/2}}, \tag{2}$$

where I_{CD} is the cathode boundary current just downstream of the POS. The flow impedance of the POS is a time-varying quantity that is related to the amount of vacuum electron flow in the POS. When the POS is closed $Z_{\text{flow}}^{\text{POS}}=0$ since $V_{\text{sw}}=0$. In this case, the entire generator current flows across the POS and nearly all the POS current is carried by the plasma electrons. It can be shown that, when there is very little electron flow in the POS (i.e., $I_{CD}\rightarrow I_G$), $Z_{\text{flow}}^{\text{POS}}$ is bounded above by the vacuum impedance of the POS gap.[3,24]

From the definition of flow impedance it can be shown that the fractional change in the POS flow impedance is related to the fractional change in the cathode current downstream by

$$\frac{\Delta Z_{\text{flow}}^{\text{POS}}}{Z_{\text{flow}}^{\text{POS}}}=\frac{\Delta I_{CD}}{I_{CD}}\left[\left(\frac{I_{CD}^2}{I_G^2}\right)\left(1-\frac{I_{CD}^2}{I_G^2}\right)^{-1}\right]\cong\frac{\Delta I_{CD}}{I_{CD}}\frac{I_{CD}^2}{I_G^2}, \tag{3}$$

where the last expression is valid in the switch limited regime where $I_{CD}\ll I_G$. This last equation shows that, when I_{CD} is small compared to I_G, $Z_{\text{flow}}^{\text{POS}}$ is insensitive to changes in the cathode current downstream of the POS. This feature makes $Z_{\text{flow}}^{\text{POS}}$ insensitive to changes in the load impedance in the switch-limited regime. Conversely, relatively small changes in $Z_{\text{flow}}^{\text{POS}}$ can lead to large increases in I_{CD}. For example, when $I_{CD}=250$ kA and $I_G=850$ kA, a fractional increase of 5% in $Z_{\text{flow}}^{\text{POS}}$ produces an increase in the cathode boundary current to $I_{CD}=380$ kA.

By assuming electromagnetic pressure balance at the anode and cathode, it can be shown that, when the electron flow uniformly fills the POS gap (critically insulated or saturated flow), the POS flow impedance is given by[25]

$$Z_{\text{flow}}^{\text{POS}}(t)=0.5Z_0(t)=30\ \ln[1+D(t)/r_c]\cong 30D(t)/r_c, \tag{4}$$

where $Z_0(t)$ is the vacuum impedance of the time-dependent gap. Equation (4) shows the relationship between the POS flow impedance and gap size in the switch-limited regime. A comparison of $Z_{\text{flow}}^{\text{POS}}$ calculated from the voltages and currents in the simulation and Eq. (4) is shown in Fig. 14. From this figure it can be seen that the POS flow impedance rises from zero to approximately 2.5 Ω during the gap opening process and remains fairly constant once the gap is fully open. Although this figure shows good agreement between Eq. (4) and the simulation, the discussion following Eq. (3) showed that small differences in $Z_{\text{flow}}^{\text{POS}}$ can lead to large changes in the cathode current downstream of the POS. Empirically it is found that a value of $0.53Z_0(t)$ produces better agreement with the simulation.

The effective impedance of the inductive region between the POS and the load can be defined as

$$Z_{\text{eff}}=L_{\text{MITL}}\frac{dI_L/dt}{I_L}\approx\frac{L_{\text{MITL}}}{\tau_{\text{rise}}}, \tag{5}$$

where τ_{rise} is the time scale for the rise in the load current. If τ_{rise} is approximately the 30 ns opening time of the POS, then, for the $L_{\text{MITL}}=160$ nH simulation results presented

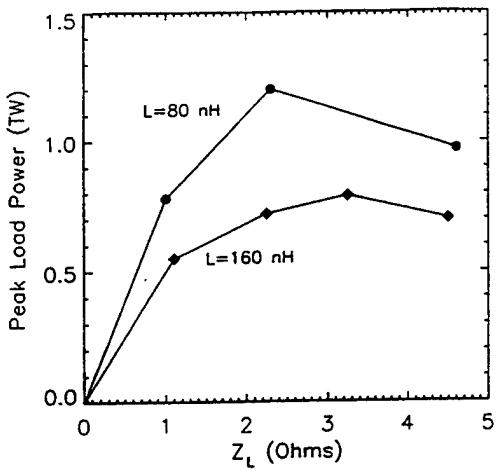


FIG. 15. A plot of the peak load power vs load impedance for MITL inductances of 80 nH (○) and 160 nH (◇).

here, Z_{eff}=5.3 Ω. Therefore, the MITL inductance imposes an impedance on the POS that is large compared to the flow impedance of the POS. When the impedance of the series combination of Z_{eff} and Z_L is large compared to $Z_{\text{flow}}^{\text{POS}}$, the POS operates in the switch-limited regime. In the switch-limited regime a significant amount of electron current is lost in the POS. As a result, many electrons lose insulation at the load end of the POS gap and those that do flow into the MITL have been accelerated to nearly the full POS potential. If electrons then **E×B** drift along equipotential lines as they enter the downstream MITL, electron flow from the POS would be concentrated near the anode as seen in the simulations.

To examine power flow when Z_{eff} is comparable to maximum POS flow impedance a second series of simulations was done with L_{MITL}=80 nH. This reduces Z_{eff} to approximately 2.65 Ω. The POS voltage predicted with the shorter MITL is very similar to that shown Fig. 5. The electron flow pattern, D_{eff}, and $Z_{\text{flow}}^{\text{POS}}$ were also very similar to the results shown in Figs. 6(a), 13, and 14, respectively. This is expected since the POS flow impedance rises from zero and hence is smaller than Z_{eff} during the opening process and the POS still operates in the switch-limited regime.

Figure 15 shows the peak load power for the two MITL inductances as a function of load impedance. Even though the POS operation is very similar for the two MITL inductances, Fig. 15 shows that the peak load power is significantly enhanced by reducing the MITL inductance. This increase in peak power at lower MITL inductance is the result of the improved current coupling from the POS to the load. This improved current coupling is produced by reducing both current losses in the POS and vacuum electron flow in the MITL.

To better understand the reduced POS losses as MITL inductance decreases consider an ideal opening switch that opens to an infinite impedance into a short-circuit load. If the capacitor voltage does not change significantly as the switch opens, then magnetic flux is conserved during the opening process and the final current in the MITL can be expressed as

$$I_{Lf}=\frac{L_S}{L_S+L_{\text{MITL}}}I_{Gi}, \tag{6}$$

where I_{Gi} is the current flowing in the storage inductor prior to switching. Equation (6) shows that reducing the MITL inductance results in higher MITL currents and reduced POS losses.

Figure 15 also shows that the power coupled into the load from a POS has a maximum at an optimum load impedance. The load power can be written as

$$P_L=V_L I_L, \tag{7}$$

where V_L is the load voltage. Since maximum load power occurs at peak load current for a resistive load, there is no inductive correction between the POS and load and the peak load power can be expressed as

$$P_L^{\text{peak}}=Z_{\text{flow}}^{\text{POS}}(I_G^2-I_{CD}^2)^{1/2}I_L, \tag{8}$$

where Eq. (2) has been used to substitute for $V_L=V_{\text{sw}}$. If it is assumed that all the vacuum electron flow from the POS is lost before the load, then $I_{CD}=I_L$ and the power coupled to the load can be written as

$$P_L^{\text{peak}}=Z_{\text{flow}}^{\text{POS}}(I_G^2-I_L^2)^{1/2}I_L. \tag{9}$$

Since the load current depends on the load impedance, Eq. (9) is a function of Z_L. If it is assumed that $Z_{\text{flow}}^{\text{POS}}$ and I_G at peak load power are weak functions of Z_L, then the load current which maximizes Eq. (9) is given by $I_L=I_G/\sqrt{2}$. Substituting this result into Eq. (9) gives

$$\max(P_L^{\text{peak}})=Z_{\text{flow}}^{\text{POS}}I_L^2. \tag{10}$$

Since P_L is also equal to $Z_L I_L^2$, Eq. (10) shows that the peak load power is maximized when the load impedance is equal to the POS flow impedance. Although more simulations need to be done to determine the exact impedance that maximizes peak load power, it is pretty clear from Fig. 15 that the optimum load impedance predicted from simulations is close to $Z_{\text{flow}}^{\text{POS}}$. It should be noted that Eqs. (9) and (10) are conservative estimates of the load power since higher load powers can be obtained if vacuum electron flow is either retrapped in the cathode along the MITL or flows directly into the load.

VI. SUMMARY AND CONCLUSIONS

An approximate model of the opening process of a plasma opening switch was used in a PIC code to study the vacuum electron flow into the downstream MITL as the POS opens. To be similar to opening switch experiments on DPM1 at Physics International, the simulations were performed with a large amount of inductance between the POS and load. When $L_{\text{MITL}}/\tau_{\text{rise}}$ is large compared to the flow impedance of the POS, this large inductance causes the POS to operate in the switch-limited regime. In the switch-limited regime, the POS voltage is insensitive to changes in the load impedance and is controlled by the generator current, the cathode radius, and the dynamic POS gap size. It is anticipated that the load-limited regime can be accessed by reducing $L_{\text{MITL}}/\tau_{\text{rise}}$ which would reduce the POS voltage and could result in better insulated electron flows in the POS and MITL. This will be the subject of future work.

In contrast to previous theories of magnetically insulated flows, there is a large amount of electron flow in the MITL concentrated near the anode. This is a result of the high effective impedance imposed on the POS by the inductive load. As a result, many electrons lose insulation at the load end of the POS gap and those that do flow into the MITL have been accelerated to nearly the full POS potential. Electrons then $\mathbf{E}\times\mathbf{B}$ drift on equipotential lines near the anode as they enter the MITL. In addition to current losses in the POS gap, current losses are also observed in the MITL due to the proximity of the electron flow to the anode. As electrons approach the resistive load, some electrons flow directly into the load and are registered as load currents while others $\mathbf{E}\times\mathbf{B}$ drift back toward the POS along the cathode surface. This is possible because the electron flow launched into the MITL from the POS is large enough to cause sufficient positive image charges on the cathode so that the electric field points out of the cathode surface. Some of the electron flow near the cathode surface is retrapped in the cathode boundary and registered as an increase in the cathode current. The net vacuum electron current in the MITL is the difference between these two oppositely directed flows.

Another series of simulations with the MITL inductance reduced by a factor of 2 show a significant increase in the peak power that is delivered to the resistive load. In addition, the peak power delivered to the load has a maximum at an optimum load impedance. Analysis and simulations show that the optimum load impedance is approximately equal to the flow impedance of the POS.

ACKNOWLEDGMENTS

Discussions of these results with Dr. Cliff Mendel, Dr. Phil Spence, Dr. Eduardo Waisman, and Dr. Bruce Weber were very valuable and greatly appreciated. Dr. Larry Ludeking is gratefully acknowledged for his technical assistance with the MAGIC code. Simulation results were obtained using the MAGIC code under the AFOSR-sponsored MAGIC User's Group. This work was supported by Defense Nuclear Agency.

[1] See guest editorial and articles in special issue on fast opening vacuum switches, IEEE Trans. Plasma Sci. **PS-15**, 629 (1987).

[2] R. J. Commisso, P. J. Goodrich, J. M. Grossmann, D. D. Hinshelwood, P. F. Ottinger, and B. V. Weber, Phys. Fluids B **4**, 2368 (1992).

[3] C. W. Mendel, M. E. Savage, D. M. Zagar, W. W. Simpson, T. W. Grasser, and J. P. Quintenz. J. Appl. Phys. **46**, 3731 (1992).

[4] B. Goplen, L. Ludeking, D. Smithe, and G. Warren, MAGIC User's Manual, Mission Research Tech. Report No. MRC/WDC-R-282.

[5] C. McDonald, P. Sincerny, and L. Schlitt, in *Proceedings of the 8th IEEE Pulsed Power Conference*, San Diego, CA, edited by R. White and K. Prestwich (IEEE. New York, 1991), p. 675.

[6] P. J. Goodrich, J. R. Boller, R. J. Commisso, D. D. Hinshelwood, J. C. Kellogg, and B. V. Weber, in *Proceedings of the 8th IEEE Pulsed Power Conference*, San Diego, CA. edited by R. White and K. Prestwich (IEEE, New York, 1991), p. 515.

[7] R. I. Lawconnell and J. M. Neri, Phys. Fluids B **2**, 629 (1990).

[8] C.-K. Ng and R. N. Sudan, J. Appl. Phys. **69**, 137 (1991).

[9] J. R. Goyer, D. Kortbawi, F. K. Childers, and P. S. Sincerny, J. Appl. Phys. **74**, 4236 (1993).

[10] B. V. Weber (private communication).

[11] E. M. Waisman, P. G. Steen, D. E. Parks, and A. Wilson, Appl. Phys. Lett. **46**, 1045 (1985).

[12] J. M. Grossmann, P. F. Ottinger, J. M. Neri, and A. T. Drobot, Phys. Fluids **29**, 2724 (1986).

[13] J. M. Grossmann, P. F. Ottinger, and R. J. Mason, J. Appl. Phys. **66**, 2307 (1989).

[14] R. J. Mason, M. E. Jones, J. M. Grossmann, and P. F. Ottinger, Phys. Rev. Lett. **61**, 1835 (1988).

[15] R. J. Mason, P. L. Auer, R. N. Sudan, B. V. Oliver, C. E. Seyler, and J. B. Greenly, Phys. Fluids B **5**, 1115 (1993).

[16] J. M. Grossmann, S. B. Swanekamp, P. F. Ottinger, R. J. Commisso, D. D. Hinshelwood, and B. V. Weber, Proceedings of the 10th International Conference on High Power Particle Beams, San Diego, 1994.

[17] W. Rix, D. Parks, J. Shannon, J. Thompson, and E. Waisman, IEEE Trans. Plasma Sci. **PS-19**, 400 (1991).

[18] D. Hinshelwood, B. V. Weber, J. M. Grossmann, and R. J. Commisso, Phys. Rev. Lett. **68**, 3567 (1992).

[19] B. V. Weber, J. R. Boller, R. J. Commisso, P. J. Goodrich, J. M. Grossmann, D. D. Hinshelwood, J. C. Kellogg, P. F. Ottinger, and G. Cooerstein, in Proceedings of the 9th International Conference on High Power Particle Beams, edited by D. Mosher and G. Cooperstein, Washington, DC, May 25–29, 1992, NTIS document No. PB92-206168, p. 375; copies can be ordered from the NTIS, Springfield, VA 22161.

[20] J. M. Creedon, J. Appl. Phys. **46**, 2946 (1975).

[21] R. V. Lovelace and E. Ott, Phys. Fluids **17**, 1263 (1974).

[22] P. F. Ottinger, S. A. Goldstein, and R. J. Meger, J. Appl. Phys. **56**, 774 (1984).

[23] R. J. Barker and S. A. Goldstein, Bull. Am. Phys. Soc. **26**, 921 (1981).

[24] C. W. Mendel, D. B. Siedel, and S. E. Rosenthal, Laser Particle Beams **1**, 311 (1983).

[25] S. E. Rosenthal, IEEE Trans. Plasma Sci. **PS-19**, 823 (1991).

Appendix AA

Gap Formation Processes in a High-Density Plasma Opening Switch

Gap formation processes in a high-density plasma opening switch

J. M. Grossmann, S. B. Swanekamp,[a] P. F. Ottinger, R. J. Commisso,
D. D. Hinshelwood,[b] and B. V. Weber
Pulsed Power Physics Branch, Plasma Physics Division, Naval Research Laboratory, Washington, D. C. 20375-5000

(Received 7 March 1994; accepted 15 September 1994)

A gap opening process in plasma opening switches (POS) is examined with the aid of numerical simulations. In these simulations, a high density ($n_e=10^{14}$–5×10^{15} cm^{-3}) uniform plasma initially bridges a small section of the coaxial transmission line of an inductive energy storage generator. A short section of vacuum transmission line connects the POS to a short circuit load. The results presented here extend previous simulations in the $n_e=10^{12}$–10^{13} cm^{-3} density regime. The simulations show that a two-dimensional (2-D) sheath forms in the plasma near a cathode. This sheath is positively charged, and electrostatic sheath potentials that are large compared to the anode–cathode voltage develop. Initially, the 2-D sheath is located at the generator edge of the plasma. As ions are accelerated out of the sheath, it retains its original 2-D structure, but migrates axially toward the load creating a magnetically insulated gap in its wake. When the sheath reaches the load edge of the POS, the POS stops conducting current and the load current increases rapidly. At the end of the conduction phase a gap exists in the POS whose size is determined by the radial dimensions of the 2-D sheath. Simulations at various plasma densities and current levels show that the radial size of the gap scales roughly as B/n_e, where B is the magnetic field. The results of this work are discussed in the context of long-conduction-time POS physics, but exhibit the same physical gap formation mechanisms as earlier lower density simulations more relevant to short-conduction-time POS.

I. INTRODUCTION

Inductive energy storage (IES) is considered a lower cost, smaller size alternative pulsed power technology compared with conventional capacitive energy storage.[1] An integral part of an IES system is the plasma opening switch (POS), which bridges the anode–cathode (A–K) gap of a section of transmission line as the storage inductor charges with current. In its simplest form, the POS consists of a plasma injected between the electrodes of this line. More complex designs incorporate external applied magnetic fields.[2] During the conduction phase of the POS, the plasma conducts all of the generator current, allowing none of it to pass to the load. During the opening phase of the POS, an increasing fraction of the generator current is allowed to pass to the load. Several physical mechanisms have been proposed to be responsible for the conduction and opening phases of POS operation. These mechanisms include the magnetohydrodynamic (MHD) displacement and deformation of plasma by $J\times B$ forces,[2–6] the penetration of magnetic field by electron-magnetohydrodynamic (EMH) effects[7,8] or anomalous resistivity,[9] and electrostatic gap formation.[10–13] Any of these mechanisms may be active singly or in concert during either the conduction phase or opening phase depending on the experimental regime. This paper focuses primarily on electrostatic gap formation processes near a cathode. This mechanism is considered primarily in the context of long-conduction-time (about 1 μs) POS, although the scaling results are applicable to short-conduction-time ($\leq$100 ns) POS.

Although the work presented in this paper is motivated by the long-conduction-time POS, its main contribution to POS understanding is the extension of earlier PIC simulations, motivated by short-conduction-time POS experiments, by more than 2 orders of magnitude in density to 5×10^{15} cm^{-3}. The applicability of the present results to long-conduction-time experiments remains to be demonstrated, pending observations that confirm or disprove them. The paper describes in detail a new 2-D model for gap formation including the dynamics of a potential hill. The concept of a preformed gap at the time of POS opening is also introduced that may explain experimental observations concerning inferred gap sizes. And finally, density and magnetic field scaling for the preformed gap size and the maximum potential associated with the potential hill are derived that can be compared with experimental observations. These scaling results agree with the scalings of the earlier, lower density simulations. What follows in this section of the paper is (1) a review of long-conduction-time POS investigations which have motivated this work, (2) a discussion of terminology which will be used in the paper, and (3) a discussion of the proposed model and results from the simulations based on it that contrasts them with previous models in order to highlight significant differences. These differences must be compared with future experiments to determine the validity of these models.

Recent results on the HAWK generator[14] at the Naval Research Laboratory (NRL) suggest that the POS conduction phase in these experiments is dominated by MHD (i.e., magnetic pressure or $J\times B$) effects.[4–6] The evidence supporting this conclusion consists of experiments investigating the scaling of the conduction current with plasma density at various cathode radii, plasma lengths, and current rise times.[5,6,15]

[a]Science Applications International Corporation, McLean, Virginia.
[b]Jaycor, Inc., Vienna, Virginia.

 1070-664X/95/2(1)/299/11/$6.00

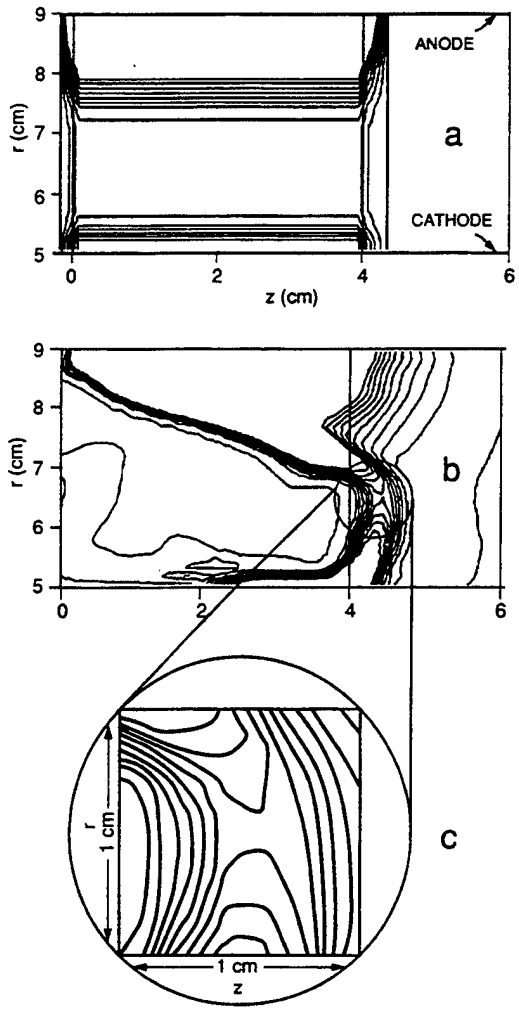


FIG. 1. Plasma density contours at the beginning (a) and end (b) of the conduction phase of a long-conduction time POS. A blowup of the density saddle is shown in (c). Typical experimental dimensions are shown. The generator is on the left, the load is on the right.

The experimental conduction current can be predicted with simple scaling laws based on center of mass motion under the action of magnetic pressure forces. Typical long-conduction-time POS experiments on HAWK have plasma densities of several times 10^{15} cm^{-3} with currents rising to 0.7 MA in 1 μs. Simulations of the POS using the two-fluid code ANTHEM[16] demonstrate the deformation of plasma by $J \times B$ forces and suggest the following picture of the conduction phase. Shown in Figs. 1(a) and 1(b) are density contours of the plasma at the beginning and end of the conduction phase. During the conduction phase, a current channel develops along the generator edge of the plasma. As magnetic pressure displaces and distorts the original plasma column, this current channel moves toward the load. The current channel front acts like a piston displacing the plasma. Both radial and axial displacement of the plasma occurs because of the radial variation of magnetic field resulting from cylindrical geometry and because of radial density gradients in the initial plasma fill.[5,6] The magnetic pressure is largest near the cathode where the plasma density is also largest because of stagnation at the surface. The current channel can also penetrate into the plasma by resistive diffusion and EMH effects, but for the parameters of interest here, MHD displacement dominates.[15]

At roughly the point where the current front reaches the load edge of the original plasma fill region, the POS is observed to open.[6] Interferometric measurements along an axial line-of-sight confirm that there is a rapid drop in the axial line-integrated density at a certain radial location just as the POS opens.[5] This suggests that the plasma has been thinned out along a large fraction of that line of sight, if not along all of it. The radial location of this thinned density region is determined by the radial gradients in both the magnetic field and density and lies somewhere in between the two electrodes. Interferometric measurements also show that on either side of the thinned region, the plasma density rises as the line of sight moves radially in either direction toward the electrodes.[5] This observation suggests that gap opening in the long-conduction-time POS does not occur at the cathode as may have been the case in the short-conduction time POS.[13] Diagnostic measurements of the line integrated density in two-dimensional (2-D) snowplow models,[17] 2-D ANTHEM simulations,[5,15] and MHD simulations[18,19] show results similar to those above. ANTHEM simulations indicate a density "saddle" in the plasma at the end of the conduction phase, as shown in Fig. 1(b) and in detail in Fig. 1(c). Since the density in the POS is lowest in the saddle, this is where gap formation is assumed to occur.

While an MHD code can be used to simulate most of the conduction phase of a long-conduction-time POS, a particle-in-cell (PIC) code is better suited to simulate the last stages of the conduction phase if the underlying physics of POS behavior changes from being dominated by MHD effects to being controlled by electrostatic gap opening processes. In this paper, PIC simulations of the gap formation process are performed by examining a very small fraction of the total plasma volume [Fig. 1(c)]. Cylindrical geometry with azimuthal symmetry is assumed. The simulation plasma is meant to represent the plasma in the density saddle generated by magnetic pressure forces during the conduction phase. This plasma region thus represents the final volume of plasma carrying the generator current just before POS opening. Radial dimensions of the simulation are chosen to be at least twice the radial gap size that develops in the simulated plasma, and axial dimensions are chosen to be many times wider than the current channel width that develops in the simulated plasma. For simplicity, uniform density profiles in r and z are chosen in the simulation because the expected gap sizes are a fraction of the dimensions of the saddle region. Nonuniform density profiles will be tested in future work. The dense plasma on the cathode side of the density saddle in Fig. 1(c) is modeled as a space-charge-limited source of electrons in the simulations and will be called the "effective cathode." In the simulations a radius of 4.5 cm is chosen for the effective cathode rather than the 6.5 cm radius of the density saddle depicted in Fig. 1(b). Either choice is representative of typical HAWK experimental parameters.

The simulations show the development of a 2-D, localized, magnetized sheath in the POS plasma. It is localized

because it occupies a limited region near the effective cathode that is a small fraction of both the axial and radial dimensions of the original POS plasma fill region. It is considered magnetized for two reasons; one is that electron orbits in the sheath are bent by the magnetic field after being accelerated across some fraction of the sheath potential. The other reason is that two-fluid simulations of the POS show that the development of a localized 2-D sheath (rather than a 1-D sheath along the entire axial length of the plasma) is related to the $v \times B$ force on the electrons.[20,21]

Scaling properties of the sheath potential and size with magnetic field and density are examined here and compared with simple analytic estimates. In addition, the self-consistent process of POS gap formation is studied. The term "gap" will be used to refer to a region largely evacuated of plasma electrons and ions. A good example of a gap is a region bounded radially above by a plasma, below by an emitting cathode, and in which emitted electrons are either magnetically insulated, or their drifts are largely in the axial direction. (Radial drifts and electron current loss from the cathode to the plasma may occur.) The term "sheath" will refer to a region still populated by ions, but locally non-neutral and across which current is carried by ions and electrons that are accelerated by electrostatic forces. In a sheath, the emission of electrons and their subsequent motion are strongly influenced by the presence of the local ion space charge. Examples of sheaths are a region of bipolar Child–Langmuir flow,[22] or a region of space-charge-limited electron emission from a cathode in a uniform ion background.[23,24]

The simulations show that a gap forms in the wake of the 2-D sheath while the POS is still in the conduction phase (i.e., before current is diverted to the load). To avoid confusion with load-driven gap formation processes which may occur during POS opening, and because the gap forms during POS conduction, this gap will be referred to as the "preformed" gap. While the 2-D sheath is limited in axial length, the axial length of the preformed gap grows and becomes as long as the POS plasma fill region as the 2-D sheath migrates toward the load. The preformed gap exists along the effective cathode between the generator side of the POS fill region up to the location of the 2-D sheath. Electrons emitted in the preformed gap do not radially cross the gap because they are magnetically insulated. Some ions are radially accelerated into the preformed gap from the plasma toward the effective cathode. The presence of this sparse population of ions in the preformed gap makes it hard to precisely determine the radial location of the plasma–gap boundary, but the size (radial extent) of the preformed gap is roughly one to two times the radial size of the 2-D sheath. The size of the preformed gap scales with plasma density and magnetic field in the same way as the 2-D sheath.

The simulations described here suggest a picture that differs from previous models of gap formation.[10,13,25–27] These models were essentially one dimensional in that uniformity in the axial direction was assumed, while gap opening occurred in the radial direction. The NRL model[10,13,27] was originally applied to the short conduction time ($\leq$100 ns) POS and assumed that during the conduction phase an equilibrium Child–Langmuir bipolar sheath developed at the cathode and grew in axial length as the generator current rose. Current in the bipolar sheath was carried in the radial direction. The end of the conduction phase was determined by the point when the sheath extended along the entire axial length of the POS plasma. At the end of the conduction phase, the current carried by the sheath was larger than it could carry in equilibrium and the now axially uniform sheath grew in radial size. Subsequent modeling of this sheath growth showed that it can be very rapid at high sheath potentials.[12] If radial sheath growth proceeded beyond a certain point, the sheath size became large enough that the magnetic field could produce significant electron orbit bending within the sheath. Following the nomenclature of the present paper, the sheath becomes a gap at this point. The gap was uniform in z and grew rapidly in radial size by enhanced ion erosion as long as the current remained near the critical current.[10] In enhanced erosion, plasma ions are rapidly accelerated into the gap by the additional electron space charge in the gap produced by electron flow at the critical current. Goyer developed a variation on the NRL model by finding conditions for which the end of the conduction phase is determined by the formation of a magnetically insulated gap rather than by the bipolar sheath limit.[25] Parks *et al.* developed a model similar to the enhanced erosion model in which the ion current density is determined by energy and momentum conservation.[26] In addition to being 1-D, all of these models assumed that sheath and gap voltage are equal to the POS voltage (the voltage measured across the A–K gap in the POS region). To the extent that the POS voltage is inductively coupled to the load voltage, the load influences gap growth in these models. Finally, these models assumed that during periods of rapid gap growth the electron flow in the gap is critically insulated, i.e., that the gap size, gap voltage, and magnetic field are related by the critical, or saturated current formula.[28] In a critically insulated gap, electron orbits extend across the entire gap so vacuum electron flow fills the gap.

In contrast to 1-D models, 2-D simulations of the end of the conduction phase show that a localized 2-D sheath forms during the conduction phase. This sheath migrates axially toward the load, creating a magnetically insulated preformed gap in its wake. Previous simulations of the entire conduction phase of a short-conduction-time POS showed similar sheath and gap development, and indicated that the axial migration of the 2-D sheath is not determined by equilibrium bipolar sheath physics.[29] Those simulations differ from the present in that the densities were 2 to 3 orders of magnitude smaller and the current was not constant but rose in time. More recent simulations of Gamble I[30] short-conduction-time POS experiments[4] with electron densities of $0.5–1.75\times10^{13}$ cm^{-3}, current rises of 250 kA in 80 ns, and with a plasma of axial length 10 cm, showed the 2-D sheath migrating from the generator to the load side of the POS plasma in 20–30 ns. This migration time is roughly consistent with the conduction times of experiments run at the same densities as those simulated. These results, and results from the present simulations, show axial penetration of magnetic field is limited to the site of the potential hill and that

the full generator current is not carried at the load side of the POS early in the conduction phase of the POS. Rapid penetration of magnetic field along the cathode and current conduction to the load end of the POS are assumed in a model of the type of POS that incorporates external applied magnetic fields.[2] The resulting magnetic field pressure in the radial direction is assumed to open a gap in the POS along the entire cathode. In contrast, the present simulation and Gamble I simulation results show gap formation that is localized to the site of a potential hill which migrates relatively slowly toward the load. In addition, the gap formation mechanism is the electrostatic acceleration of ions predominantly toward the cathode, rather than the $J\times B$ acceleration of ions toward the anode.

In the present simulations the radial size of the preformed gap remains fairly uniform in z and constant while the POS conducts current. Since the preformed gap is created before current is diverted to the load, it is concluded that the radial size of the preformed gap is independent of the load. In addition, 2-D simulations show that the electrostatic potential measured in the sheath is much larger than the POS voltage measured across the A–K gap. Electron orbits in the preformed gap behind the sheath are confined close to the cathode, so that electron flow fills only a small fraction of the gap. This demonstrates that the preformed gap is strongly insulated (unsaturated), and not just critically insulated. The preformed gap exists before the beginning of the POS opening phase, and thus represents the smallest gap that can form in the POS, since the presence of a nonzero impedance load (or large inductance in the load transmission line) could result in further gap growth during the POS opening phase. A nonzero impedance load will increase the POS voltage and could produce electron flows that are more saturated than the strongly insulated flows seen in these short-circuit load simulations. This saturated flow can cause further opening of the gap by enhanced erosion.[10] Thus the final gap size, meaning the gap size after the POS is fully open, could be larger than the preformed gap. In fact, for large enough load impedances (or large downstream inductances), it may be sufficient to model gap growth from the preformed to the final gap by assuming that the gap is critically insulated and that the POS voltage and gap voltage are the same. These possibilities will be the subject of future work. In the present simulations, where the load is a low inductance short circuit, the final gap and the preformed gap are of the same size, and both are strongly insulated.

In Sec. II, the design of the simulations is described. These are the first high density (up to 5×10^{15} cm^{-3}) simulations of the POS extending previous results at lower densities ($<2\times10^{13}$ cm^{-3}). Results of the simulations are also discussed in Sec. II, including descriptions of the formation of the 2-D sheath and preformed gap. In Sec. III scaling relations, in terms of electron density and magnetic field, are derived for the magnitude of the potential in the sheath and the radial size of the sheath. The scaling relations are then compared with simulation results. Finally, conclusions drawn from this work are discussed in Sec. IV.

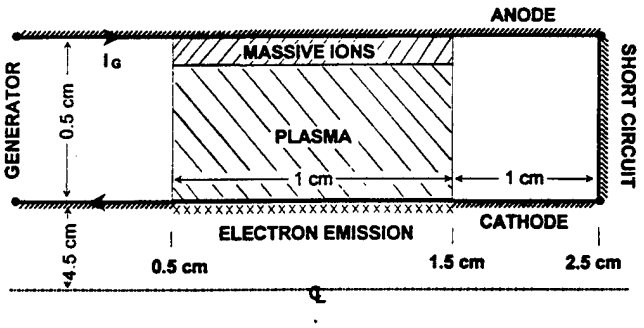


FIG. 2. Schematic diagram of the simulation region. The effective cathode radius is 4.5 cm. The plasma ions consist of C^{+4}, but a massive ion species is introduced near the anode to prevent gap opening there.

II. SIMULATION SETUP AND RESULTS

A schematic diagram of the numerical simulations is shown in Fig. 2. A section of transmission line is connected to a series LC circuit designed to mimic the primary energy source (generator) in a typical long-conduction-time POS experiment. The A–K gap of a portion of the axial length of the transmission line is filled with plasma. A carbon plasma is used, with the carbon ions doubly ionized.[27] On the load side of the POS, the transmission line is terminated with a short circuit (a conductor connects the anode and cathode of the transmission line).

Typical dimensions used in the simulations are shown in Fig. 2. Because these simulations are motivated by the long-conduction-time POS problem and only the final stages of the POS conduction phase are being modeled, the simulation is started by preloading the generator circuit and transmission line on the generator side of the POS with the full generator current. Since the focus of this paper is on the gap formation processes in the POS at the end of the conduction phase rather than on the interaction of the POS with a load, the POS-to-load distance is kept short.

As described in Sec. I, the simulation represents a small subregion of the original POS plasma volume [see Figs. 1(b) and 1(c)]. An emitting cathode boundary condition (the effective cathode) is used at the lower boundary of the simulation and models the dense plasma on the cathode side of the subregion. A modification of a conducting anode boundary condition is used at the upper boundary of the simulation. This modification consists of placing a small layer of immobile ions near the conducting anode surface. The actual anode in an experiment is far removed from the simulation region. Therefore the boundary condition at the upper surface of the simulation region should represent a continuation of the plasma rather than a conducting surface. Experiments and ANTHEM simulations show that the plasma at the anode side of the density minimum being simulated has a positive radial density gradient.[5] Simulations of such a plasma show axial penetration of the magnetic field and current into the plasma (by EMH effects) with electron currents in the radial direction. Simulations of a conducting anode on the other hand, also allow axial penetration of the magnetic field and current, but with the current in the axial direction near the anode.[31–34] This axially directed current produces a radial

TABLE I. Simulation setup and results. Rows 2–6 are input parameters, while rows 7–9 are simulation results. Input parameters shown are the density, current, magnetic field at the cathode, radial anode–cathode spacing, and axial plasma length. Results are the maximum sheath potential, radial sheath size, and an estimate of the preformed gap size. These results are dependent on the input density and magnetic field (current) only, not on the geometric parameters of rows 5 and 6. The cathode radius in all cases is 4.5 cm.

Case	1	2	3	4	5
$n_e(\times 10^{14}\ \mathrm{cm}^{-3})$	10	10	10	1	50
I (MA)	1.2	0.6	2.4	1.2	1.2
$B(\times 10^4\ \mathrm{G})$	5.33	2.66	10.66	5.33	5.33
Δ_{A-K} (cm)	0.5	0.5	1.0	2.5	0.28
l (cm)	1.0	1.5	1.0	3.0	0.5
φ (MV)	0.2	0.06	0.5	1.2	0.03
d_{2D} (cm)	0.07	0.05	0.15	0.7	0.01
d_{PF} (cm)	0.1	0.1	0.15	0.7	0.02

$J\times B$ force on the plasma that opens a gap near the anode. For this reason the behavior of a conducting anode boundary is unlike the behavior of the dense plasma on the anode side of the density minimum. To make the conducting anode boundary condition behave more like the appropriate plasma boundary, a massive ion species is introduced in the region immediately adjacent to the simulation boundary. The massive ion plasma still allows the magnetic field to penetrate, but prevents gap formation at the anode boundary because the ions do not move. The intent here is to remove the possibility of extraneous gap opening at this boundary so that it will not obscure the gap function physics on the cathode side. In addition, the upper simulation boundary is placed far enough from the effective cathode and cathode potential hill that the upper boundary condition has no effect on the processes being studied in this paper.

The results reported in this paper are obtained using the $2\frac{1}{2}$-D electromagnetic particle-in-cell (PIC) code MAGIC[35] in cylindrical geometry. A series of numerical simulations is run at various plasma densities and current levels. Simulation input parameters are shown in Table I. The anode–cathode spacing, Δ_{A-K}, and plasma length, l, are varied somewhat both to reflect the different gap sizes expected when the density and magnetic field are varied, and to keep the problem numerically tractable. However, the effective cathode radius, r_c, is kept the same and the results reported here are independent of the geometrical variations used.

To illustrate the main features of the simulations, the case of a $10^{15}\ \mathrm{cm}^{-3}$ plasma carrying 1.2 MA is used as an example. Shown in Fig. 3 is a series of snapshots in time of contours of $2\pi rB/\mu_0$ in increments of 100 kA (μ_0 is the permeability of free space). These contours represent surfaces of constant enclosed current and are referred to as current streamlines. In Figs. 3–5 the initial plasma fill occupies the region 0.5 cm$\leq z\leq$1.5 cm and the load is at z=2.5 cm. Figure 3 shows current flowing in a wedge-like pattern near the effective cathode and that the wedge migrates axially in time. The radial extent of the wedge, that is the radial distance between the effective cathode and where the current flows axially back toward the generator in the plasma above the cathode, is about 1–2 mm. This distance can be associated later with approximately twice the radial size of the electrostatic sheath that develops in the plasma (see Fig. 5). The POS begins to open at about 3 ns. At about 3.5 ns, the POS allows half the generator current, or 600 kA to reach the load. At about 4 ns, the POS is fully opened. Notice that at t=3 ns there are very few current streamlines in the gap between 4.5 cm$<r<$4.6 cm for $z<$1.3 cm. This indicates that electrons emitted in the gap are well insulated before the POS opens. After opening, the streamlines also show little or no vacuum electron current launched on the load side of the POS. Because there is little or no vacuum flow launched near the POS, these simulations demonstrate that a gap can form in the plasma that is significantly larger than the minimum necessary to magnetically insulate the electron flow before and after opening. Here, strongly insulated flow after opening is due to the proximity of the short circuit load to the POS.

In Fig. 4, a series of particle snapshots shows the positions of the carbon ions at various times during the creation of the preformed gap. The snapshots show the removal of ions from a region of plasma near the effective cathode. Electron particle snapshots are very similar to the ion plots and are not shown. For ease of viewing only one in ten of the simulation ions is plotted in Fig. 4. The development of a preformed gap (or region of low ion density) along the effective cathode in the POS region is seen. Ions are accelerated across the preformed gap from the plasma to the effective cathode. Referring back to Fig. 3, little or no electron current exists in the gap on the generator side of the 2-D sheath. Electrons emitted into the preformed gap from the effective cathode are thus strongly magnetically insulated until they reach the region of large radial current density (see Fig. 3) at the load side of the gap. In order to understand the process that creates this preformed gap, the potentials that develop in the plasma will be examined.

Figure 5 shows a series of time snapshots of the electrostatic potential contours that develop in the POS plasma just before POS opening. The minimum contour level is 75 kV with increments of 25 kV between contour levels. The development of a localized 2-D sheath near the cathode is seen. These contours should be compared with the current streamlines of Fig. 3. Electrostatic potential contours are shown because the voltages that develop in the plasma during the

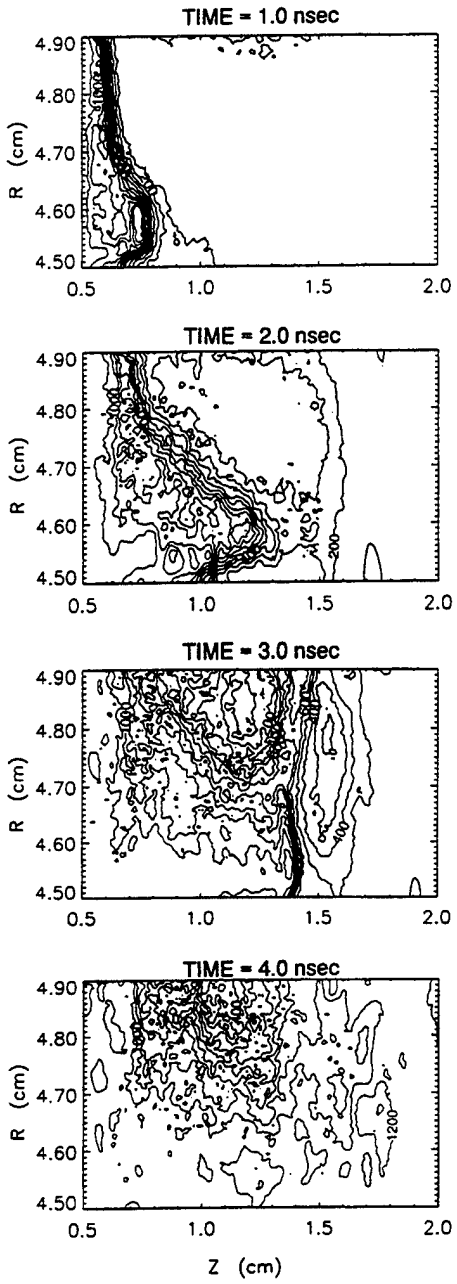


FIG. 3. Current streamline contours. The minimum contour level is 100 kA and the increment between contour levels is 100 kA. The initial plasma fill is between 0.5 cm$\leq z \leq$1.5 cm. The figure shows the penetration of a wedge-like current structure near the effective cathode.

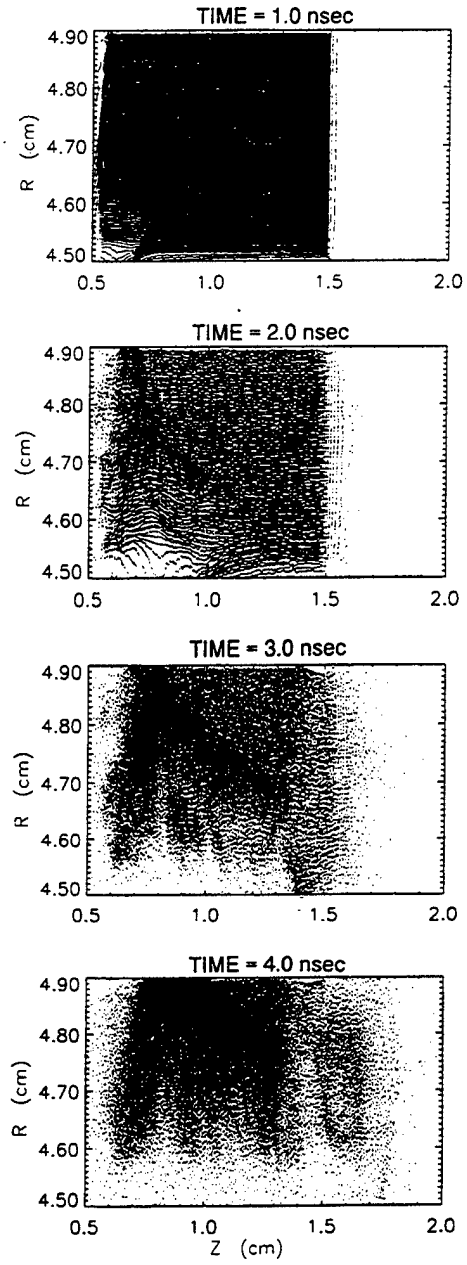


FIG. 4. Ion particle positions at various times. Only one in ten of the total number of simulation ions is shown. The opening of a gap and the progressive increase in its axial length is seen.

conduction phase are predominantly electrostatic in nature—the inductive contribution is negligible. Notice that large potentials and potential gradients develop near the main electron emission site at the effective cathode where the large current densities exist (see Fig. 3). The POS voltage (the A–K voltage in the POS region) is very close to zero while the POS is in the conduction phase ($t \leq 30$ ns). In contrast, the sheath floats to potentials up to about 200 kV. These potentials are positive, indicating the presence of non-neutralized ion space charge. Because the electron emission site and the region of ion space charge are non-neutral and because of its limited axial and radial extent, this phenomenon is collectively referred to as the 2-D sheath. Just on the generator side of the main emission site, the potential contours form closed concentric loops that indicate a local maximum. This portion of the 2-D sheath has been referred to as the potential hill.[11] Near the potential hill, large electrostatic electric fields draw electrons from the effective cathode, and also accelerate ions out of the hill, mostly toward the effective cathode. The electrostatic fields also allow current-carrying electrons to $E \times B$ drift around the hill after they are accelerated radially across a fraction of the 2-D sheath. As ions are accelerated out of the hill, a region evacuated of plasma is left behind (see Fig. 4). The potential hill migrates

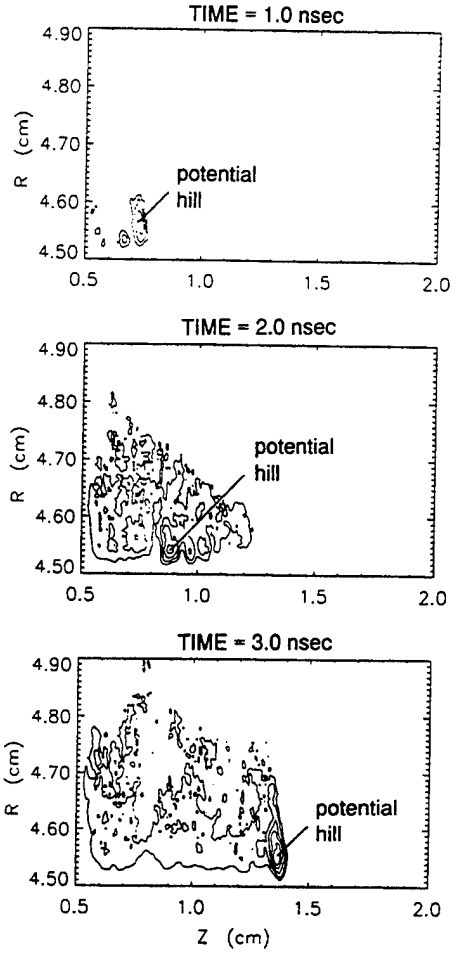


FIG. 5. Electrostatic potential contours before POS opening. These isopotential contours are shown in increments of 25 kV between contours with minimum contour level of 75 kV. Comparing with Fig. 3, the development of a potential hill immediately on the generator side of the main emission site is seen. The potential reaches a maximum of 200 kV at $t=3$ ns.

toward the load as new ions are uncovered on the load side to replace those accelerated out of the potential hill on the generator side. In this way, both the potential hill region and emission site migrate toward the load, leaving a preformed gap in their wake. For the cases examined in this paper, there is a rough correlation between the gap size, defined as the radial extent of the evacuated region, and the sheath size d_{2D}, defined as the radial distance between the effective cathode and the sheath potential maximum. By examining Figs. 4 and 5 before opening ($t \leq 3.0$ ns) it is found that, for this particular case, the potential maximum is about 200 kV, the sheath size is about 0.7 mm, and the gap size is roughly 1 mm. Radial electric fields near the potential hill are close to 2–4 MV/cm and current densities of emitted electrons of about 0.8 MA/cm^2 are seen.

The preformed gap size and potential hill are independent of the load because they form before the POS has any information about the load. As noted before, as the POS opens, further gap growth by enhanced erosion is possible especially in the case of nonzero impedance or large inductance loads. However, in this paper only the case of a short-circuit load is considered and, on the time scales considered here (few ns), no further gap growth is seen as the current is diverted to the load. Another consequence of the short-circuit load is that very little or no vacuum electron flow is launched downstream of the POS after it is opened. The effect of nonzero load impedance on further gap opening as well as power flow to the load[36] will be the subject of future work.

The transition from the end of the POS conduction phase to the beginning of the POS opening phase in these simulations occurs when the 2-D sheath reaches the load end of the POS and is accompanied by a rapid rise in the load current. The opening transition can be best understood in terms of the rapid removal of the remaining ions out of the preformed gap. When the potential hill, or 2-D sheath reaches the load end of the POS, a block of ions still fills the radial extent of the gap at the load end. This block is small in axial length compared with the axial length of the POS. Initially there is enough ion space charge to allow electrons to flow across the gap so that no current reaches the load. As these ions are removed by the large electrostatic forces in the potential hill, insulation of the electron flow begins. Eventually all these ions are removed and the electron flow is magnetically insulated in the gap.

In the simulation described above, the size of the preformed gap and the maximum potential that develops in the 2-D sheath was found at one particular POS plasma density and current. To determine the scaling of the 2-D sheath potential, the sheath size, and the preformed gap size with density and current, a series of numerical simulations were run with different plasma densities and currents at fixed effective cathode radius. At the current level of 1.2 MA, three different electron densities, 10^{14}, 10^{15}, and 5×10^{15} cm^{-3}, were used for the initial plasma fill. For an initial plasma with electron density of 10^{15} cm^{-3}, three different current levels were simulated, 600 kA, 1.2 MA, and 2.4 MA. The results of these simulations are shown in Table I. In each case, the 2-D sheath size that forms during the gap opening process is recorded, together with the maximum potential that develops in the sheath structure near the effective cathode. A rough estimate of the preformed gap size, d_{PF}, that develops in the plasma is also shown. The 2-D sheath size is much easier to identify than the preformed gap size because the location of the sheath potential maximum is well defined (see Fig. 5) while the boundary between the preformed gap and the plasma above it can only be approximated (see Fig. 4). Scaling of the sheath size and sheath potential with density and magnetic field are compared to simple analytic estimates in the next section.

III. SCALING RELATIONS

Attempts to understand the fundamental physical mechanisms underlying the formation of potential hills and some of their properties have been made in earlier work.[11,20,21,37] Some qualitative similarity exists between the potential hills seen here and the 1-D potential structures found in simulations of plasma-filled diodes (PFD),[38–41] although the magnetic field plays a more important role in the development of the 2-D potential hills. For simplicity, it is assumed in the following discussion that potential hills are magnetized

sheaths that develop near plasma/electrode interfaces. An approximate scaling law for the magnitude and radial size of these sheaths can be derived from the following 1-D analysis.

In this analysis, it is assumed that the ion background density at the electron emission site is roughly uniform. This is justified by the observation that as soon as a substantial number of ions are accelerated out of the sheath at one axial location, the sheath migrates axially into a region where ions have not yet moved. So even though ions are eventually accelerated, it is assumed that the sheath forms on a time scale short compared to the ion response time. Assuming space-charge-limited emission of electrons in a uniform, stationary background of ions, a relation between the plasma ion density, n_i, sheath potential, φ, and electron current density, j_e, can be found by a simple modification of the technique developed by Child.[42] Poisson's equation is integrated once in the relativistic limit using the boundary condition that the electric field vanishes at both ends of the sheath, yielding

$$j_e = n_i Zec\left(\frac{e\varphi/mc^2}{2+e\varphi/mc^2}\right)^{1/2}, \tag{1}$$

where Z, e, c, and m are the ion charge state, electron charge, speed of light, and electron mass, respectively. For the purposes of deriving simple scaling relations and because the relativistic correction to Eq. (1) is not large for the cases considered in this paper, the nonrelativistic limit of the equation above will be used. This relation is

$$j_e = \alpha_\varphi \frac{n_i Ze}{2}\left(\frac{2e\varphi}{m}\right)^{1/2}, \tag{2}$$

where the term α_φ is a multiplier of the formula. A value of $\alpha_\varphi<1$ will be used later because empirically, electron magnetization prevents electrons from achieving the maximum sheath potential. Notice that the term under the square root is the velocity of the electron beam accelerated in the full sheath potential, so Eq. (2) (with $\alpha_\varphi=1$) implies that the beam electron density after acceleration is half the original plasma electron density, $n_e=n_iZ$. In the nonrelativistic limit, Poisson's equation can be integrated once more to yield the following modification of the Child–Langmuir formula:[23,24]

$$j_e = \alpha_\varphi \frac{\pi}{16}\left(\frac{2e}{m}\right)^{1/2}\frac{\varphi^{3/2}}{d^2}, \tag{3}$$

where d is the radial sheath size and the multiplier α_φ is used as it was in Eq. (2). The current density in Eq. (3) (with $\alpha_\varphi=1$) is $9\pi^2/16$ times larger than the unipolar and about three times larger than the bipolar Child–Langmuir formulas. Note that equating Eqs. (2) and (3) and solving for d yields an expression for the sheath size that scales like the Debye length with an effective temperature given by $e\varphi$.

If it is assumed that the beam electrons carry all of the current and that the axial width of the current channel scales as the collisionless skin depth associated with the beam density, then

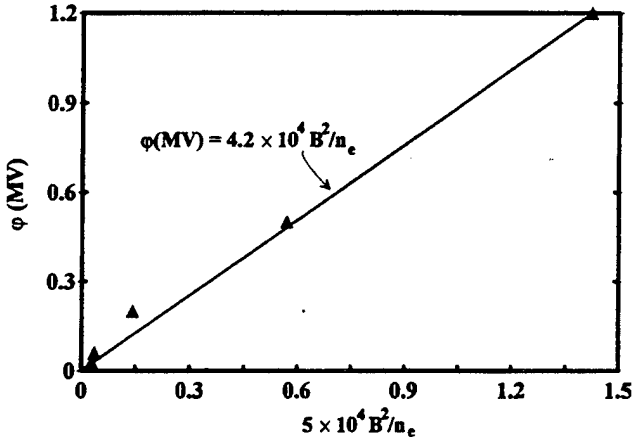


FIG. 6. The scaling of the maximum sheath potential as a function of the normalized parameter $5\times10^4\ B^2/n_e$ is shown. This parameter is the right-hand side of Eq. (6) with multiplication factors set=1.0. Simulation results of the potential are plotted, together with a straight-line fit of the data.

$$j_e \approx \frac{I}{2\pi r_c(\alpha_\omega c/\omega_{p_b})} = \left(\frac{n_b}{\pi m}\right)^{1/2}\frac{eB}{2\alpha_\omega}, \tag{4}$$

where B is the magnetic field at the effective cathode, n_b is the beam density of the electrons in the current channel, ω_{p_b} is the plasma frequency associated with the beam, and α_ω is a multiplier of the collisionless skin depth. A value of $\alpha_\omega>1$ will be used because current channel widths seen in simulations are typically a few collisionless skin depths. Eliminating j_e from Eqs. (2) and (4), and assuming $n_b\approx Zn_i/2$ [consistent with Eq. (2)], we find in Gaussian units that,

$$\varphi \approx \frac{B^2}{8\pi n_b e\alpha_\omega^2\alpha_\varphi^2}, \tag{5}$$

or in mixed units, B (G), φ (MV), n_b (cm^{-3}),

$$\varphi(\mathrm{MV}) \approx \frac{5.0\times10^4}{\alpha_\omega^2\alpha_\varphi^2}\frac{B^2}{n_e}. \tag{6}$$

The last approximation is obtained by again assuming that the plasma electron density, $n_e\approx2n_b$.

An estimate of the radial size of the 2-D sheath, d_{2D}, can be made by assuming that the 2-D sheath scales as a multiplicative factor, α_D, times the 1-D sheath size from Eq. (3). A value of $\alpha_D>1$ will be necessary to obtain agreement with simulation results. Combining Eqs. (2), (3), and (5) to find the 1-D sheath size in terms of B and n_e, the following is obtained:

$$d_{2D}\approx\alpha_D d = 3.7\times10^8\frac{\alpha_D}{\alpha_\omega\alpha_\varphi}\frac{B}{n_e}. \tag{7}$$

The scaling relation derived from the 1-D analysis above for the voltage and sheath size are now compared to the 2-D simulation results. In Figs. 6 and 7, simulation results for the potential and sheath size are plotted, together with straight line fits of the data. An analytic expression for the straight line fit is given in each figure. The horizontal axis parameter in Fig. 6 corresponds to the right-hand side of Eq. (6) with

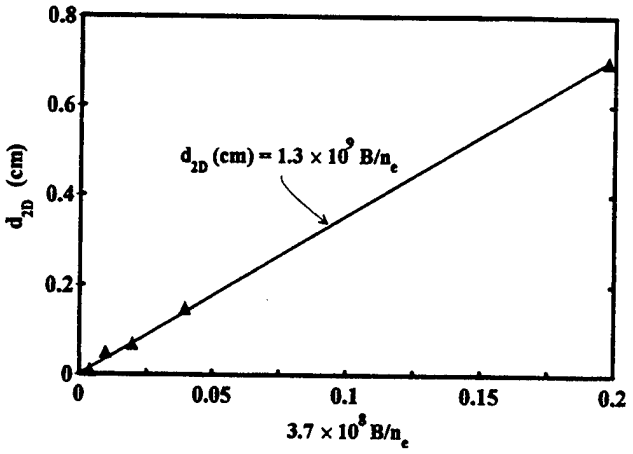


FIG. 7. The scaling of the 2-D sheath size as a function of the normalized parameter $3.7\times10^8\ B/n_e$ is shown. This parameter is the right-hand side of Eq. (7) with multiplication factors set=1.0. Simulation results of the sheath size are plotted, together with a straight-line fit of the data.

the multiplication factors set to 1. The horizontal axis parameter in Fig. 7 corresponds to the right-hand side of Eq. (7) with the multiplication factors set to 1. Figure 6 demonstrates that the simulation data fits the scaling relations of Eq. (6) with the product $\alpha_\omega\alpha_\varphi\approx1$. Checking with simulations, it is found that values of $\alpha_\varphi\approx0.3$, and $\alpha_\omega\approx3$ are typical. A value of $\alpha_\varphi<1$ may be a manifestation of the magnetization effect on the electrons since electron orbit bending in the magnetic field prevents them from penetrating to the core (maximum potential) of the potential hill region. (If they could reach the core, they would neutralize the ion space charge and the potential hill would cease to exist.) Assuming $\alpha_\omega\alpha_\varphi\approx1$, Fig. 7 demonstrates that the good agreement with simulation can be found with $\alpha_D\approx3.5$. Earlier MASK PIC code[43] simulations of the POS[29] in the lower density regimes of $n_e=4\times10^{12}$–10^{13} cm^{-3} and with magnetic fields comparable to those used here obtained results that agree with the general scaling laws shown in Figs. 6 and 7 to within 20%.

It should be noted that the simple analytic model described above is not unique. Arguments based on the assumption that the electron current density is related to electron $E\times B$ drift can be used to derive the same scaling for the potential. Another method based on magnetic and electric pressure balance[44] can also be used to arrive at the scaling, $\varphi\propto B^2/n$. These analyses, together with the model presented here, provide some confidence that basic plasma physics principles apply, but do not illuminate the fundamental physical origin of potential hills.

IV. CONCLUSIONS AND DISCUSSION

In a series of simulations of a POS coupled to a low inductance transmission line terminated with a short-circuit load, the formation of a gap in the plasma near the effective cathode is observed. This gap is associated with the development and migration of a localized 2-D sheath in the plasma during the POS conduction phase. The process of sheath formation and ion evacuation out of the sheath occur before information about the load or downstream transmission line is available to the POS plasma, so that development of this preformed gap in the POS is independent of the load. The simulation results also differ from previous models of sheath and gap formation in that the sheath potential can be much larger than the POS voltage and that the preformed gap that forms is strongly insulated during the conduction phase. This suggests that the description of the initial gap formation process in previous models is not appropriate. Previously it was assumed that the gap forms during the POS opening process effectively starting with zero radial size at the end of the conduction phase. These simulations suggest that a preformed gap has formed at the end of the conduction phase, so that during POS opening a gap of nonzero size already exists. Because of the voltage that a load imposes on the POS, further gap growth can occur by the interaction of the POS with a nonzero impedance load if the electron flow in the gap is not strongly insulated.

Scaling of the sheath and preformed gap with density and magnetic field shows that the maximum potential that develops in the sheath scales as φ (MV)$\approx4\times10^4\ B^2/n_e$, while the sheath size scales as d_{2D} (cm)$\approx10^9B/n_e$, where B is in G and n_e is measured in cm^{-3}. The preformed gap size is up to a factor of two larger than the sheath size and scales similarly. These simulations show that when the POS is closely coupled to a short-circuit load, the vacuum electron flow launched by the fully opened POS is negligible and that the final gap is significantly larger than the critically insulated gap. The final gap is the same size as the preformed gap because of the proximity of the short-circuit load to the POS. The preformed gap represents the minimum gap size of the opening POS since POS/load interactions can drive this gap to larger sizes during opening by processes such as enhanced erosion.

The results presented here extend previous short-conduction-time POS simulations in the $n_e=10^{12}$–2×10^{13} cm^{-3} density regime to the higher density regimes of 10^{14}–5×10^{15} cm^{-3} expected in long-conduction-time experiments. In the short-conduction-time POS simulations, the axial length of the POS plasma was 10 cm, the current rises to 250 kA in 80 ns, and the 2-D sheath structure migrated axially through the plasma on a time scale of 20–30 ns. These low-density simulations also showed the development of a 2-D sheath and a preformed gap. Scaling of the gap size and sheath potentials with magnetic field and density was the same as that reported here.

For long-conduction-time POS experiments, a conceptual model is presented describing the transition of POS dynamics from the MHD-dominated conduction phase to the erosion-dominated opening phase. The simulations described here are intended to model the end of the conduction phase after MHD redistribution of the plasma. An assumption is made in this paper that the dense plasma at the cathode side of the density saddle in Fig. 1(c) can be modeled as a cathode (the effective cathode). More work is needed to verify this assumption. In future work the final stages of the POS conduction phase will be simulated using a density profile like that shown in Fig. 1(c), with a smaller density in the middle of the POS and much larger densities near the electrodes. These simulations could reveal problems with the ef-

fective cathode assumption. For example, gap opening may always occur at the actual cathode, not in the radial middle of the POS. Experimental observations seem to contradict this, and the scaling results of this paper suggest that a gap at the actual experimental cathode would be too small to hold off much voltage because of the large plasma density there.

In experiments, an inferred gap size[6] can be defined from the POS voltage and generator current at the time of peak load power by using a critical,[10] or saturated[28] current formula relating these three parameters. However, when the electron flow in the real gap is well insulated (unsaturated), the inferred gap size is related more closely to the electron gyro radius, or the electron sheath size within the gap, than the gap size itself.[45] Experimental studies on HAWK have been made of the parametric dependence of the inferred gap size on the load impedance.[6,46] In the switch-limited regime,[6,46] where the POS voltage is limited by the POS rather than the load, the inferred gap size is weakly dependent on the load impedance. In the load-limited regime,[6,46] where the POS voltage is limited by the load impedance, the inferred gap size increases from a minimum size as the load impedance, Z_L, is increased from 0 Ω. As the load impedance increases further, the value of the inferred gap reaches a peak at the transition between the load and switch-limited regimes. For the POS parameters studied on HAWK (e.g., r_c=2.5 cm, l=8 cm, conduction current I=0.65 MA, conduction time τ=0.85 μs), the inferred gas rises from a minimum of about 0.5 mm for a short-circuit load to about 3 mm for Z_L >1.5 Ω. The nonzero minimum inferred gap at zero load voltage results from the inductive voltage drop between the POS and the load (POS to load inductance ≈25 nH). Because the inferred gap size could be interpreted either as the real gap size or as the electron gyroradius within a larger gap, the inferred gap is a lower bound on the true gap size. It is therefore possible that in the load-limited regime, the true gap size remains close to its switch-limited value (i.e., it does not fall from 3 to 0.5 mm) as the load impedance decreases. This would suggest that the true gap size is only weakly dependent on the load impedance in the load-limited regime. The simulation results presented here provide one possible explanation of this interpretation by suggesting a mechanism for the formation of a gap that is independent of the load in the limit of low load impedances.

Experimental values of the inferred gap[6,46] described above depend on the load impedance but seem insensitive to the effective density[6] during the POS conduction phase or to the corresponding conduction current. This is true despite a variation in the effective density from $n_e=10^{15}$ to 9×10^{15} cm^{-3} (this corresponds to a variation in the conduction current from about 0.4 to 0.8 MA). During the conduction phase, the plasma is redistributed by $J\times B$ forces so that the plasma density in the region of opening at the time of opening could be considerably different from the effective density during conduction. The relationship (if any) between these two densities remains to be established. Neither of them has yet been directly measured. The axially line-averaged density throughout POS conduction and opening has been measured, but the effective axial length of the plasma during conduction is inferred from the aperture limiting the axial length of the plasma injection region. The effective axial length of the plasma at opening is unknown. In Fig. 1, the plasma length as predicted by ANTHEM when the POS is close to opening is almost an order of magnitude less than the initial plasma length. Using this, and assuming that the line-averaged density at opening drops roughly 2 orders of magnitude from its value during conduction,[5,6,15] it can be estimated that the density at opening is on the order of $n_c=10^{14}$–10^{15} cm^{-3}. For assumed densities at opening in the range $n_e=10^{14}$–10^{15} cm^{-3}, and with corresponding conduction currents ranging from 0.4–0.8 MA, the scaling relation of Fig. 7 (assuming the gap size=d_{2D} and that r_c=5 cm) predicts preformed gaps ranging from 2.0–0.4 mm, respectively. This range of gap sizes gives an estimated range of preformed gaps that might be predicted by simulations; they should not be interpreted to suggest a gap size dependence on conduction density or conduction current. Such a dependence would only hold if the density at opening were related to the effective density during conduction. If the simple estimates made above are valid, the predicted results from the scaling relation of Fig. 7 are reasonably close to the inferred gaps from experiments (d =0.5 mm at Z_L=0 Ω, d=3 mm for Z_L>1.5 Ω).

Future work will study the relation between the effective density during conduction and the density at opening, the nature of the electron flow launched from the POS with more realistic load geometries, the effect of nonuniform density profiles and nonzero impedance loads on gap formation, and the validity of the effective cathode assumption. The scaling of the preformed gap with B and n_e found here suggests that improvements in POS performance (i.e., larger gaps) would be obtained either with greater thinning of the plasma at the end of the conduction phase and/or with higher magnetic fields. It remains to be seen whether these parameters can be controlled independently in actual experiments, where experiments at higher magnetic fields (currents) also require higher injected plasma densities to obtain the appropriate conduction time.

ACKNOWLEDGMENTS

The assistance of Larry Ludeking at Mission Research Corporation with MAGIC is gratefully acknowledged.

Simulation results in this work were obtained using the Mission Research Corporation MAGIC code under the AFOSR-sponsored MAGIC User's Group. This work was supported by the Defense Nuclear Agency.

[1] A good overview is provided in the Special Issue on Fast Opening Vacuum Switches, IEEE Trans. Plasma Sci. **PS-15**, 629 (1987).
[2] C. W. Mendel, Jr., M. E. Savage, D. M. Zagar, W. W. Simpson, T. W. Grasser, and J. P. Quintenz, J. Appl. Phys. **71**, 3731 (1992).
[3] W. Rix, D. Parks, J. Shannon, J. Thompson, and E. Waisman, IEEE Trans. Plasma Sci. **PS-19**, 400 (1991).
[4] B. V. Weber, R. J. Commisso, P. J. Goodrich, J. M. Grossmann, D. D. Hinshelwood, J. C. Kellogg, and P. F. Ottinger, IEEE Trans. Plasma Sci. **PS-19**, 757 (1991).
[5] D. D. Hinshelwood, B. V. Weber, J. M. Grossmann, and R. J. Commisso, Phys. Rev. Lett. **68**, 3567 (1992).
[6] R. J. Commisso, P. J. Goodrich, J. M. Grossmann, D. D. Hinshelwood, P. F. Ottinger, and B. V. Weber, Phys. Fluids B **4**, 2368 (1992).
[7] A. S. Kingsep, Yu. V. Mokhov, and K. V. Chukbar, Sov. J. Plasma Phys. **10**, 495 (1984).
[8] A. Fruchtman and Y. Maron, Phys. Fluids B **3**, 1546 (1991).

[9]R. M. Kulsrud, P. F. Ottinger, and J. M. Grossmann, Phys. Fluids **31**, 1741 (1988).
[10]P. F. Ottinger, S. A. Goldstein, and R. A. Meger, J. Appl. Phys. **56**, 774 (1984).
[11]J. M. Grossmann, P. F. Ottinger, D. Mosher, and R. J. Mason, Bull. Am. Phys. Soc. **35**, 2055 (1990).
[12]J. M. Grossmann, S. B. Swanekamp, and P. F. Ottinger, Phys. Fluids B **4**, 44 (1992).
[13]B. V. Weber, R. J. Commisso, R. A. Meger, J. M. Neri, W. F. Oliphant, and P. F. Ottinger, Appl. Phys. Lett. **45**, 1043 (1984).
[14]P. J. Goodrich, J. R. Boller, R. J. Commisso, D. D. Hinshelwood, J. C. Kellogg, and B. V. Weber, *8th IEEE International Pulsed Power Conference*, San Diego, CA, edited by R. White and K. Prestwich (Institute of Electrical and Electronic Engineers, Piscataway, NJ, 1991), p. 515, IEEE Catalog number: 91CH3052-8.
[15]B. V. Weber, J. R. Boller, R. J. Commisso, P. J. Goodrich, J. M. Grossmann, D. D. Hinshelwood, J. C. Kellogg, P. F. Ottinger, and G. Cooperstein, *9th International Conference on High Power Beams*, Washington, D.C., May 25–29 1992, edited by D. Mosher and G. Cooperstein, p. 375. See National Technical Information Service Document No. PB92-206168. Copies may be ordered from the National Technical Information Service, Springfield, Virginia 22161. All orders must be prepaid.
[16]R. J. Mason, M. E. Jones, J. M. Grossmann, and P. F. Ottinger, Phys. Rev. Lett. **61**, 1835 (1988).
[17]R. Ingermanson, D. Parks, and E. Waisman, *9th IEEE International Conference on Pulsed Power*, Albuquerque, NM, 1993, June 21–23, edited by K. Prestwich and W. Baker (Institute of Electrical and Electronic Engineers, Piscataway, NJ, 1993), p. 165.
[18]J. D. Huba, J. M. Grossmann, and P. F. Ottinger, Phys. Plasmas **1**, 3444 (1994).
[19]J. J. Watrous and M. Frese (private communication, 1993).
[20]B. V. Oliver, L. I. Rudakov, R. J. Mason, and P. L. Auer, Phys. Fluids B **4**, 294 (1992).
[21]R. J. Mason, P. L. Auer, R. N. Sudan, B. V. Oliver, C. E. Seyler, and J. B. Greenly, Phys. Fluids B **5**, 1115 (1993).
[22]I. Langmuir and K. B. Blodgett, Phys. Rev. **22**, 347 (1923); I. Langmuir, Phys. Rev. **2**, 450 (1913).
[23]T. A. Voronchev, Elektronnaya Tekhnika, Ser. 3, Gazorazryadnye Prib. **3**, 36 (1971).
[24]V. I. Koshelev and N. A. Ratakhin, Sov. Phys. Tech. Phys. **23**, 84 (1978).
[25]J. Goyer, IEEE Trans. Plasma Sci. **PS-19**, 920 (1991).
[26]D. Parks, E. Waisman, R. Ingermanson, and E. Salberta, in Ref. 15, p. 553. See National Technical Information Service Document No. PB92-206168. Copies may be ordered from the National Technical Information Service, Springfield, Virginia 22161. All orders must be prepaid.
[27]B. V. Weber, R. J. Commisso, G. Cooperstein, J. M. Grossmann, D. D. Hinshelwood, D. Mosher, J. M. Neri, P. F. Ottinger, and S. J. Stephanakis, IEEE Trans. Plasma Sci. **PS-15**, 635 (1987).
[28]J. M. Creedon, J. Appl. Phys. **46**, 2946 (1975).
[29]J. M. Grossmann, R. Kulsrud, J. M. Neri, and P. F. Ottinger, J. Appl. Phys. **64**, 6646 (1988).
[30]G. Cooperstein, J. J. Condon, and J. R. Boller, J. Vac. Technol. **10**, 961 (1973).
[31]E. Waisman, P. G. Steen, D. E. Parks, and A. Wilson, Appl. Phys. Lett. **46**, 1045 (1985).
[32]J. M. Grossmann, P. F. Ottinger, J. M. Neri, and A. T. Drobot, Phys. Fluids **29**, 2724 (1986).
[33]R. J. Mason, J. M. Wallace, J. M. Grossmann, and P. F. Ottinger, IEEE Trans. Plasma Sci. **PS-15**, 715 (1987).
[34]P. L. Auer and R. J. Mason, Phys. Plasmas **1**, 481 (1994).
[35]B. Goplen, L. Ludeking, D. Smithe, and G. Warren, Mission Research Corporation (private communication).
[36]S. G. Swanekamp, J. M. Grossmann, P. F. Ottinger, R. J. Commisso, and J. R. Goyer, J. Appl. Phys. **76**, 2648 (1994).
[37]J. M. Grossmann, P. F. Ottinger, and R. J. Mason, J. Appl. Phys. **66**, 2307 (1989).
[38]R. J. Kares, J. Appl. Phys. **71**, 2155 (1992).
[39]R. J. Kares, J. L. Geary, and J. M. Grossmann, J. Appl. Phys. **71**, 2168 (1992).
[40]S. B. Swanekamp, J. M. Grossmann, P. F. Ottinger, and J. L. Geary, Phys. Fluids B **4**, 3608 (1992).
[41]S. B. Swanekamp, S. J. Stephanakis, J. M. Grossmann, B. V. Weber, J. C. Kellogg, P. F. Ottinger, and G. Cooperstein, J. Appl. Phys. **74**, 2274 (1993).
[42]C. D. Child, Phys. Rev. **32**, 492 (1928).
[43]A. T. Drobot, Bull. Am. Phys. Soc. **29**, 1379 (1984).
[44]C. W. Mendel, Jr. and J. P. Quintenz, Comments Plasma Phys. Controlled Fusion **8**, 43 (1983).
[45]R. V. Lovelace and E. Ott, Phys. Fluids **17**, 1263 (1974).
[46]P. J. Goodrich, R. C. Fisher, D. D. Hinshelwood, J. R. Boller, R. J. Commisso, and B. V. Weber, in Ref. 15, p. 609. See National Technical Information Service Document No. PB92-206168. Copies may be ordered from the National Technical Information Service, Springfield, Virginia 22161. All orders must be prepaid.

Appendix BB

Particle-In-Cell Simulations of Fast Magnetic Field Penetration into Plasmas Due to the Hall Electric Field

PARTICLE-IN-CELL SIMULATIONS OF FAST MAGNETIC FIELD PENETRATION INTO PLASMAS DUE TO THE HALL ELECTRIC FIELD

S.B. Swanekamp[#], J.M. Grossmann, A. Fruchtman[*], B.V. Oliver[+], and P.F. Ottinger
Plasma Physics Division
Naval Research Laboratory
Washington, DC 20375

[#]JAYCOR, Vienna, VA 22812.
[*]SAIC, McLean, VA and Princeton Plasma Physics Laboratory, Princeton, NJ.
[+]National Research Council Research Associate at the Naval Research Laboratory.

PACS Numbers: 52., 52.30.-q, 52.65.Kj, 52.65.Rr, 52.75.Kq

Abstract

Particle-in-cell (PIC) simulations are used to study the penetration of magnetic field into plasmas in the electron-magnetohydrodynamic (EMHD) regime. These simulations represent the first definitive verification of EMHD with a PIC code. When ions are immobile, the PIC results reproduce many aspects of fluid treatments of the problem. However, the PIC results show a speed of penetration that is between 10% and 50% slower than predicted by one-dimensional fluid treatments. In addition, the PIC simulations show the formation of vortices in the electron flow behind the EMHD shock front. The size of these vortices are on the order of the collisionless electron skin depth and are closely coupled to the effects of electron inertia. An energy analysis shows that one-half the energy entering the plasma is stored as magnetic field energy while the other half is shared between internal plasma energy (thermal motion and electron vortices) and electron kinetic energy loss from the volume to the boundaries. The amount of internal plasma energy saturates after an initial transient phase so that late in time the rate that magnetic energy increases in the plasma is the same as the rate at which kinetic energy flows out through the boundaries. When ions are mobile it is observed that axial magnetic field penetration is followed by localized thinning in the ion density. The density thinning is produced by the large electrostatic fields that exist inside the electron vortices which act to reduce the space-charge imbalance necessary to support the vortices. This mechanism may play a role during the opening process of a plasma opening switch.

I. Introduction.

Magnetic field penetration into plasma is one of the most fundamental issues studied in plasma physics. Although the results of this paper are presented in the context of the plasma-opening switch (POS),[1,2] the results have a wide range of other applications including space plasmas,[3] fast Z-pinches,[4] charged particle beam propagation in plasmas,[5,6] and basic laboratory experiments[7]. Much of the early work on this subject was performed within the realm of magnetohydrodynamic (MHD) theory where ions move under the influence of J×B forces and magnetic field is convected with the ion fluid.[8] In this case, the characteristic ion speed is given by the Alfven speed, $V_A = B/(4\pi n_i m_i)^{1/2}$, where B is the magnetic field strength, n_i is the ion density, and m_i is the ion mass. If the plasma is collisional, magnetic field can penetrate due to resistive diffusion at a rate given by $V_D = c^2\eta/(4\pi L_D)$, where c is the speed of light, η is the resistivity (arising from either classical or anomalous collisions), and L_D is the characteristic diffusion length.[9] Recent theoretical studies have shown that magnetic field can penetrate into initially unmagnetized plasma on a time scale faster than either the Alfven speed or resistive diffusion.[10] This penetration results from the addition of the Hall electric field in Ohm's law in ideal MHD theory. In this case the magnetic field is convected with the electron fluid with a characteristic speed given by $V_H = cB/(4\pi n_e e L_H)$, where n_e is the electron density, e is the magnitude of the electron charge, and L_H is the characteristic scale length for Hall penetration. The ratio of the Alfven speed to the Hall speed can be written as $V_A/V_H = L_H/(c/\omega_{pi})$, where $\omega_{pi} = (4\pi Z n_i e^2/m_i)^{1/2}$ is the ion plasma frequency and c/ω_{pi} is the collisionless ion skin depth. Therefore, Hall penetration is faster than magnetic field convection due to ion motion when the characteristic Hall length is small compared with the ion collisionless skin depth. Similarly, it can be shown that Hall penetration is faster than resistive diffusion provided $\nu < \Omega_\theta L_D/L_H$ where ν is the collision frequency, $\Omega_\theta = eB_\theta/m_e c$ is the electron cyclotron frequency, and m_e is the electron rest mass.

This paper presents the first verification of Hall penetration using the particle-in-cell (PIC) method.[11] The simulations are performed using the MAGIC code developed at Mission Research Corporation.[12] Many of the earlier studies of Hall penetration were performed with a fluid approach with various approximations.[13,14,15,16] Some of the approximations that are necessary to close the fluid equations, such as the form for the pressure tensor and the equation of state, make solutions of the fluid equations questionable. In addition, much of the past work required a small amount of collisionality to remove singularities. However, even this small amount of collisionality can lead to difficulties. For example, a small amount of collisionality can lead to unphysically large plasma heating for the drive currents expected in a POS. The PIC method can provide answers to these questions since it is a collisionless, kinetic approach, that requires no assumptions regarding the pressure tensor or plasma equation of state. When there are no collisions, it would be interesting to understand how energy is partitioned between magnetic field energy and internal plasma energy.

Previous PIC simulations of gap formation in the POS have shown evidence of magnetic field penetration and vortices in the electron flow during the gap formation process.[17] However, because several competing mechanisms associated with ion motion occur simultaneously, it has been difficult to verify that this penetration is connected to the Hall effect. Vortices in the electron flow have also been predicted analytically[18,19,20,21] and observed with the two fluid code Anthem[22].

The PIC results presented in this paper show that vortices in the electron flow accompany magnetic field penetration and these vortices are a natural consequence of the electron inertia. These vortices are similar in many aspects to the vortices observed in PIC simulations of vacuum electron flow in magnetically insulated transmission lines (MITL).[23] Much of the theory developed in Sec. II applies to vacuum flow in MITLs as well.

The outline of the paper is as follows. In Sec. II a brief discussion of the fluid treatment of Hall penetration is presented as well as some of the key results from the fluid approach. In Sec. III results are presented from several PIC simulations that show magnetic field penetration into the plasma. It is shown that the observed penetration agrees very well with the fluid description in many aspects. However, the PIC simulations also show paramagnetic vortices in the electron flow and the penetration speed is slower than that predicted by the fluid theory. Section IV presents a detailed discussion of the partition of energy in the plasma between magnetic energy, thermal and directed electron energy, and energy loss to the boundaries. Because the main concern of the paper is the regime where Hall penetration dominates ($L_H << c/\omega_{pi}$), the ion mass is taken to be infinite for much of the paper. However, the role of ion motion is addressed in Sec. III by comparing simulation results for a case where ions can move with the case where ion mass is infinite.

II. Fluid theory of magnetic field penetration into plasmas due to the Hall electric field.

In this section, the theory of magnetic field penetration into a plasma is presented using a fluid approach. To get a solution to these equations it is necessary to make several questionable assumptions regarding the form of the pressure tensor, the role of displacement current, and the importance of electron inertia. As the basic equations for magnetic field penetration are developed, the key assumptions of the fluid approach are highlighted and compared with the physics contained in the PIC approach. In addition, the main results of the fluid theory are developed for comparison with the PIC simulations presented in the next section.

The geometry considered in this paper is that of the plasma opening switch shown in Fig. 1. In this geometry, a plasma of axial length l is placed between the anode and cathode of a coaxial transmission line. Current (I_G) is driven through the plasma in the form of a transverse electromagnetic (TEM) wave that is applied from the left boundary. In the analysis presented below azimuthal symmetry is assumed and only the TM mode set is retained with field components $E_r(r,z,t)$, $E_z(r,z,t)$, and $B_\theta(r,z,t)$.

The evolution of the electron fluid is governed by the continuity and momentum balance equations which are given by

$$\frac{\partial n_e}{\partial t} + \nabla \cdot n_e \mathbf{V}_e = 0 \quad , \tag{1}$$

$$\frac{\partial \gamma \mathbf{V}_e}{\partial t} + \mathbf{V}_e \cdot \nabla \gamma \mathbf{V}_e = -\frac{e}{m_e}\left(\mathbf{E} + \frac{\mathbf{V}_e \times \mathbf{B}}{c}\right) - \frac{\nabla \cdot \mathbf{P}}{n_e m_e} + \frac{e}{m_e}\eta \mathbf{J} \quad , \tag{2}$$

where $\gamma = (1 - V_e^2 / c^2)^{-1/2}$ is the relativistic mass factor, $\mathbf{V}_e$ is the electron fluid velocity, $\mathbf{P}$ is the pressure tensor, and $\mathbf{J}$ is the current density. In fluid treatments it is common to neglect the off diagonal terms in the pressure tensor and treat the pressure as a scalar. Since the PIC method allows for orbit crossings, no assumption regarding the form of the pressure tensor is needed. PIC simulations of the POS have shown that the off diagonal terms can be important in regions of the plasma where there is magnetic field penetration.[24]

When azimuthal symmetry is assumed, Eqs. (1) and (2) can be combined with Faraday's Law,

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \quad , \tag{3}$$

to give the following equation describing the evolution of the electron fluid vorticity and magnetic field,

$$n_e r \frac{D}{Dt}\left(\frac{\omega_\theta - \Omega_\theta}{n_e r}\right) = -(\nabla \frac{1}{m_e n_e} \times \nabla p)_\theta + \frac{e}{m_e}(\nabla \times \eta \mathbf{J})_\theta \quad . \tag{4}$$

In this last equation $D/Dt = \partial/\partial t + V_e \cdot \nabla$ is the convective derivative, p is the scalar pressure, $\Omega_\theta = eB_\theta/m_e c$ is the electron cyclotron frequency, and $\omega_\theta = (\nabla \times \gamma V_e)_\theta$ is the electron fluid vorticity. If the scalar pressure is a function of the density, then the first term on the right hand side of Eq. (4) is zero. If collisions are also ignored, then Eq. (4) describes the constancy of $\Gamma/n_e r$ along an electron trajectory, where $\Gamma = \omega_\theta - \Omega_\theta$ is the generalized electron vorticity. For example, the electrons in the plasma prefill are initially unmagnetized and at rest so that their initial generalized vorticity is zero. As magnetic field penetrates into the plasma these electrons seek regions of the plasma that keep their generalized vorticity zero along their trajectory.

The appearance of vorticity in solutions to the fluid equations is a direct consequence of electron inertia. In much of the previous work on Hall theory, electron inertia is ignored. This is equivalent to assuming that electrons are **E**×**B** drifting and results in the neglect of the fluid vorticity ($\omega_\theta = 0$) in Eq. (4). When electron inertia is ignored, Eq. (4) can be written as[25]

$$\frac{D}{Dt}\left(\frac{r\Omega_\theta}{n_e r^2}\right) = 0 \quad , \tag{5}$$

where the pressure and collision terms have been dropped. Equations (4) or (5) apply regardless of the details of the ion motion. Therefore, they are appropriate not only to electron flows in plasmas but also apply to vacuum electron flows in magnetically insulated transmission lines. If there is ion motion, then the electron density will evolve in a complicated manner as ions redistribute themselves. If the electron fluid remains in quasi-equilibrium, then Eq. (5) implies that the electron fluid will evolve to keep $r\Omega_\theta/n_e r^2$ constant along the electron trajectories as the ion density evolves in time.

In electron magnetohydrodynamics (EMHD) it is customary to ignore ion motion and assume that the plasma is space charge neutral ($n_e = Zn_i$). This simplifies the fluid equations because the electron density is constant and determined by the initial ion distribution. With these assumptions, an equation describing Hall penetration can be derived by expanding the convective derivative in Eq. (5) and using Ampere's Law (displacement current is neglected) to eliminate the electron fluid velocity ($\mathbf{V_e} = -\frac{c}{4\pi Z n_i e}\nabla \times \mathbf{B}$). This equation can be written as

$$\frac{\partial B_\theta}{\partial t} - \frac{c r B_\theta}{4\pi e}\left(\nabla r B_\theta \times \nabla \frac{1}{Z n_i r^2}\right)_\theta = 0 \quad . \tag{6}$$

A stable, non-linear solution to Eq. (6) predicts a shock wave for rB_θ which propagates along the $n_i r^2$ contours.[14, 26]

As a special case of Eq. (6), consider the situation where the ion density is independent of z [i.e. $n_i(r,z) = n_i(r)$]. In this case the $Zn_i r^2$ contours are parallel to the z axis and shock

propagation is expected to occur in the z direction. For this special case, Eq. (6) can be expressed as[13]

$$\frac{\partial B_\theta}{\partial t} - \frac{c\,B_\theta}{4\pi Z n_i e L_H}\frac{\partial B_\theta}{\partial z} = 0 \quad , \tag{7}$$

where

$$L_H = \left[(1/Zn_i r^2)\, d(Zn_i r^2)/dr\right]^{-1} \tag{8}$$

is the characteristic length for Hall penetration. The penetration speed of the shock front can be expressed as[13,27]

$$V_H(r) = \alpha \frac{c(-rB_\theta)}{4\pi Z n_i e r L_H} \quad , \tag{9}$$

where α is a constant on the order unity that depends on the time dependence for the applied current, $I_G(t)=crB_\theta(r,z=0,t)/2$. Two cases of practical importance include the constant applied current and the linearly rising current for which $\alpha=1/2$ and 3/8 respectively.

It is important to note that magnetic field will penetrate only in regions of plasma where the Hall speed is positive. When $V_H \le 0$, Eq. (7) predicts evanescent waves that do not propagate. Furthermore, if the plasma is initially magnetized then expulsion of magnetic field is expected in regions where $V_H \le 0$.[15,22] For a POS in negative polarity ($-rB_\theta>0$), Eq. (7) predicts magnetic field penetration in regions where $d(Zn_i r^2)/dr > 0$. Exclusion of magnetic field is expected in regions where $d(Zn_i r^2)/dr \le 0$. In positive polarity ($rB_\theta>0$) the regions of penetration and exclusion are reversed. In general it can be shown that magnetic field penetration occurs in the $\mathbf{B}\times\nabla(Zn_i r^2)$ direction.[16]

III. Simulation Results

Since Hall penetration dominates ion motion when $L_H/(c/\omega_{pi}) << 1$, ion motion is assumed negligible in the majority of this section by taking ions to be infinitely massive. The role of ion motion will be addressed later in this section by comparing the infinitely-massive-ion case with a case where the ions can move.

One of the fundamental weaknesses of the PIC method is the so called "grid" instability that produces numerical plasma heating. This instability saturates when[11]

$$\lambda_D \sim \Delta x/\pi \quad , \tag{10}$$

where Δx is the grid size, $\lambda_D=(k_B T_e/4\pi n_e e^2)^{1/2}$ is the plasma Debye length, k_B is the Boltzmann constant, and T_e is the electron temperature. Artificial plasma heating from the grid instability severely limits the plasma densities that can be accurately simulated with the PIC method. To limit the numerical heating to 10-100 eV, grid sizes of 0.005 to 0.01 cm were used with plasma densities ranging between 10^{12} cm^{-3} and 3×10^{13} cm^{-3}. For these densities the electron mass was reduced from the normal electron mass to keep the collisionless electron skin depth small compared to the dimensions of the plasma (typically a few cm). At the end of this section, it is demonstrated that the results from the reduced electron mass are equivalent to those with normal electron mass and a corresponding higher magnetic field and electron density.

The simulations presented in this section are set-up as shown in Fig. 1. Space-charge-limited emission of electrons is allowed from the entire cathode surface. The emitted electrons are given an initial velocity of 10^8 cm/s. This corresponds to an initial energy of energy of about 3 eV. Plasma is loaded between z=2 cm and z=5 cm between the cathode radius of r_c=2.5 cm and the anode radius of r_a=5.0 cm. A linearly rising current ramp with dI_G/dt=10 kA/ns is

applied on the left hand boundary. The electron mass used in these simulations was 1/10 the normal electron mass. Two slightly different density profiles were used. These profiles are shown Fig. 2a. The first profile is a parabolic density profile with the minimum density of 10^{12} cm^{-3} at r=3.75 cm rising to 3×10^{13} cm^{-3} at the electrodes. For this density profile the characteristic Hall scale length is 0.2 cm. The second profile is a uniform-parabolic profile that is uniform at 10^{12} cm^{-3} between the cathode and r=3.75 cm and then rises parabolically to 3×10^{13} cm^{-3} at the anode. The instantaneous Hall speed profiles for these two density profiles in negative polarity ($-rB_\theta>0$) are shown in Fig. 2b. Recall from the discussion in Sec. II that magnetic field penetration is expected only in regions where the Hall speed is positive. Therefore, exclusion of magnetic field is expected for $r\leq3.75$ cm in negative polarity for the parabolic profile. Magnetic field penetration is expected for $r>3.75$ cm with maximum penetration occurring near r=4 cm. In positive polarity the regions of penetration and exclusion are reversed. For the uniform-parabolic profile, magnetic field is expected to penetrate the entire plasma in negative polarity. Since the regions of penetration and exclusion are reversed in positive polarity ($rB_\theta>0$), no magnetic field penetration is expected in positive polarity for the uniform-parabolic profile since the Hall speed profile is negative everywhere.

The fluid theory of Hall penetration is tested with several PIC simulations. One of the most useful diagnostics from these simulations are the contours of $rB_\theta(r,z)$. From Ampere's Law $rB_\theta(r,z)$ is proportional to the current enclosed within a radius r at axial location z and which is expressed as $I(r,z)=2\pi rB_\theta(r,z)/\mu_0$. With this definition, the difference between the contour levels of I(r,z) represents current flow in the plasma parallel to the contours.

Current enclosed contours at t=2 ns are shown in Fig. 3 for the parabolic profile in negative polarity. As predicted, rapid magnetic field penetration takes place at a radial position where the maximum Hall speed occurs and no magnetic field penetration takes place in regions where the Hall speed is negative. Simulation results (not shown) also show that the regions of penetration and exclusion are reversed in positive polarity. However, the penetration speed predicted by the PIC method is about 50% slower than predicted by fluid theory. A one dimensional analysis of Eq. (4) indicates that electron inertia does not change the penetration speed.[28] Therefore, it is speculated that the slower penetration is caused by either strong two-dimensional effects or by the effects of electron pressure that are not included in the fluid analysis.

Figure 4 shows the current enclosed contours for the uniform-parabolic density profile in positive polarity. As predicted by theory, no magnetic field penetration into the bulk plasma deeper than a collisionless skin depth is observed. Rapid penetration of magnetic field is observed at the anode in both Figs. 3 and 4. Anode penetration occurs because the conductor boundary condition ($E_{||}=0$) causes electrons to $\mathbf{E}\times\mathbf{B}$ drift parallel to the anode.[29] Therefore, this penetration is strongly coupled to the anode boundary conditions and is different from the Hall penetration observed in the body of the plasma.

Figure 3 also shows that magnetic field penetration is accompanied by a train of vortices in the electron flow. The radius of the vortices is about 0.25 cm which is comparable to both the characteristic Hall scale length (L_H) and the collisionless skin depth (c/ω_{pe}). Magnetic flux is compressed inside each vortex producing eddy currents in the center that are larger than the drive current. A one-dimensional analysis of Eq. (4) predicts oscillations in the magnetic field behind the shock front that can also be larger than the drive field.[28] The wavelength of these oscillations is proportional to the collisionless electron skin depth.

The center of a vortex is charged positively which is produced by a reduction in the number of electrons in the center of the vortex. This space charge imbalance produces an electric field that allows the electrons to **E×B** drift in a counter-clockwise direction around the positively charged center. The detailed electron dynamics inside a vortex is beyond the scope of this paper. However, since the vortices are observed to be paramagnetic, the conservation of generalized vorticity prohibits the vortices from being comprised entirely of electrons from the plasma prefill when pressure and collision terms are neglected in Eq. (4). To see this assume the that pressure term can be neglected so that Γ=0 for the electrons in the plasma prefill. For these electrons

$$(\nabla \times \gamma V_e)_\theta = \Omega_\theta \ . \tag{11}$$

If Eq. (11) is integrated across the area of a vortex then it is possible to write

$$R\gamma(R)V_\varphi(R) = \int_0^R \Omega_\theta R' dR' \ , \tag{12}$$

where R is vortex radius, and φ is the coordinate direction around the vortex. Since Ω_θ<0 in these simulations, Eq. (12) predicts that V_φ<0. This direction of rotation is such as to produce diamagnetic vortices. However, the vortices are observed to be paramagnetic which implies that our assumption that the pressure term is negligible and the vortices contain only electrons from the initial plasma prefill is false. Therefore, either the pressure terms are important or the vortices must contain a sufficient number of emitted electrons with positive initial generalized vorticity. These issues will be investigated in a future paper.

To better understand the role of ion motion, a simulation was run with the parabolic density profile in negative polarity and an ion mass ¼ the proton mass. For this choice of parameters, c/ω_{pi}=8 cm at the radial position where the most rapid penetration of magnetic field is expected. Because $c/\omega_{pi} >> L_H$ the simulation is still in a regime where Hall penetration is expected to dominate. The current enclosed contours for this simulation at t=2 ns are shown in Fig. 5. Figure 5 demonstrates that, when ions can move, the vortices exhibit a smaller degree of paramagnetism than the infinitely massive case shown in Fig. 3. In the simulations depicted in Fig. 3, the maximum electric field in a vortex at t=2 ns is 250 kV/cm. This electric field is sufficient to move ions several mm in this time and thus significantly reduces the space charge imbalance necessary to support the vortices. Magnetic field penetration followed by the removal of ions due to the electrostatic forces has been shown to lead to gap opening in a POS.[17] The details of this process and the role of the Hall electric field will be the subject of a future paper.

As mentioned above, the size (radius) of the vortices in Fig. 3 is comparable to both the collisionless skin depth and the characteristic Hall scale length. It is of interest to understand whether the vortex size scales with the characteristic Hall scale length, L_H, or the collisionless electron skin depth, c/ω_{pe}. To examine this question, a simulation was run with a density profile for which the Hall speed is independent of r. This profile is given by

$$n(r) = \frac{n_c r_c^2}{r^2} \frac{1}{1 - f \ln(r / r_c)} \ , \tag{13}$$

where $0 < f < 1/\ln(r_a/r_c)$ is a parameter that sets the Hall speed. The penetration speed for the density distribution given by Eq. (13) can be written as

$$V_H = \alpha \frac{c(-r_c B_c)}{4\pi n_c e r_c^2} f_{-} \ , \tag{14}$$

where α is defined in the text following Eq. (9). The solid curve in Fig. 6 shows the density profile for f=1/2, r_c=1 cm, r_a=5 cm, and $n_c=5\times10^{12}$ cm^{-3}. For this choice of parameters the characteristic Hall scale length is about 2 cm while the collisionless electron skin depth is about 0.3 cm. The density profile used in the simulations is given by the dashed curve in Fig. 6 which shows a slightly modified density distribution near the cathode. This slightly altered profile was chosen to avoid problems associated with magnetic field penetration right at the cathode boundary. Since $d(Znr^2)/dr < 0$ near the cathode, field penetration is not expected there. The simulations for the density profile shown in Fig. 6 were done in negative polarity with an electron mass 1/90 the normal electron mass. The applied current rose from 0 to 5 kA in approximately 1.5 ns and then was held constant.

The current enclosed contours at t=5 ns and t=10 ns from a simulation with the density profile depicted in Fig. 6 are shown in Fig. 7. The current enclosed contours at t=5 ns (Fig. 7a) show the formation of a line of vortices along the plasma/magnetic field boundary. Each of these vortices has a diameter of about two collisionless electron skin depths. These results show that the vortex size scales with c/ω_{pe} and not L_H. The penetration speed in this simulation is only 10% slower than that predicted by fluid theory. The current enclosed contours at t=10 ns (Fig. 7 b) show that this initial line of vortices propagates into the plasma at approximately the same axial speed. As the initial line of vortices propagates into the plasma, additional lines of weaker vortices form along the plasma/magnetic field boundary. Each line of the additional lines of vortices propagates with the same axial speed as the initial line of vortices but exhibit a smaller degree of paramagnetism. Notice that the vortices drift slightly radially upward as they propagate into the plasma. This radial drift may be caused by either interactions with other vortices or by interactions with a boundary.

The simulations presented thus far have been with an electron density that is much smaller than those expected in a real POS. This density was chosen to keep the numerical plasma temperature as low as possible. In addition, the electron mass was reduced to keep the electron collisionless skin depth small compared to the size of the plasma. In the remainder of this section it is demonstrated that the simulations with reduced density, magnetic field and electron mass are essentially equivalent to simulations with normal electron mass with density and magnetic field increased by $\kappa=m_e/m$ where m is the artificially light electron mass. That this is true can be seen by multiplying Eqs. (1)-(3) and Ampere's Law by $1/\kappa$. These new equations describe the dynamics of a species with mass $m=m_e/\kappa$. The dynamics of this new species is unchanged provided the drive magnetic field and density are increased by the factor $1/\kappa$.

The relative importance of the numerical plasma heating can be estimated by comparing the electron velocity in the current channel with the thermal speed. To first order the electron speed in the current channel can be estimated from Ampere's Law to be

$$|V_e| = \frac{c}{4\pi e n_e}|\nabla\times B| \cong \frac{c|B|}{4\pi e n_e}\frac{\omega_{pe}}{c} \quad , \tag{15}$$

where it has been assumed that the width of the current channel in the plasma is c/ω_{pe}. The ratio of the electron speed in the current channel to the electron thermal speed $[V_{th}=(2k_BT/m_e)^{1/2}]$ can be written as

$$V_{th}/V_e = \left(\frac{n_e k_B T}{B^2/8\pi}\right)^{1/2} \quad . \tag{16}$$

Equation (16) shows that the relative importance of numerical plasma heating is small provided the ratio of kinetic pressure to magnetic pressure is small. It is important to note that, for a fixed grid, the numerical temperature is proportional to the density [see Eq. (10)]. Therefore, the relative importance of numerical plasma heating is unchanged if B and n_e are increased by the same factor.

To test these ideas a simulation was run with the density profile given by Eq. (13) using normal electron mass, n_c=4.5×10^{14} cm^{-3}, and I_G=450 kA. This simulation should be similar to the simulation with m=m_e/90, n_c=5×10^{12} cm^{-3}, and I_G=5 kA shown in Fig. 7. Current enclosed contours for this simulation at t=5 ns and 10 ns are shown in Fig. 8. Although there are minor differences between the simulation results depicted in Figs. 7 and 8, the size of the vortices and the speed of penetration remain unchanged. This shows that the simulations with reduced electron mass are essentially equivalent to simulations with normal electron mass with a corresponding higher density and magnetic field.

IV. Energy considerations.

It is of considerable interest to investigate the flow of energy as the magnetic field penetrates into the plasma. For the simulations described in this paper, conservation of energy over the plasma volume can be expressed as

$$E_{IN} = E_B + E_E + E_P + E_A \tag{17}$$

where $E_{IN} = \frac{c}{4\pi}\int dt \int \mathbf{E}\times\mathbf{B}\cdot\hat{n}dA$ is the energy flowing into the plasma through the generator boundary, $E_B = \int \frac{B^2}{8\pi} d^3x$ is the magnetic-field energy, $E_E = \int \frac{E^2}{8\pi} d^3x$ is the electric-field energy, $E_P = \sum_{i=1}^{N_p} (\gamma_i - 1)mc^2$ is the total internal energy of the plasma, $E_A = \sum_{i=1}^{N_A} (\gamma_i - 1)mc^2$ is the particle energy that flows out through the anode, N_p is the number of plasma particles in the simulation at time t, and N_A is the total number of particles leaving the volume through the boundaries. For the simulations described here, the energy flow out of the plasma occurs by the electrons in the current channel that flow to the anode. Terms describing the flow of electromagnetic energy out of the plasma and the flow of particle energy into the plasma from the cathode have been omitted in Eq. (17) since the simulations show that these terms are negligible.

For an applied current that is constant in time the Hall model predicts that the rate at which energy flows into the plasma is[30]

$$\frac{dE_{IN}}{dt} = \frac{c}{2}\int E_r B_\theta dr = \frac{cr_c B_c(t)}{16\pi Z e n_e}\left(1 - \frac{n_c r_c^2}{n_a r_a^2}\right) . \tag{18}$$

It can also be shown from fluid theory that half the energy that flows into the plasma goes into magnetic field energy.[31] The other half of the energy that flows into the plasma is either dissipated in the plasma or flows out of the plasma to the boundaries. In Ref. 31 it was shown that, if there is a small amount of resistivity, electrons are heated in the shock front where large magnetic field gradients exist. In this case the dissipated energy stays in the plasma and goes into electron thermal motion. For the parameters expected in a POS, this picture would lead to unrealistically large electron temperatures. Therefore, it is likely that a large fraction of the energy is convected out of the plasma to the anode by the electrons in the current channel.[32,33] In addition to plasma heating and convection to the anode, the simulation results shown in Figs. 3,

7, and 8 show that energy can also go into the electron vortex motion (i.e. the kinetic energy associated with the electron rotation around the center of a vortex).

The partition of energy for the simulation depicted in Fig. 7 is shown in Fig. 9. It is well known that PIC codes do not exactly conserve energy.[11] In this simulation great care was taken so that energy conservation was obeyed to within 10%. As Fig. 9 shows, nearly half the energy that flows into the plasma goes into magnetic field energy. The majority of the remaining energy is convected to the anode by the electrons in the current channel. However, Fig. 9 also shows that a significant amount of energy goes into increasing the internal plasma energy. The internal plasma energy appears primarily in the form of vortices. The internal plasma energy initially rises rapidly as the initial line of vortices is created but increases much more slowly once the initial vortices get established. Once the initial line of vortices is established, the rate at which energy is convected to the anode is about the same as the rate at which magnetic energy increases in the plasma. The electric field energy is not shown in Fig. 9 since the simulations show it is much smaller than both the magnetic field energy and the internal energy of the plasma. In Sec. III it was shown that large electrostatic electric fields exist inside a vortex. However, the total electrostatic energy associated with this field is small because $|\mathbf{E}|/|\mathbf{B}| \sim |\mathbf{V}_e|/c < 1$ inside a vortex and the electric field is confined to this region which occupies only in a small fraction of the entire plasma volume.

V. Conclusions

This paper has used a PIC code to demonstrate fast magnetic field penetration into a plasma associated with the Hall term in fluid theory. This is an improvement over traditional fluid treatments since the PIC method makes no *a priori* assumptions about the plasma equation of state, form of the pressure tensor, or the importance of displacement currents. In addition, fluid treatments generally require a small amount of collisionality to remove singularities in the solutions. Since the PIC technique is inherently collisionless, the PIC results described in this paper treat the collisionless transport of magnetic field into the plasma.

The simulations reproduce many aspects of magnetic field penetration that are consistent with fluid treatments. However, the speed of penetration observed in the simulations is somewhat slower than analytic estimates. This is caused by either strong two-dimensional effects or by the effects of the pressure tensor ignored in the fluid analysis. The simulations show the formation of vortices behind the EMHD shock front that are a natural consequence of electron inertia. These vortices are paramagnetic in nature and their size is a few collisionless electron skin depths.

Most of the results described in this paper are for the case of infinitely massive ions. When ions can move, axial magnetic field penetration is followed by ion motion produced by the large electrostatic force that exists in the center of a vortex. In this case, the vortices in the electron flow are much weaker since the ion motion acts to reduce the space-charge separation that is necessary to support the vortices. The reduction of ion space-charge in regions where the magnetic field penetrates may also lead to gap formation in a POS.

It is observed that half the energy flowing into the plasma goes into magnetic energy. The majority of the remaining energy is convected to the anode by the electrons in the current channel. A significant amount of energy also goes into the internal energy of the plasma while is the initial line of vortices are being established. However, once the initial line of vortices is created, very little additional energy appears as internal plasma energy and the rate at which

energy is convected to the anode is approximately the same as the rate of increase in the magnetic field energy in the plasma.

VI. Acknowledgments

The authors would like to acknowledge Dr. Joe Huba of the Naval Research Laboratory (NRL) for many interesting discussions pertaining to this problem. In addition, many thanks to Drs. Bob Commisso and Bruce Weber at NRL for their continued interest in this work. The authors also profited greatly from the careful reading and technical discussions of this manuscript provided by Mr. David Rose. The first author also benefited from discussions with Drs. Larry (Lars) Ludeking and David Smithe of Mission Research Corporation. Simulation results were obtained using the MAGIC code under the Air Force Office of Scientific Research sponsored MAGIC User's Group. This work was supported by Defense Nuclear Agency.

LIST OF FIGURES

---

[1] R.J. Commisso, P.J. Goodrich, J.M. Grossmann, D.D. Hinshelwood, P.F. Ottinger, and B.V. Weber, Phys. Fluids **B4**, 2368 (1992).

[2] See IEEE Trans. Plasma Sci. special issue on Plasma Opening Switches **PS-15**, (1987).

[3] P.A. Bernhardt, Phys. Fluids **B4**, 2249 (1992).

[4] P. Sheehey, J.E. Hammel, I.R. Lindemuth, D.W. Scudder, J.S. Schlachter, R.H. Lovberg, and R.A. Riley, Phys. Fluids **B4**, 3698 (1992).

[5] A. Fruchtman and Y. Maron, Phys. Fluids **B3**, 1546 (1991).

[6] K. Gomberoff, and A. Fruchtman, Phys. Plasmas **1**, 2480 (1994).

[7] R.L. Stenzel, J.M. Urrutia, and C.L. Rousculp, Phys. Rev. Lett. **74**, 702 (1995).

[8] M. Rosenbluth, in Plasma Physics and Thermonuclear Research, eds. C.L. Longmire, J.L. Tuck, and W.B. Thompson, (Pergamon Press, London, 1963), p.217.

[9] J.D. Jackson, *Classical Electrodynamics*, John Wiley and Sons, New York, pp. 472-473 (1975).

[10] A.S. Kingsep, K.V. Chukbar, and V.V. Yankov, in Reviews of Plasma Physics, edited by B.B. Kadomstev (Consultants Bureau, New York, 1990), Vol. 16, p. 243.

[11] C.K. Birdsall, and A.B. Langdon, *Plasma Physics via Computer Simulation*, (McGraw Hill, New York, 1985), p. 179.

[12] B. Goplen, L. Ludeking, D. Smithe, and G. Warren, Comp. Phys. Comm. **87**, 54 (1995).

[13] A.S. Kingsep, L.I. Rudakov, and K.V. Chukbar, Sov. Phys. Dokl. **27**, 140 (1982).

[14] A.S. Kingsep, Y.V. Mokhov, and K.V. Chukbar, Sov. J. Plasma Phys. **10**, 495 (1984).

[15] A. Fruchtman, Phys. Fluids **B3**, 1908 (1991).

[16] J.D. Huba, J.M. Grossmann, and P.F. Ottinger, Phys. Plasmas **1**, 3444 (1994).

[17] J.M. Grossmann, S.B. Swanekamp, R.J. Commisso, P.J. Goodrich, D.D. Hinshelwood, J.D. Huba, P.F. Ottinger, and B.V. Weber, Proceeding of the 10th International Conference on High Power Particle Beams, edited by W. Rix and R. White, San Diego, CA, June 20-24 1994, NTIS document No. PB95-14437, p. 280. Copies can be ordered from NTIS, PO Box 1425, Springfield, VA 22151.

[18] L.I. Rudakov, C.E. Seyler, and R.N. Sudan, Comments on Plasma Phys. and Controlled Fusion **14**, 171 (1991).

[19] P.M. Bellan, Phys. Fluids **B5**, 1955 (1993).

[20] M.B. Isichenko, and A.M. Marnachev, Sov. Phys. JETP **66**, 702 (1987).

[21] A. Fruchtman, and K. Gomberoff, Phys. Fluids **B5**, 2371 (1993).

[22] R.J. Mason, P.L. Auer, R.N. Sudan, B.V. Oliver, C.E. Seyler, and J.B. Greenly, Phys. Fluids **B5**, 1115 (1993).

[23] B. Church and R.N. Sudan, Phys. Plasmas **2**, 1837 (1995).

[24] J.M. Grossmann and D. Mosher private communication (1991).

[25] R. Kulsrud. P.F. Ottinger, and J.M. Grossmann, Phys. Fluids **31**, 1741 (1988).

[26] K. Gomberoff, and A. Fruchtman, Phys. Fluids **B 5**, 2841 (1993).

[27] B.V. Weber, R.J. Commisso, P.J. Goodrich, J.M. Grossmann, D.D. Hinshelwood, P.F. Ottinger, and S.B. Swanekamp, Phys. Plasmas **2**, 299 (1995).

[28] Y.L. Kalda and A.S. Kingsep, Sov. J. Plasma Phys. **15**, 508 (1989).

[29] J.M. Grossmann, S.B. Swanekamp, P.F. Ottinger, R.J. Commisso, D.D. Hinshelwood, and B.V. Weber, Phys. Plamsas **2**, 299 (1995).

[30] A. Fruchtman, Phys. Rev. A 45, 3938 (1992).

[31] A. Fruchtman and K. Gomberoff, Phys. Fluids B4, 117 (1992).

[32] A. Fruchtman and L.I. Rudakov, Phys. Rev. E 50, 2997 (1994).

[33] J.M. Grossmann, P.F. Ottinger, and R.J. Mason, J. Appl. Phys. **16**, 2307 (1989).

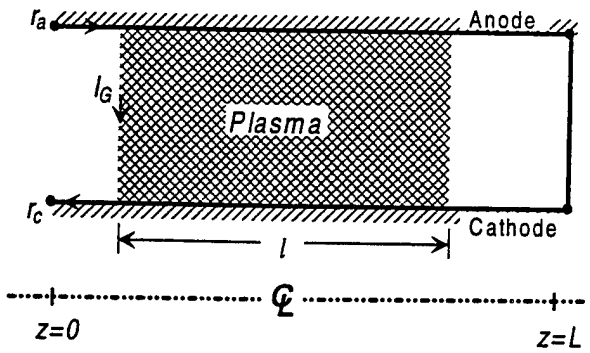
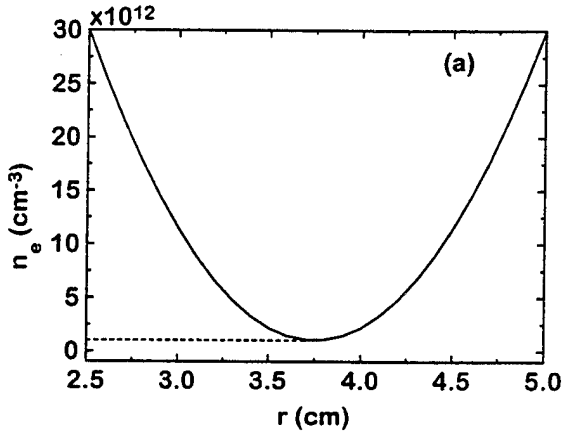


Figure 1
Swanekamp et. al



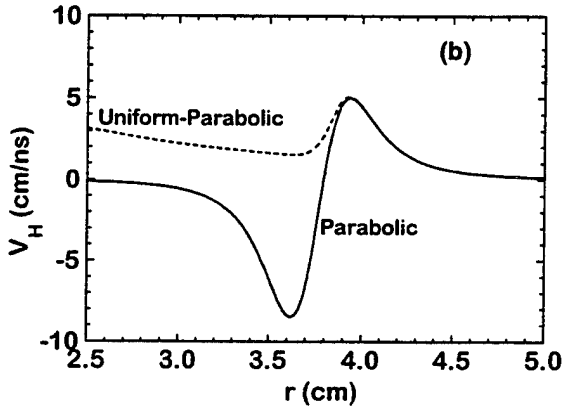


Figure 2
Swanekamp et. al

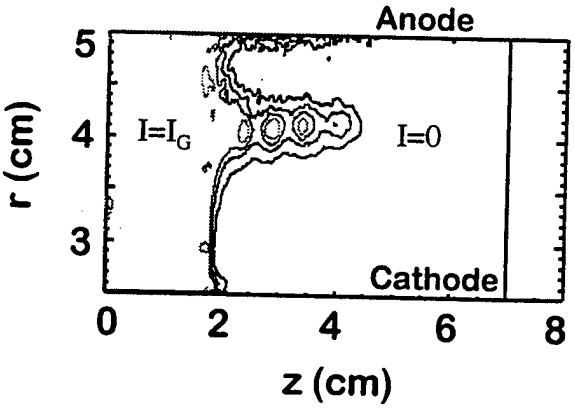


Figure 3
Swanekamp et. al

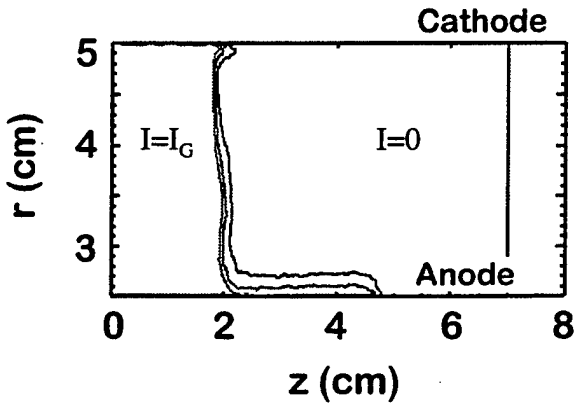


Figure 4
Swanekamp et. al

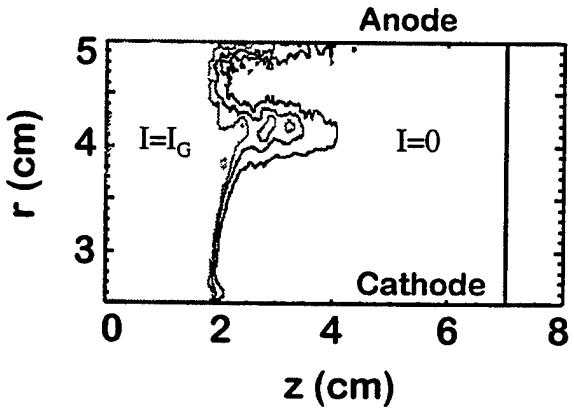


Figure 5
Swanekamp et. al

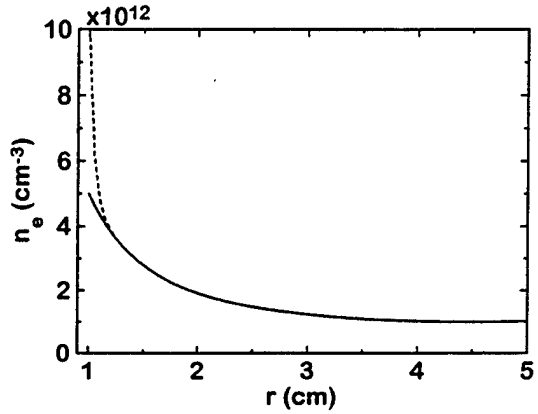
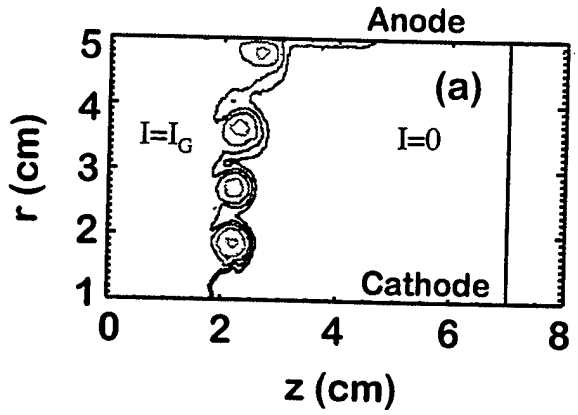


Figure 6
Swanekamp et. al



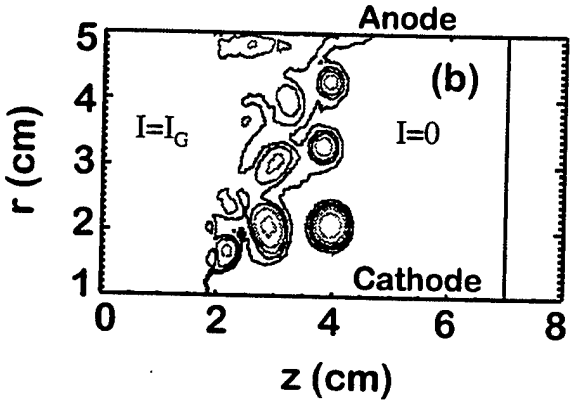
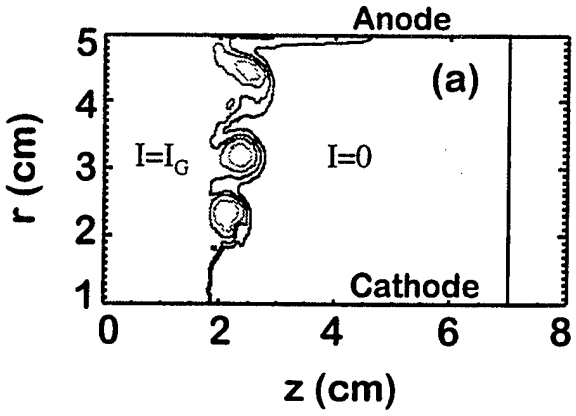


Figure 7
Swanekamp et. al



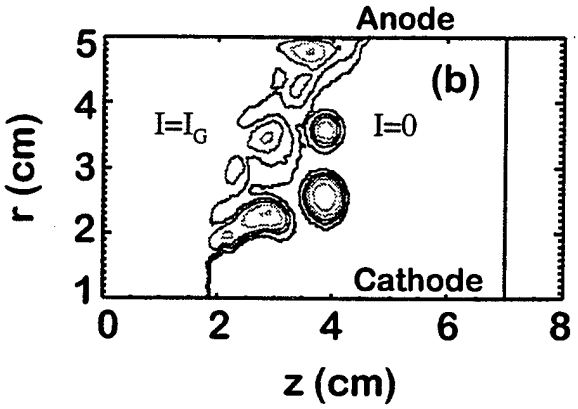


Figure 8
Swanekamp et. al

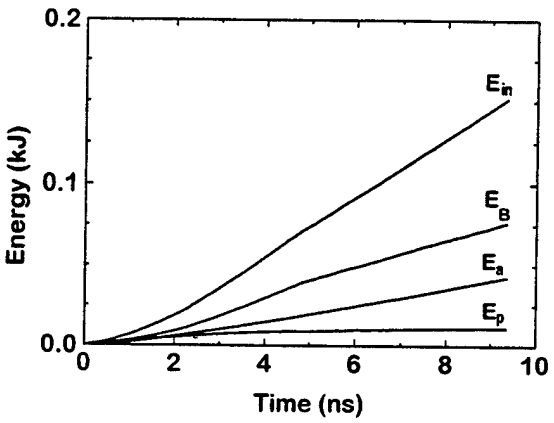


Figure 9
Swanekamp et. al

Appendix CC

Controlling Unstable States in Reaction-Diffusion Systems Modeled by Time Series

PHYSICAL REVIEW E VOLUME 50, NUMBER 4 OCTOBER 1994

Controlling unstable states in reaction-diffusion systems modeled by time series

Ira B. Schwartz and Ioana Triandaf
U.S. Naval Research Laboratory, Special Project for Nonlinear Science, Code 6700.3, Plasma Physics Division, Washington, D.C. 20375-5000
(Received 4 January 1994)

We introduce an algorithm for controlling unstable states of a spatiotemporal system modeled by a time series. Control is achieved by adjusting an external parameter at the boundary. Our time series is taken as the concentration from an experiment modeled by a reaction-diffusion system. It is shown that unstable states can be maintained by performing fluctuations of the concentration at the boundaries, while monitoring the dynamics from an interior spatial point.

PACS number(s): 05.45.+b

I. INTRODUCTION

Recent developments in the theory of nonlinear dynamical systems have provided experimentalists with new tools for exploring a wide range of aspects in the dynamics of real systems based on analyzing a single time series. Methods such as embedding techniques allow one to reconstruct the geometric model of the attractor and recover all its essential properties from time series measurements alone [1,2]. These methods have led to a control algorithm [known as the Ott-Grebogi-Yorke (OGY) algorithm)] for stabilizing unstable orbits inside a chaotic attractor [3] by applying small, carefully computed perturbations of an accessible system parameter.

Recently, the authors have designed an algorithm that stabilizes unstable orbits and also tracks them as a function of a system parameter, thus extending the region over which control can be achieved [4]. This algorithm also applies to the time series itself and makes use of embedding techniques. As the parameter is varied, control is maintained by a predictor-corrector technique. The correction step incorporates the OGY technique or any analogous form of linear control. The tracking algorithm was implemented for maps as well as flows, and has been successfully applied to experiments [5,6].

Tracking an unstable state of spatiotemporal processes usually modeled by partial differential equations is also possible and will constitute the subject of a future paper. Tracking and control along unstable branches as a function of a parameter can lead to interesting new stable patterns that do not form spontaneously in an experiment [8], possibly leading to new experimentally realizable regimes. As a first step in this direction, we present a method of stabilizing an unstable state, which achieves control spatially as well as temporally. The method is applied to a reaction-diffusion system that models pattern formation in Couette flow reactors. This system exhibits both small amplitude chaos and chaotic bursting [7]. Our goal is to stabilize an unstable periodic orbit when the dynamics exhibits periodic bursting or chaotic behavior.

In our method, the numerical solution of the system is generated by a partial-differential-equation (PDE) solver, and control is applied by adjusting the boundary data referencing the dynamics at a fixed spatial point. We simulate the solution from an experiment where only the time series at a spatial point is accessible.

II. MODEL

We consider the following one-dimensional reaction-diffusion system:

$$\frac{\partial u}{\partial t}=D\frac{\partial^2 u}{\partial x^2}+\frac{1}{\epsilon}[v-f(u)]\,,$$
$$\frac{\partial v}{\partial t}=D\frac{\partial^2 v}{\partial x^2}-u+\alpha,\quad x\in[0,1]\,, \tag{1a}$$

subject to Dirichlet boundary conditions,

$$u(x=0,0)=u_0,\quad v(x=0,0)=v_0\,,$$
$$u(x=1,0)=u_0,\quad v(x=1,0)=v_0\,. \tag{1b}$$

The reaction term is a two-variable Van der Pol–like equation, which accounts for the excitable bursting character of the dynamics. We remark that in the absence of diffusion, chaotic solutions are not possible.

This is a formal model that does not completely meet the experimental conditions and the requirements of chemical kinetics laws in the Couette flow reactor. However, it reproduces most of the phenomena associated with the observed front patterns in a chlorite-iodide reaction [7]. The interaction of reaction and diffusion terms gives rise to a variety of sustained patterns such as stationary periodic structures, nonlinear waves, or chaotic spatiotemporal structures of large amplitude. In [9], numerical evidence for chaotic intermittent bursting was reported and analyzed taking the transport rate D as a bifurcation parameter. In our study, we noticed extreme sensitivity of the solution with respect to the diffusion coefficient D as well as with respect to α, which did not allow the use of these parameters for control. (Changes in the seventh significant digit were sufficient to change the attractors.) Instead we looked at the solution as one of the Dirichlet boundary conditions is varied and observed transition to chaos via an intermittency route, the details of which will be presented elsewhere.

III. ALGORITHM

The algorithm we present is meant for stabilizing unstable orbits of a spatiotemporal process modeled by a time series; a desired unstable orbit is maintained by adjusting an external system parameter. For Eqs. (1a) and (1b), the time series is measured at one spatial point in the interior region. Control is achieved by appropriately choosing fluctuations in the boundary conditions, introduced as one integrates in time. In fact, it is sufficient to use only one of the Dirichlet boundary conditions as an accessible parameter, which we will refer to from now on as our control parameter.

In order to maintain the system on the unstable state, we measure a time series at $x=x_0$, namely $v(x_0,t)$. The fixed space value is taken anywhere in the middle region of the interval (approximately the middle third of the interval) where the most severe bursting occurs. The system is reaction dominated in this region.

One way to form a discrete dynamics of the time series is sampling the variable v at successive minima. (Due to the strong coupling between the variables u and v, control applied to the v variable leads to controlling the u variable as well.) If we denote successive minima in the time series for $v(x,t)$ by v_n, we obtain a map denoted by f:

$$v_{n+1}=f(v_n,p)\ , \tag{2}$$

where p stands for the control parameter, which in our case is one of the boundary conditions. An unstable orbit of this map is then controlled by using any of the linear control methods.

Control is extended in time as follows. At each iterate of the map (2), the computed value of the solution v_n is used to determine the fluctuation δp_n in the parameter according to a linear control method. The change in the parameter will be proportional to the deviation from the unstable state to be controlled. The new value of the parameter $p_n+\delta p_n$ is fed back into the PDE solver. The solution v_{n+1} at $p_n+\delta p_n$ is obtained, and a new evaluation of the fluctuation in the boundary condition then follows based on v_{n+1}, and so on. Control of the unstable orbit is thus extended in time.

To fix ideas, we assume from here on that the state we are interested in is a period-1 fixed point v_0 of the map f; i.e., $v_0=f(v_0,p)$. Such a fixed point corresponds to a period-1 time series at the spatial point x_0.

As our linear control algorithm, we used the OGY control method, which amounts to ensuring at each time iteration that the next iterate of the map will fall on the stable manifold of the unstable state we are controlling. In the case of a two-dimensional map, if we denote by λ_s,λ_u the eigenvalues of the unstable state v_0 and by f_u the contravariant vector corresponding to the unstable direction, then the above mentioned condition applied to the linear approximation of the map yields that the control parameter p must be modified by

$$\delta p_n \equiv \frac{\lambda_u[\mathbf{v}_n-\mathbf{v}_0(p)]\cdot\mathbf{f}_u}{(\lambda_u-1)\mathbf{g}\cdot\mathbf{f}_u}\ , \tag{3}$$

at each iteration of the map [3]. Notice that Eq. (3) depends on the spatial point at which the time series is measured. The vector g is the derivative of the unstable state v_0 with respect to p. In the case of weak diffusion, the map is in fact nearly one dimensional at a spatial point. This amounts to having $\lambda_s=0$, in which case the formula becomes

$$\delta p_n \equiv \frac{\lambda_u[\mathbf{v}_n-\mathbf{v}_0(p)]}{(\lambda_u-1)\mathbf{g}}\ . \tag{4}$$

Equation (4) is a traditional control method known as occasional proportional feedback [10,11]. The eigenvalues and eigenvectors involved in the formulas (2), (3), or (4) associated with the saddle $v_0(p)$ can all be calculated based on the reconstructed attractor.

Summarizing, our method achieves control of a spa-

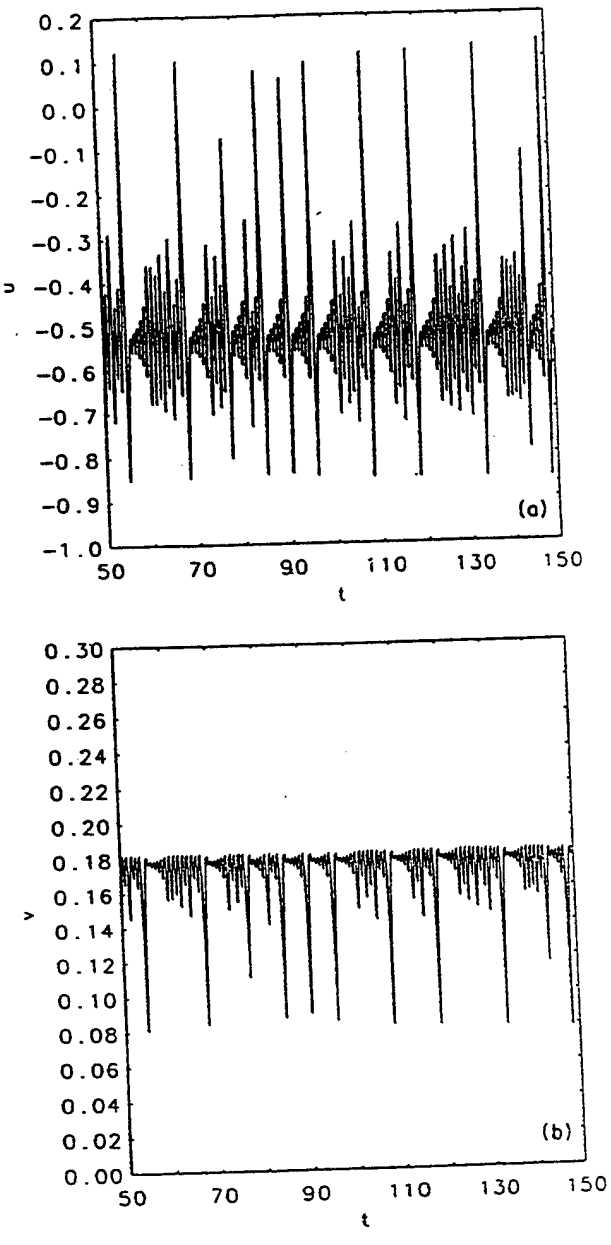


FIG. 1. (a) and (b) Time series for u and v, respectively, recorded at $x=\frac{1}{2}$ and $p=-2.0$ in the absence of control.

tiotemporal process as follows. After reconstructing a discrete map from the data sampled at a fixed spatial point, control is implemented by fluctuating the boundary conditions proportional to the deviation from the state we want to maintain. The method designed in this way has the advantage in that it can be applied directly to experimental data for which an analytical model is unavailable, and requires no mode expansion as in [12].

IV. NUMERICAL RESULTS

In our numerical tests we took in the model (1a) $f(u)=u^2+u^3$, as in [7]. Before illustrating our scheme, we examined the bifurcation of solutions by taking the control parameter to be $p=u(0,t)$. Let us take $D=0.032\,249$ and $\alpha=0.01$, values at which a stable period one orbit exists for p between $p=-0.5$ and $p=-0.6$. As p is decreased past -0.6 the period-1 orbit becomes unstable and bifurcates into a period-2 orbit. At about $p=-1.15$ the period two destabilizes, giving rise to an intermittent bursting regime, which becomes chaotic as we further decrease p past $p=-1.6$.

Figures 1(a) and 1(b) show the time series for the solution at $x=\frac{1}{2}$ and $p=-2.0$, without control. The time series is chaotic and exhibits three distinct types of oscillations. First there is a large amplitude burst, which occurs on a fast reaction time scale. Following the burst, there occurs an exponentially growing small amplitude oscillation. This is followed by small amplitude chaotic oscillations whose length is random in time.

For a time series of 14 000 points at $x=\frac{1}{2}$, we compute an information dimension of 2.1, having one positive Lyapunov exponent [14]. Spectral analysis reveals that most of the energy is contained in four spatial modes. Note that time series sampled at other spatial points may have no positive Lyapunov exponent. We have not found any other chaotic solutions having more than one positive Lyapunov exponent.

From the solution, we form a map for which the iterates are the successive minima of the variable v. Figure 2 shows the successive minima of v when $D=0.032\,249$, $\alpha=0.01$, and $p=-2.0$, values of the parameters at which the solution exhibits chaos. Simultaneously, we display a stable period-6 solution in the intermittent regime at $D=0.032\,249$, $\alpha=0.01$, and $p=-1.4$. From the picture we see that this orbit obeys nearly the same nonlinear law as the chaotic attractor at $p=-2.0$. Similar results of periodic bursting hold in the Belovsov-Zhabotinsky (BZ) continuously stirred tank reactor [13]. The chaotic map in Fig. 2 is nearly one dimensional, justifying our use of the occasional proportional feedback in

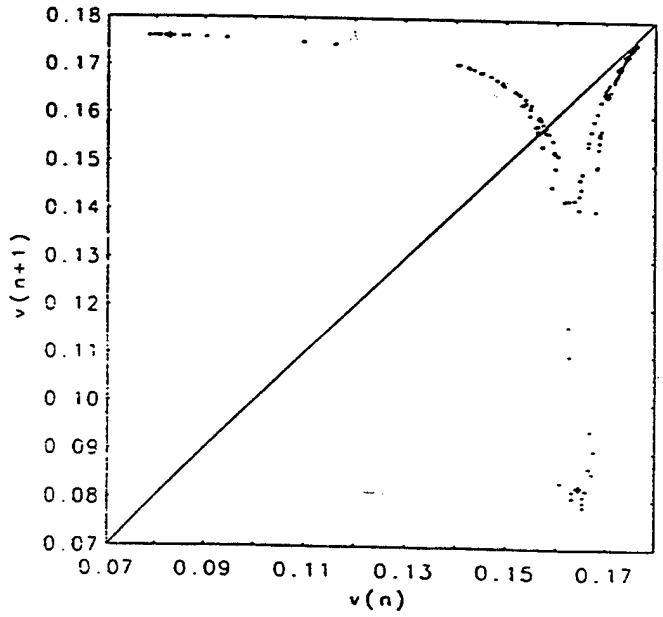


FIG. 2. v_{n+1} vs v_n at $D=0.032\,249$, $\alpha=0.01$, and $p=-2$. Large dots indicate the period-6 orbit at $p=-1.4$. Notice that both parameter values of p obey nearly the same dynamics law.

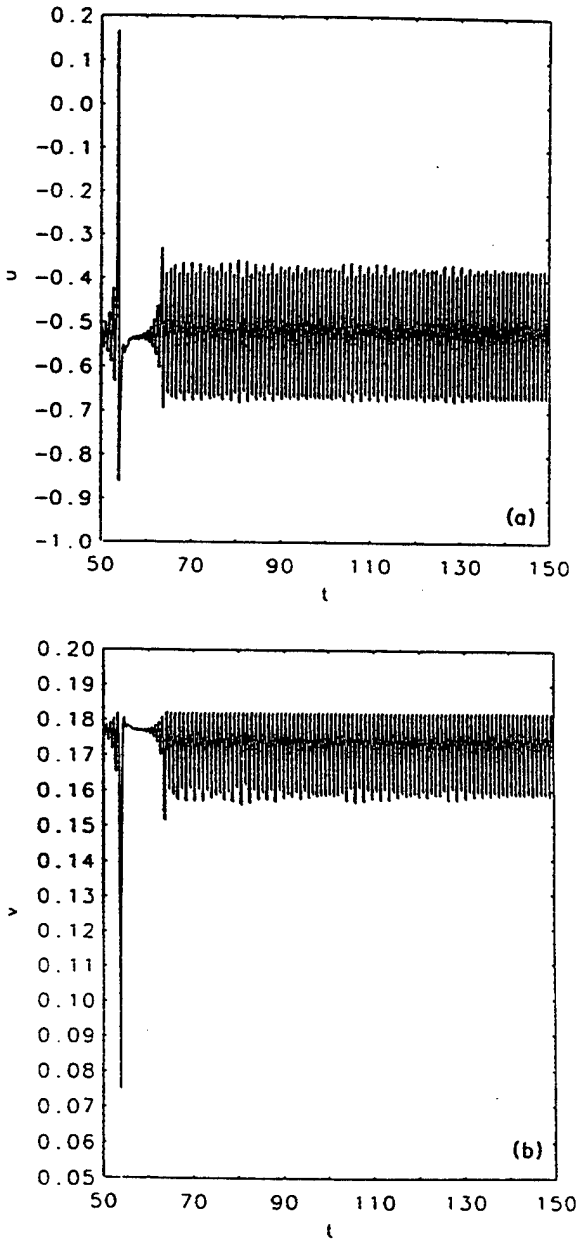


FIG. 3. (a) and (b) Time series for u and v, respectively, recorded at $x=\frac{1}{2}$ and $p=-1.4$, when control is applied to the corresponding time series at $x=\frac{1}{2}$.

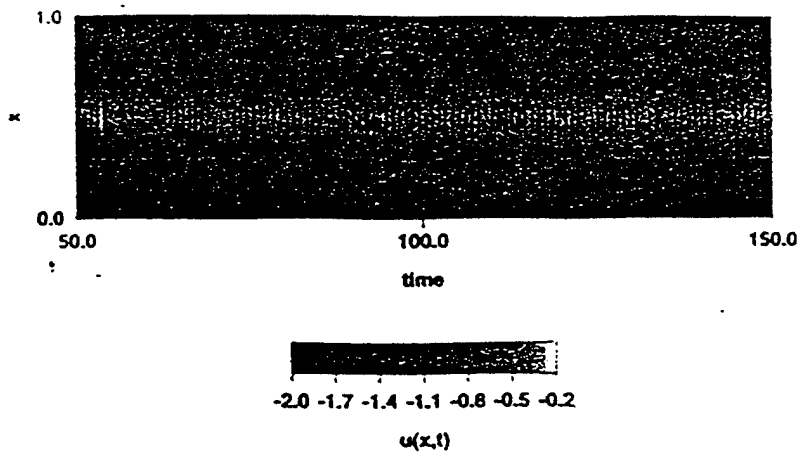


FIG. 4. Spatiotemporal pattern when control is applied to the time series at $x=\frac{1}{3}$.

our algorithm. Notice that the three oscillation types are clearly evident in the one-dimensional attractor reconstruction, the large bursts are near the peak of the map, the exponentially growing solutions form the beginning of the left branch, and the small amplitude chaos is formed around the period-1 fixed point. Also notice that at $p=-1.4$, the period-1 fixed point is inaccessible, since the stable period 6 is attracting.

Since control of period-1 orbits in chaotic attractors has been done elsewhere [11], we now describe control of inaccessible period-1 points. We consider a parameter value $p=-1.4$, where a stable periodic intermittent bursting solution exists. This orbit consists of a large burst followed by five small growing oscillations and is shown on the graph mapped in Fig. 2. We remark that at $p=-1.4$, the period-1 fixed point and its local neighborhood are inaccessible. That is, the dynamics does not enter a neighborhood of the period-1 fixed point, since there is no chaos. For control to occur, we fluctuate the parameter so that the dynamics enters the neighborhood of the period-1 fixed point. We do this by looking at the first preimage of a controllable neighborhood of the fixed point in the bursting regime, on the right branch of the map. The parameter is adjusted at the bursts so the dynamics enters a neighborhood of the period-1 fixed point, at which point control is implemented.

We generate a time series by sampling at $x=\frac{1}{3}$. In order to choose the reference value $v_0(p)$ at $p=-1.4$ in formula (4), we notice that the orbit we want to control has a fixed point that lies on the $y=x$ line in Fig. 2. Because the dynamics at $p=-1.4$ is approximated by the map at $p=-2$ (as shown in Fig. 2), we approximate $v_0(-1.4)$ by $v_0(-2)$.

Figures 3(a) and 3(b) show the stabilized solution at $x=\frac{1}{2}$, when the control is based on the solution sampled at $x=\frac{1}{3}$. The amplitude of this solution agrees with the amplitude of the stable solution at $p=-0.5$, where the period-1 orbit is stable.

During control of the periodic solution, the amplitude of the control at the boundaries is approximately 30% of the signal at $x=\frac{1}{2}$. The control perturbations are larger than in previous applications [4,5], since they must overcome weak diffusion to be effective in the interior.

In Fig. 4 the whole stabilized spatiotemporal pattern is shown. The period-1 solution in this example is indeed unstable. To see this, in Fig. 5 we show the same time series as in Fig. 3 where control was removed after 200 iterates. This results in the reappearance of the intermittent pattern after a short delay. In this example, the time series was sampled at $x=\frac{1}{2}$ and we started with an initial condition that is constant and equals the boundary conditions at $x=0$. The control parameters were not optimized to minimize fluctuations about the period-1 reference state.

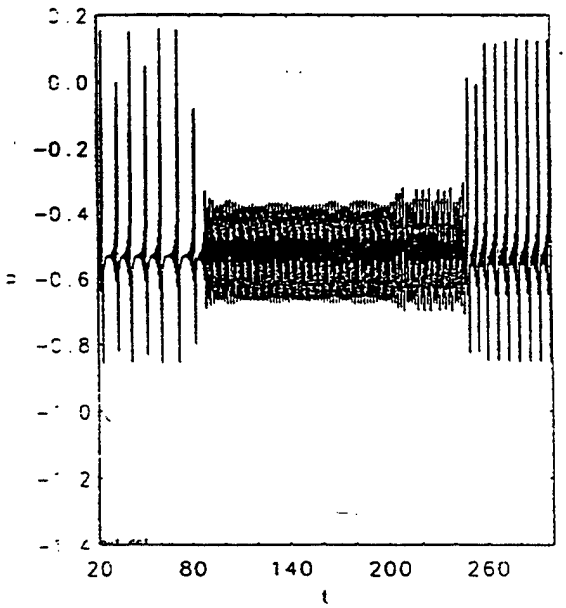


FIG. 5. Time series for u, respectively recorded at $x=\frac{1}{2}$ and $p=-1.4$, when control is applied at $x=\frac{1}{2}$ for the first 200 iterates. Control is removed afterwards.

V. CONCLUSIONS

We have introduced and tested an algorithm that applies to stabilizing unstable states of spatiotemporal processes occurring in reaction-diffusion processes. The novelty of this procedure consists in the fact that it applies

to the time series directly, combining nonlinear analysis embedding techniques with classical linear control. Compared to similar techniques, it has the advantage that it achieves control spatially as well as temporally. Since it applies to the time series directly, the method is suitable for experimentalists. Furthermore, by using the targetting of intervals in periodic regimes, previously inaccessible control points are now achievable by making use of the global nonlinear dynamics.

ACKNOWLEDGMENT

Dr. I. Triandaf gratefully acknowledges the support of the Office of Naval Research for conducting this research.

[1] F. Takens, in *Detecting Strange Attractors in Turbulence*, edited by D. S Rand and L. S. Young, Lecture Notes in Mathematics Vol. 898 (Springer-Verlag, Berlin, 1981).
[2] Tim Sauer, James A. Yorke, and Martin Casdagli, J. Stat. Phys. 65, 579 (1991).
[3] Edward C. Ott, Celso Grebogi, and James A. Yorke, Phys. Rev. Lett. 64, 1196 (1990).
[4] Ira Schwartz and Ioana Triandaf, Phys. Rev. A 46, 7439 (1992).
[5] Thomas L. Carroll, Ioana Triandaf, Ira Schwartz, and Lou Pecora, Phys. Rev. A 46, 6189 (1992).
[6] Zelda Gills, Christina Iwata, Rajarshi Roy, Ira Schwartz, and Ioana Triandaf, Phys. Rev. Lett. 69, 3169 (1992).
[7] J. Elezgaray and A. Arneodo, Phys. Rev. Lett. 68, 714 (1992).
[8] John A. Vastano, John E. Pearson, W. Horsthemke, and Harry L. Swinney, Phys. Rev. A 124, 320 (1987).
[9] A. Arneodo, J. Elezgaray, J. Pearson, and T. Russo, Physica D 49, 141 (1991).
[10] E. R. Hunt, Phys. Rev. Lett. 67, 1953 (1991).
[11] Raj Roy, T. W. Murphy, Jr., T. D. Maier, Z. Gills, and E. R. Hunt, Phys. Rev. Lett. 68, 1259 (1992).
[12] Hu Gang and He Kaifen, Phys. Rev. Lett. 71, 3794 (1993).
[13] J. Rinzel and I. B. Schwartz, J. Chem. Phys. 80, 5610 (1984).
[14] E. J. Kostelich and H. L. Swinney, *Chaos and Related Nonlinear Phenomena*, edited by I. Procaccia and M. Sapiro (Plenum, New York, 1987).

Appendix DD

Detecting Motion on Center Manifolds from a Time Series-An Example

Detecting Motion on Center Manifolds from a Time Series-An Example

Ira B. Schwartz, Ioana Triandaf
Special Project for Nonlinear Science
U.S. Naval Research Laboratory
Plasma Physics Division, Code 6700.3
Washington, DC 20375

ABSTRACT

Recent progress in control from a geometric viewpoint using nonlinear dynamics techniques has focussed primarily on linear control. When using embedding methods to locate local properties of controllable states, hyperbolicity is always assumed. However, many problems which contain symmetries also contain nonlinear instabilities. It is the purpose of this paper to give a procedure for detecting these nonlinear directions from time series using embedding methods.

I. Introduction

Recent advances in the theory of dynamical systems have resulted in geometric analysis of low dimensional phenomena from the study of single time series measurements [BK]. The main technique which has been advocated and has been quite useful is the technique of embedding a time series in a reconstructed phase space [T,SYC]. The idea allows one to take a single time series measurement and create a geometric representation of the dynamics by using coordinates which are delayed in time. The technique of embedding has been used to capture quantitative information about the global behavior in chaotic attractors [BS], but more recently, it has been used to reveal local structure about periodic orbits contained within the attractor [LK]. When examining chaotic attractors, it is known that the attractors contain an infinite number of periodic orbits [GH]. Moreover, these orbits are all unstable, containing both contracting and expanding directions. By analyzing these orbits, local behavior may be stabilized by performing control with the use of an accessible parameter [AGOY, H, OGY, RGOD, RMMGH]. By allowing the parameter to vary, nonlocal pictures of the controlled unstable orbits may now be found [CTSP, ST], where the experimentalist can implement methods which are similar in spirit to that of model dependent continuation [DR]. That all of this analysis for an experiment has derived from the theory of embedding is quite profound. The main success of all the methods has relied on the fac that the local orbits are hyperbolic. That is, all of the in stabilities reside in the linearization about the orbit, an are easily measured. In general, however, there are sys tems which possess orbits of interest and which contai neutral directions; i.e., the stability along these direction must be determined by knowing the nonlinear terms. Certain systems which possess symmetries have such behavior. For example, Sirovich and Xhou [SX] derive a mode Karhune-Loeve system from a fully developed turbulen flow of the Navier Stokes equations. The equations are invariant with respect to reflection and rotation of the complex amplitudes. For various values of Reynold's number many types of dynamical behavior are observed, and are summarized in [SX]. The one interaction of interest is that of two oblique waves and one roll wave, and appears as an unstable periodic orbit of the model. It can be shown this interaction is always linearly unstable whenever it exists. It is seen that in addition to being linearly unstable, there is also one component which is always neutrally stable. The other components are stable with complex eigenvalues. Another example of neutral stability in a real system comes from globally coupled arrays of Josephson junctions [WS, TS], which also possesses attracting and repelling directions. When measuring local properties of an unstable periodic orbit, such as the one found in the models mentioned, where there is one unstable direction, one neutral direction, and the rest stable, one can acquire information only about the linear part. Since the terms contributing to the local stability are nonlinear, their contribution is weak. Therefore, embedding methods used to extract the local dynamics are incomplete, since the linear terms are the ones which are measurable. In this paper, we offer an alternative to discover the nonlinear terms. We combine control with the local dynamics to show how to discover the nonlinear instability. This will be done by demonstrating the procedure on a specific map. Specifically, we will detect nonlinear motion on a center manifold. Once

0-7803-2570-2/95 $4.00

a local model is known which is complete, nonlinear control theory may be applied to design a dynamical system which is stable. The full theory and applications to the problems cited above will appear in a journal paper.

II. Detecting nonlinear motion

We demonstrate the theory of detecting motion on center manifolds from a time series by considering an example. The example is a discrete map which is designed to be uncontrollable. Furthermore, it also will have a nonlinearly unstable direction. The example mimics the local stability from an orbit of a model of wall bounded turbulence, where it is of interest to control three dimensional flow.

Consider the following three dimensional map,

$$z \to F(z,p) \tag{1}$$

where, $z=(u,\nu,w)$ and p is a scalar parameter, and

$$F(z,p) \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} u \\ \nu \\ w \end{bmatrix} + \begin{bmatrix} \nu w \\ u^2 \\ u\nu \end{bmatrix} + p \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{2}$$

In this map, when $p=0$, the origin is a fixed point, and the linear part of the map is diagonal for ease of exposition. (The ideas presented are no less general if we assume the linear parts are block diagonal.) Let

$$\begin{aligned} A = D_z F(0,0) &= diag(1, -\tfrac{1}{2}, \tfrac{3}{2}) \\ B = D_p F(0,0) &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \end{aligned} \tag{3}$$

Clearly, $rank\,(B|AB|A^2B) < 3$, implying the origin is uncontrollable [RGOD]. However, since B is contained in the span of the unstable eigenspace, we can control the linearly unstable subsystem. To control the unstable direction, we project off of the unstable manifold. Letting $e_u = (0,0,1)^t$, we require that the parameter be chosen so that

$$e_u^t z_{n+1} = 0 \tag{4}$$

Considering the linear dynamics only, where $z_{n+1} = Az_n + Bp_n$, Eq. 4 implies

$$p_n = -\frac{3}{2} \begin{bmatrix} 0 \\ 0 \\ w_n \end{bmatrix}. \tag{5}$$

The dynamics of the controlled nonlinear system is now given by

$$z_{n+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} z_n + \begin{bmatrix} \nu w \\ u^2 \\ u\nu \end{bmatrix}. \tag{6}$$

The effect of control on the system is such that the dynamics is constrained to lie in the product of the attracting and neutral manifolds. The eigenvalue of unity implies the existence of a center manifold, $W_c(0)$, defined by

$$\begin{aligned} W_c(0) \; &= \{(u,\nu,w) \in R^3 | \nu = h_1(u), w = h_2(u) \\ &\quad h_i(0) = 0,\; Dh_i(0) = 0, i = 1,2;\; for |u| < \delta\} \end{aligned} \tag{7}$$

The map restricted to the center manifold is given by

$$u \to u + h_1(u) h_2(u). \tag{8}$$

Stability along this manifold will determine the stability of the controlled system, Eq. 6. We analyse Eq. 6 by explicitly computing the leading nonlinear terms of h_1 and h_2. To do this, notice that the dynamics restricted on this manifold obeys the following:

$$\begin{aligned} h_1(u_{n+1}) &= -\tfrac{1}{2} h_1(u_n) + u_n^2 \\ h_2(u_{n+1}) &= h_1(u_n) u_n \end{aligned} \tag{9}$$

We assume an analytic form of h_i, $i=1,2$:

$$h_i(u) = a_i u^2 + b_i u^3 + O(u^4), i = 1,2. \tag{10}$$

Inserting Eq. 10 into Eq. 9, we equate coefficients of powers of u, and find that

$$\begin{aligned} a_1 &= \tfrac{2}{3}, \quad b_1 = 0 \\ a_2 &= 0, \quad b_2 = a_1 \end{aligned} \tag{11}$$

Equation 8 for the dynamics on the center manifold becomes

$$u_{n+1} = u_n + \frac{4}{9} u_n^5 \tag{12}$$

which, for any initial perturbation from the origin $u_0 = \epsilon$, results in a magnification; i.e., $|u_1| = |\epsilon|\,|\,(1 + \frac{4}{9}\epsilon^4)\,| > |\epsilon|$. Therefore, Eq. 8 is nonlinearly unstable, which implies that the controlled system Eq. 6 is nonlinearly unstable as well.

Computations of the controlled system Eq. (2,5) with noise added were done to test the nonlinear instability. The results are presented in Figs. 1 and 2. Each component is plotted as a function of iterate for 1000 iterates. In Fig. 2, it is observed that the first component, u, diverges from the origin slowly, which is expected from the analysis on the center manifold given in Eq. 12. In Fig. 2, components ν and w are held fixed close to the origin since they correspong to directions which stable. To check the precitability of the nonlinear terms, a fit for the nonlinear exponent in Fig. 3 shows a log-log plot of $u_{n+1} - u_n$ vs. u_n . A straight line reveals an approximate slope of 4.98, which is in good agreement with Eq. 3.11.

III. Nonlinear control

Once a model of the local linear hyperbolic structure and nonlinear motion along the center manifold has been found, the control law may be adjusted to incorporate nonlinear terms. We assume the control law is of the following form:

$$p(z) = k \cdot z + \Lambda \cdot zz, \tag{13}$$

where $k \in R^3$ and is a bounded linear operator from R^3 to R^3. Using the same linear control as Eq. 5, we now have the map:

$$z \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} z + \begin{bmatrix} \nu w \\ u^2 \\ uw \end{bmatrix} + \Lambda \cdot zz \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{14}$$

Using center manifold theory as above, we now compute motion on the center manifold incorporating as an unknown. As before, stability is determined by Eq. 8, and the equations describing the manifolds are

$$\begin{aligned} h_1(u_{n+1}) &= -\tfrac{1}{2} h_1(u_n) + u_n^2 \\ h_2(u_{n+1}) &= h_1(u_n) u_n + \\ &\quad (u_n h_1(u_n), h_2(u_n))^T \Lambda (u_n, h_1(u_n), h_2(u_n)) \end{aligned} \tag{15}$$

After some manipulation, we find that one possible solution of Λ which stabilizes the nonlinear direction is $\Lambda_{12} = \Lambda_{21} = -\frac{8}{3}$, and the rest of the components are equal to zero. Therefore, the control law given by

$$p(z) = -\frac{3}{2} w - \frac{16}{3} uw \tag{16}$$

will stabilize the nonlinear instability. We illustrate the results nonlinear control in Fig. 4, where it is seen that nonlinear instability of the u component has been controlled.

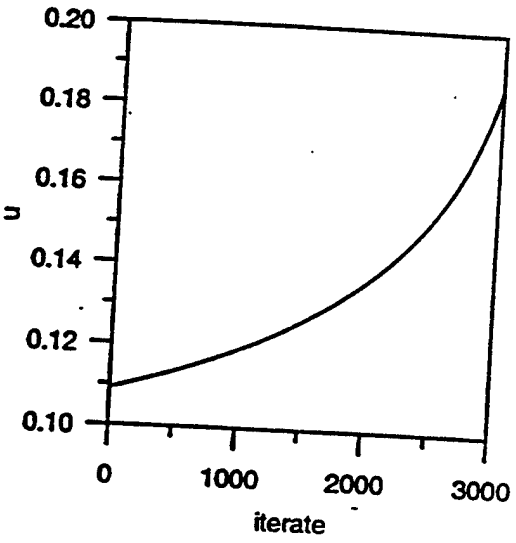


Figure 1: Nonlinear instability along center manifold.

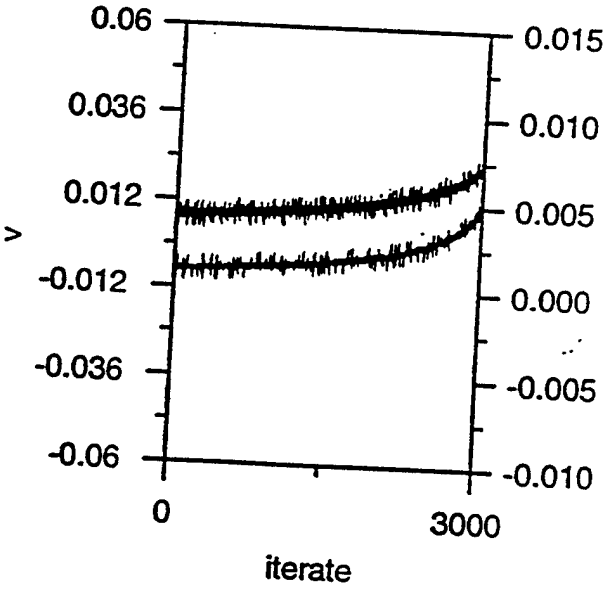


Figure 2: Linear control in ν and w.

References

[AGOY] Auerbach D., Grebogi C., Ott E. and York J. A., "Controlling Chaos in High Dimensional Systems", Phys. Rev. Lett. 69, (1992), 3479-3482.

[BK] Broomhead D.S. and King G.P., "Extracting qualitative dynamics from experimental data", Physica D 20, (1986), 217-236.

[BS] Brandstater A. and Swinney H.L., "Strange Attractors in Weakly Turbulent Couette-Taylor Flow", Phys. Rev. A, 35, (1987), 2207-2220.

[CTSP] Carroll T.L., Triandaf I., Schwartz I., and Pecora L., "Tracking Unstable Orbits in an Experiment", Phys. Rev. A, 46, (1992), 6189-6192.

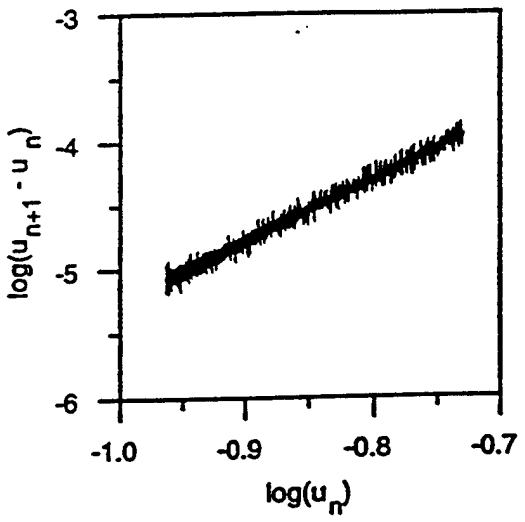


Figure 3: Log-Log Plot of Nonlinearity.

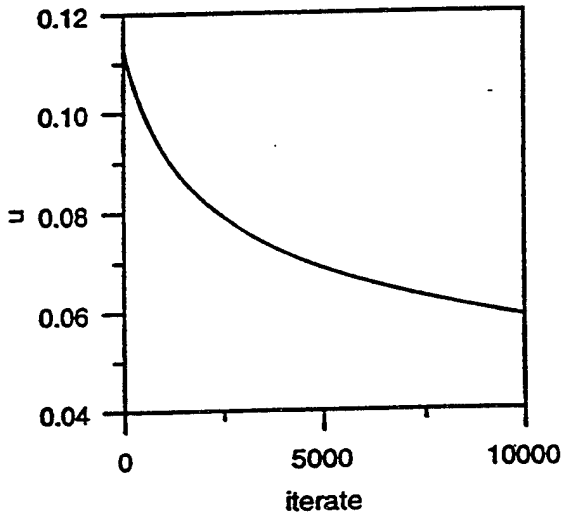


Figure 4: Nonlinear control of u.

[DR] Den Heijer C. and Rheinboldt W.C., "On Steplenght Algorithms for a Class of Continuation Methods", SIAM. J. Numer. Anal. Vol. 18, (1981), 925-948.

[GH] Guckenheimer J. and Holmes P., "Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Field", Springer-Verlag, 1983.

[H] Hunt E.R., "Stabilizing High-Period Orbits in a Chaotic System: The Diode Resonator", Phys. Rev. Lett. 67, (1991), 1953-1955.

[LK] Lathrop D.P., and Kostelich E.J., "Characterization of an Experimental Strange Attractor by Periodic Orbits", Phys. Rev. A 40, (1989), 4028-4031.

[OGY] Ott E.C., Grebogi C. and Yorke J.A., "Controlling Chaos", Phys. Rev. Lett. 64, (1990), 1196-1199.

[RGD] Romeiras F., Grebogi C., Ott E. and Dayawansa W.P., "Controlling Chaotic Dynamical Systems", Physica D 58 (1992), 165-192.

[RMM] Roy R., Murphy T.W. Jr., Maier T.D., Gills Z. and Hunt E.R., "Dynamical Control of a Chaotic Laser: Experimental Stabilization of a Globally Coupled System", Phys. Rev. Lett. 68, (1992), 1259-1262.

[ST] Schwartz I.B. and Triandaf I., "Tracking Unstable Orbits in Experiments", Phys. Rev A, 46, (1992), 7439-7444.

[SYC] Sauer T., Yorke J.A. and Casdagli M., "Embedology", J. Stat. Phys. Vol.65, (1991), 579-616.

[SZ] L. Sirovich and Xhou, A dynamical model of wall-bounded turbulence, Phys. Rev. Letts. 74, p34. (1994)

[T] Takens F.,"Detecting Strange Attractors in Turbulence" in Dynamical Systems and Turbulence edited by D.A. Rand and L.-S. Young Lecture Notes in Mathematics 898, Springer-Verlag, Berlin (1981). 366-381.

[TS] K. Tsang and I. B. Schwartz, Interhyperhedral diffusion in Josephson Junction arrays, Phys. Rev. Letts. 68, pp. 2265-2268 (1992)

[WS] S. Watanabe and S. H. Strogatz, Integrability of a globally coupled oscillator array, Phys. Rev. Letts 70, pp. 2391-2394 (1989)

Appendix EE

Chaos and Intermittent Bursting in a Reaction-Diffusion Process

Chaos and intermittent bursting in a reaction-diffusion process*

Ira B. Schwartz
Special Project in Nonlinear Science, U.S. Naval Research Laboratory Code 6700.3, Plasma Physics Division, Washington, D.C. 20375-5000

Ioana Triandaf[a)]
Science Applications International, Corporation, Applied Physics Operation, McLean, Virginia 22102

(Received 10 January 1996; accepted for publication 7 March 1996)

Karhunen–Loeve decomposition is done on a chaotic spatio-temporal solution obtained from a nonlinear reaction-diffusion model of a chemical system simulating a chemical process in an open Couette-flow reactor. Using a Galerkin projection of the dominant Karhunen–Loeve modes back onto the nonlinear partial differential system, we obtain an ordinary differential equation model of the same process. Major features such as intermittent and chaotic bursting of the nonlinear process as well as the mechanism of transition to chaos are shown to exist in the low-dimensional model as well as the PDE model. From the low-dimensional model the onset of intermittent bursts followed by small amplitude oscillations is shown to arise due to a sequence of saddle-node bifurcations. [S1054-1500(96)00802-6]

I. INTRODUCTION

In this paper we consider a nonlinear reaction-diffusion model in which one observes spatio-temporal patterns in many respects similar to those observed experimentally in the Couette flow reactor for a chlorite-iodide reaction and are characteristic of a wide class of systems.[1] This class of chemical reactions provides a remarkable variety of sustained intermittent bursting and chaotic patterns that organize in an open chemical system due to the interaction of the diffusion process with a chemical reaction which itself would proceed in a stationary manner if diffusion was negligible. Transitions to oscillatory and bursting spatio-temporal patterns that are generic in experiments[2] were presented in Refs. 3–6. The transition to intermittent and chaotic bursting patterns as diffusion is varied was presented in Ref. 7. In Ref. 8 a discrete control method was introduced in this model by fluctuating the concentration at the boundary based on monitoring the concentration at a single spatial point inside the central region of the spatial domain where reaction dominates the chemical process. When studying the dynamics as a function of the concentration at the boundary the authors observed transitions to chaos as one of the Dirichlet boundary conditions was varied.

In this paper we are interested in taking a dynamical systems approach to analyzing spatio-temporal data generated by a nonlinear system which evolves chaotically in time, and is spatially inhomogeneous. Our aim is to obtain a qualitative picture of the dynamics with a model involving fewer degrees of freedom. In addition, we would like to predict the onset of intermittent bursting as a function of bifurcations parameters. To this end, we perform Karhunen–Loeve (KL) decomposition on a chaotic solution of the reaction-diffusion model. The advantage of using KL modes over other eigenfunction expansion methods is that, in general, the same energy is captured in fewer modes. KL modes approximate the data with minimal least square error and they are decorrelated in space. Due to spatial decorrelation, the KL modes account for physically independent features in the process[9–11] which, as defined in fluid dynamics, are called coherent structures. Using the Karhunen–Loeve eigenmodes in a Galerkin projection procedure, we can obtain a low-dimensional model for the process and explore its dynamics. This type of mode analysis was pioneered by Lumley for the study of fluid turbulence problems,[12,13,9] and it has been also used for analyzing chaotic data.[14] In this paper, we apply KL modal analysis to a reaction-diffusion process exhibiting rich dynamical behaviour such as bistability, oscillations, pattern formation and spatio-temporal chaos, all of which appear in actual chemical experiments.[2,15]

The KL procedure consists in finding best fitting functions in the sense of the least square error for the representation of a given spatio-temporal pattern obtained as a solution of the reaction-diffusion model. The solution is assumed to be a sum of amplitude modulated modes. Projecting the KL eigenfunctions on the reaction-diffusion model, one obtains a model consisting of a system of ordinary differential equations for the amplitudes of the modes resulting in a lower dimensional representation of the process. The KL model is derived from data with respect to a fixed set of parameter values, namely the parameter values corresponding to the spatio-temporal pattern used to generate the modes. However, the KL generated model retains all of the physical parameters from the spatio-temporal model. Therefore, one can observe the bifurcation structure as parameters are changed in the KL model. The KL model gives remarkably good agreement with the original model in capturing

*See AIP Document No. E-PAPS file via: http://www.aip.org/epaps/epaps.html E-CHAOE-229-14.5 MB for color versions of Figs. 1, 6(a), 6(c), and 7. There is no charge for retrieval of electronic PAPS document files from the FTP server. For further information: e-mail: paps@aip.org or fax: 516-576-2223.

[a)]Current mailing address: Special Project in Nonlinear Science, U.S. Naval Research Laboratory, Code 6700.3, Plasma Physics Division, Washington, D.C. 20375-5000.

essential features such as bursting and small amplitude chaos along with the mechanism producing it. In some cases a complete agreement can be obtained between the two models over a whole range of parameter values; such is the case for the Ginzburg–Landau equation.[9] In other cases, such as for the reaction-diffusion process presented in Ref. 16, only part of the dynamics carries over. In this paper we investigate to what extent the lower-dimensional model can qualitatively capture dynamical features of the chosen reaction-diffusion process.

The layout of the rest of the paper is as follows. In section II we present the original reaction-diffusion system,[7] where we extract the five highest energy KL modes from a spatio-temporal chaotic solution exhibiting two-front behavior. We use these modes to generate a low-dimensional model by performing the above-mentioned Galerkin procedure. In section III we discuss the dynamical features of the two models. We end the paper with a conclusion section.

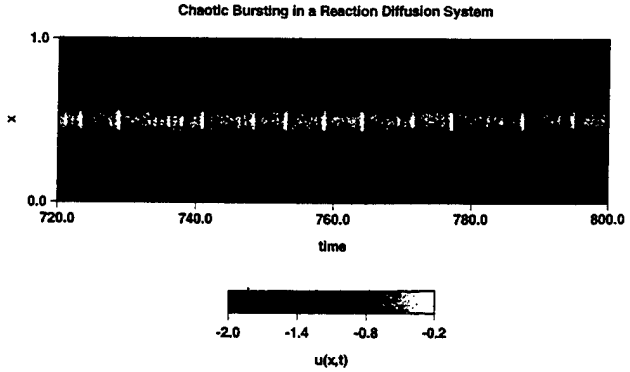


FIG. 1. Spatio-temporal pattern for the u-variable in the reaction-diffusion model. This data was expanded in KL modes to generate a lower-dimensional model. Parameter values used here are $D=0.032249$, $\alpha=0.1$, $\epsilon=0.1$. Boundary conditions are $u(0,t)=u(1,t)=-2$, $v(0,t)=v(1,t)=-4$. (A color version of this figure is available via E-PAPS: http://www.aip.org/epaps/epaps.html.)

II. THE KARHUNEN–LOEVE MODEL

A. The reaction-diffusion model

We consider a reaction-diffusion model of a chemical reaction, which is a variant of the chlorite–iodide reaction where bursting and chaotic spatio-temporal patterns are obtained by imposing parameters and boundary conditions to create a concentration gradient close to the two boundaries. Our model reads as

$$\begin{aligned} \frac{\partial u}{\partial t} &= D\frac{\partial^2 u}{\partial x^2}+\frac{1}{\epsilon}[v-f(u)], \\ \frac{\partial v}{\partial t} &= D\frac{\partial^2 v}{\partial x^2}-u+\alpha, x\in[0,1], \end{aligned} \tag{1}$$

and is subject to Dirichlet boundary conditions. Here D is a diffusion constant, and ϵ and α are fixed.

The reaction term is a two-variable Van der Pol-like equation, which accounts for the excitable bursting character of the dynamics. Notice that in the absence of diffusion, chaotic solutions are not possible, since the resulting system is two-dimensional.

Equation (1) is a formal model that generically approximates the dynamics in many reaction-diffusion systems. In particular, it reproduces the experimental conditions and requirements of chemical kinetic laws in a Couette flow reactor. It reproduces most of the phenomena associated with the observed front patterns in a chlorite-iodide reaction, namely bistability, excitability and relaxation oscillations. Furthermore, as shown in Ref. 11, the spatio-temporal patterns found in Eq. (1) are characteristic of a wide class of reaction-diffusion systems. We let $f(u)=u^2+u^3$, a generic form of the reaction term. The reaction term then ensures the existence of an S-shaped manifold, $v=f(u)$, consisting of three branches: two of them are attracting trajectories in a time of order $1/\epsilon$ and are separated by an unstable branch. The state variables u and v are taken as scaled concentration variables, and one refers to the upper and lower branches of the slow manifold as analogous to the oxidized and reduced state branches of the chlorite-iodide reaction. In Ref. 7 the authors have elucidated the mechanism for the periodic and nonperiodic appearance of spatially localized bursts of reduced states in an oxidized medium imposed from the boundaries. A concentration gradient was obtained at certain values of diffusion depending on the shape of the slow manifold, the choice of α and boundary conditions. The solution patterns for Eq. (1) consist of a central spatial region of reduced states (upper branch), where temporal intermittent bursting and chaos occurs, and two smooth regions of oxidized states (lower branch), close to the two boundaries, giving rise to a two front spatial solution. The interaction of the reaction and diffusion terms give rise to a variety of sustained patterns such as stationary periodic structures, nonlinear waves or chaotic spatio-temporal structures of large amplitude. We restrict our analysis to a two-front pattern which exhibits chaotic bursting in the central reduced state region.

For future reference we denote $\mathbf{u}=(u,v)$ as the solution of (1) and we write Eq. (1) in compact form:

$$\frac{d\mathbf{u}}{dt}=\mathbf{RD}(\mathbf{u}).$$

B. Derivation of the Karhunen–Loeve model

In this section we obtain a low-dimensional model for the process described by (1). We start with a chaotic set of data which is a solution of (1). The spatio-temporal pattern for the u-variable is shown in Fig. 1. Notice the two regions in space of central reduced states and oxidized states near the boundaries. A similar spatial distribution occurs in the v-variable. The temporal chaos occurs primarily in the central region and is nonhomogeneous in space. In Fig. 2 we show the time series for the same solution at $x=\frac{1}{2}$. The data consists of large bursts separated by reinjection to small exponentially growing solutions, followed by small amplitude chaos. The time duration of the small amplitude chaos prior to the burst appears to be random. An approximate discrete model may be constructed as a one-dimensional map by plot-

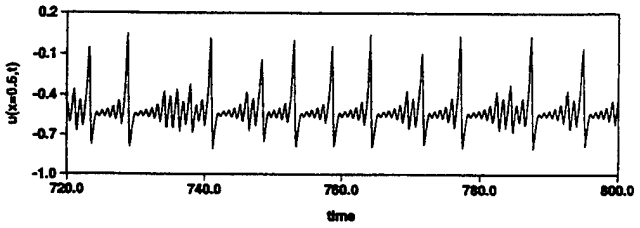


FIG. 2. The time series at $x=\frac{1}{2}$ for the pattern in Fig. 1.

ting successive maxima against each other.[8] This process works well for the bursts, but is too low-dimensional to predict the small amplitude chaos. A more realistic approach to model a process such as intermittent bursting requires at least three dimensions, and due to the small amplitude chaos possibly more. For our model, one will therefore retain the five dominant modes of the KL expansion. This data is decomposed in Karhunen–Loeve modes from which we retain the five highest mean-square energy modes. These modes contain about 95% of the total energy and they represent an optimal representation for this solution in the sense that they approximate this solution with minimal least square error. The basis functions generated in this way will be used in a Galerkin procedure to generate a dynamical system. As an example of the compactness of KL modes compared to Fourier mode decomposition, we show in Fig. 3 the Fourier modes contained in the two highest energy KL modes. Notice that there are multiple harmonics contained in each KL mode, showing that KL modes derived from the data are more efficient in capturing information about the spatial structure.

Following Ref. 13, the database for performing the KL procedure is an ensemble of M time snapshots of the solution $\mathbf{u}=(u,v)$, where u is represented in Fig. 1, at uncorrelated times $\{t_n\}$ which we denote by

$$\{\mathbf{u}^n(x)\}=\{\mathbf{u}(x,t_n)\}_{n=1,M}. \tag{2}$$

The vectors in (2) can be written explicitly as

$$\mathbf{u}(x,t_i)=(\mathbf{u}(x_1,t_i),\ldots,\mathbf{u}(x_p,t_i)),$$

for fixed mesh points $x_1,\ldots,x_p$. The highest energy KL mode will be the one structure which, when properly normalized and projected onto the snapshots in (2), yields the maximum mean square energy. Repeated application of this procedure yields a complete hierarchy of structures ranked by their mean-square energy content which constitute the KL basis functions, $\{\psi^{(n)}(x)\}$. Such a set of eigenmodes describe, in a mean-square sense, an optimal coordinate system. If we approximate the solution $\mathbf{u}$ by $\mathbf{u}_N(x,t)=\Sigma_{k=1}^{N}a_k(t)\psi^{(k)}(x)$, where $a_k(t)$ are defined by $a_k=(\psi^{(k)},\mathbf{u})$ then, as shown in Ref. 13, the KL procedure determines a complete, orthonormal set of basis functions $\{\psi^{(n)}(x)\}$. If we denote the error by $\epsilon_N(t)=\|\mathbf{u}-\mathbf{u}_N\|^2$, then the time average over the entire data set, denoted by $\langle\epsilon_N\rangle=\langle\|u-u_N\|^2\rangle$ is minimized, for any N over all possible sets of orthonormal functions $\psi(x)$.

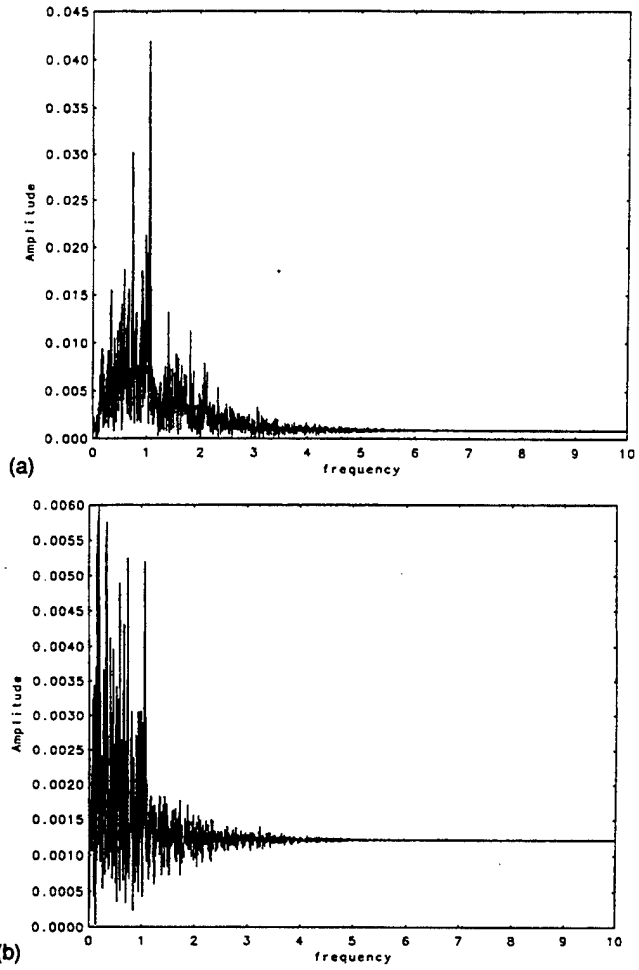


FIG. 3. (a) Fourier mode expansion of the first KL mode derived from the time series in Fig. 2. (b) Fourier mode expansion of the second KL mode derived from the time series in Fig. 2.

It has been shown in Ref. 13 that the KL basis functions are eigenfunctions of the kernel K, defined by

$$\mathbf{K}(x,x')=(K_{ij})=(\langle\mathbf{u}(x,t)\mathbf{u}(x',t)\rangle)$$
$$=\frac{1}{M}\sum_{i=1}^{M}\mathbf{u}_i^T(x)\mathbf{u}_i(x'), \tag{3}$$

where $\langle\cdot,\cdot\rangle$ denotes time average. Thus the basis functions which represent the solution $\mathbf{u}=(u,v)$ with the smallest average least square error, are the eigenfunctions of the correlation matrix (3). The eigenfunctions determined in this way are called empirical eigenfunctions or coherent structures. We remark that this basis is determined by the spatio-temporal data $\mathbf{u}=(u,v)$ itself. The eigenmodes thus obtained are statistically orthogonal:

$$\langle a_n(t)a_m(t)\rangle=\lambda_n\delta_{nm}, \tag{4}$$

where the left term is a time average. So, in this representation, the Fourier coefficients $a_n(t)$ are decorrelated. As a result, when projecting the data (2) back onto each eigenmode we get uncorrelated time series from each mode, representing statistically independent phenomena. We remark

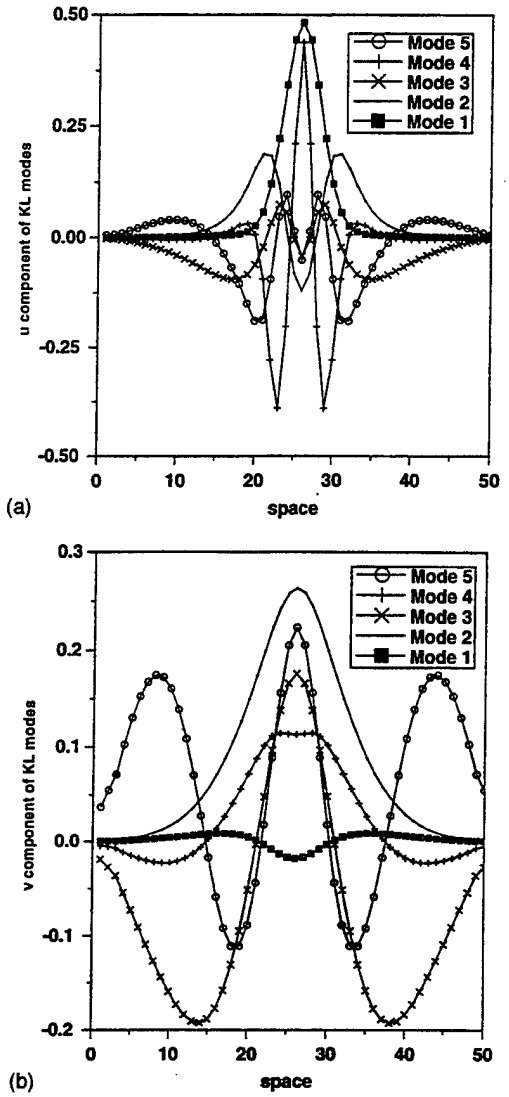


FIG. 4. (a) The five-highest energy KL modes for u. (b) The five-highest energy KL modes for v.

here that an alternate method to compute the KL eigenvectors is the method of snapshots due to Sirovich.[9] When the data is spatially one-dimensional this method is equivalent to the regular KL procedure. In higher dimensions it has the advantage that the resulting eigenfunction determination is one-dimensional. It can be shown[9] that the correlation matrix $\mathbf{K}$ has eigenvectors of the form

$$\Psi=\frac{1}{M}\sum_{k=1}^{M}\alpha_k(t)\mathbf{u}_k(x),$$

where the coefficient $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_M)$ is determined by requiring that

$$\mathbf{C}\alpha=\lambda\alpha,$$

where $\mathbf{C}=(C_{ij})$:

$$C_{ij}=(\mathbf{u}_i\cdot\mathbf{u}_j).$$

where the dot stands for dot product.

Using the data from Fig. 1 to generate the snapshots $\{\mathbf{u}^n(x)\}$, the KL modes for u and v were computed and are shown in Figs. 4a and 4b. Notice that the first mode in both cases picks up information about the central region in Fig. 1. This is because the central region contains the most active dynamics. Higher modes then pick up information about the dynamics of the boundary layer. That is, although much lower in energy, the dynamics of the layer separating the central region from the diffusion dominated region has an active structure. In terms of energy content, it can be seen that the first two modes contain 94.9% of the energy, while the other three contain less than 0.5×10^{-4}. The effect of the higher order modes in reconstructing the profile from the ODE model will be discussed below.

In Fig. 5a we show the projection of the solution $\mathbf{u}$ on the first two KL eigenmodes. If we compare this with Fig. 2, we see that the projection on the first KL mode carries information about the u concentration variable, whereas the second KL mode contains information about the v-variable. In Fig. 5b the projection of the data on the other three modes is shown carrying considerably less contribution to the chaotic pattern, as it is expected from the amount of energy contained in these modes.

The eigenfunctions discussed above can be used to generate a dynamical system from the equations (1). To this end we express the solution in the form

$$\mathbf{u}(x,t)=\sum_{n=1}^{5}a_n(t)\psi_n(x). \tag{5}$$

This is used in a Galerkin procedure;[13] i.e., we project (5) and the equations (1) onto a subset of the eigenmodes. This yields the equations

$$\left(\psi_m, RD\left(\sum_{n=1}^{5}a_n(t)\psi_n(x)\right)\right)=0,\quad m=1,\ldots,5. \tag{6}$$

Equations (6) represent a system of five ordinary differential equations for the amplitudes a_n of the modes. These equations together with the expansion (5) are our low-dimensional model whose dynamics will be explored in the next section. The parameters D, α and ϵ are all represented explicitly in Eq. (6), so we can evaluate the dynamics as a function of parameters.

III. ANALYSIS OF THE KARHUNEN–LOEVE MODEL

In this section we describe the dynamics of the Karhunen–Loeve model and compare it to the dynamics of the nonlinear model (1). The system (1) exhibits extreme sensitivity of the solution with respect to the diffusion coefficient D as well as with respect to α and ϵ. Changes in the 7th significant digit are sufficient to change the stability of the attractors. This makes it difficult to match accurately the parameters at which certain attractors occur in the two models. So in what follows we will compare the two models qualitatively.

For the partial differential equations (1), when the diffusion coefficient is low enough one notices a two-front pattern

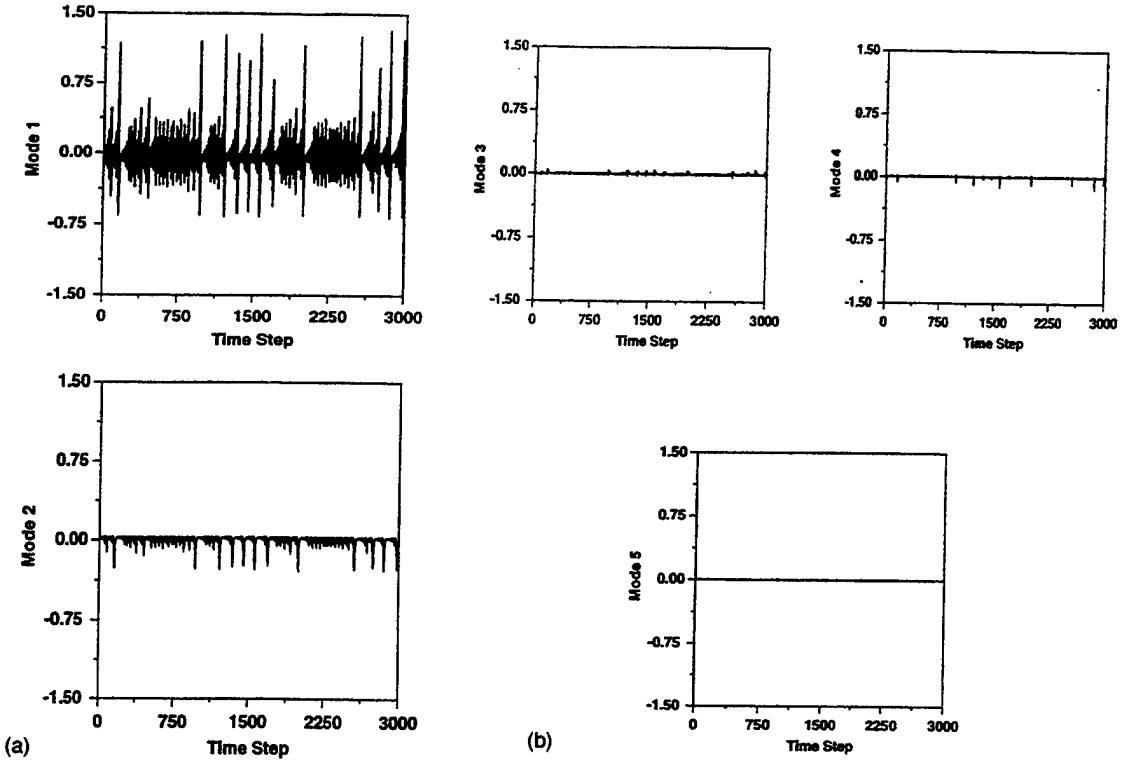


FIG. 5. (a) Projection of the highest energy KL modes back on the solution which generated them. These mode amplitudes recover the dynamics of the u and v variables, respectively. (b) Projection of the lower energy KL modes back on the solution which generated them. Much less structure is carried by the these three modes.

for the concentrations u and v. Two diffusion dominated fronts at the boundaries separate a central region where reaction dominates (Refs. 7 and 8). Spatially the solutions are smooth, but nonhomogeneous. As D is varied, the variation in time in the reaction dominated region exhibits small amplitude periodic behavior, and chaotic or periodic intermittent bursting.

In the case of the KL model we find that this type of spatio-temporal pattern is preserved remarkably well. To compare the types of dynamical behavior between the reaction-diffusion and KL models, we labeled the observed bursting patterns. We denote by L_mS_n a periodic pattern formed by m large amplitude bursts and n small-amplitude bursts. In Fig. 6a we show a spatio-temporal pattern where the variable u was reconstructed using the first three modes with amplitudes determined by Eq. (6). Ten periods are shown for an L_1S_1 pattern. Notice that using only three modes captures the spatially distinct reaction and diffusion dominant regions.

In Fig. 6c we reconstruct the variables u and v using the first four modes. Notice the presence of the same smooth regions separated by a bursting regime. In Figs. 6a and 6b, the bursting is periodic, large bursts being separated by small amplitude bursts, behavior that we denote as L_1S_1. In space, for both the reaction-diffusion and KL models, the concentrations u,v increase from the boundaries toward the central region, as a result of chemicals diffusing from the boundaries. Due to extreme sensitivity in parameters, when passing from one model to another, it would be difficult to compare amplitudes in the examples below. We notice however, that in both models the ratio of the u and v variables is preserved, the v variable (not shown) has amplitude one order of magnitude smaller than the u variable. One also sees that the ODE model (6) preserves the excitable character of the dynamics which in the PDE model is due to the existence of the "S"-shaped slow manifold, $v=f(u)$. In the solution represented in Fig. 1, the boundary data was chosen on the lower branch and the parameter $\alpha=0.1$ was taken to generate states on the upper branch of this manifold. In this way a concentration gradient appears (even with uniform feeding from the boundaries) which generates a variety of bursting patterns in the central region for different values of the diffusion coefficient D.[7] In the lower-dimensional model (6) we also found that intermittent bursting patterns appear as well, as the parameter D is varied, which we discuss below.

Analyzing further the spatial profile in the ODE model, we show in Fig. 6b a transversal section of this pattern which agrees with the shape of the solution of the RD model in the intermittent bursting regime. The bell shaped profile is preserved with just three modes. (We remark that by using only the first two KL modes to reconstruct the solution essentially the same pattern is observed.) For completness we show in Fig. 6c the pattern generated when the first four modes are used, along with the corresponding spatial profile in Fig. 6d. The four mode reconstruction gives similar behavior as the three mode model in that there are distinct diffusion and reaction dominated regions. However, in the four mode case, the boundary layer is amplified in the KL model. We con-

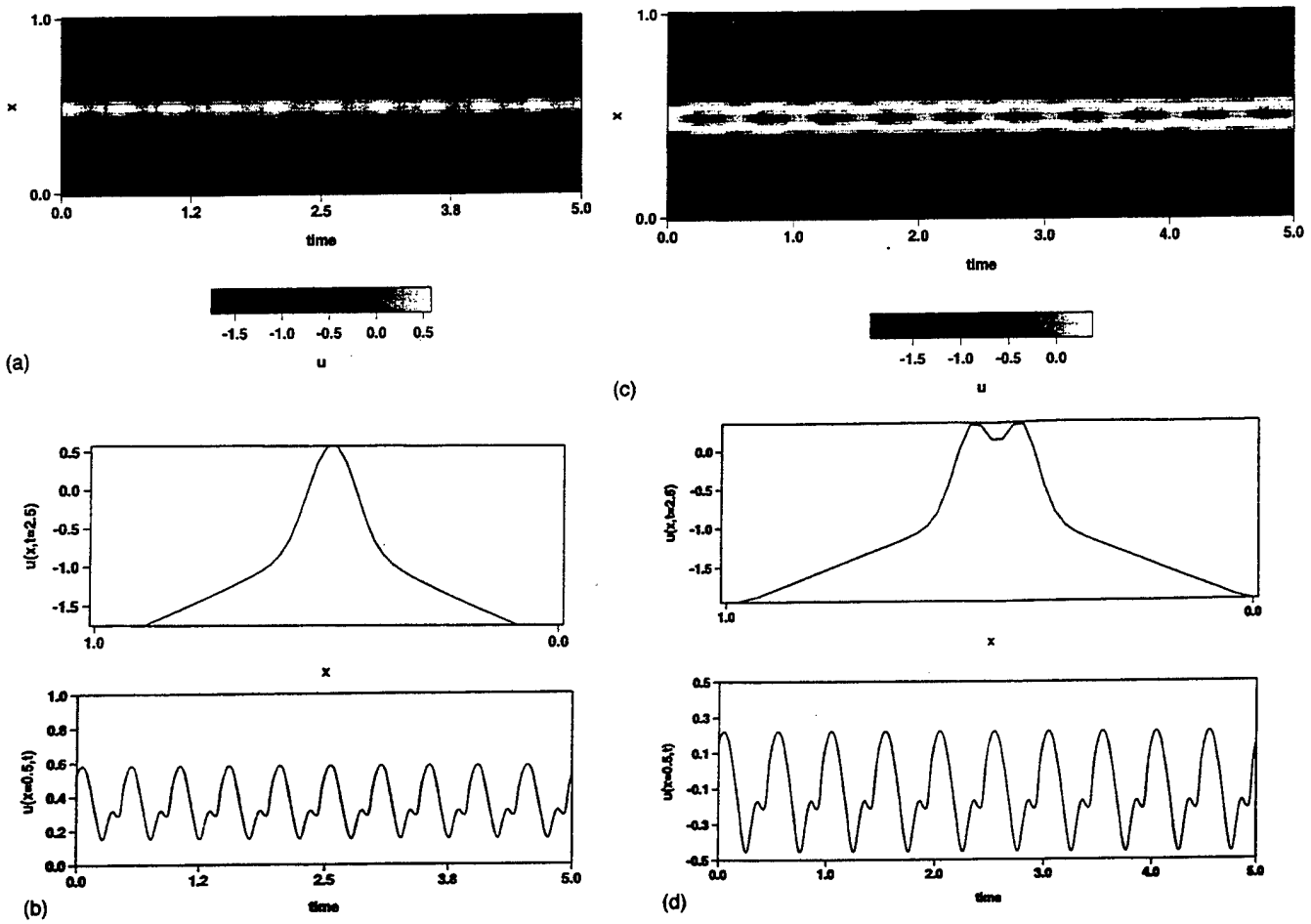


FIG. 6. (a) Bursting spatio-temporal pattern L_1S_1 obtained using the first three KL modes with amplitudes derived from the KL model. Parameter values are the same as in Fig. 1. (A color version of this figure is available via E-PAPS: http://www.aip.org/epaps/epaps.html.) (b) Spatial and temporal profiles for the pattern in (a). (c) Bursting spatio-temporal pattern obtained using the first four KL modes with amplitudes derived from the KL model. (A color version of this figure is available via E-PAPS: http://www.aip.org/epaps/epaps.html.) (d) Spatial and temporal profiles for the pattern in (c).

clude that the dynamical spatial structure is well represented by the three highest KL modes since three is the largest number to give a correct profile. In Fig. 4a we see that spatially the first mode contains a lot of the structure in the u-variable whereas the second mode (Fig. 4b) carries information on the v-variable. Referring to Fig. 5 we see that the behavior in time of the u-variable is governed by the amplitude of the first modes whereas the behavior in time of the v-variable is faithfully represented by the amplitude of the second mode. The KL modes distinguish the dynamics of the two concentration variables as well as their temporal and spatial behavior. Therefore the KL modes account for separate features in the dynamics. Similar insights using KL analysis on an array of semiconductor lasers were found in Ref. 17. The number of KL modes needed to capture the essential dynamics can thus be regarded as the number of fundamental degrees of freedom in a complex spatio-temporal pattern. In Ref. 18 the authors found that for an experiment on Rayleigh-Bénard convection the number of needed KL modes containing most of the energy is about one integer larger than the fractal dimension. In our case, three modes is close to the information dimension of a time series extracted from this pattern which is 2.1.[8]

The data used to perform KL expansion presented in Fig. 1 represents the solution of (1) at fixed parameter values. It is customary in this kind of procedure[13] to compare the two models for values of the parameters close to the values that generated the original data. In our case, due to extreme sensitivity of the solution to changes in the parameters, we look for qualitative agreement throughout the phase-space corresponding to the diffusion coefficient D. We found that certain dynamical features of the reaction-diffusion model can be seen in the KL model as described below. We examine the bifurcations of bursting patterns by using tracking or continuation methods[19–21] for the KL model to follow solution curves as parameters are varied in a similar range of values as those considered in Ref. 7. In particular, we consider the diffusion coefficient D, where [0,0.04].

In Fig. 7 we show a solution curve for the system (6) with respect to the diffusion coefficient. We represent in this figure the L^2-norm of a solution (the blue curve) which initially has period one and period doubles twice (the red and

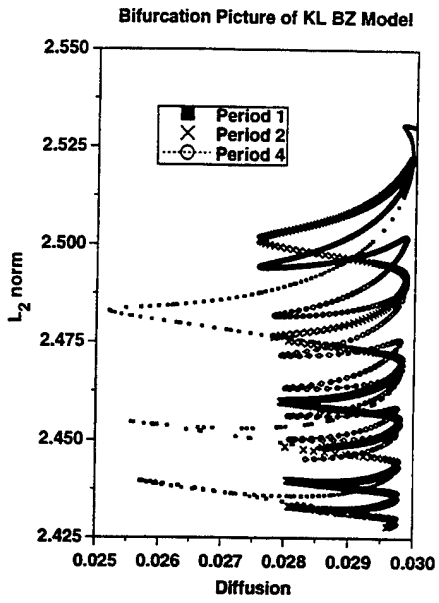


FIG. 7. L_2-norm of solution curves for the KL model emanating from a period one solution curve (the □ curve) for the KL model, which period-doubles twice as diffusion is decreased. The limit points on the right of the □ curve are saddle-nodes where L_1S_k $k=\{1,2\}$ orbits emerge. Limit points on the right of the period-two branch (the × curve) represent the onset of solution branches of type L_2S_k, $k=\{1,2,3,4\}$. (A color version of this figure is available via E-PAPS: http://www.aip.org/epaps/epaps.html.)

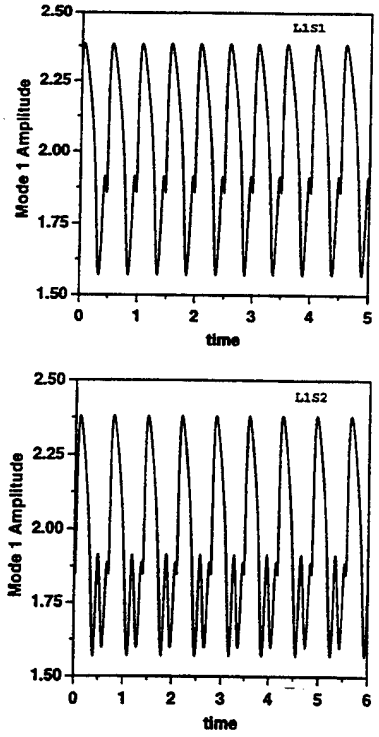


FIG. 8. The highest energy KL mode is shown for orbits of type L_1S_1 and L_1S_2 along the blue solution curve in Fig. 7.

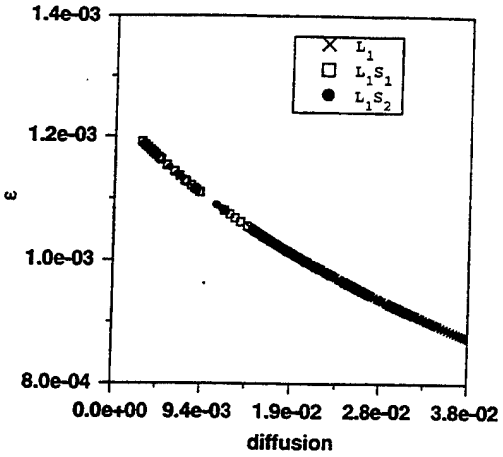


FIG. 9. Onset of intermittent behavior of type L_1,L_1S_1 and L_1S_2 in the $\epsilon-D$ parameter space.

green curves). At the top of the blue curve the solution has only period one large bursts (followed by no small amplitude oscillations). As we follow the blue path in Fig. 7, at each limit point on the right of the curve one small spike in the motif of the pattern is added. This behavior in the solution is reflected in the amplitudes of the KL eigenmodes, shown in Fig. 8, which correspond to L_1S_1, and L_1S_2 patterns along the blue solution curve in Fig. 7. In Fig. 9 we graphed the curve of saddle-node bifurcation points which are the onset of intermittent bursting of type L_1S_k ($k=1,2$) in the space of the parameters ϵ and D. Onset here refers to a saddle-node bifurcation point opening to the left. Below the curve, L_1S_k patterns are observed, while no L_1S_k bursting is observed above the curve.

Returning to Fig. 7, as diffusion is decreased along the blue curve, the period one solution gives rise to a period-two solution, in which the basic motif has two large bursts. That is, there are no small amplitude oscillations. Tracking this solution as a function of D we obtain the red curve in Fig. 7, where as before at each limit point on the right an extra small spike is added, giving rise to bursting sequences of the form L_2S_k. The corresponding amplitudes of the first KL mode along this solution curve are shown in Fig. 10. Similar behavior along the green curve is observed where each limit point on the right denotes the onset of a branch of orbits of the form L_4S_k.

As we further decrease D along the blue curve, the solution undergoes a period-doubling cascade leading to small-amplitude chaotic behavior. In Fig. 11 the time series for the amplitude of the first mode inside the chaotic regime is shown. In this case, there are no bursts. In Fig. 12 we present a bifurcation diagram for the solution described above. On the figure we draw the maxima and minima of the first KL mode. For the first KL mode, the minima clearly oscillate chaotically, and go through a period doubling sequence. In Ref. 7, for the partial differential equation model, period-doubling transition to small amplitude chaos was observed. (The small amplitude chaos for the PDE has dimension approximately equal that of the KL model chaos, which

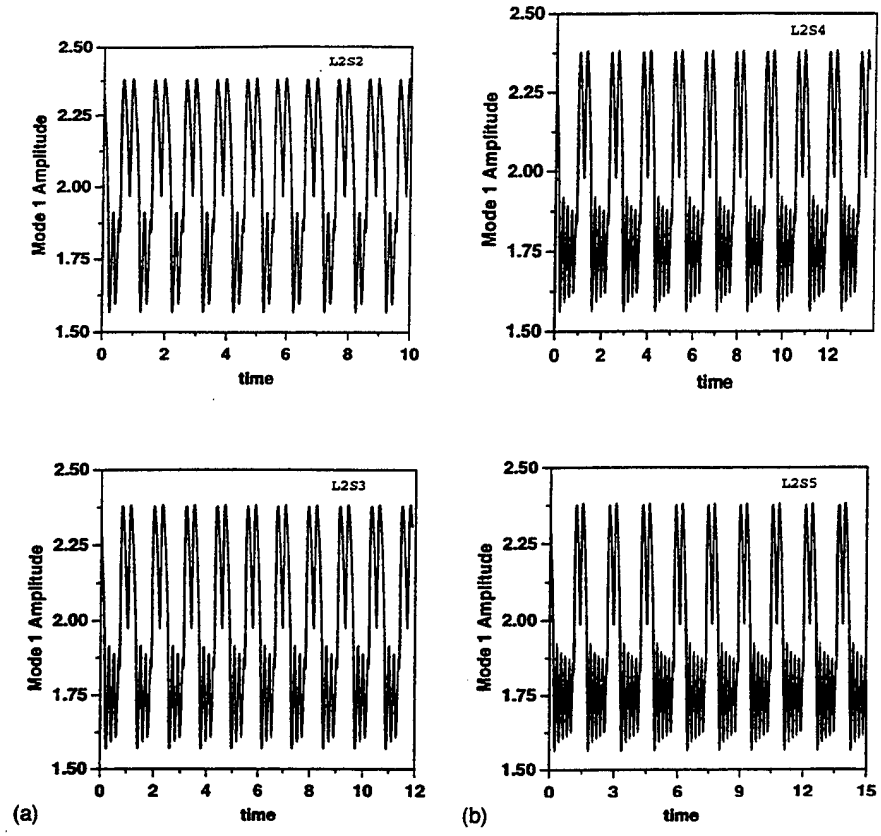


FIG. 10. (a) The highest energy KL mode is shown for orbits of type L_2S_k for $k=\{2,3\}$ along the red solution curve in Fig. 7. (b) The highest energy KL mode is shown for orbits of type L_2S_k for $k=\{4,5\}$ along the red solution curve in Fig. 7.

was found to be 1.1. Bursting chaos, as shown in Fig. 1, found in the PDE, but not in the KL model, was found to have a dimension of 2.1.[8]) So the KL model recovers small amplitude chaotic behavior, as well as the mechanism of transition to chaos observed in the partial differential equation. This is not always the case when using the KL procedure. In Ref. 16 the authors analyzed a one-dimensional ionic Brusselator reaction-diffusion system forced with an electric field. In that case the KL procedure accurately captures oscillatory and chaotic dynamics but does not give a clear insight in the mechanism for transition to chaos.

For the PDE model an alternating sequence of chaotic (C_1^m) and periodic (L_1S^m, where $m=7$) bursting patterns was found as the diffusion coefficient is decreased.[7] For the KL model we have found a similar sequence of periodic patterns and we conjecture that the corresponding windows of chaos still exist but are over such a small parameter range

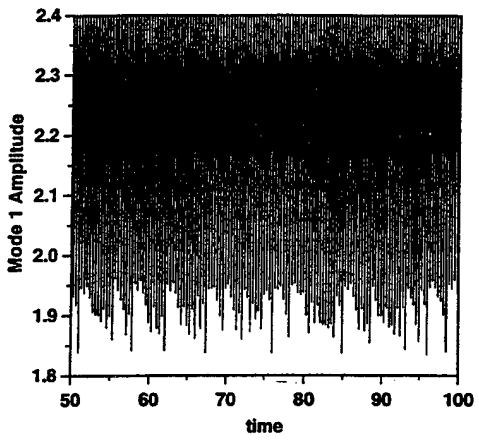


FIG. 11. Amplitudes of the highest energy KL modes inside a small amplitude chaotic regime.

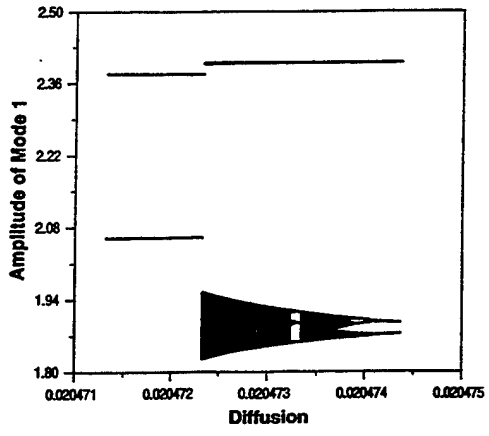


FIG. 12. Bifurcation diagram for the amplitude of the highest energy KL mode. On the picture we represent the maxima and minima of the amplitude and notice that the minima undergo a period-doubling cascade to chaos.

that they could not be detected. In addition, for the KL model we identified a sequence of periodic bursting patterns of type L_2S_m and L_4S_m. Other sequences of the type L_jS_k, where j follows windows of periodic behavior in the cascading sequence are also observed. We do not display them here.

In Ref. 7, the authors have shown the existence of chaos in the PDE model that satisfies the symbolic dynamics of Shil'nikov homoclinic chaos. They conjectured (Refs. 7, 22, and 23) the existence of a cascade of saddle-node bifurcations originating the oscillating pattern of type L_1S_m (where m is less than or equal to 7), along with a cascade of subharmonic bifurcations giving the C_1^m behavior, both of which accumulate at a locus of homoclinicity. The above discussion shows that in our KL model we see some of the predicted saddle-node bifurcations originating L_1S_1 and L_1S_2 behavior as predicted in Ref. 7.

IV. CONCLUSIONS

We have explored the bifurcation structure of a low-dimensional model for a reaction-diffusion system, based on performing Karhunen-Loeve mode expansion on the chaotic spatio-temporal solution of a partial-differential equations model. The KL model simulates remarkably well the two-front pattern of the solution observed in the nonlinear model consisting in two diffusion dominated regions close to the boundaries and a reaction-dominated region in the center. We have found that the global behavior in the nonlinear model is partly described by the dynamics in the KL model. Intermittent bursting and small amplitude chaos can be observed in both models. Moreover in both models small amplitude chaos appears as a result of a period-doubling cascade as the diffusion parameter is continuously varied. The bursting we identified in the KL model appears due to a sequence of saddle-node bifurcations which behave like the sequence of L_1S_m orbits converging to a point of homoclinicity which was conjectured in Ref. 7. This phenomenon may still take place in the KL model over a very narrow parameter window which would make it difficult to detect. Bursting patterns not seen in the PDE model were observed in the low-dimensional model. Thus the KL model gives an accurate qualitative description for part of the phase-space of the original model and gives insight in the mechanism which generates the observed patterns in the PDE model.

ACKNOWLEDGMENTS

The authors are very grateful to Dr. Larry Sirovich and Dr. Robert Handler for fruitful discussions on the Karhunen–Loeve procedure.

[1] J. Elezgaray and A. Arneodo, J. Chem. Phys. **95**, 323 (1991).
[2] Q. Ouyang, V. Castets, J. Boissonade, J. C. Roux, P. DeKepper, and H. L. Swinney, J. Chem. Phys. **95**, 351 (1991).
[3] A. Arneodo and J. Elezgaray, Phys. Lett. A **143**, 25 (1990).
[4] J. Elezgaray and A. Arneodo, in *New Trends in Nonlinear Dynamics and Pattern Forming Phenomena: The Geometry of Nonequilibrium*, edited by P. Coullet and P. Huerre (Plenum, New York 1990), p. 21.
[5] A. Arneodo and J. Elezgaray, in *Spatial Inhomogeneities and Transient Behavior in Chemical Kinetics*, edited by P. Gray, G. Nicolis, F. Baras, P. Borckmans, and S. K. Scott (Manchester University Press, Manchester, 1990), p. 415.
[6] A. Arneodo and J. Elezgaray, in *Chemical Waves and Patterns*, edited by R. Kapral and K. Showalter (Kluwer Academic, Dordrecht, 1995), p. 517.
[7] J. Elezgaray and A. Arneodo, Phys. Rev. Lett. **68**, 714 (1992).
[8] I. Schwartz and I. Triandaf, Phys. Rev. E **50**, 2548 (1994).
[9] L. Sirovich, Q. Appl. Math. **XLV**, 3561 (1987).
[10] L. Sirovich, K. S. Ball, and R. A. Handler, Theor. Comput. Fluid Dyn. **2**, 307 (1991).
[11] L. Sirovich, B. W. Knight, and J. D. Rodriguez, Q. Appl. Math. **XLVIII**, 3535 (1990).
[12] J. L. Lumley, in *Atmospheric Turbulence and Radio Wave Propagation*, edited by A. M. Yaglom and V. I. Tatarski (Nauka, Moscow, 1967), p. 166.
[13] J. D. Rodriguez and L. Sirovich, Phys. Rev. A **120**, 211 (1987).
[14] D. S. Broomhead and G. P. King, Physica D **20**, 217 (1986).
[15] Q. Ouyang, V. Castets. J. Boissonade, J. C. Roux, P. De Keeper, and H. L. Swinney, J. Chem. Phys. **95**, 351 (1991).
[16] P. Hasal, A. F. Munster, and M. Marek, Chaos **4**, 531 (1994).
[17] D. Merbach, O. Hess, H. Herzel, and E. Scholl, Phys. Rev. E **52**, 2 (1995).
[18] S. Ciliberto and B. Nicolaenko, Europhys. Lett. **14**, 303 (1991).
[19] E. J. Doedel, A. D. Jepson, and H. B. Keller, "Numerical methods for Hopf bifurcation and continuation of periodic solution paths," in *Computing Methods in Applied Sciences and Engineering VI*, edited by R. Glowinski and J. L. Lions (North-Holland, Amsterdam, 1984), pp. 127–136.
[20] I. Schwartz and I. Triandaf, Phys. Rev. A **46**, 6189 (1992).
[21] W. C. Rheinboldt, University of Maryland, Technical Report No. ICMA-79-04 (unpublished).
[22] P. Gaspard, R. Kapral, and G. Nicolis, J. Stat. Phys. **35**, 697 (1984).
[23] P. A. Glendenning and C. Sparrow, J. Stat. Phys. **35**, 645 (1983).

Appendix FF

From Chemical Structure to Viscoelastic Properties of Polymers

Macromol. Symp. 90, 95-129 (1995)

FROM CHEMICAL STRUCTURE TO VISCOELASTIC PROPERTIES OF POLYMERS

K.L. Ngai*
Naval Research Laboratory, Washington, DC 20375-5320

Kwok Yeung Tsang
Naval Research Laboratory, Washington, DC 20375-5320
Science Application International Corp., McLean, VA 22102

Abstract: The basic result of the *coupling model* consists of a cross-over from $\exp-(t/\tau_o)$ to $\exp-(t/\tau^*)^{1-n}$ at a temperature insensitive microscopic time t_c. We indicate how this basic result can be derived from chaotic, interacting Hamiltonian systems which include densely packed polymer molecules. Recent quasielastic neutron scattering experiments and molecular dynamics simulations are discussed and the results are shown to support this result as well. An application of the *coupling model* to find how the viscoelasticity of a polymer depends on the chemical structure of the monomer through the coupling parameter of the local segmental motion is given to illustrate the utility of the model.

INTRODUCTION

The coupling model was introduced fifteen years ago to deal with the cooperative dynamics of relaxation in densed packed interacting systems (Refs.1-4). It is based on notions that were, and still are, not familiar to workers interested in the relaxation of various kinds of complex materials. Being formulated in a general manner, the coupling model is deemed to be applicable to many different fields, including complex fluids and polymer viscoelasticity which are the subjects of interest to the majority of the readers of this volume. The first version of the coupling model (Refs.1-4) was based on a quantum mechanical approach which might have estranged workers of relaxations in systems which classical mechanics suffice for a description. In retrospect one of the authors (KLN) finds that this quantum approach is much

 CCC 1022-1360/95/$ 04.00

easier to formulate than the corresponding classical counterpart as we shall discuss below. Thus without first doing the problem by the easier though not the most appropriate way, serendipidy would not have intervened and the coupling model would not have been discovered. Fortunately at this early times there are two saving graces of the first approach that enabled the author to go on further.

First the results of this quantum mechanical approach do not contain $\hbar$ and therefore are immediately applicable to classical systems. Second the results offer predictions that one of the authors can use to falsify his own theory if the theory turns out to be inadequate. That important prediction that spawns many other predictions when applied to different fields is the "second relation".

$$\tau^* = [\,(1-n)\,\omega_c^{n}\tau_o\,]^{\frac{1}{1-n}} \qquad (1)$$

that relates the independent relaxation time τ_o to the measured slowed down cooperative relaxation time τ^*. Here ω_c is the cross over frequency and n is the coupling parameter that appear in the exponent of the Kohlrausch relaxation function (Refs.5,6),

$$C(t) = \exp-\left(\frac{t}{\tau^*}\right)^{1-n} \qquad (2)$$

which is also predicted and called the "first relation" of the coupling model. The fact that the Kohlrausch function gives a good though by no means perfect description of the time dependence of many relaxation processes is comforting, but this is not enough to convince one of the authors in the early stage of development of the research. For him it would be far more convincing to find examples in which the second relation is also verified simultaneously. At that time the exact magnitude of the crossover frequency ω_c is not known but the theory says that it is temperature independent and depends only on the interaction Hamiltonian in these densed packed systems. Nevertheless there are systems in which the dependence of τ_o on a variable U, $\tau_o \sim f(U)$, is known and we can critically test the predicted dependence of τ^* on the same variable U:

$$\tau^* \sim [\,f(U)\,]^{\frac{1}{1-n}} \qquad (3)$$

with n determined from fitting the dispersion of the relaxation process by Eq.2. (Refs.7-12) If the experimental data have indeed the predicted U-dependence given by Eq.3 then naturally this provides strong indication that the coupling model may be correct afterall. We found several such remarkable examples (7-12) in the early days that gave some confidence for

continuing this pursuit further. When starting out one of the authors had very little familarity with any of the fields now he works nowadays. Much time has been spent on learning the problems and phenomenologies of each field before important application of the coupling model can be made. As time progresses repeated successful applications (by a combination of Eqs.1 and 2) have been discovered and by now the relevance of the coupling model for describing the dynamics of interacting or cooperative systems is no longer in doubt (see most recent review, Ref.12). Colleagues who are familiar with these applications and recognize the achievements of the coupling model have inquired into the theoretical basis of the model.

However with much time and energy spent on seriously applying the coupling model to many areas, the task of strengthening the theoretical base has not been attended to as much as the author wishes. There are several other intermittent attempts (Refs.13-15) to derive the coupling model theoretically by other methods different from the original ones (Refs.1-4). But admittedly these later efforts are even less fundamental than the original approach and they all lack punch. The stagnation encountered in the theoretical front is perhaps not too surprising because the problem at hand is a difficult one and cannot be easily solved on a fundamental level, like starting from the Hamiltonian of the system. The right theoretical approach was actually recognized a year after the publication of the first version of the coupling model in 1979. The original version was based on energy level spacing distribution that has the form given by the Gaussian Orthogonal Ensemble (GOE) from random matrices (Refs.16-20).In 1980 M.V. Berry (Refs.21,22) has shown that an nonintegrable (chaotic) Hamiltonian in classical mechanics when quantized will have energy levels distributed according to the GOE. Berry's finding has subsequently been shown to be true repeatedly in many Hamiltonian systems (Ref.23). Since the results of the original coupling model was obtained from the GOE in a quantum approach, in view of Berry's finding it is natural to expect that the same results will be obtained in a classical approach from chaos in phase space of the Hamiltonian. It is easy to surmise (Refs.24,25) that the ultimate fundamental theory that will provide the same predictions as the original coupling model must involve chaos, but it is difficult to construct the theory. Although today occurrence of chaos in physical science is known to be prevalent and has become a household word (26) and there are many popular books written on the subject, its mastery is limited to the experts. One of the authors was educated in the era in which the standard text of classical mechanics was written by Goldstein (Ref.27) in which chaos is a foreign concept. Thus there are plenty of reasons why the theory of the coupling model based on chaotic Hamiltonian, though conceived as early as the author had read Berry's work of 1980, has not been forthcoming until very recently (Refs.28-30). These recent works

can only be considered as a start in this new frontier research area of deriving the dynamics of relaxation in interacting systems from chaotic Hamiltonians.

The brief sketch of the history of development of the coupling model given above shows that the model brings with it interesting predictions and applications but also concepts foreign to the community of researchers in polymers, viscous liquids, vitreous ionic conductors, etc. The concepts and theoretical basis of the coupling model being foreign have hampered the understanding of the true meaning and appreciation of the value of the model by other workers. To make matters worse the papers are published in a variety of scientific journals, some of which may not be familiar to workers in the fields. The purpose of this article is to remove as much as possible the gap of understanding and provide a road map for the reader to find out for himself or herself the status of the theoretical foundation of the coupling model and its experimental verifications.

RELEVANCE OF A CHAOTIC HAMILTONIAN APPROACH

Earlier in the Introduction, in connection with the original version of the coupling model, a reason has been given for why we believe the chaotic Hamiltonian approach to the dynamics of interacting systems is the right choice. Here another reason will be given which is based on a common characteristic of most if not all interacting systems that we are interested to describe their dynamics. The part of the Hamiltonian that describes the interactions between the molecular moieties (molecules in small molecule viscous liquids and glasses, monomer units in polymers, and ions in vitreous ionic conductors) in these systems do not correspond to harmonic forces. For example the pairwise interaction between molecules in small molecules glass-forming liquids such as *ortho*-terphenyl (Ref.31) and the nonbonded interaction between CH_2 groups on different polyethylene chains and between those separated by more than three bonds on the same chain (Ref.32) modelled by the Lennard-Jones potential $V(r)=4\epsilon[(\sigma/r)^{12}-(\sigma/r)^{6}]$ are anharmonic. The presence of anharmonic interactions in classical mechanics has the consequence that the solutions to the Newtonian equations of motion cannot be always described as periodic trajectories in phase space resulting in chaotic behavior (33). Similarly the screened Coulomb interactions between ions in vitreous ionic conductors being anharmonic will also lead to chaos. Thus all the dense packed interacting systems that we are interested (polymers, complex fluids, glass-forming liquids, vitreous ionic conductors, etc.) are naturally chaotic in nature. Since the manifestation of chaos whenever it arises is very general irrespective of the origin and the system (Ref.33), we can expect universal properties to

follow. If indeed chaos has a profound effect on relaxation process then it is highly probable that the effect can be expressed as general laws with the same flavor as the predictions of the coupling model. Molecular dynamics simulations (MDS) have made tremendous progress in recent years that the results of these numerical experimental data describe very well relaxation processes in systems of our interest. As we shall discuss later on in another Section, the results of MDS are in accord with the basic feanures of the coupling model. The starting points of these MDS and the proposed theoretical approach to relaxation based on chaos have something in common, i.e. the same Hamiltonian which is chaotic. Therefore from this connection between MDS and chaos we have reason to believe that the proposed theoretical approach will likely to succeed as the computer experiments have shown.

It will not be an easy task to sort out the effects that chaos will have on the relaxation process. Although chaos or nonlinear dynamics is now an expanding field of research and many high power theoretical results have been derived by mathematical physicist, immediate applications of these results to physical problems like ours are mostly not practical. For our specific purpose of looking for the effects that chaos has on relaxation, there is little we can find in the current literature of nonlinear dynamics (i.e. chaos) that can offer immediate help in this problem. This research is in unchartered territory which means great opportunities for research but also the journey may be long and difficult. Seemingly there are two ways to attack the problem. One is entirely theoretical and general consideration of relaxation in a Hamiltonian system with nonintegrable interactions between the fundamental units that constitute the system. Attempts of this kind have been made (29), but these treatment still lack the rigor needed for a satisfactory solution. Let us first consider the simpler problem of relaxation in the same Hamiltonian system but *without* the nonintegrable interactions between the units. The solution of this simpler problem is already known from irreversible or nonequilibrium statistical mechanics (Refs.34,35) (which should be distinguished from classical nonlinear dynamics in the full phase space) and are described by Langevin equation, master equation or Fokker-Planck equation. The solution of these well known equations leads usually to an exponential coorelation function or relaxation function, $\exp-(t/\tau_o)$. However a survey of published literature shows that there are only a very few instances the same results have been derived from nonlinear dynamics for the simpler problem, giving us little to start with. The literature provides absolutely no guidance to how to proceed in solving the full problem after putting back the nonlinear interactions.

FERMI MAP AND THE FERMI-STADIUM MAP

The other way to attack the problem is to develop simpler prototype models that can be solved numerically without and with the nonlinear interactions between the units. One example without interactions is the Fermi map which was proposed by Fermi (Ref.36) to model a realistic problem of cosmic ray acceleration in which charged particles are accelerated by collisions with moving magnetic structures. In the model a particle of mass M bounces back and forth between a fixed dissipative wall on one side and on the other side at a distance of L an elastic wall oscillating sinusoidally with time described by $x = a\cos(\omega t)$, see Fig.1. The

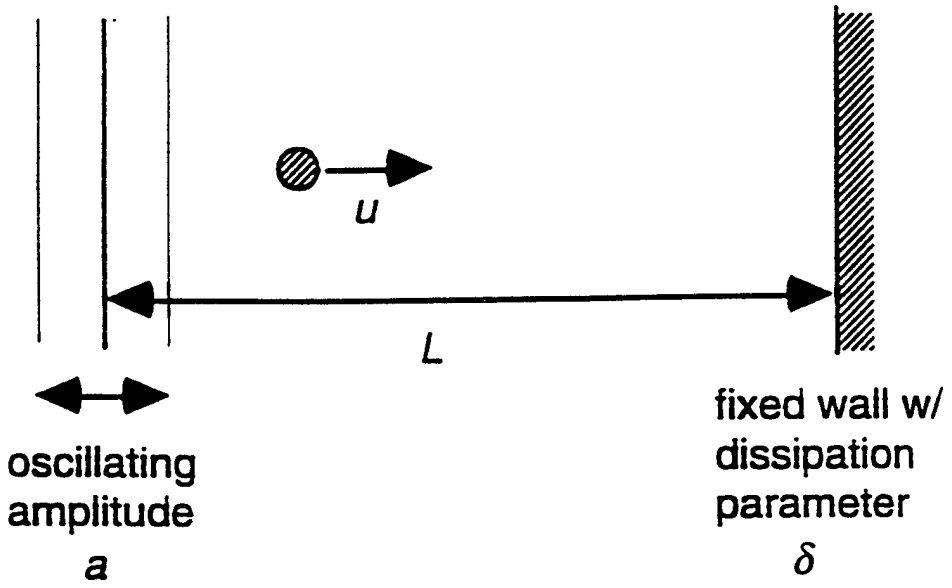


Fig. 1. One-dimensional dissipative Fermi map

motion of the particle can be described by the Fermi map. The map determines the velocity and phase of discrete time $n+1$ from the information at time n. Iterating maps makes the computation much simpler than solving differential equations, without losing relevant information of the system. Tsang and Lieberman (Refs.37-40) calculated this map to find that almost all initial distributions in velocities evolve to the invariant Maxwell distribution, $f(u)$–$\exp(-Au^2)$, where u is the velocity. The phase-averaged distribution evolves with time as a Markov process according to the Fokker-Planck equation. The normalized difference, Φ, between the mean energy and the equilibrium energy when plotted against time decays approximately exponentially with time. These results can now be found in a chapter of a standard text of nonlinear dynamics (33). Naturally it would be nice to put in many more particles and introduce nonlinear interactions between them and solve this many-particle

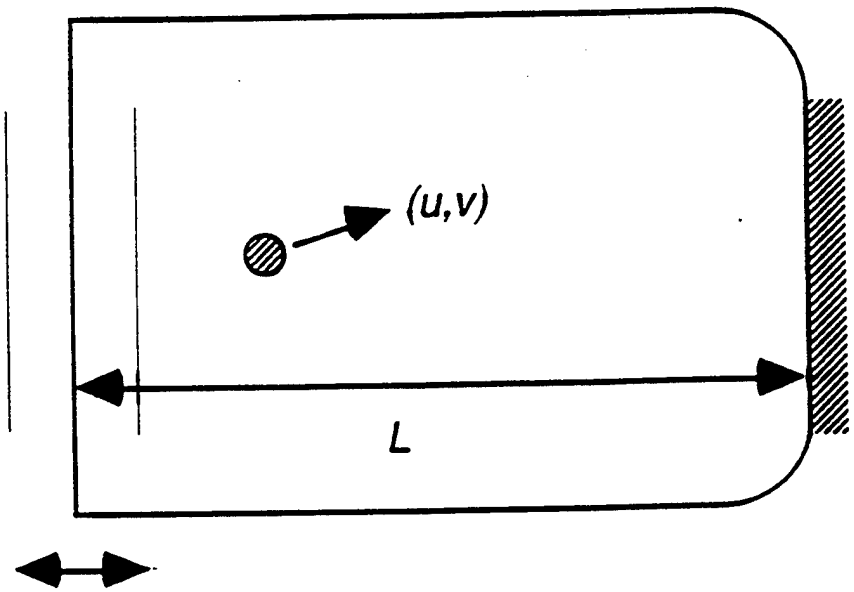


Fig. 2(a). Two-dimensional Fermi-stadium map

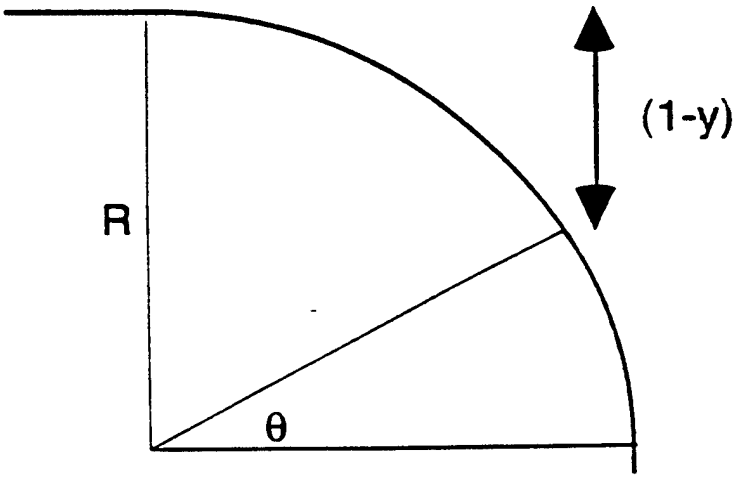


Fig. 2(b). A curved corner of the added stadium portion

generalized Fermi problem again. The calculation involved is however technically too difficult to perform at this time. There is another way of introducing additional nonlinearity instead of nonintegrable interactions into the Fermi map that can be solved at this time. It is well known in billiard systems (Refs.21-23, 41) that boundaries partly made up by arcs of circles will cause chaotic behavior in the motion of a particle. One such billiard system that is quite well known is the "stadium" whose boundary looks like a stadium in real life. This

fact has led us to first add two more elastic walls orthogonal to the two already present in the Fermi map making it two dimensional, and make two corners of the newly formed rectangle into circular arcs. The new system is a combination of a Fermi map and half a stadium (see Figs.2a and 2b). Naturally we call this new system the Fermi-stadium map. It is straightforward to construct the new map which includes also specular reflections from the additional elastic walls and the circular arc of the stadium (Ref.29).

First let us consider the effect of adding one more dimension to the Fermi map by adding two more elastic walls to the existing ones of the Fermi map. The width between the new walls is 2, which is much smaller than L. Although the motion of the ball is now two-dimensional, the Fermi system with the two horizontal walls added is just as simple as the one-dimensional Fermi map (since v remain constant forever) before we add nonlinearity to the system by replacing each of the two right-angled corners by quadrants of a circle with radius R which is smaller than 1 as indicated in Fig.2. The map is written as:

$$u'=(1-\delta)\,u_n+\sin\varphi_n$$

$$\varphi'=\varphi_n+2\pi M/u'$$

$$v'=v$$

$$y'=y+vL/u\ (\mathrm{mod}4)$$

The modulo 4 restricts y to be between -2 and 2, in view of the fact that one collision at the ceiling and another at the floor are equivalent to a translation of 4 in the coordinate y. For an extra collision, the value of y would be larger than 1 or smaller than -1. To adjust y again so that it is between -1 and 1, we do the following:

$$\text{if } y' > 1, \text{ then } y'' = 2-y;$$
$$\text{if } y' < -1, \text{ then } y'' = -2-y.$$

We then add the effect due to collisions at the curved corners. If

$$y'' \begin{matrix} >1-R \\ <-(1-R) \end{matrix},$$

then

$$\begin{pmatrix} u'' \\ v'' \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \mp\sin 2\theta \\ \pm\sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix},$$

where $\theta = \sin^{-1}[\ (|y''| - 1 + R) / R]$. The new values at the $(n+1)$st collision with the oscillating wall is then:

$$u_{n+1} = |u''|$$
$$v_{n+1} = v''$$
$$y_{n+1} = |y''|$$
$$\varphi_{n+1} = \varphi'.$$

To study the relaxation of the average energy in the Fermi-stadium map, we observe the time evolution of 10,000 initial conditions. The initial distribution $f_0(u)$ is such that $u_0 << 1$ and $v_0 << u_0$. In order to smooth the high frequency fluctuations in time, we use a running average over 50 timesteps. We then compare the relaxation in the Fermi-stadium map with different R's to that in the Fermi map.

Figure 3 shows the numerical result of the Fermi-stadium map with parameters δ=0.01, L and M =100, for various values of R (0, 0.02, 0.1, and 0.5). The normalized difference, Φ,

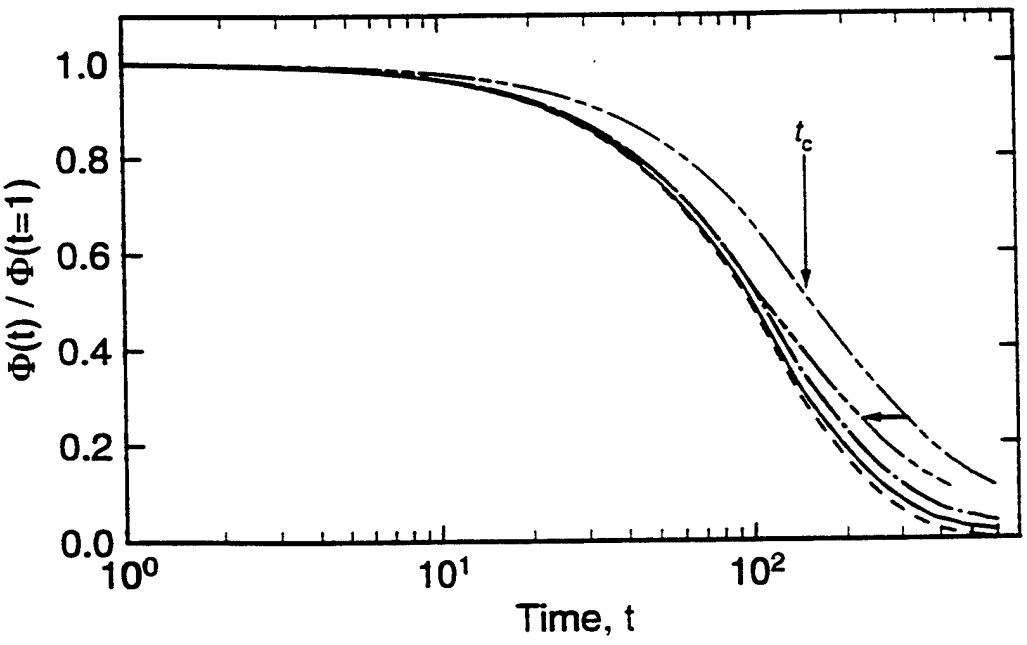


Fig. 3. Numerical result of the Fermi-stadium map with parameters δ=0.01, L and M =100, for various values of R (0, 0.02, 0.1, and 0.5, from left to right). The normalized difference, Φ, between the mean energy and the equilibrium energy, is plotted against time t. The arrow indicates a rescaling of the R=0.5 curve to show a longer coincidence with the Fermi curve.

between the mean energy and the equilibrium energy, is plotted against time t. The relaxation is shown retarded for different values of R. The larger R is, the more chaos is introduced and the more the retarding. Note that for R=0 (the dashed line at the bottom), the map is reduced to the Fermi map. At t=140, the solid curve, corresponding to R=0.02, begins to significantly deviate from the Fermi curve; at t=120, the next curve above, corresponding to R=0.1, begins to significantly deviate from the Fermi curve. For R=0.5, since the change from the Fermi map is substantial, we rescale the curve horizontally so that it coincides for a significant portion with the Fermi curve. The rescaling is shown by the arrow in Fig.3. At t=85, the curve, corresponding to R=0.5, begins to significantly deviate from the Fermi curve. One can observe a crossover time t_c (which depends on R), after which the Fermi-stadium curve deviates from the Fermi curve.

To study the nature of the relaxation, we plot Log(-Log Φ) against t in Fig. 4. After a

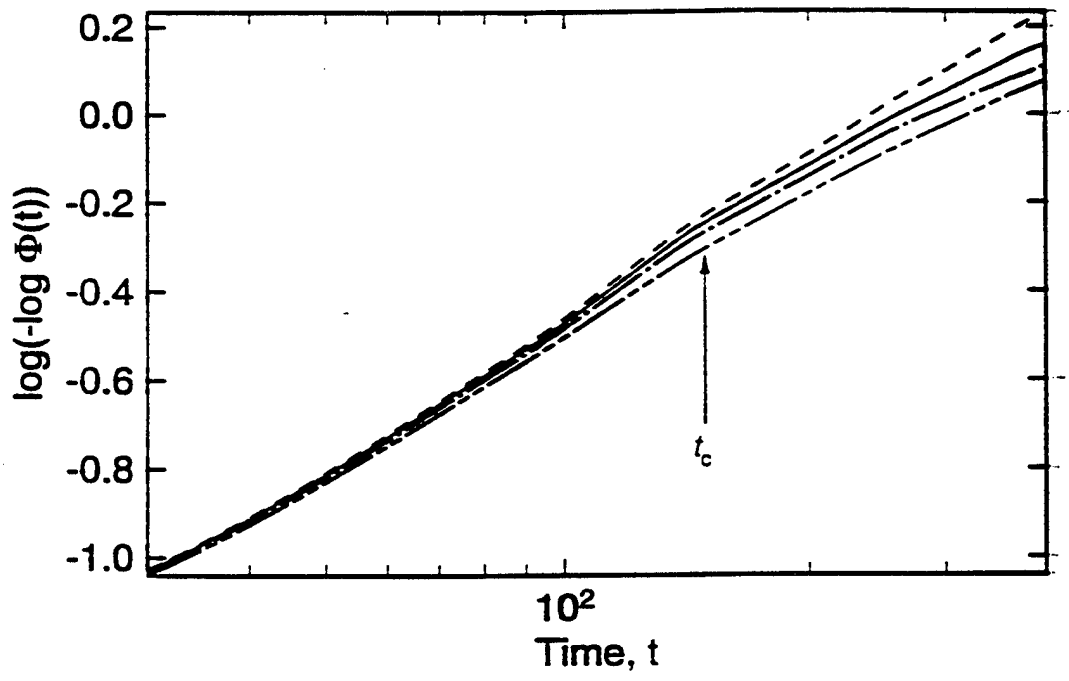


Fig. 4. Log(-Log Φ) against t. The curves become nearly straight after a crossover time t_c which depends on R, and can be fitted to straight lines of slopes β = 1.0, 0.86, 0.8, and 0.78, respectively for R=0, 0.02, 0.1, and 0.5 (from left to right).

crossover time t_c (which depends on R), the curves become nearly straight. The portion of the curves can be fitted to straight lines of slopes β= 1.0, 0.86, 0.8, and 0.78, respectively for

R=0, 0.02, 0.1, and 0.5. Figure 5 shows that the portion of curves corresponding to large t can be approximated by simple functions. In particular, the Fermi curve (solid curve below) is shown approximated by the (dotted) exponential function. The Fermi-stadium curve, on the other hand, is shown approximated by a (dotted) stretched-exponential function. The dashed curve is an exponential indicating a failure to fit the Fermi-stadium curve.

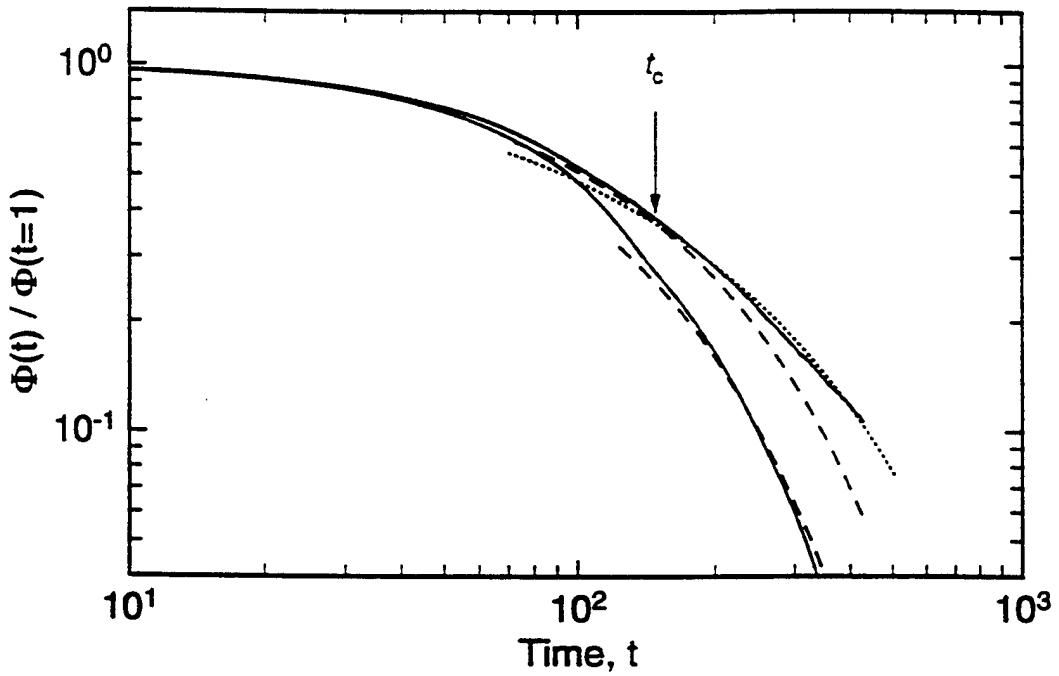


Fig. 5. The Fermi curve (solid curve below) approximated by the (dotted) exponential function; and the Fermi-stadium curve, with R=0.5, approximated by a (dotted) stretched-exponential function. The dashed curve is an exponential indicating a failure to fit the Fermi-stadium curve.

The effect of additional nonlinearity, in form of half a stadium, to the Fermi map, has been studied. It is found that for large t, $\Phi \sim \exp(-t/\tau)$ for R=0, indicating that the Fermi map has exponential relaxation. For other R's, it is found that $\Phi \sim \exp(-t/\tau^*)^\beta$ for $t > t_c$. Thus the relaxation of the Fermi-stadium map proceeds with a stretched exponential time dependence for $t > t_c$. The nonlinearity from the stadium is observed to slow down the relaxation of the original system. As expected, the fractional exponent β decreases further from unity with increasing nonlinearity or larger R's.

COUPLED NONLINEAR OSCILLATORS

Lately in the community of research in chaos there is considerable interest in nonlinearly

coupled oscillators. Various models of a globally coupled oscillator array (Refs.42-46). It was conjectured and later proved (though with conditions) that for coupling strength below a certain threshold, such systems would relax to an incoherent state (Ref.47). Recently the relaxation process in such an array is found to be exponentially fast (Ref.48). We consider an array of coupled oscillators governed by the map:

$$\varphi_i' = \varphi_i + \frac{K}{N}\Sigma_{j=1}^{N}\sin(\varphi_j - \varphi_i)$$

for $i, j=1,\ldots,N>>1$, which is closely related to the equations of motion studied by Strogatz and coworkers (48). Like Strogatz et al. we are interested in the decay of the phase coherence, r, which is the absolute value of the order parameter, $re^{i\psi}$, defined by

$$r = |re^{i\psi}| = |\frac{1}{N}\Sigma_j e^{i\varphi_j}|$$

We find that our array of coupled oscillators decay to an incoherent state, i.e. $r=0$, for $K<0$. Numerical calculation shows (Ref.30) that the decay of r is almost exactly an exponential function of time at least up to the longest time we have carried out so far and 99.9% of the initial value of r has decayed. Now if we allow a number of these arrays to interact to obtain an even more complex system, the interesting problem is to find out what modification these interactions will have on the relaxation of r for each array towards incoherence. The interactions betweeen the arrays $\alpha=1,\ldots,M$ can be chosen in several ways. One choice we have made is indicated by the new maps (Ref.30):

$$\varphi_{i\alpha}' = \varphi_{i\alpha} + \frac{K}{N}\Sigma_{j=1}^{N}\sin(\varphi_{j\alpha} - \varphi_{i\alpha}) - \frac{K'}{N^2}\Sigma_{\beta=1}^{M}\Sigma_{j=1}^{N}\sin(\varphi_{j\beta} - \varphi_{i\alpha})$$

The strength of the interactions between the arrays is measured by K'/K.

We have iterated the maps to obtain the evolutions of the coupled arrays system numerically. From the result we calculate the decay of r for each array. Figures 6-8 show the decay of r for interaction strength of 0.80, 1.0 and 1.2. In each of these three figures, the dashed curve is the almost exponential decay without inter-array interaction, i.e. K'/K=0. It has been shifted horizontally along the time axis towards longer times to make it coincident with the calculated curve (a string of beads) at short times. It is clear that the relaxation is slowed down by the inter-array interactions, and the degree of slowing down is proportional to the interaction strength K'/K. In all these calculations N is fixed at the value of 28. The degree of slowing down depends also on M, increasing with M but levels off for $M>4$. The results shown in

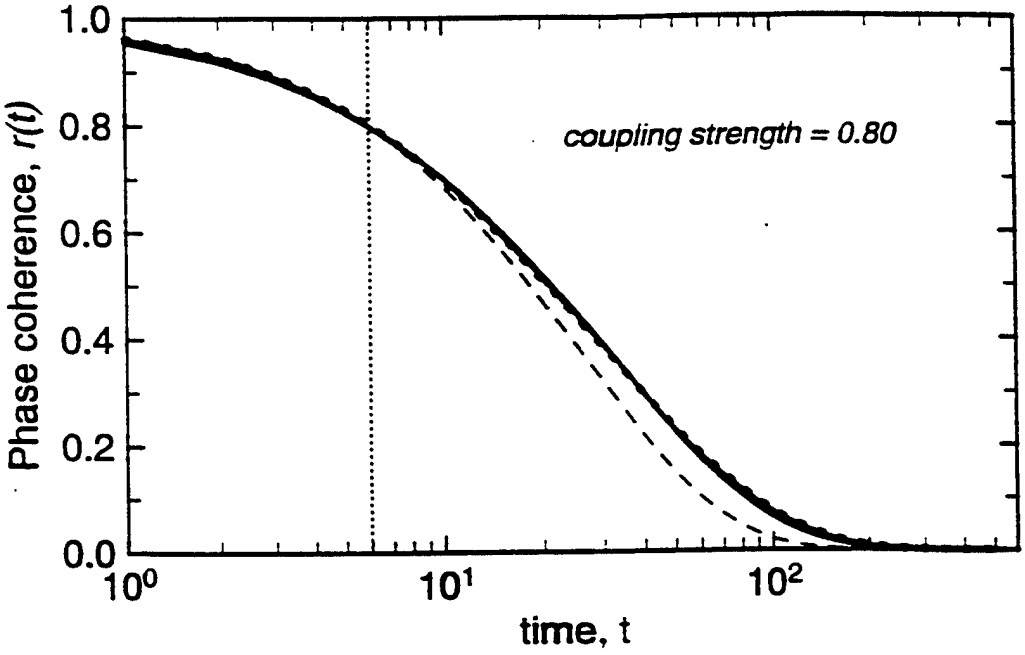


Fig. 6. Decay of phase coherence, $r(t)$, calculated for interaction strength K'/K=0.8 (beaded curve). The thick curve is the stretched exponential function, $\exp-(t/\tau^*)^{1-a}$, and the dashed curve is the $\exp-(t/\tau_o)$ function. The cross-over time t_c is indicated by the vertical dotted line. The calculated $r(t)$ conforms well to the $\exp-(t/\tau_o)$ for $t<t_c$ and the $\exp-(t/\tau^*)^{1-a}$ for $t>t_c$

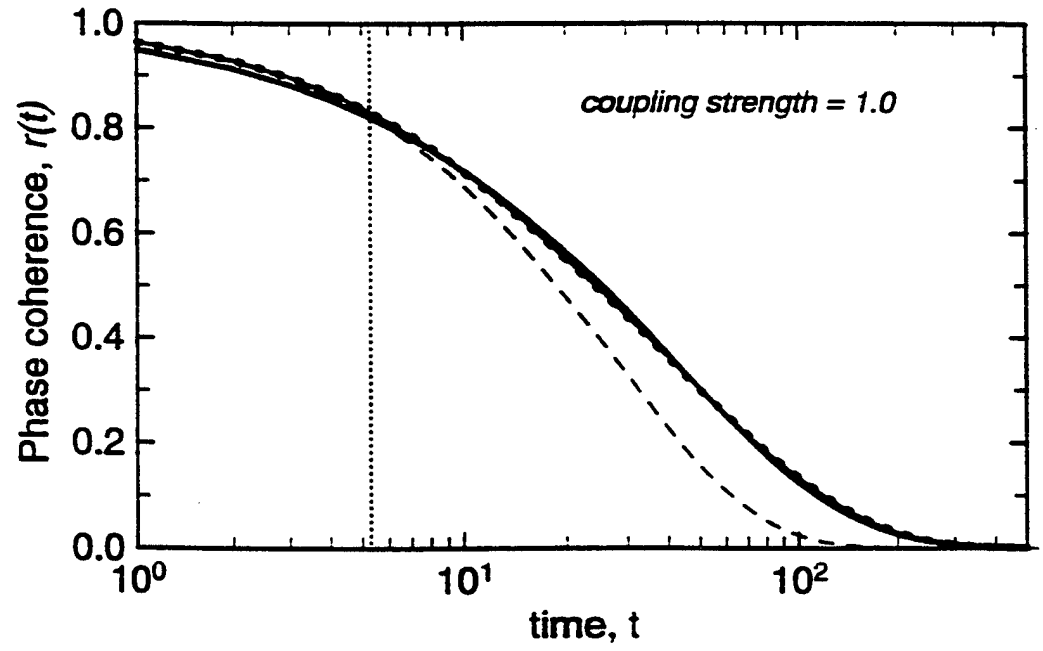


Fig. 7. Same as for Fig.6 except the interaction strength is 1.0

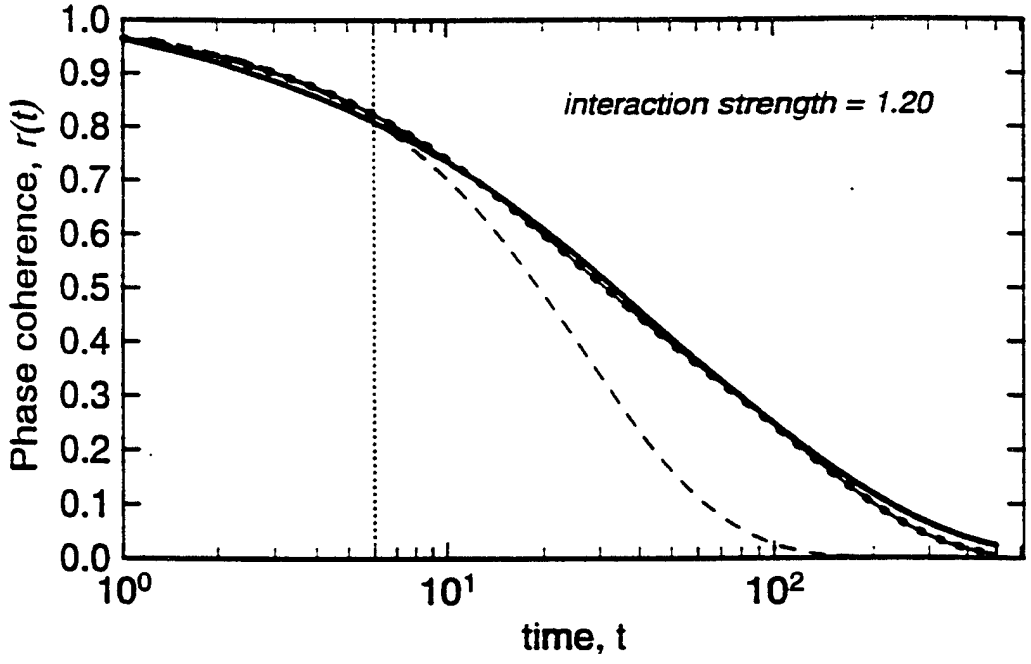


Fig. 8. Same as for Fig.6 except the interaction strength is 1.2

Figs.6-8 are calculated for M=3. The results are consistent with an initial exponential decay, exp-(t/τ_o), and crossover at some t_c to a stretched exponential decay, exp-$(t/\tau^*)^{1-n}$ (solid curve) as indicated by the vertical dotted line drawn in the figures. Fitting the calculated results this way we find that the relaxation time τ^* increases and the exponent $1-n$ decreases with the interaction strength, K^*/K, between the arrays (see Figs.9 and 10). These behaviors are in accord with the results of the coupling model. Continuity of the two pieces, exp($-t/\tau_o$) and exp-$(t/\tau^*)^{1-n}$ at t_c guarantees the validity of the relation between τ^* and τ_o given by:

$$\tau^* = [\, t_c^{-n} \tau_o \,]^{\frac{1}{1-n}} \qquad (4)$$

which is the same as Eq.1 provided the identification

$$t_c = (1-n)^{\frac{1}{n}} \omega_c^{-1} \qquad (5)$$

is made. The reciprocal of the ω_c^{-1} in Eq.1 is the time at which the relaxation rate defined by $-(1/C(t))dC(t)/dt$ are equal for the two correlation functions, exp-(t/τ_o) and exp-$(t/\tau^*)^{1-n}$; while t_c is the time at which they are continuous. Thus the relaxation of the nonlinearly coupled arrays of coupled nonlinear oscillators has verified the basic features of the coupling model.

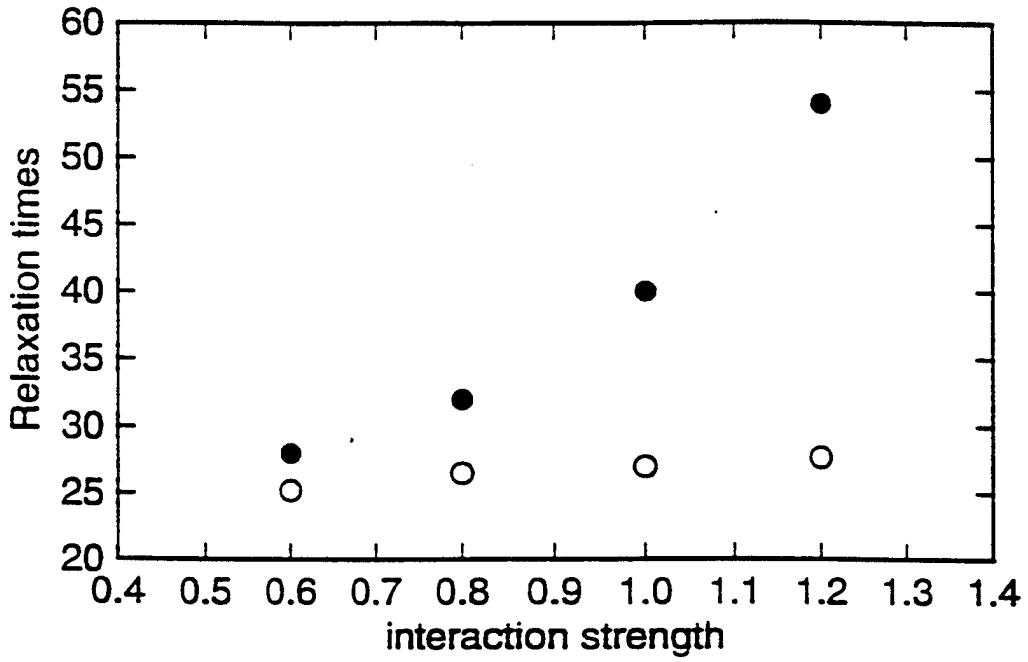


Fig. 9. Relaxation times, τ_o (open circles) and τ^* (filled circles), as a function of the interaction strength

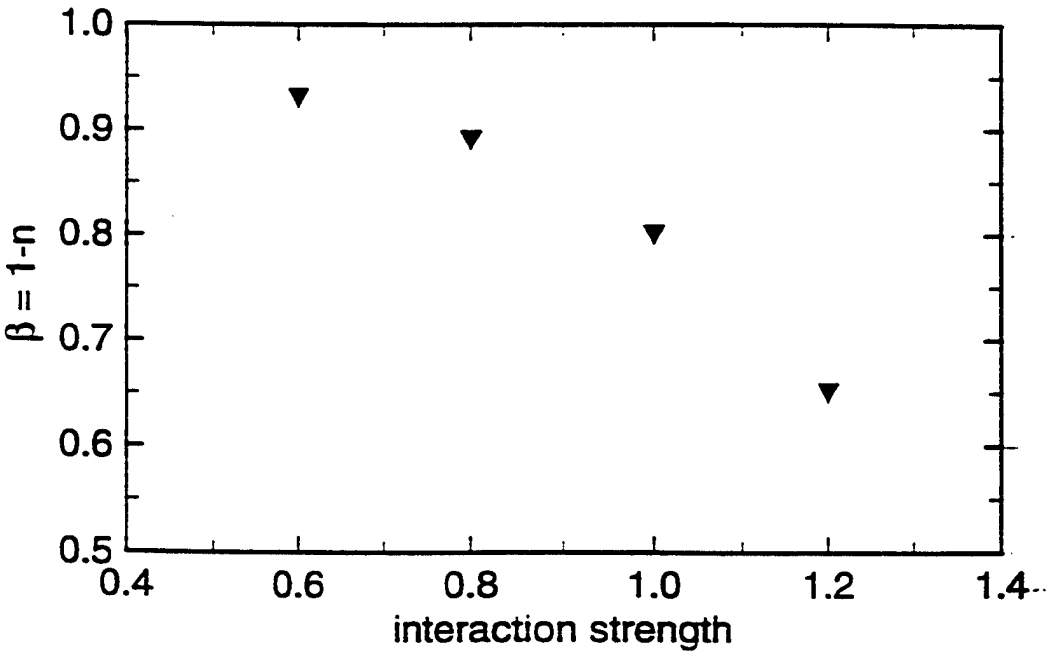


Fig. 10. The Kohlrausch exponent, β, determined as a function of interaction strength

The good correspondence between relaxation of simpler prototype chaotic systems and the key features of the coupling model shows that a theory of the coupling model is in the making. The encouraging results obtained will provide impetus for future study of more complex chaotic Hamiltonian systems that bear closer resemblance to polymers, small molecule glass-forming liquids and vitreous ionic conductors.

QUASIELASTIC NEUTRON SCATTERING

The quasielastic neutron scattering measurements were made by Colmenero and coworkers (Refs.49,50) using the TOF spectrometer IN6 at the Institut Laue-Langevin, Grenoble, France. The incoherent scattering function $S(Q,\omega)$ was obtained for various scattering wavevectors Q in the range $0.2 < Q < 2$ Å^{-1} and energy transfer $\hbar\omega$ up to 5 meV at constant temperatures in the range $T_g - 8 < T < T_g + 100$ K. The poly(vinylchloride) (PVC) sample studied has a glass temperature T_g=358 K and a number-averaged molecular weight M_n =4.55x10^4. The relaxation contribution, $S^{relax}(Q,\omega)$, to the scattering was isolated after removal of the harmonic vibrational contribution from the measured $S(Q,\omega)$, taking into account the instrument resolution. Fourier transform of $S^{relax}(Q,\omega)$ gives the normalized intermediate scattering function $I_o(Q,t)=F_{relax}(Q,t)$. The results are reproduced in Figure 11a for different temperatures at a constant Q value of 1.5 Å^{-1}. It can be seen that, independent of T and Q, all the intermediate scattering functions

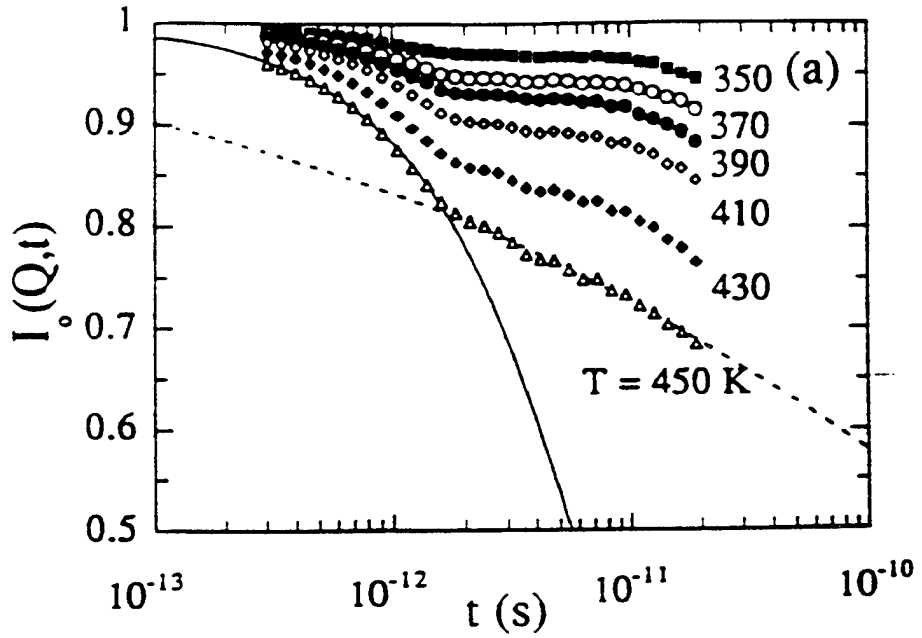


Fig. 11a. Normalized intermediate scattering function for PVC at Q=1.5 Å^{-1} at different temperatures. After Colmenero, Arbe and Alegria

exhibit a break in curvature at a time of about 1.7 ps, signifying a crossover of $I_o(Q,t)$ from one regime at short times (t < 1.7 ps) to another at long times (t > 1.7 ps). This crossover between two different dynamic regimes is more clearly seen in the mean squared displacement of the scattering centers, $\langle r^2(t)\rangle$, which can be obtained from the $I_o(Q,t)$ data by solving the equation $I_o(Q,t) = \exp[-Q^2\langle r^2(t)\rangle/6]$. The values of $\langle r^2(t)\rangle$ so obtained (Refs.49,50) at different temperatures are plotted against time in Figure 11b, and the existence over a considerably large temperature range of a T-independent crossover time (approximately equal to 1.7 ps) is clear. In the short time regime of t < 1.7 ps, $\langle r^2(t)\rangle$ is proportional to t. While for t > 1.7 ps, $\langle r^2(t)\rangle \propto t^{1-n}$ with n = 0.77. This sublinear time dependence of $\langle r^2(t)\rangle$ was found previously using a longer time window of $10^{-11} < t < 10^{-8}$ s (Refs.51,52). It has been shown to originate from the local segmental (alpha) relaxation of PVC, with a stretched exponential form, $\exp[-(t/\tau^*(Q,T))^{1-a}]$, for the intermediate scattering function.

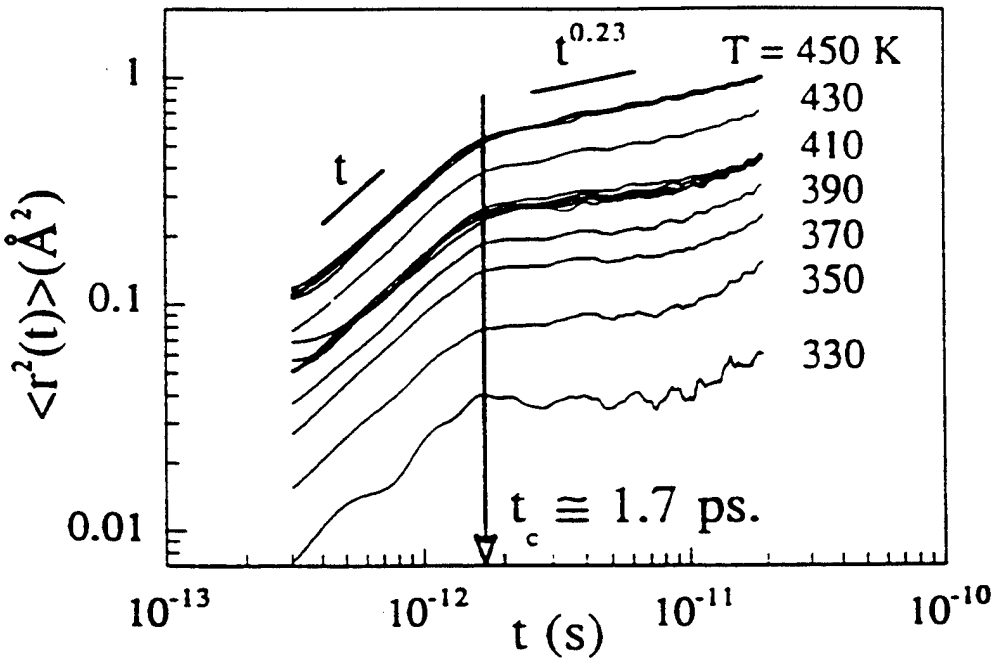


Fig. 11b. Mean square displacement obtained from data in Fig.1. For 450 K, data from different Q's are plotted together (After Colmenero, Arbe and Alegria)

To see if this crossover corresponds to the one proposed by the coupling model, the intermediate scattering function data in the short time regime was fitted by (Refs.49,50)

$$I_o(Q,t) = \exp-(t/\tau_o(Q,T)) \text{ for } t < 1.7 \text{ ps} \qquad (6)$$

and the data in the long time regime by the stretched exponential (see Fig.11a)

$$I_o(Q,t) = \exp-[t/\tau^*(Q,T)]^{1-n} \text{ for } t > 1.7 \text{ ps} \quad (7)$$

with the exponent $1-n$ taken to be that obtained from $\langle r^2(t)\rangle \propto t^{1-n}$, (i.e., $1-n = 0.23$ for T=430 and 450 K (see Fig.11b)). While fits to Eq 6 (solid curve in Fig.11a) in the short time regime have been carried out successfully in the entire temperature range, fits to Eq.7 in the long time regime can be carried out with confidence at only the two highest temperatures, 430 and 450 K, where the $I_o(Q,t)$ data decrease significantly with time, before being cut off by the long time edge of the time window. These fits, shown as dashed curves in Fig.11a, describe well the experimental $I_o(Q,t)$ at T=430 and 450K.

From these fits of the experimental data in the two separate time regimes using eqs 4 and 5, τ_o and τ^* were obtained as a function of Q and T. For τ_o it was found that

$$\tau_o(Q,T) \propto Q^{-2}\exp[5.8\,(\text{kcal/mol})/RT]. \quad (8)$$

The Q^{-2} dependence indicates simple diffusion of the scattering centers, consistent with independent relaxation of local segments without the intermolecular interaction expected for $t < t_c$. The Arrhenius temperature dependence is valid over an enormous temperature range and has an activation enthalpy of about 6 kcal/mol. This is comparable to the internal rotation isomerism energy barrier of a PVC chain deduced before from different measurements (Ref.53). Such a temperature dependence provides additional support for the coupling model interpretation of the short time ($t < t_c$) dynamics. On the other hand, for τ^*, combining the current TOF neutron scattering data with the previously published neutron backscattering data, we find $\tau^* \propto Q^{-9}$, similar to earlier results (Refs.51,52) based on backscattering data. Substituting $\tau_o(Q,T)$ given by Eq.8 into Eq.4, recalling the fact that t_c is independent of Q and T and $n = 0.77$, the coupling model predicts

$$\tau^*(Q,T) \propto Q^{-2/(1-0.77)}\exp\{25(\text{kcal/mol})/RT\} \quad (9)$$

These predicted temperature and Q dependencies of τ_o and τ^* are summarized and illustrated respectively in Figs.12a and 12b. The predictions are in good agreement with the experimental data. Similar results have been obtained for other polymers including poly(vinylmethylether), polybutadiene and polyisoprene (Ref.54). Thus, we conclude that neutron scattering data

directly corroborate the coupling model.

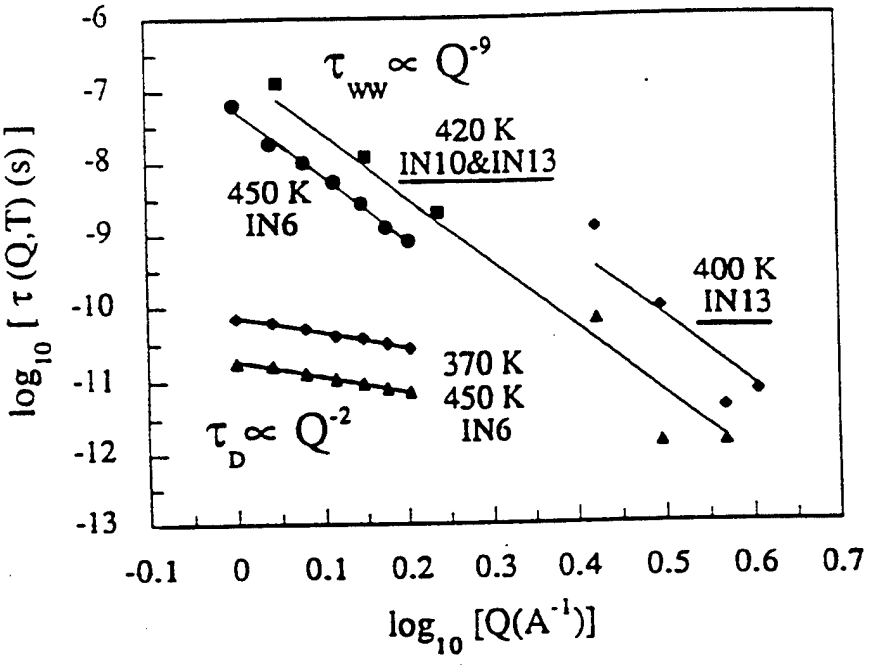


Fig. 12a. *Q*-dependences of τ_o and τ^*. After Colmenero et al.

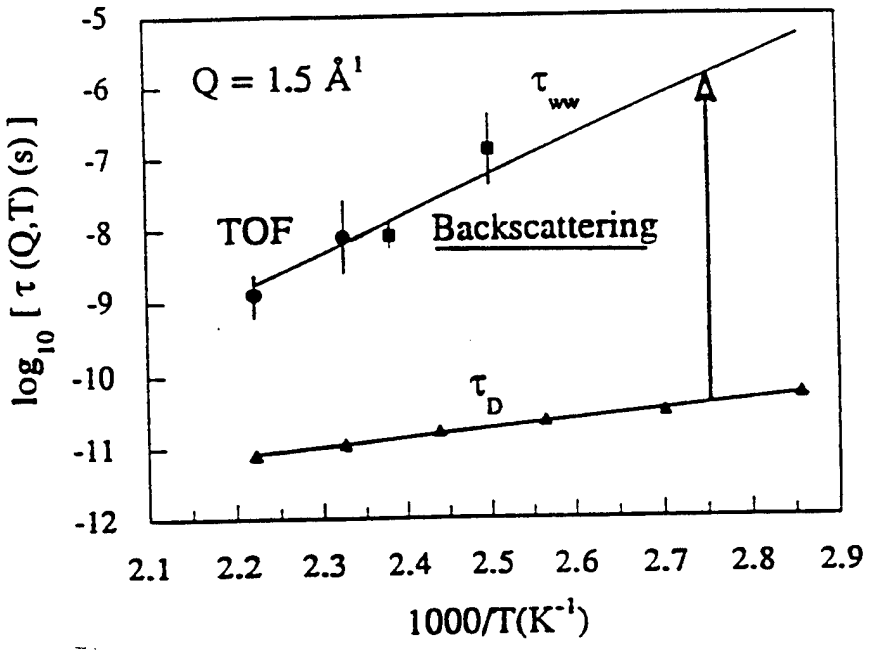


Fig. 12b. *T*-dependences of τ_o and τ^*. After Colmenero et al.

Questions can be raised on the plausibility of the observation of diffusive motion (consistent with the Q^{-2} dependence of τ_o in Eq.8) in the short time regime of $3x10^{-13} < t < t_c$. In fact, the existence of a multitude of high frequency vibrational (bond bending and stretching) modes provides the fast variables (i.e., heat bath) that enable "diffusion" of the hydrogen as seen by the neutrons. In the next section we discuss molecular dynamics simulation of small molecule liquids and polymers. In these simulations (Refs.55-57) the same cross-over phenomenon seen by incoherent neutron scattering are obtained, with similar t_c for both $F_s(q,t)$ and the reorientational time correlation function, $M_k(t) = \langle P_k[\cos\theta(t)]\rangle$, where $P_k(x)$ is the Legendre polynomial of order k, and $\theta(t)$ is the reorientation angle of a vector at time t. For reorientational motion, the rotational diffusion relaxation time does not have any Q-dependence, and it is obvious that rotational diffusion can occur on time scales of the order of a picosecond.

MOLECULAR DYNAMICS SIMULATION

In recent years improvement in the technique has made molecular dynamics simulation (MDS) a powerful tool to investigate the dynamics of molten salts (Ref.55), glass-forming small molecule liquids (Ref.56) and polymers (Ref.57). Realistic potentials have been used to represent the interactions between the molecular units, making the results of computer experiments as quantitatively accurate as actual experiments. In addition, informations extracted from MDS are often richer than that from real experiments. Thus, these MDS data provide additional tests of the theoretical basis of the coupling model. In this Section we shall analyze one set of MDS data (Ref.57) and show that the results confirm the coupling model's dexcription of the dynamics.

Roe (Ref.56) has performed MDS of polyethylene (PE). The intermolecular interactions are of the short range van der Waals kind, represented by Lennard-Jones potentials. $F_s(Q,t) = \langle \exp\{-iq\cdot[r_i(t)-r_i(0)]\rangle$, where $r_i(t)$ is the position of the i-th segment at time t, has been evaluated for different temperatures and several values of the wavevector Q. Results for Q=1.38 Å^{-1} are shown in Ref.56. At temperatures below 108K there is an initial decay of $F_s(Q,t)$ which can be ascribed to harmonic phonons, with no additional relaxation process observed in the time window. The assignment of the initial decay to harmonic phonons is consistent with the plateau value $F_s(Q,t\rightarrow\text{large})$ being well described by the Debye-Waller factor, $\exp(-WQ^2T)$, where W is a constant independent of Q and T (Ref.58). A semilog plot of the Debye-Waller factor obtained from the low temperature $F_s(Q,t)$ data yields a straight

line which can be used to extrapolate the phonon contribution to higher temperatures. This extrapolation clearly indicates that the contribution of phonons to $F_s(Q,t)$ has to be taken into consideration at all temperatures. At temperatures equal to and higher than 132 K, additional relaxation processes appear in the time window of Figs.13a and 13b, making it difficult to isolate the individual components. We can account for the harmonic phonon contribution to $F_s(Q,t)$ at higher temperature by Fourier transformation of the $F_s(Q,t)$ data at low temperatures (say T < 108 K), where only the harmonic phonons contribute, followed by scaling of the resultant dynamic structure factor $S(Q,\omega)$ by the Bose factor and the Debye-Waller factors (Ref.58). The inverse Fourier transform of the scaled $S(Q,\omega)$ yields $F_s^{ph}(Q,t)$, the density-density correlation function from phonons at higher temperatures. Assuming that scattering by harmonic phonons and relaxation are statistical independent processes, we write the intermediate scattering function as a product $F_s(Q,t) = F^{phonon}(Q,t)F^{relax}(Q,t)$. This equation can be solved for $F^{relax}(Q,t)$ at any T. The results can be compared with the predictions of the coupling model in a similar manner as done previously by Colmenero and coworkers using his QENS data. The $F_{phonon}(Q,t;T)$ at all temperatures for which Roe has made his simulations can be calculated. We can now assess the coupling model by examining whether the experimental $F_s(Q,t;T)$ can be represented at each temperature by:

$$F_s(Q,t;T) = F_{phonon}(Q,t;T) \times \begin{cases} \exp-(t/\tau_o(T)) & \text{for } t<t_c \\ \exp-(t/\tau^*(T))^{1-n} & \text{for } t>t_c \end{cases} \qquad (10)$$

for a temperature independent t_c such that the continuity condition, $\exp-(t/\tau_o(T)) = \exp-(t/\tau^*(T))^{1-n}$ is satisfied always at $t=t_c$. We start by first choosing a t_c and then for each temperature find two independent parameters, τ_o and n (the third parameter τ^* is automatically fixed by Eq.4), such that the products on the right hand side of Eq.10 give a good fit to the experimental $F_s(Q,t;T)$. In carrying out this, we find that for good fits at all temperatures can only be obtained for t_c's limited to a narrow range around 2 ps.

The results of the best fits for t_c=2 ps for two temperatures are shown in Figs.13a and 13b. The $F_{relax}(Q,t;T)$ used to obtain these fits are displayed also in these figures. At each temperature $F_{relax}(Q,t;T)$ is comprised of two pieces: $\exp-(t/\tau_o(T))$ for $t<t_c$ and $\exp-(t/\tau^*(T))^{1-n(T)}$ for $t>t_c$. The relaxation times, τ_o and τ^*, and the coupling parameter, n, are plotted as a function of temperature in Fig.14 and Fig.15 respectively. The temperature dependences of τ_o is approximately Arrhenius with an activation enthalpy of about E_a = 2.4

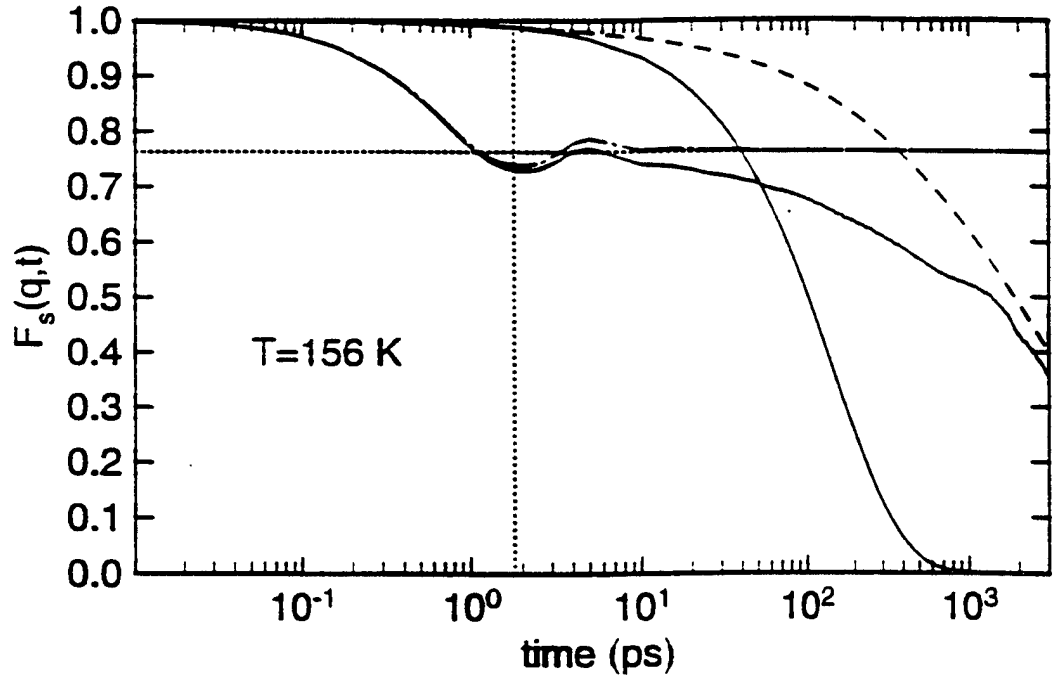


Fig. 13a. Theoretical fit (thick long dashed-short dashed curve) to experimental data $F_s(Q,t;T)$ of polyethylene (solid curve) showing the cross-over time t_c (vertical dotted line), the phonon part, $F_{phonon}(Q,t;T)$,(dashed dotted curve) and the good fit of the relaxation part, $F_{relax}(Q,t;T)$, by exp-(t/τ_0) (solid curve) for $t<t_c$ and exp-$(t/\tau^*)^{1-n}$ (dashed curve) for $t>t_c$ at T=156K.

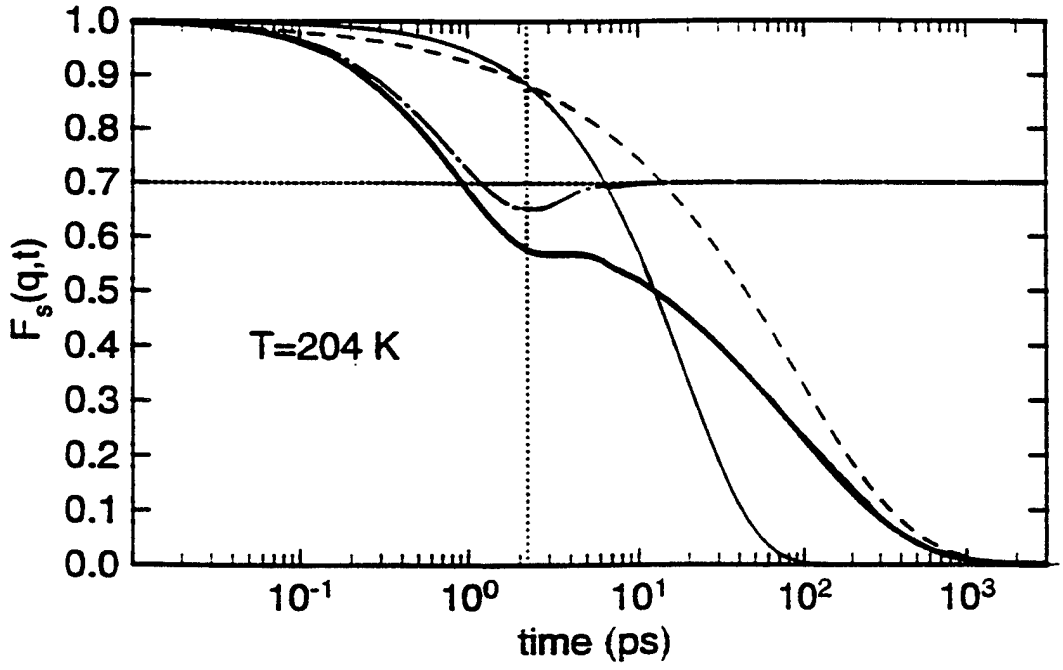


Fig. 13b. Same as for Fig.13a but now T=204K

kcal/mole. It is apparent that the temperature dependence of τ_o becomes milder at high temperatures ≥ 252 K. which makes the overall activation energy appear smaller. Excluding the three highest temperatures, the activation enthalpy is close to 3 kcal/mole.

The coupling parameter has the value of 0.40 and exhibits a decrease at high temperatures. This decrease of n at high temperatures may be correlated with the corresponding milder

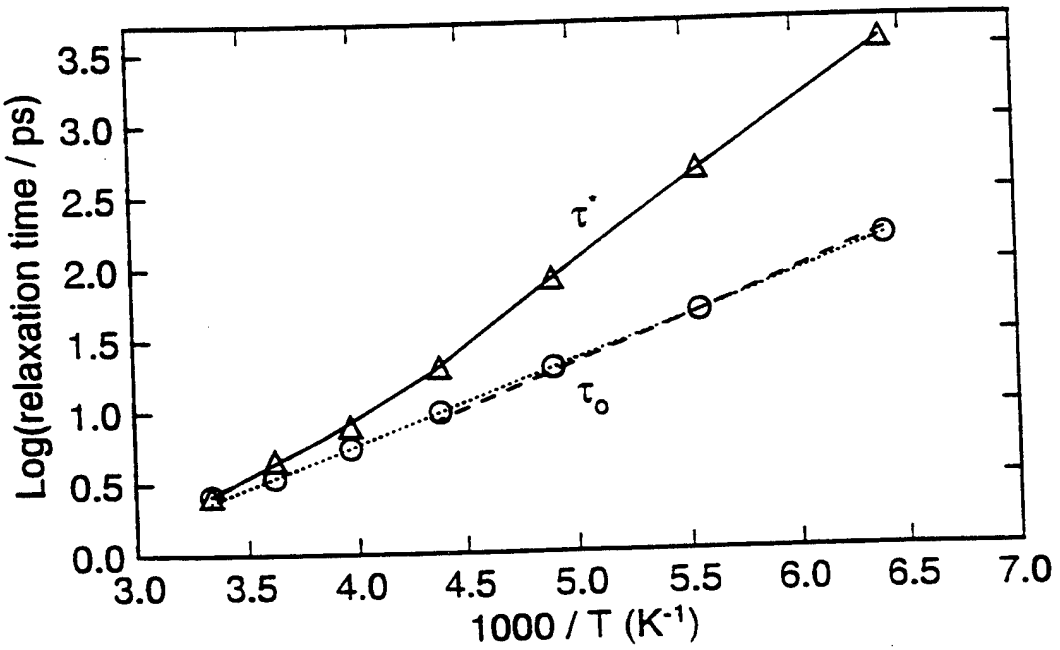


Fig. 14. Arrhenius plot of τ_o and τ^* of polyethylene

temperature dependence of τ_o seen there. The value for E_a is roughly what is expected for the activation energy of the conformational energy barrier for a single polyethylene chain. As modelled by Roe, this is about 3 kcal/mole (Ref.57). This is a striking agreement for the microscopic activation energy which governs local segmental motion of a single chain. The stretch exponent, $\beta \equiv 1-n$, determined for polyethylene is close to that expected from the empirical correlation between β and the steepness of the polymer's cooperative plot (Ref.58). The coupling parameter n= 0.40 for polyethylene is smaller than that of all other polymers with bulkier monomer structure. The much larger coupling parameter (n=0.77) found for polyvinylchloride (PVC) from QENS as well as dielectric and mechanical mesurements, arises from its larger intermolecular interaction due to the polar nature of the PVC backbone. The disparity of the values of n in PVC and PE is directly responsible for the different appearance

of the $F_{relax}(Q,t;T)$'s obtained for these two polymers (see Fig.11a and Figs.13a and 13b). In

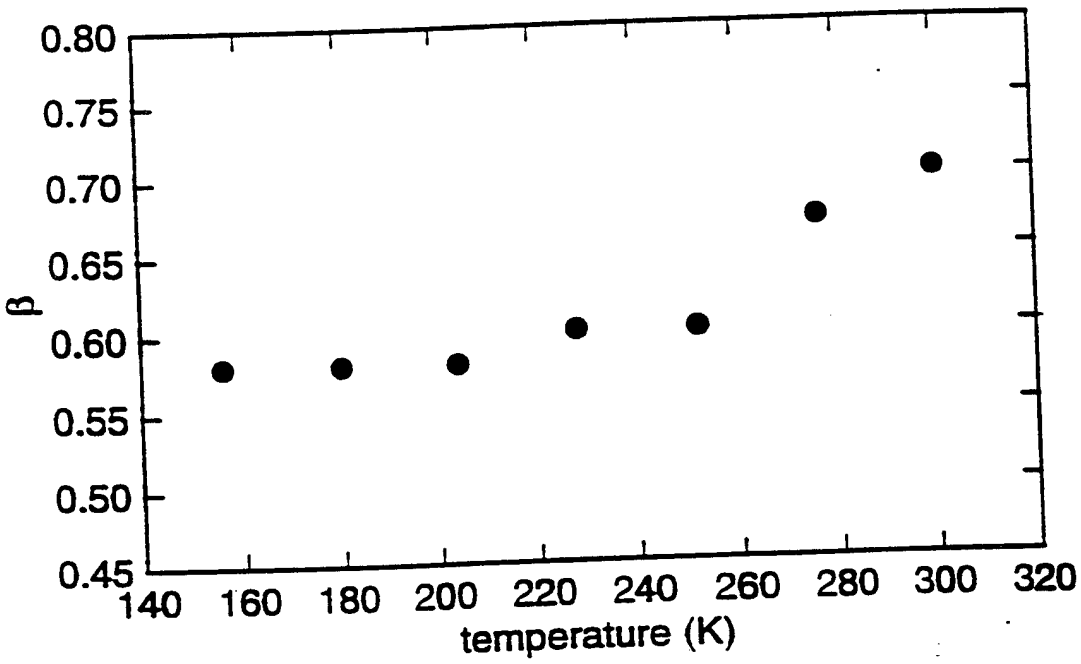


Fig. 15. The coupling parameter, n, as a function of temperature

PVC we see clearly a break in the $F_{relax}(Q,t;T)$ at the cross-over time t_c. For PE the break at the cross-over is less obvious (Figs.13a and 13b) due to the smaller n. Nevertheless, in PE the initial exp-(t/τ_o(T)) decay for $t<t_c$ is necessary to explain the departure of $F_s(Q,t;T)$ from the phonon contribution $F_{phonon}(Q,t;T)$ at short times.

FROM MICROSCOPIC DYNAMICS TO VISCOELASTICITY OF POLYMERS

In previous sections we have shown microscopically from theory (chaotic Hamiltonian), real experiment (QENS), and computer experiment (molecular dynamics simulation) that the basic result of the coupling model seems to be correct. This microscopic basis justifies the application of the coupling model to many practical problems in relaxations of complex systems. As mentioned already in the Introduction, we have made many efforts to apply the coupling model to macroscopic relaxation processes in the last ten years, and the results have been remarkably successful. Many representative achievements were the subjects of a very recent review (Ref.12) and will not be repeated here. The interested reader can consult Reference 12 and the papers cited therein for details. In the remainder of this paper we shall

briefly discuss a more recent development (Ref.60) that is concerned with the application of the coupling model to viscoelastic properties of amorphous polymers, particularly in the glass-rubber transition region.

The local segmental motion involves a few monomers. Although the number of monomers and hence the length scale of the motion can be a function of temperature, we can expect from the coupling model that the characteristics of the local segmental motion will depend on the chemical structure of the polymer. First t_c may vary somewhat with chemical structure. However, if the nature of the intermolecular interaction is the same we do not expect t_c, the onset of the effect due to chaos from the nonintegrable intermolecular interactions, to differ much from one polymer to another. After the onset time t_c the coupling parameter n_α ,which determines the segmental relaxation time τ_α via Eq.5 , will also depend on the chemical structure. Intuitively we may expect that monomer that is more complex , sterically constrained and having backbone that is more inflexible will have a larger coupling parameter, even the nature of the intermolecular interaction has not changed. Motion of such a monomer will be further slowied down (i..e. larger n) because the degree of chaos caused by the intermolecular interactions is enhanced in considering the local segmental motion of the polymer. This expectation will be demonstrated by future calculations based on either the chaotic Hamiltonian or the molecular dynamics simulation. At this time the coupled nonlinear oscillators model discussed above provides an illustration. Keeping the interaction strength K'/K constant, increase of the number, M, of arrays from 2 to 4 resembles making the monomer more complex. Indeed we find that n increases with M . If molecular dynamics simulation performed by Roe (Ref..57) on poly(ethylene) were extended to bulkier and more inflexibe polymers, we predict before the numerical experiment is done that the coupling parameter found will be increased above the value determined for poly(ethylene) (see Fig.15).

Experimental measurements of the time/frequency dispersion of the local segmental motion in many amorphous polymers are available in the literature. The most direct measurements of local segmental motion are by dielctric and phton correlation spectroscopic techniques. These techniques usually do not see the Rouse modes in the glass-rubber transition region neighboring in time scale to the local segmental motion. Thus the local segmental motion can be isolated, its dispersion fit to a Kohlrausch function (Eq.2), and the coupling parameter n_α determined. A survey of the values n_α so determined for many amorphous polymers show that they vary over a wide range (Ref.61). Amongst amorphous polymers with carbon backbones so far the smallest value of about 0.45 for n_α is found for polyisobutylene

(PIB)(Ref.62) and the largest value of 0.77 is found for poly(vinylchloride) (PVC) (Refs.49-53). The reason why PVC has a larger n_α than PIB is clearly due to the stronger intermolecular interaction in PVC which has an hydrogen replaced by a chlorine. Polystyrene (PS), is related in chemical structure to PIB by being more complex and having a bulkier ring attached to the backbone. Indeed it has a coupling parameter of 0.64, larger than that of PIB. In the treatise on viscoelasticity of polymers (Ref.60), many different polymers are compared. In this Lecture we shall limit the comparison to between PS and PIB only. As has been shown previously as a consequence of the coupling model in several publications (Refs.63-65) and a review article (Ref.12), the viscoelastic properties of two polymers like PS and PIB having a large difference in n_α will necessarily be very different . We shall not repeat in detail the arguments used before (Refs.12,60, 63-65). The interested reader can find them in these References. However we shall briefly indicate how the coupling model can account for the drastically different viscoelastic properties exhibited by PS and PIB that will be summarized below:

(1) Width and Dispersion of the Glass-Rubber Transition

It is well known since the beginning of the development of viscoelasticity by Tobolsky and coworkers (Ref.66), and by Ferry and coworkers (Ref.67) that the width and dispersion of the

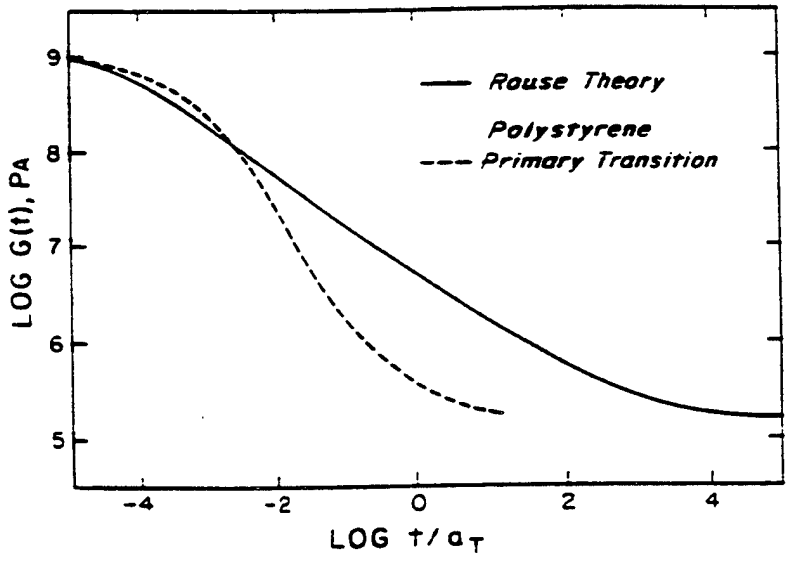


Fig. 16. The comparison between stress relaxation master curve of polystyrene and predictions of the modified Rouse theory by Ferry in the glass-rubber transition region (after Aklonis and MacKnight, Introduction to Polymer Viscoelasticity, reproduced by permission)

glass rubber transition zone for PS and PIB are very different. PS has a much narrower width and sharper dispersion than those of PIB. This can be seen by comparing stress relaxation data of PS (Fig.16) and PIB (Fig.17). The predicted stress relaxation curve from the Rouse theory modified by Ferry (Ref.67) for undiluted polymer also shown in Fig.16 is totally different from that of PS but about the same as that of PIB (Fig.17). The large deviation of PS from the modified Rouse theory originates from its large coupling parameter n_α which makes the temperature dependence of the local segmental relaxation time τ_α stronger than that of the modified Rouse modes (see Eqs.1 and 3) (Refs.12,60). As a consequence the local segmental relaxation encroaches on the time scale towards the relaxation time spectrum of the modified

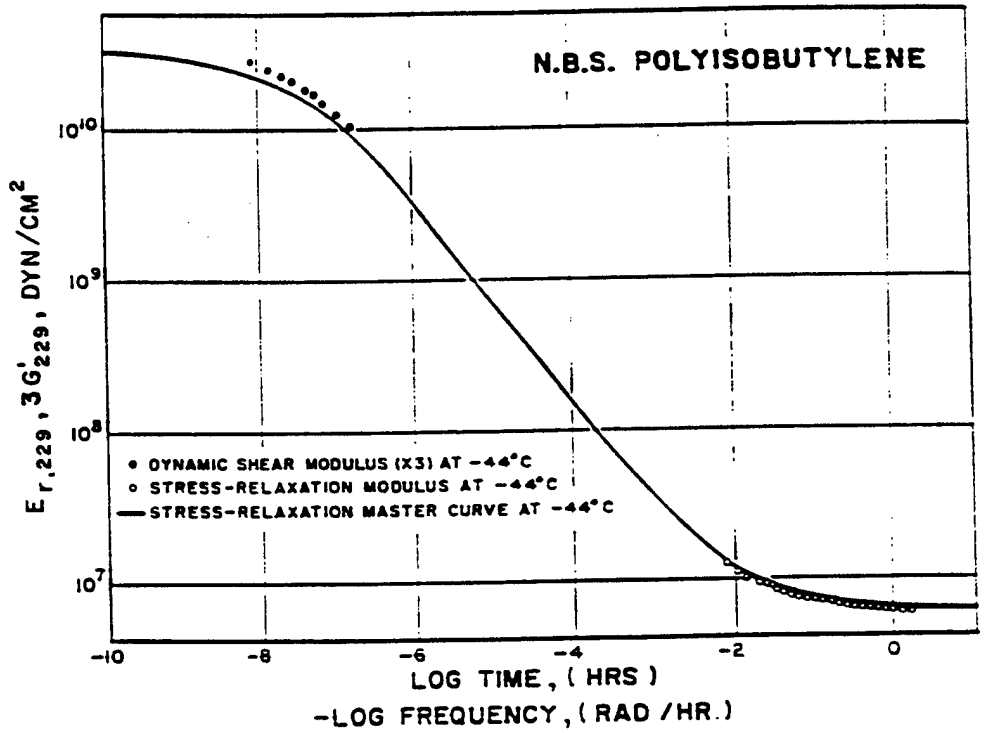


Fig. 17. Master stress relaxation curve for NBS polyisobutylene at -44°C. (After A.V. Tobolsky, J.Apply.Phys.,27,673(1956)).

Rouse modes, making the width of the glass-rubber transition of PS much narrower than the modified Rouse theory would predict. This effect is much reduced in PIB because it has a significantly smaller n_α, and hence the time dependence of stress relaxation in PIB resembles that predicted by the modified Rouse theory.

(2) Normalized Temperature Dependencies of τ_α

An objective way to compare the temperature dependences of τ_α is to plot $\log(\tau_\alpha)$ against normalized temperature variable such as T_g/T or $(T-T_g)/T_g$., where T_g is the "glass transition temperature" operationally defined by the temperature at which the measured τ_α has the value of an arbitrarily time, say 100 seconds. The T_g used here may not be the same as the glass temperature obtained by DSC measurement.. Angell (Ref.68) and others (Refs.61, 63,69) have demonstrated the utility of this normalized temperature dependence plot. Plazek et al. (Ref.63) have found that there exists a strong correlation between the steepness of this normalized temperature plot of τ_α and the coupling parameter n_α for the family of amorphous polymers. This correlation has been expanded to more materials in Reference 61. Such a correlation leads to the suggestion (Ref.69) that the plot of τ_α against T_g/T should be referred to as the "cooperativity plot". The name "fragility plot" for the same thing commonly used by other workers is not appropriate for polymers because fragility as originally used by Angell (Ref.68) is to describe degradation of structure with increase in temperature, but the structure of polymer is not changing with temperature. In the spirit of comparing only PS with PIB the cooperativity plot of these two polymers are presented in Fig.18. PS having a larger n_α does exhibit a steeper normalized temperature dependence in the cooperativity plot. We have already shown in Reference 61 and 63 that such a correlation between n_α and the steepness index is expected by Eqs.1 and 3 of the coupling model.

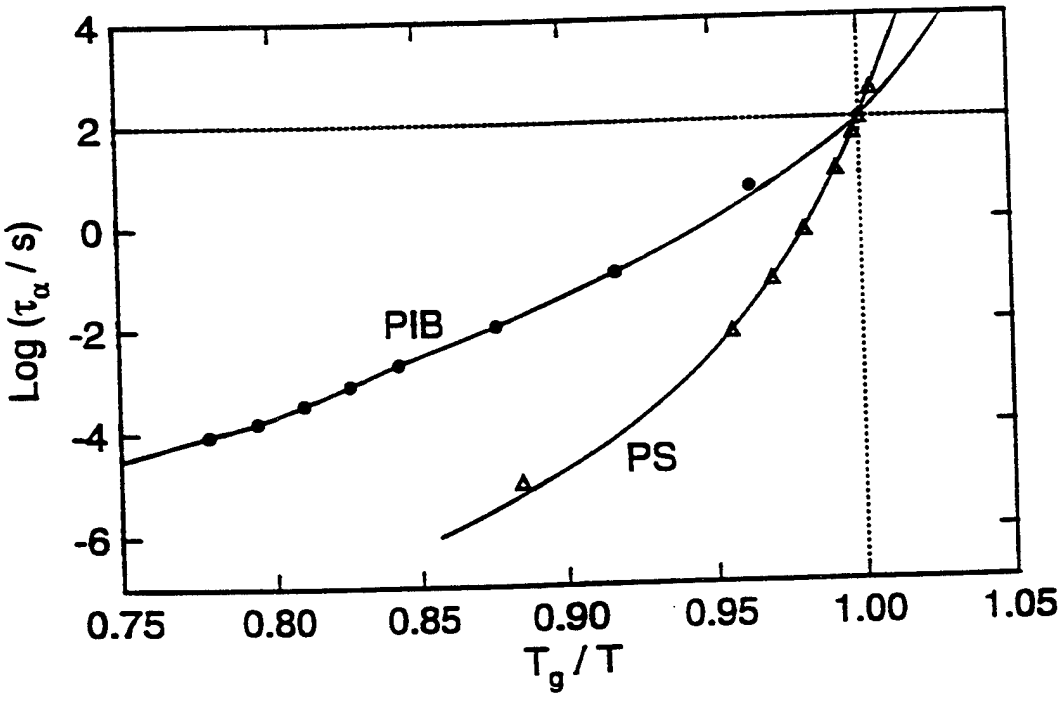


Fig. 18. A "cooperative" plot of log τ_α against normalized reciprocal temperature, T_g/T, where T_g is the glass temperature at which the mechanical relaxation time determined from creep compliance data of Plazek and Ngai, τ_α, equals 10^2s.

(3a) Degree of Thermorheological Complexity: High Molecular Weight

Plazek have shown by recoverable creep measurement on PS that the terminal relaxation (related to the viscosity η) has a weaker temperature dependence than that of the local segmental dispersion in a *common* temperature range near T_g (Ref.69) (see Fig.19). This outstanding feature of thermorheological complexity was explained by Eqs.1 and 3 of the coupling model (Ref.70) from the fact that in PS coupling parameter of local segmental motion, n_α , is larger than the coupling parameter of terminal motion, n_η, from chain entanglement interactions. The entanglement interaction involves length scale much larger than that of the monomer. As a consequence n_η is independent of chemical structure of the polymer and has the value of about 0.41 for linear polymers (Refs.8,9,12). The condition

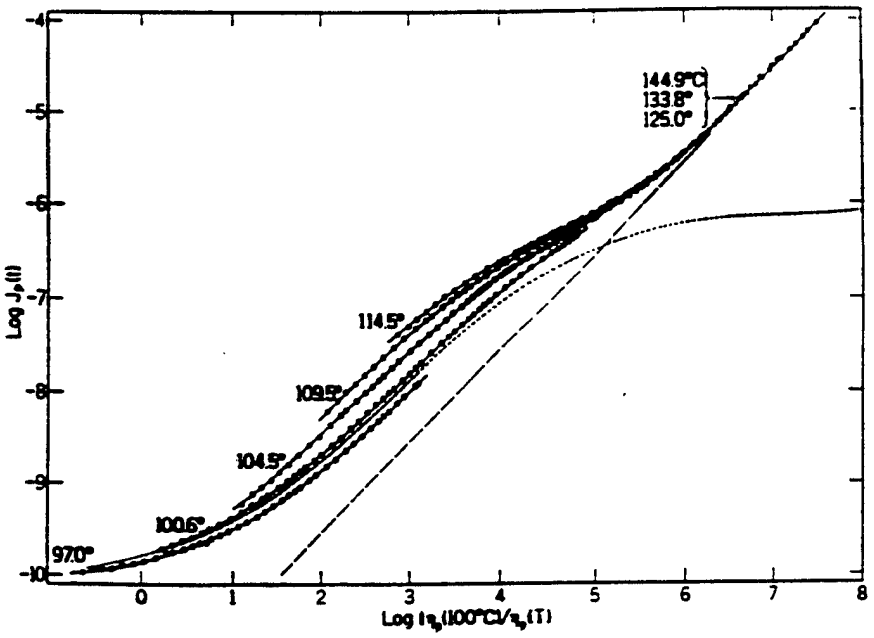


Fig. 19. Creep measurements on a polystyrene with molecular weight 46,900, reduced from different temperatures as indicated to 100°C with shift factors calculated from steady flow viscosity. (After Plazek.) Subscript p denotes multiplication by $T\rho/T_o\rho_o$. This plot shows the stronger temperature dependence of the segmental motion compared with that of viscosity as predicted by the coupling model

$n_\alpha > n_\eta$ is satisfied for PS. However PIB with a chemical structure having less capacity for intermolecular coupling has a smaller n_α =0.45. For PIB $n_\alpha \approx n_\eta$ and hence the degree of thermorheological complexity seen in PS should be much reduced in PIB. This clear prediction

is verified by creep measurements by Plazek and coworkers in a high molecular weight PIB. The data can be reduced successfully to a master curve (see Fig.20) in contrast to the failure found in PS (Fig.19).

(3b) Degree of Thermorheological Complexity: Low Molecular Weight

In unentangled low molecular weight PS Plazek and O'Rourke (Ref.71) have found that the steady state recoverable compliance J_e decreases as temperature is lowered towards T_g. The decrease of J_e is tantamount to loss of contribution to viscoelastic response from the longer time Rouse modes, leading to spectacular breakdown of thermorheological simplicity. This anomaly has been explained by the large n_α of PS leading to a higher temperature sensitivity of the shift factor of the local segmental motion than that of the Rouse modes (Ref.72,65). The smaller n_α of PIB reduces the difference between the temperature sensitivity of the two

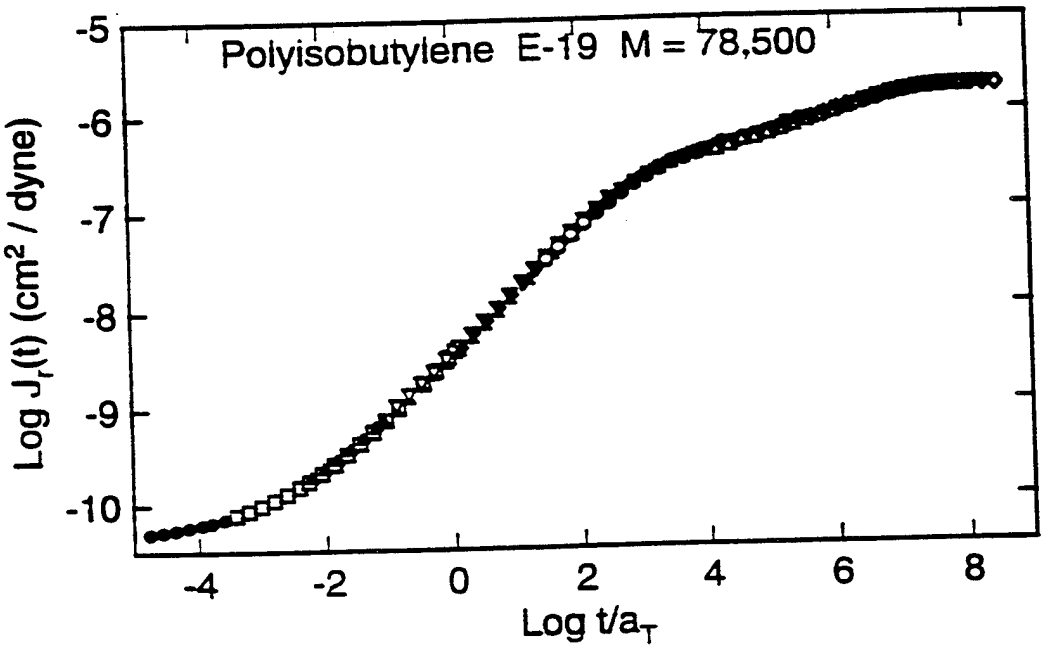


Fig. 20. Logarithm of the recoverable compliance data of PIB E-19 plotted against the logarithm of the reduced time scale with the reference temperature T_o=-72.6°C.

shift factors and as a consequence the decrease of J_e with falling temperature is less severe in PIB than for PS. In Fig.21 the creep data of PIB with molecular weight equal to 10,700 is compared with that of PS with molecular weight equal to 16,400. These samples are

compared because they have comparable maximum J_e value and about the same number of backbone carbon atoms. By inspection of Fig.21 it is clear that the decrease of J_e found in PS is much reduced in PIB as expected.

CONCLUSION

The coupling model is better known to workers in the field of polymers for the many successes in applications to problems. Not so well known is its theoretical foundation. One of the objectives of this paper is to point out that densely packed ineracting systems like polymers have Hamiltonians that are chaotic in character, and hence their dynamics are most appropriately described by a model that is based on chaos. A simple chaotic Hamiltonian and an interacting nonlinearly coupled oscillators are used to show that the basic results of the coupling model can be derived from first principles. These results indicate that theoretical

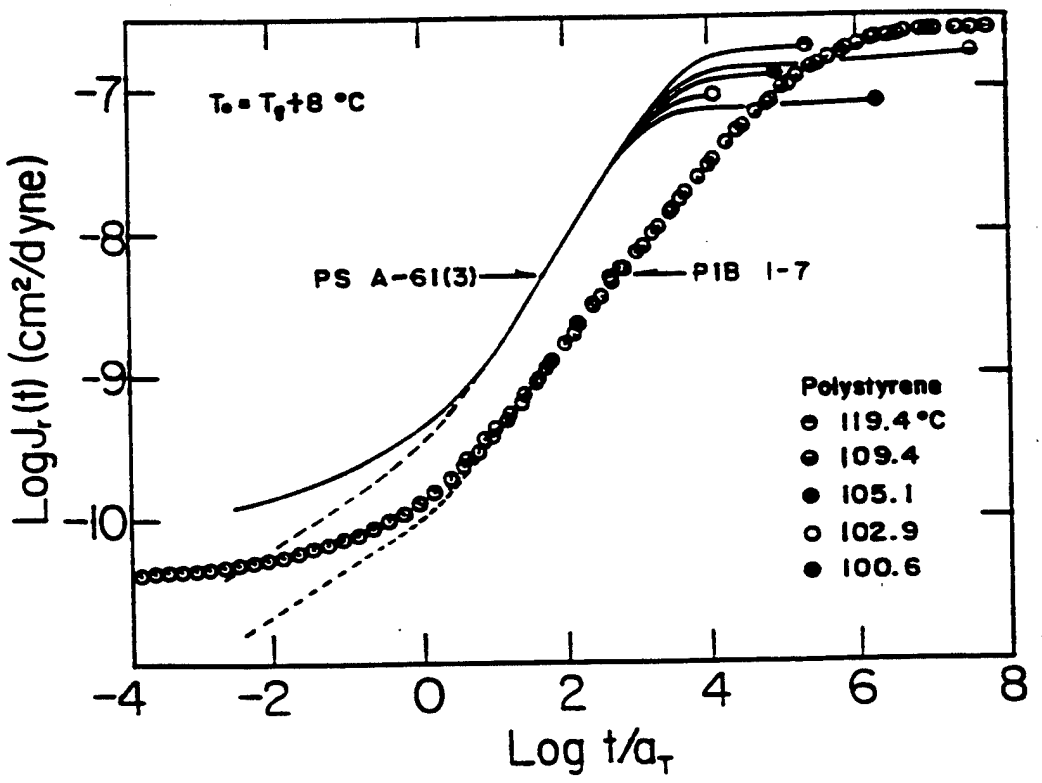


Fig. 21. Double logarithmic plot of the recoverable compliance, $J_r(t)$, as a function of reduced time, t/a_T, for PIB I-7 and PS A-61(3). Temperatures where data points for I-7 were obtained are indicated. The reference temperatures for reduction T_o are T_g +8°C for both specimens. The dashed lines are log $[J_r(t)-J_g]$, where J_g is the glassy compliance. For polystyrene, the five levels of steady-state compliance in descending order correspond to temperatures of 119.4, 109.4, 105.1, 102.9, and 100.6°C, respectively

approach based on chaos seems promising. However more work is needed particularly on larger systems with more realistic interactions before the theoretical problem can be considered to be solved.

In spite of the fact that a basic theory of the coupling model is still at large, microscopic experimental data including quasielastic neutron scattering and molecular dynamics simulation have remarkably confirmed the basic results of the coupling model. There exists indeed a temperature insensitive cross over time from independent to coupled dynamics. The key predictions of the coupling model are also verified by the experimental data.
The momentum we have built over the last decade in applying the coupling model to various problems of relaxation is substantial. We are able to solve many difficult problems in the relaxation of polymers. This capability is demonstrated by showing how we can now understand the different viscoelastic properties of polymers with different chemical structures. Polystyrene and polyisobutylene are used as examples with different molecular structures and viscoelastic properties.

ACKNOWLEDGEMENT

KLN is supported in part by ONR Contract N0001494WX23010. We thank A. Arbe, Juan Colmenero, and R.J. Roe for making their published data available to us and permission to reproduce some of these published data.

REFERENCES

(1) K.L. Ngai, *Comments Solid State Phys.*, **9**, 127 (1979)
(2) K.L. Ngai and C.T. White, *Phys. Rev. B*, **20**, 2475 (1979)
(3) A.K. Rajagopal, K.L. Ngai and S. Teitler, *J. Phys. C.*
(4) K.L. Ngai and A.K. Rajagopal, *"Non-Debye Relaxation in Condensed Matter"*, eds. T.V. Ramakrishnan and M. Raj Lakshmi, World Scientific, p. 387 (1987)
(5) R. Kohlrausch, *Ann. Phys. (Leipzig)*, **12**, (1847)
(6) G. Williams and D.C. Watts, *Trans. Faraday Soc.*, **66**, 80 (1970)
(7) K.L. Ngai, R.W. Rendell and H. Jain, *Phys. Rev. B.*, **30**, 2133 (1984)
(8) K.L. Ngai and D.J. Plazek, *J. Polym. Sci. Polym. Phys.*, **Ed.23**, 2159 (1985)
(9) G.B. McKenna, K.L. Ngai and D.J. Plazek, *Polymer*, **26**, 1651 (1985)
(10) A.A. Jones, J.F. O'Gara, P.T. Inglefield, J.T. Bendler, A.F. Yee and K.L. Ngai, *Macromolecules*, **16**, 685 (1983)

(11) K.L. Ngai and D.J. Plazek, *J. Polym. Sci.*, Polym. Phys. Ed. **24**, 619 (1986)

(12) For a recent review see K.L. Ngai, Chapter in *Disorder Effects on Relaxation Processes*, R. Richert and A. Blumen (ed.), Springer-Verlag, Berlin, pp. 89-150 (1994)

(13) K.L. Ngai, A.K. Rajagopal and S. Teitler, *Physica A*, **133**, 213 (1985)

(14) A.K. Rajagopal, R.W. Rendell, K.L. Ngai and S. Teitler, *Physica A*, **149**, 358 (1988)

(15) K.L. Ngai, A.K. Rajagopal and S. Teitler, *J. Chem. Phys.*, **88**, 5086 (1988)

(16) C.E. Porter, *Statistical Theories of Spectra: Fluctuations*, Academic, New York (1965)

(17) M.L. Mehta, *Random Matrices and the Statistical Theory of Energy Levels*, Academic, New York (1967)

(18) W.P. Reinhardt in G. Casati, (ed) *Chaotic behavior in Quantum Systems*, NATO ASI Series, Vol. 120, Plenum, New York (1985)

(19) R.U. Hag, A. Pandey and O. Bohigas, *Phys. Rev. Lett.*, **48**, 1086 (1982)

(20) R.H.G. Helleman, in *Fundamental Problems in Statistical Mechanics*, **Vol. 5**, E.G.D. Cohen (ed.), North Holland, Amsterdam, p. 165 (1980)

(21) M.V. Berry, *Ann. Phys.*, **131**, 163 (1981)

(22) M.V. Berry, *Proc.Roy.Soc.London* A**423**, 219 (1989).

(23) M. Gutzwiller, *Chaos in Classical and Quantum Mechanics*, Springer, Berlin Heidelberg New York (1990)

(24) K.L. Ngai, A.K. Rajagopal, R.W. Rendell and S. Teitler, *IEEE Trans. Electrical Insulation*, **E121**, 313 (1986)

(25) K.L. Ngai and R.W. Rendell, *Handbook of Conductivity Polymers*, **Vol. 2**, T.A. Skotheim, editor, Marcel Dekkar, New York (1986)

(26) *Der Spiegel*, cover page of Nr38/47, Jahrgang, 27 September 1993

(27) H. Goldstein, *Classical Mechanics*, 2nd ed., Addison-Wesley, Reading, MA (1980).

(28) K.L. Ngai, S.L. Peng, and K.Y. Tsang, *Physica A*, **191**, 523 (1992)

(29) K.Y. Tsang and K.L. Ngai, submitted to *Phys. Rev. E.*

(30) K.Y. Tsang and K.L. Ngai, submitted to *Phys. Rev. E.*

(31) L.J. Lewis and G. Wahnström, *Phys. Rev. E*, (1994) in press

(32) R.J. Roe, *J. Chem. Phys.*, **100**, 1610-1619 (1994)

(33) A.L. Lichtenberg and M.A. Lieberman, *Regular and Chaotic Dynamics*, Springer, New York (1992)

(34) R. Kubo, *Statistical Mechanics*, Springer Verlag, New York 1987.

(35) R. Balescu, *Equilibrium ans Nonequilibrium Statistical Mechanics*, John Wiley and Sons, New York (1975).

(36) E. Fermi, *Phys. Rev.*, **75**, 1169 (1949)

(37) K.Y. Tsang and M.A. Lieberman, *Physica D*, **11**, 147 (1984)

(38) K.Y. Tsang and M.A. Lieberman, *Phys. Lett. A*, **103**, 175 (1984)

(39) M.A. Lieberman and K.Y. Tsang, *Phys. Rev. Lett.*, **55**, 908 (1985)

(40) K.Y. Tsang and M.A. Lieberman, *Physica D*, **21**, 401 (1986)

(41) S.W. McDonald and A.N. Kaufman, *Phys. Rev. Lett.*, **42**, 1189 (1979)

(42) K. Wiesenfeld and P. Hadley, *Phys. Rev. Lett.*, **62**, 1335 (1989)

(43) K.Y. Tsang and K. Wiesenfeld, *Appl. Phys. Lett.*, **56**, 495 (1990)

(44) K.Y. Tsang, R.E. Mirollo, S.H. Strogatz and K. Wiesenfeld, *Physica D*, **48**, 102 (1991)

(45) K.Y. Tsang and I.B. Schwartz, *Phys. Rev. Lett.*, **68**, **2265-2268** (1992)

(46) K.Y. Tsang and I. Schwartz, *Int. J. Bifurcation and Chaos*, 3, 773 (1993).

(47) A. T. Winfree, *J. Theor. Biol.*, **16**, 15 (1967); *The Geometry of Biological Time*, (Springer, New York) (1980)

(48) S.H. Strogatz, R. Mirollo and P.C. Matthews, *Phys. Rev. Lett.*, **68**, 2730 (1992)

(49) J. Colmenero, A. Arbe and A. Alegría, *Phys. Rev. Lett.*, **71**, 2603-2606 (1993)

(50) J. Colmenero, *Physica A*, **201**, 38-51 (1993)

(51) J. Colmenero, A. Arbe, A. Alegría and B. Frick, *Phys. Rev. Lett.*, **69**, 478-482 (1992)

(52) K.L. Ngai, J. Colmenero, A. Arbe and A. Alegría, *Macromolecules*, **25**, 6727-6729 (1992)

(53) K.L. Ngai and A.F. Yee, *J. Polymer Sci.:, Part B:Polymer Physics*, **29**, 1493-1501 (1991)

(54) J. Colmenero, private communication (1994)

(55) G.F. Signorini, J.L. Barrat, and M.L. Klein, *J. Chem. Phys.*, **92**, 1294-1303 (1990)

(56) L.J. Lewis and G. Wahnström, *Phys. Rev. E*, (1994) in press.

(57) R.J. Roe, *J. Chem. Phys.*, **100**, 1610-1619 (1994)

(58) N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, (New York, NY: Holt, Rinehart and Winston, (1976)), Appendix N

(59) K.L. Ngai and C.M. Roland, *Macromolecules*, **26**, 6824-6830 (1993)

(60) K.L. Ngai and D.J. Plazek, to appear in *Handbook of Polymer Properties*, ed.J.E. Mark (American Institute of Physics, New York, NY),

(61) R. Böhmer, K.L. Ngai, C.A. Angell and D.J. Plazek, *J.Chem.Phys.*,**99**, 4201 (1993).

(62) R. Rizos, T. Jian, G. Fytas and K.L. Ngai, *Macromolecules*, in press.

(63) D.J. Plazek and K.L. Ngai, *Macromolecules*, **24**, 1222 (1991).

(64) D.J. Plazek, Z.X. Zheng and K.L. Ngai, *Macromolecules*, **25**, 4920 (1992).

(65) K.L. Ngai, D.J. Plazek and C. Bero, *Macromolecules*, **26**, 1605 (1993).

(66) A.V. Tobolsky, *Properties and Structure of Polymers*, (John Wiley & Sons, New York, NY(1960)).

(67) D.J. Ferry, *Viscoelastic Properties of Polymers 3rd ed.*, (Wiley, New York, NY (1980)).

(68) C.A. Angell in *Relaxations in Complex Systems*, ed. K.L. Ngai and G.B. Wright, (US Government Printing House, Naval Research Laboratory, Washington, D.C. (1984), p.3; *J.Non-Cryst.Solids*, **131-133**, 13 (1991).

(69) K.L. Ngai and C.M. Roland, *Macromolecules* **26**,6824(1993).

(70) K.L. Ngai and D.J. Plazek, *J.Polym.Sci.Polym.Phys.Ed.*, **24**, 619 (1986).

(71) D.J. Plazek and V.M. O'Rourke, *J.Polym.Sci.Polym.Phys.Ed* **9**, 209 (1971).

(72) K.L. Ngai, D.J. Plazek and S.S. Deo, *Macromolecules*, **20**, 3047 (1987).

Appendix GG

Relaxation in Interacting Arrays of Oscillators

RAPID COMMUNICATIONS

PHYSICAL REVIEW E

STATISTICAL PHYSICS, PLASMAS, FLUIDS, AND RELATED INTERDISCIPLINARY TOPICS

THIRD SERIES, VOLUME 54, NUMBER 3 PART A SEPTEMBER 1996

RAPID COMMUNICATIONS

The Rapid Communications section is intended for the accelerated publication of important new results. Since manuscripts submitted to this section are given priority treatment both in the editorial office and in production, authors should explain in their submittal letter why the work justifies this special handling. A Rapid Communication should be no longer than 4 printed pages and must be accompanied by an abstract. Page proofs are sent to authors.

Relaxation in Interacting Arrays of Oscillators

Kwok Yeung Tsang[1,2] and K.L. Ngai[1]

[1]Naval Research Lab., Washington, DC 20375

[2]Science Applications International Corp., McLean, VA 22102

0. ABSTRACT

We analyze a system of interacting arrays of globally coupled nonlinear oscillators. The relaxation in the interacting arrays with different interaction strengths is compared to that in an array not subject to interaction with others. The relaxation of the latter is found to be an exponential function of time. On the other hand the relaxation of the interacting arrays is slowed down and departs from an exponential of time. There exists a cross-over time, t_c, before which relaxation of the interacting arrays is still an exponential function. However, beyond t_c, relaxation is no longer exponential but well approximated by a stretched exponential, $\exp-(t/\tau)^{\beta}$. The fractional exponent β decreases further from unity with increasing interection strength. The result bears strong similarity to the basic features suggested by the coupling model and seen experimentally by neutron scattering for relaxation in densely packed interacting molecules in glass forming liquids.

PACS numbers: 05.45.+b, 82.20.Rp, 31.70.Hq, 05.40.+j

1. INTRODUCTION

Recent results [1] have shown that studies of nonlinear dynamical systems can enhance the understanding of some fundamental problems in physics, such as stability of the solar system, phase transitions, turbulance, and the ergodic problems in statistical mechanics. One common characteristic of such systems is the irreversibility of the dynamics due to sensitive dependence on initial conditions. Lately in the community of research in chaos there is considerable interest in nonlinearly coupled oscillators. Various models of a globally coupled oscillator array [2-6] have been studied. It was conjectured and later proved (though with conditions) that for coupling strength below a certain threshold, such systems would relax to an incoherent state [7,8]. Recently the relaxation process in such an array is found to be exponentially fast [9].

It would be of interest to study the relaxation of a more complex system, consisted of a number of such arrays coupled together by nonlinear interactions. Such study may be beneficial to the understanding of relaxation processses in glass-forming viscous liquids, polymers and ionic conductors, to name a few. These problems in condensed matter physics, physical chemistry and materials science involve irreversible proceeses in densely packed interacting systems [10]. The interactions in these systems come from nonlinear potentials such as that of Lennard-Jones in polymers and of Coulomb in vitreous fast ionic conductors.Each array is considered as a relaxing molecular unit. The nonlinear coupling between the arrays mimics the effect that interactions between the molecular units has on the relaxation of the individual molecular units. What we learn from the solution of this problem should be promising to shed some light on relaxation of nonlinearly coupled many body problems in condensed matter physics.

2. GLOBALLY COUPLED OSCILLATORS

We consider a simplified version of an array of coupled phase-only oscillators studied by Strogatz and coworkers [9]. The governing equations are

$$\frac{d}{dt}\varphi_i = \frac{K}{N}\Sigma_{j=1}^{N}\sin(\varphi_j-\varphi_i)$$

for $i, j=1,...,N>>1$. The problem was originally motivated by the study of the biological phenomenon of mutual synchronization [7]. Since the state of the system is a point on an N-fold torus, we reset φ when it exceeds 2π. We further simplify the problem to a map by picking appropriate timesteps and rescaling the time. The system becomes a map:

$$\varphi_i' = \varphi_i + \frac{K}{N}\Sigma_{j=1}^{N}\sin(\varphi_j-\varphi_i) \qquad (1)$$

Like Strogatz et al. we are interested in the decay of the phase coherence, r, which is the absolute value of the order parameter, $re^{i\psi}$, defined by

$$r = |re^{i\psi}| = |\frac{1}{N}\Sigma_j e^{i\varphi_j}| \qquad (2)$$

We find that our array of coupled oscillators decays to an incoherent state, i.e. $r=0$, for $K<0$. Numerical calculation shows that the decay of r is almost exactly an exponential function of time at least up to the longest time we have carried out so far and 99.9% of the initial value of r has decayed (see Fig.1). This resembles the (Debye) relaxation of an isolated molecule in dilute solution which is also exponential.

3. SYSTEMS OF INTERACTING ARRAYS

Now if we allow a number of these arrays to interact to obtain an even more complex system, the interesting problem is to find out what modification these interactions will have on the relaxation of r for each array towards incoherence. The interactions betweeen the arrays $\alpha=1,...,M$ can be chosen in several ways. One choice we have made is indicated by the new maps:

$$\varphi_{i\alpha}' = \varphi_{i\alpha} + \frac{K}{N}\Sigma_{j=1}^{N}\sin(\varphi_{j\alpha}-\varphi_{i\alpha}) - \frac{K'}{N^2}\Sigma_{\beta=1}^{M}\Sigma_{j=1}^{N}\sin(\varphi_{j\beta}-\varphi_{i\alpha}) \qquad (3)$$

The strength of the interactions between the arrays is measured by K'/K. The interacting arrays mimics an assembly of molecules densely packed together.

4. NUMERICAL RESULTS: RELAXATION OF PHASE COHERENCE

We have iterated the maps to obtain the evolutions of the coupled arrays system numerically. From the result we calculate the decay of the phase coherence r for each array. Figures 1 shows the decay of r is slowed down by the interactions between the arrays. By inspection (and as shown in the inset) it is clear that the degree of slowing down varies directly as the interaction strength K'/K which has been chosen to have the values of 0.6, 0.8, 1.0 and 1.2. In all these calculations, M=3 and N is fixed at the value of 32. The results shown in Fig.1 are calculated for M=3. The results are consistent with an initial

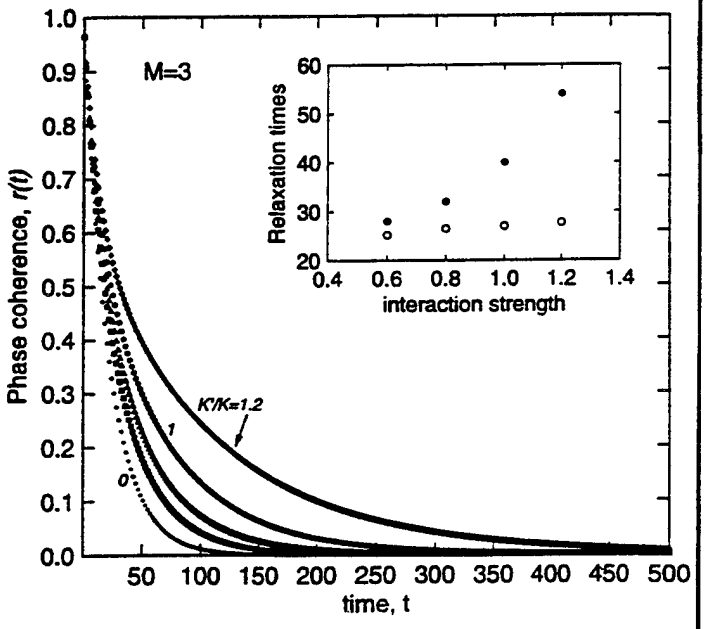


Fig. 1. Decay of phase coherence, $r(t)$, calculated numerically for K=-0.03, M=3, and interaction strength K'/K= 0 (dot), 0.6 (square), 0.8 (triangle), 1.0 (circle), and 1.2 (diamond). The calculated $r(t)$ for K'/K=0, corresponding to a single array with dynamics described by Eqn. (1), conforms well to the exponential decay. The inset shows the relaxation times τ_0 (circles) and τ^* (dots), as functions of interaction strength.

exponential decay, $\exp-(t/\tau_o)$, and crossover at some t_c to a stretched exponential decay, $\exp-(t/\tau^*)^\beta$ as indicated by the curves drawn in Fig. 2. The Kohlrausch exponent [11] β is actually the slope of the solid line fits in Fig. 2. Fitting the calculated results this way we find that the relaxation time τ^* increases and the exponent β decreases with the interaction strength, K'/K, between the arrays (see insets of Figs.1 and 2). These behaviors are in accord with the results of the coupling model [12] for relaxation of densely packed interacting molecular systems. Continuity of the two pieces, $\exp(-t/\tau_o)$ and $\exp-(t/\tau^*)^\beta$ at t_c guarantees the validity of the relation between τ^* and τ_o given by:

$$\tau^* = [t_c^{-n}\tau_o]^{\beta} \tag{4}$$

which is the same as

$$\tau^* = [(1-n)\omega_c^n \tau_o]^{\frac{1}{1-n}} \tag{5}$$

written down in the original version of the coupling model, provided the identifications

$$t_c = (1-n)^{\frac{1}{n}}\omega_c^{-1} \tag{6}$$

and $\beta = 1-n$ are made. The reciprocal of the ω_c in Eq.(5) is the time at which the relaxation rate defined by $-(1/C(t))dC(t)/dt$ are equal for the two pieces of correlation functions: $C(t) = \exp-(t/\tau_o)$ for $t < t_c$, and $\exp-(t/\tau^*)^{1-n}$ for $t < t_c$; and t_c is the time at which they are continuous. Figure 1 shows that t_c decreases with increasing inter-array interaction strength. Figure 2 also shows that t_c (indicated by arrows) decreases with increasing inter-array interaction strength. Thus the relaxation of the nonlinearly coupled arrays of coupled nonlinear oscillators has verified the basic features of the coupling model.

The degree of slowing down in relaxation due to the inter-array interaction depends also on M, increasing with M but levels off for $M>4$. Figure 3 shows the results for different M's. Note that $M=1$ is the case without any interaction.

The good correspondence between relaxation of simpler prototype chaotic systems and the key features of the coupling model shows that a theory of the coupling model is in the making. The encouraging results obtained will provide impetus for future study of more complex chaotic Hamiltonian systems that bear closer resemblance to polymers, small molecule glass-forming liquids and

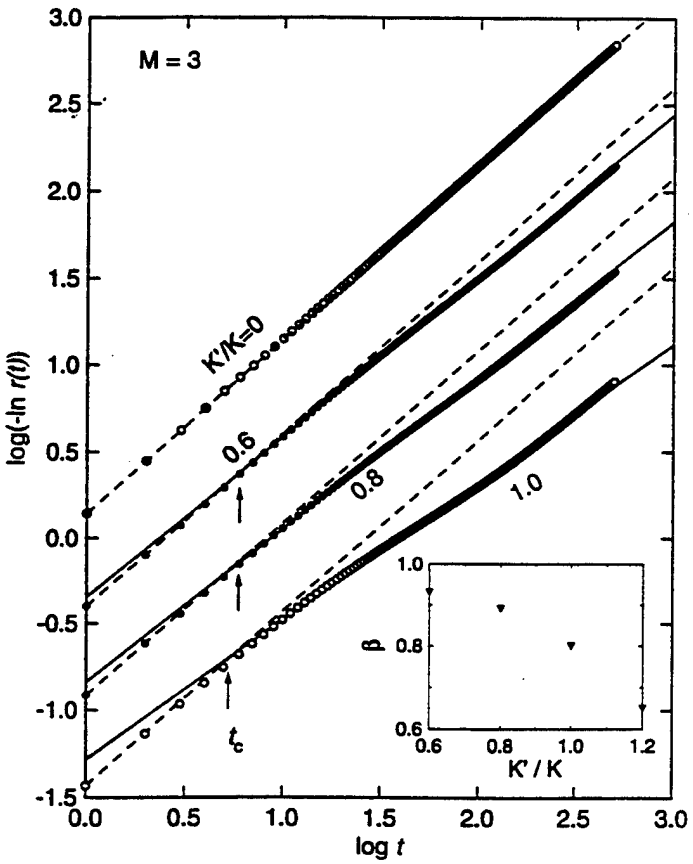


Fig. 2. Log[-ln $r(t)$] verses log t, calculated numerically for K=-0.03, M=3, and interaction strength K'/K= 0, 0.6, 0.8, and 1.0. The dashed lines are the exponential fits and the solid are the stretched exponential fits. Curves for K'/K = 0, 0.6, and 0.8 are shifted vertically up in multiples of 0.5 to avoid overlapping on the graph. The cross-over time t_c is indicated by a vertical arrow. The calculated $r(t)$ for other interaction strengths conforms well to the $\exp[-(t/\tau_0)]$ for $t<t_c$ and the $\exp[-(t/\tau_0\tau^*)^\beta]$ for $t>t_c$. The inset shows the Kohlrausch exponent βas a function of interaction strength. The exponent β for K'/K=1.2 case is also included in the inset.

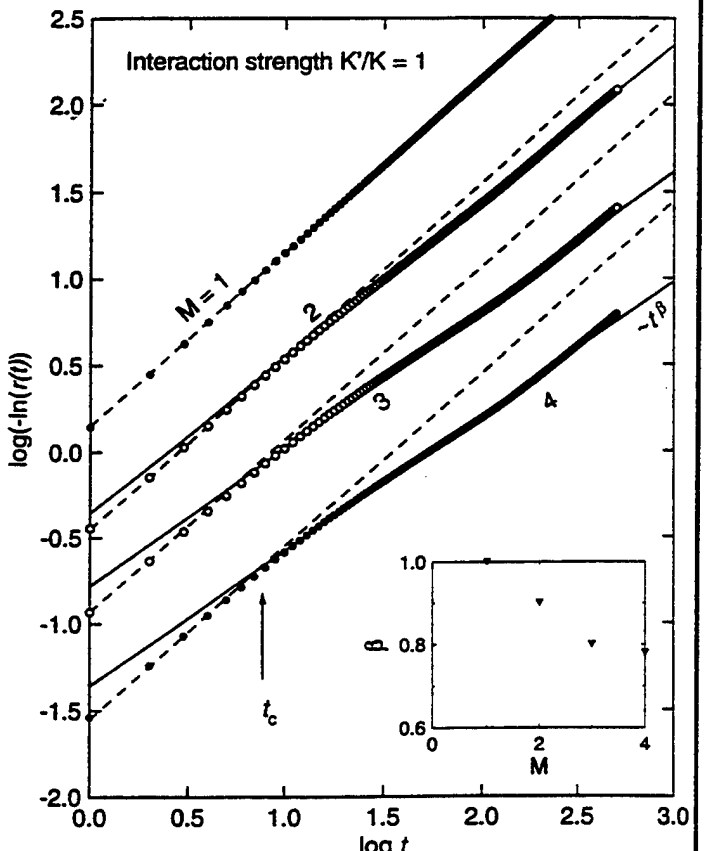


Fig. 3. Log[-ln $r(t)$] verses log t, calculated numerically for K=-0.03, interaction strength K'/K= 1, and different M's. Curves for M = 1, 2, and 3 are shifted similarly to curves in Fig. 2. The calculated $r(t)$ for M=1, actually the same as for K'/K=0 in Fig. 2, corresponding to a single array with dynamics described by Eqn. (1), conforms well to the exponential decay. The calculated $r(t)$ for other M's conforms well to the $\exp[-(t/\tau_0)]$ for $t<t_c$ and the $\exp[-(t/\tau^*)^\beta]$ for $t>t_c$. The inset shows β as a function of M.

vitreous ionic conductors.

5. CONCLUSION AND DISCUSSION

The effect of interactions between M arrays of coupled nonlinear oscillators has been studied. It is found that for sufficiently large t, $r \sim \exp(-t/\tau)$ for $M=1$, indicating that the coupled array has exponential relaxation. For other $M > 1$ with nonzero interaction strength, it is found that there exists a time t_c such that r is still an exponential function of time, $\exp(-t/\tau_0)$, for $t < t_c$, but $r \sim \exp -(t/\tau^*)^\beta$ for $t > t_c$. Thus the relaxation of the interacting arrays proceeds with a stretched exponential time dependence for $t > t_c$ and $\tau^* >> \tau$. Thus the interaction between the arrays is observed to slow down the relaxation of single arrays. As expected, the fractional exponent β decreases further from unity with increasing nonlinearity or larger interaction strengths.

In the study of relaxation in real systems in physics, chemistry and materials science, it is found that a system without interactions (e.g., an isolated molecule in dilute solution) usually relaxes exponentially. However when such systems are densely packed and interacting with each other the relaxation proceeds differently and exhibits many fascinating properties. A coupling model has been very successful in explaining these properties [12]. This model is based on the hypothesis that an interacting system relaxes initially exponentially until a temperature independent microscopic time, t_c, but stretched exponentially afterwards with continuity of the correlation function at the time of crossover [13,14]. A recent neutron scattering measurement on a polymer has shown direct experimental evidence for this hypothesis [15]. In our present work, the addition of inter-array interactions introduces additional nonlinearity which has similar effects on the relaxation towards equilibrium as many body interactions have on relaxation in densely packed molecular systems. Since the numerical result obtained bears strong similarity to the basic features suggested by the coupling model [13] and seen in neutron scattering experiment [15], the interacting array model provides a useful first step in applying nonlinear dynamical models to the study of irreversible processes of real physical systems in physics, chemistry and materials science.

6. ACKNOWLEDGMENTS

This work is supported by ONR contract N0001495WX20203.

7. REFERENCES

[1] A.L. Lichtenberg and M.A. Lieberman, Regular and Chaotic Dynamics, (Springer, New York, 1992).

[2] K. Wiesenfeld and P. Hadley, Phys. Rev. Lett. 62, 1335 (1989).

[3] K.Y. Tsang and K. Wiesenfeld, Appl. Phys. Lett. 56, 495 (1990).

[4] K.Y. Tsang, S.H. Strogatz, and K. Wiesenfeld, Phys. Rev. Lett. 66, 1094 (1991).

[5] K.Y. Tsang and I.B. Schwartz, Phys. Rev. Lett. 68, 2265 (1992).

[6] I.B. Schwartz and K.Y. Tsang, Phys. Rev. Lett. 6738, 2797 (1994).

[7] A. T. Winfree, J. Theor. Biol. 16, 15 (1967).

[8] A. T. Winfree, The Geometry of Biological Time, (Springer, New York, 1980).

[9] S.H. Strogatz, R. Mirollo and P.C. Mattews Phys. Rev. Lett. 68, 2730 (1992).

[10] See papers presented in Proceedings of the International Discussion Meeting on Relaxations in Complex Systems, J. Non-Cryst. Solids 131-133 (1991); 172-174 (1994).

[11] R. Kohlrausch, Pogg. Ann. Physik. (3) 12. 393 (1847); (4) 1, 179 (1854).

[12] For a recent review see K.L. Ngai, chapter in Disorder Effects on Relaxation Processes, R. Richert and A. Blumen (ed.), Springer-Verlag, Berlin (1994), pp 89-150.

[13] K.L. Ngai, Comments Solid State Phys. 9. 127 (1979).

[14] K.L. Ngai, S.L. Peng, and K.Y. Tsang, Physica A191, 523 (1992).

[15] J. Colmenero, A. Arbe and A. Alegria, Phys. Rev. Lett. 71, 2603 (1993); R. Zorn, A. Arbe, J. Colmenero, B. Frick, D. Richter, and U. Buchenau, Phys. Rev. E 52, 781 (1995).

Appendix HH

Numerical Simulation of Combustion in Fire Plumes

NUMERICAL SIMULATION OF COMBUSTION IN FIRE PLUMES

Kuldeep Prasad

Science Applications International Corporation
Applied Physics Operations
1710 Goodridge Drive
McLean, VA 22102

May 9, 1996

Numerical Simulation of Combustion in Fire Plumes

1 ABSTRACT

This study deals with a fundamental numerical investigation of chemically reacting fluid flows through 2-D burners. The goal of this research is to develop a combustion model appropriate for buoyancy driven fire plumes and thus study the structure and energetics of hydrocarbons burning in a pool fire configuration. We make use of finite chemical kinetic rate equations to numerically simulate a laminar diffusion flame. The code has been constructed to consider the viscous effects in a mixing layer, heat conduction, the multi-component diffusion and convection of important species, the finite rate reactions of these species, and the resulting interactions between the fluid mechanics and the chemistry. The numerical model has been used to obtain a detailed description of laminar diffusion flames obtained above 2-D methane/air burners. It is shown that the leading edge flame, a flame holding point for the rest of the diffusion flame, is dominated by the kinetic aspects of the fuel/oxidizer species and is mainly responsible for heat transfer to any upstream boundary surface. Results are compared with experimental thermocouple measurements for methane - air burners with similar geometries and flow configurations.

2 INTRODUCTION

Many common fire scenarios can be classified as pool fires. These include fires ranging in size from a candle flame, where D is approximately 10^{-3} m, to a forest fire, where D can be as large as 10^{5} m. A pool fire is defined as a buoyant diffusion flame in which the fuel is configured horizontally. Although the name implies that the fuel is a liquid, it may be a gas or a solid. The fuel bed may be of an arbitrary geometry, but for simplicity, most studies consider a circular configuration characterized by a single geometrical scale, the pool diameter (D). The overall combustion is governed by

a complex interplay of chemical reactions, transport and gas dynamic processes that are strongly dependent on physical boundary conditions and type of chemical system. The ability to predict the coupled effects of various complex transport processes and chemical kinetics in these systems is critical in predicting flammability limits, stability criteria and extinction limits.

A numerical model was constructed for the purpose of understanding the interplay of diffusion, convection and chemical reactions, and the resulting flame complex when a laminar mixing flow undergoes exothermic chemical reactions. Fig. 1 provides a schematic diagram of the burner geometry and the computational domain in which the solution is desired. The classical description of a diffusion controlled flame in such a configuration pictures a thin flame in the mixing region that extends all the way to the burner surface. In the present study there was particular concern with that region of the flame near the burner surface, where heat loss from the reaction region to surrounding flow limits the temperature rise and thus limits reaction rates. In this region close to the initial point of fuel-oxidizer contact the assumption of diffusion limited flames are inapplicable, and detailed consideration of coupled thermal and species diffusion and chemical reaction rates is required for a realistic description of the flame complex.

Near the burner surface, the concentration gradients of primary reactants are high, and mixing rates correspondingly high. Chemical reaction rates are low because of the relatively low temperatures, so that a small region of partially premixed reactants develops, which is increasingly heated as it moves outwards towards a flame that stabilizes in the mixing region. The leading edge of this flame (referred to here as the "leading edge flame" or "LEF") differs from the trailing diffusion limited flame in that it stands in a region of premixed gases that are continuously premixed in the approach flow and lead to very concentrated reaction because of the premixing. The flame cannot move upstream all the way to the burner surface, and must stabilize at some location where the temperature dependent reaction rate can provide heat release rate consistent with multi-dimensional heat outflow. The details of this complex flame region are not well known because it is usually very small, and difficult to observe experimentally. It is

equally difficult to describe analytically because of the necessity to consider coupled convective flow, species diffusion, heat transfer and reaction processes in at least two space dimensions.

As a process, fire can take many forms, all of which involve chemical reaction between combustible species and oxygen from the air. Properly harnessed, it provides great benefit as a source of power and heat to meet our industrial and domestic needs, but unchecked, can cause untold material damage and human suffering. A detailed understanding of fire dynamics and the various physical and chemical processes involved therein, will be crucial in designing fire protection systems and combating wild fires.

The process of flame propagation over a fuel bed such as a forest, a densely built up area or a combustible liquid pool is basically the same. It involves:

- Preheating of the fuel bed ahead of the flame.
- Evaporation or evolution of the volatile components of the fuel.
- Fuel-oxidant mixing and
- Ignition and combustion of the gaseous mixture.

Flame spread is controlled by the slowest of the above processes, i.e. by preheating the fuel bed. Energy input to the fuel bed is the most difficult to resolve since it requires treatment of spatially varying flame properties. This energy transfer is dependent on radiation, convection and heat conduction in both the gas and solid phase. Two approaches can be applied to deal with this problem.

1. A complete description based on computer modeling of all the flame processes involved, including soot generation and burnout for accurate radiative heat transfer treatment.
2. A simplified description in which the flame is assumed to be homogeneous and isothermal.

Research in fire dynamics will provide concepts and techniques which may be used by the practising fire protection engineer to predict and quantify the likely effects of fire. These areas include radiation from flames, response of ceiling mounted fire detectors and interaction between sprinkler sprays and fire plumes. The use of water, foams and gases such as Halon for fire fighting is well established. However, the quantities of these used for suppressing large fires are out of proportion to real needs, usually exceeding by two orders of magnitude those that theory predicts should be sufficient to extinguish a fire. Most of the effective suppressants in current use have a very high Ozone Depletion potential (ODP). Because of this factor, the manufacture of Halon based fire suppressants have been abolished by International treaty. So there is a pressing need to identify and evaluate new alternatives to Halon for fire fighting.

In one form on another, water is an environmentally sound alternative to Halon for fire suppression . The problem is however complicated because of the fact that the amount of suppressant such as water required for firefighting is a strong function of the location of the nozzles that inject water in a typical compartment fire. Furthermore, there is some evidence that the size of the water droplets play a major role in the effectiveness of water. A detailed understanding of the rate controlling processes in a compartment or pool fire would help in designing the optimum location of the sprinklers for injection of water mist that would most efficiently control the spread and extinguish these fires. Firefighting in naval ship fires is further constrained by lack of fresh water supply, which again demands an accurate mathematical model for describing the shape and structure of pool fires.

As opposed to water mist, the use of flame inhibitors (retardents) added in trace quantities to a flame affect the combustion process at a molecular level, by increasing or reducing the flame velocity. A sufficiently high concentration of an inhibitor leads to the extinction of the flame. Most flame retardents such as halogenated organic compounds affect combustion through reduction in radical concentration by scavenging and by atom-assisted recombination. So there is a need for new non-halogenated chemical agents, that have low ODP.

Much of our present knowledge on fire spread is through observation of real life fires and construction of empirical relations based on these observations. The detailed mechanisms by which the fire suppressants, including water work are not well known. Very little exists in terms of robust mathematical model that can predict not only the global trends such as direction of flame propagation, ignition and production and movement of smoke, but also the detailed structure of the fire itself. Therefore, mathematical models will be developed that will predict the diffusion flame structure observed above typical pool fires, and the effect of suppressants on these fires. These mathematical models will utilize various levels of combustion chemistry to describe the reactions between the fuels and the oxidizers. With such models, one can explore the various mechanisms such as surface cooling, oxygen depletion, radiation block and thermodynamic cooling that may all play a role. For example, by performing careful simulations of jet diffusion flames and pool fires, the role of radiative feedback to the fuel surface can be isolated and studied. By adding inert diluents into the coflow of a jet diffusion flame, the effects of oxygen dilution can be explores. These and similar numerical experiments will be performed to isolate the various mechanism that are responsible for fire suppression. With the understanding gained from such experiments, a more effective fire suppressant can be developed.

Detailed numerical modeling of typical pool fire requires us to resolve a wide range of temporal and spatial scales, resulting in prohibitively excessive memory and computational costs. Domain-decomposition techniques are being developed, that subdivide the computational domain into simpler subdomains which admit a more easily constructed mesh. Several strategies are being explored to subdivide the domain and establish communications among the subdomains. One group of approaches, the grid-patching or zonal methods, uses common or shared boundaries and another uses embedded or overset grids to subdivide the domain. In this manner, a fire in a large enclosure can be modelled cost effectively.

3 MATHEMATICAL FORMULATION

Modeling gas phase reactive flows is based on a generally accepted set of time dependent, coupled partial differential equations maintaining conservation of total density, momentum, total energy and individual species density. These equations describe the convective motion of the fluid, the chemical reactions among the constituent species, and the diffusive transport processes such as thermal conduction and molecular diffusion.

3.1 Governing Equations

A strong conservation form of the two dimensional, unsteady, compressible Navier Stokes equations, used to describe gas phase reactive flows for a system containing N species undergoing M elementary chemical reactions, can be written as follows

$$\partial_t \vec{q} + \partial_x \vec{E} + \partial_y \vec{F} = \partial_x \vec{R} + \partial_y \vec{S} + \vec{K} \tag{1}$$

where $\vec{q}, \vec{E}, \vec{F}, \vec{R}, \vec{S}$ and $\vec{K}$ are vectors defined as

$$\vec{q} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho e_t \\ \rho_k \end{bmatrix} \quad \vec{E} = \begin{bmatrix} \rho u \\ \rho u^2 + P \\ \rho uv \\ (\rho e_t + P)\, u \\ \rho_k u \end{bmatrix} \quad \vec{F} = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + P \\ (\rho e_t + P)\, v \\ \rho_k v \end{bmatrix}$$

$$\vec{R} = \begin{bmatrix} 0 \\ \tau_{xx} \\ \tau_{xy} \\ -q_x + u\tau_{xx} + v\tau_{xy} \\ -\rho_k U_k \end{bmatrix} \quad \vec{S} = \begin{bmatrix} 0 \\ \tau_{xy} \\ \tau_{yy} \\ -q_y + u\tau_{xy} + v\tau_{yy} \\ -\rho_k V_k \end{bmatrix} \tag{2}$$

$$\vec{K} = \begin{bmatrix} 0 \\ \sum_{k=1}^{N} \rho_k f_{x,k} \\ \sum_{k=1}^{N} \rho_k f_{y,k} \\ \sum_{k=1}^{N} \rho_k \left(u + U_k\right) f_{x,k} + \sum_{k=1}^{N} \rho_k \left(v + V_k\right) f_{y,k} \\ \dot{\omega}_k \end{bmatrix}$$

Here ρ and ρ_k are the total mass density and the individual species density; u, v are the bulk fluid velocity components in the x and y direction respectively and e_t is the total energy per unit mass. $f_{x,k}$ and $f_{y,k}$ are the body forces per unit mass acting on the k^{th} species in x and y directions respectively. The viscous stress terms τ_{xx}, τ_{xy} and τ_{yy} appearing in the conservation equations are given by

$$\begin{aligned} \tau_{xx} &= \left(\kappa - \frac{2}{3}\mu\right)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + 2\mu\frac{\partial u}{\partial x} \\ \tau_{xy} &= \mu\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \\ \tau_{yy} &= \left(\kappa - \frac{2}{3}\mu\right)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + 2\mu\frac{\partial v}{\partial y} \end{aligned} \tag{3}$$

The terms q_x and q_y appearing in the total energy equation are the net rate of heat flux in the two co-ordinate directions and can be expressed as

$$\begin{aligned} q_x &= -\lambda\frac{\partial T}{\partial x} + \sum_{k=1}^{N} \rho_k h_k U_k \\ q_y &= -\lambda\frac{\partial T}{\partial y} + \sum_{k=1}^{N} \rho_k h_k V_k \end{aligned} \tag{4}$$

The thermodynamic pressure is defined as

$$P = R^{\circ} T \sum_{k=1}^{N} \frac{\rho_k}{W_k} \tag{5}$$

where W_k is the molecular weight of the k^{th} species, and R° is the universal gas constant. The caloric equation of state is used to define the enthalpy of the individual species, which in turn is used to define the total energy as follows

$$\begin{aligned} \rho e_t &= \rho e + \frac{\rho}{2}\left(u^2 + v^2\right) \\ &= \sum_{k=1}^{N} \rho_k \left(h_k^{\circ} + \int_{T^{\circ}}^{T} C_{p,k} dT\right) - P + \frac{\rho}{2}\left(u^2 + v^2\right) \end{aligned} \tag{6}$$

The specific heat for each species is obtained by using a sixth order polynomial in temperature.

$$C_{p,k} = a_{1,k} + a_{2,k}T + a_{3,k}T^2 + a_{4,k}T^3 + a_{5,k}T^4 + a_{6,k}T^5 + a_{7,k}T^6 \quad k = 1, \ldots, N \qquad (7)$$

The various specific heat coefficients and the heat of formations for the various species are obtained from the JANNAF thermochemical tables. The various diffusion coefficients such as thermal conductivity, binary diffusion coefficients, viscosity and thermal diffusion coefficient are obtained from a rigorous treatment of kinetic theory potential is constructed to model the inter-molecular potential function based on which the collision integrals are evaluated. The various diffusion coefficients are then obtained using detailed kinetic theory and are functions of the temperature, pressure and the various species properties.

3.2 Diffusion velocity model

The diffusion velocities in a multi-component reacting flow mixture may arise because of concentrations gradients, pressure gradients, differential body forces and due to the Soret effect. The diffusion velocities for each of the N species in both the x and y direction are obtained by solving the exact diffusion equation given by

$$\frac{\partial X_k}{\partial x} = \sum_{j=1}^{N} \left(\frac{X_k X_j}{\mathcal{D}_{kj}} (U_j - U_k) \right) + \frac{(Y_k - X_k)}{P} \frac{\partial P}{\partial x} + \frac{\rho}{P} \sum_{j=1}^{N} Y_k Y_j \left(f_{k,x} - f_{j,x} \right) \qquad (8)$$

$$\frac{\partial X_k}{\partial y} = \sum_{j=1}^{N} \left(\frac{X_k X_j}{\mathcal{D}_{kj}} (V_j - V_k) \right) + \frac{(Y_k - X_k)}{P} \frac{\partial P}{\partial y} + \frac{\rho}{P} \sum_{j=1}^{N} Y_k Y_j \left(f_{k,y} - f_{j,y} \right) \quad k = 1, \ldots, N$$

Here X_k and Y_k are the mole and mass fraction of the k^{th} species respectively. Since only N-1 equations of the above N equations are independent of each other, the above equations are solved subject to the constraint, that the diffusion velocities introduce no net momentum to the fluid flow i.e.

$$\sum_{k=1}^{N} \rho_k U_k = \sum_{k=1}^{N} \rho_k V_k = 0 \qquad (9)$$

3.3 Reaction kinetic model

The types of reactions of importance in combustion include unimolecular decomposition reactions, bimolecular exchange and dissociation reactions, and three-body recombination reactions. The following 7 types of chemical reactions have been considered.

1. Unimolecular decomposition $AB \rightleftharpoons A + B$
2. Bimolecular dissociation $AB + M \rightleftharpoons A + B + M$
3. Bimolecular $A + B \rightleftharpoons C + D$
4. Bimolecular exchange/dissociation $AB_2 + C \rightleftharpoons AC + B + B$
5. Recombination reaction $A + B + B \rightleftharpoons AB + B$
6. Two body recombination $C + D \rightleftharpoons CD$
7. Three body recombination $C + D + M \rightleftharpoons CD + M$

Here M is a third body molecule, and A, B, C, D are representative species.

For a system containing N species undergoing a set of M elementary chemical reaction, the general i^{th} reaction can be expressed as

$$\sum_{k=1}^{N} \nu'_{k,i} S_k \rightleftharpoons \sum_{k=1}^{N} \nu''_{k,i} S_k \quad i = 1, \ldots, M \tag{10}$$

The rate of production of the k^{th} species due to the i^{th} reaction is given by

$$\dot{\omega}_{i,k} = \left(\nu''_{k,i} - \nu'_{k,i}\right) \left[k_{f,i} \prod_{j=1}^{N} (C_j)^{\nu'_{j,i}} - k_{b,i} \prod_{j=1}^{N} (C_j)^{\nu''_{j,i}} \right] \quad k = 1, \ldots, N, i = 1, \ldots, M \tag{11}$$

and the total rate of production of the k^{th} species is obtained by using

$$\dot{\omega}_k = \sum_{i=1}^{M} \dot{\omega}_{i,k} \quad k = 1, \ldots, N \tag{12}$$

The $k_{f,i}$ and $k_{b,i}$ are the forward and reverse reaction rate constants. Each $k_{f,i}$ is a function of temperature usually given by the modified Arrhenius expression

$$k_{f,i} = A_i T^{n_i} \exp\left(\frac{-E_i}{RT}\right) \quad i = 1, \ldots, M \tag{13}$$

The reverse reaction rate constant $k_{b,i}$ is calculated from $k_{f,i}$ and the equilibrium constant(in concentration units) $K_{c,i}$ by the laws of microscopic reversibility.

$$k_{b,i} = \frac{k_{f,i}}{K_{c,i}} \quad i = 1, \ldots, M \tag{14}$$

The equilibrium constant $K_{p,i}$ for the i^{th} general reaction is defined by using the standard change in Gibb's free energy as follows :

$$R \ln K_{p,i} = -\frac{\Delta G_i^\circ}{T} \quad i = 1, \ldots, M \tag{15}$$

The equilibrium constant $K_{c,i}$ in terms of concentrations is then obtained from

$$K_{c,i} = K_{p,i} (RT)^{\sum_{k=1}^{N} (\nu''_{k,i} - \nu'_{k,i})} \quad i = 1, \ldots, M \tag{16}$$

4 Algorithm Development and Implementation

Producing a code that describes low-speed flames required the development of several new numerical methods as well as finding new ways to implement existing algorithms. An important requirement of any convection algorithm is that the numerical diffusion not be larger than the physical diffusion processes that must be resolved. The FCT method meets this criterion, but it is inherently an explicit method; so that the small timesteps it requires makes it inefficient for low-speed flows. The BIC-FCT method was developed specifically to combine the accuracy of FCT with the efficiency of an implicit method for low speed flows. BIC-FCT allows timesteps up to a hundred times larger than FCT and yet the calculation time of one BIC-FCT timestep is approximately equal to the calculation time of one timestep in the standard FCT module.

The codes uses an Eulerian representation of the convective transport instead of the Lagrangian representation. A Lagrangian formulation, though preferable, is exceedingly difficult in multidimensions. However, the present BIC-FCT code can be readily used to describe two-dimensional or three-dimensional flows. Most common methods for solving convection problems use algorithms that produce ripples near steep gradients such as in a flame or shock front. The first high-order, nonlinear, monotone algorithm,

Flux-Corrected Transport (FCT), was designed to prevent these ripples by maintaining local positivity near steep gradients while keeping a high order of accuracy elsewhere. Other nonlinear methods have been reviewed by Woodward and Collela. Although these methods are explicit there are recent reports on implicit, nonlinear methods. A major problem with applying these implicit methods to low-speed flows is that they are expensive even though they can be very accurate. The Barely Implicit Correction to Flux-Corrected Transport, BIC-FCT, was designed to overcome the problem of numerical diffusion in low-velocity implicit methods. BIC-FCT combines an explicit high-order, nonlinear FCT method with an implicit correction process. This method removes the timestep limit imposed by the speed of sound on explicit methods, retains the accuracy required to resolve the detailed features of the flow, and keeps the computational cost as low as possible.

Thermo-physical properties of the individual species and the mixture are required throughout the computation. These properties are modelled with high-order curve fits to values derived from more accurate calculations. The individual properties are combined where needed to obtain mixture properties through mixing rules. This simplified method is highly efficient yet sufficiently accurate.

The submodels representlying the various physical processes are in independent modules that are coupled together. Several modifications have been made to the usual timestep splitting method in order to increase the stability limits and improve the efficiency of the code.

4.1 Diffusive Transport Processes

Diffusive transport processes express the transfer of mass, momentum, or energy due to gradients in concentration, velocity, or temperature. In general, the diffusive transport processes to be considered are molecular diffusion, thermal conduction, viscosity and thermal diffusion. The molecular diffusion and thermal conduction terms in the energy equation are calculated explicitly using central difference approximations and coupled to convection using timestep-splitting techniques.

The molecular diffusion is described by a generalized Fickian scheme in which the diffusion velocities are defined by

$$Y_k v_k = -D_k \nabla Y_k + Y_k \sum_i D_i \nabla Y_i \tag{17}$$

where the D_k are effective diffusion coefficients and the second term on the right in included in order to explicitly ensure that the zero-mass-flux constraint

$$\sum_i Y_i v_i = 0, \tag{18}$$

is satisfied. In general, when there are more than two species, the effective diffusion coefficients for the species k through the mixture, D_k are different from binary diffusion coefficients and may be defined in terms fo mixtures rules.

4.2 Chemistry

Solving a detailed set of chemical kinetic rate equations in conjunction with the unsteady, three-dimensional fluid flow equations is beyond current computer resources. Therefore, we have adopted the approach of using simplified global-chemistry models to understand the effects of chemical energy release on the complex three-dimensional flow field. The simplest reasonable global-chemistry model one can used is a single-step irreversible chemical reaction model. For the methane air system under consideration, we have

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \tag{19}$$

with the reaction rate given by an Arrhenius form:

$$k_r = A_r exp[-E_{act}/kt] \tag{20}$$

For a global chemistry model to be truly useful, it should be able to simulate at least some aspects of the detailed chemistry model. Since the emphasis of this work is on the effects of heat release, in our simplified model we have attempted to fit the heat release profile from a detailed chemistry model for an idealized one-dimensional problem.

4.3 Inflow and Outflow Boundary Conditions

Inflow and outflow boundary conditions are imposed where appropriate in the streamwise direction (x), fre-stream boundary conditions in the cross-stream direction (y). The total mass density, species molar concentration, and velocities are specified at the inflow guard cells. A zero-slope condition on the energy allows the inflow pressure to vary in response to acoustic waves generated by events downstream. At the outflow, linear advection equations are imposed for the flow variables using the local velocity near the boundary, and the pressure is relaxed towards a specified ambient value. The inflow and outflow boundary conditions used in the model were developed and tested in previous shear-flow simulations. Initial step function profiles are specified for the streamwise velocity and for the chemically reacting species, constant profiles for the background species and product species.

5 Results and Discussion

We solve the convective transport part of the Navier Stokes equations using an explicit fourth-order phase accurate FCT algorithm and direction time step splitting techniques on structure grids. The molecular diffusion, thermal conduction and chemical reaction processes are couple to convective transport using time step splitting techniques. The scheme is suitable for the description of the flow-features in the regime of initially laminar reacting fluid flows. Modeling reactive flows such as considered here involves dealing with a variety of spatial and temporal scales, and these impose definite resolution requirements for their numerical simulation. Related resolution requirements are also imposed by the need to ensure the stability of the numerical procedures used to couple the various physical processes described by the system of equations. It is thus of considerable importance to calibrate the numerical model for the flow regimes studied.

Validation of the diffusive transport models was performed on nonreactive flow cases, by comparing temporal sequences of temperature and species concentration profiles with the results of calculation using a fully detailed model for the binary and

mixture diffusion coefficients, as well as for the molecular diffusion process. For these test problems, the inflow velocities were set to be identically zero, which constrained the flow to be essentially on-dimensional. Comparison were then performed between temporal sequences of temperature and species concentrations profiles. Calibration and validation of the chemistry model was performed in the same one-dimensional regime, by comparing the flame thickness and flame temperatures with those predicted by one dimensional laminar diffusion flame calculations including full chemistry and diffusive transport effects.

5.1 Computational Grid

The gridding is uniform in the regions where entrainment takes place. The computational cells are stretched in the cross-stream direction as the free stream boundaries in the transverse direction are approached, but are otherwise uniformly spaced in the flow region of interest. The number of computation cells varies in the ranges 60-100 in the cross-stream direction, 100-200 in the streamwise direction. The grid is held fixed in time. Courant numbers in the range 0.3 - 0.5 are used in the numerical simulation. Resolution studies are performed doubling and coarsening the mesh and examining the convergence of calculated quantities. The codes are fully vectorized and optimized for processing on CRAY-YMP computers.

5.2 Calibration and Comparison with Experimental Results

Figure 1 shows a schematic diagram of the initial flow and boundary conditions for the computations. The full computational domain consists of 96*192 cells and covers an are of 4 cm radially and 10 cm axially. Cells of 0.05 cm * 0.05 cm are concentrated around the lower and left hand boundaries (where the flame is located) and the grid is then stretched both axially and radially. To match the experimental conditions, the fuel and air cold flow velocities are both 2.8 and 25.41 cm/sec respectively for the steady flow calculation, with the fuel half duct of 0.5 cm and the air duct of 3.5

cm. The fuel and aire are preheated to 550 K and 330 K respectively, based upon experimental thermocouple measurements for a series of heights just above the burner. These data were extrapolated to estimate the temperature at the burner exit. The conditions modeled include parabolic pipe flow at the inflow boundary, symmetry at the left hand boundary, free-slip wall at the right-hand boundary, and zero-gradient outflow condition at the top of the computational domain.

Although a time-dependent simulation was conducted, the results from the computations successfully matched experimental stead-state conditions after 20000 timesteps. Figure 2 shows profiles of temperature and heat release rate above the burner surface. For the methane-air flame, a peak temperature of about 2050 K is observed, which is in agreement with analytical estimations of the adiabatic flame temperature for this case. Notice the relatively large regions of energy release near the lip of the burner which could be due to significant partial premixing. The rest of the energy release region is narrow but well defined. The energy release distributions provide a good idea of the height of the flame and will also be a valuable diagnostic tool when the interaction between additives such as water mist and the flame are studied. Based on the energy release profiles, the height of the flame is about 5.3 cm.

Figure 3 and 4 show profiles of reactant species CH_4 and O_2 and product species CO_2 and H_2O above a planar methane air burner. The velocity vectors and streamline pattern has been shown in figure 5. Figure 6 through 13 show comparison between numerically computed results and experimentally obtained thermocouple measurements at various heights above the burner surface.

The goal of this research effort was to investigate the effects of inert diluents such as nitrogen and water vapour on the structure and extinguishment behavior of hydrocarbon jet diffusion flames. Both methane and propane diffusion flames were studied but because the results are similar and since more extensive studies have been carried out with methane, only those results will be discussed. Figure 14 shows comparison of temperature profiles obtained above an axisymmetric diffusion flame burner under three conditions

- Normal methane - air flame.
- Addition of 20 % water vapour in the air flow.
- Addition of 20% water mis in the air flow.

For purposes of comparison, the maxima in the color bars for the various quantities have been kept the same. The flame is noticeably cooler as water vapour is added to the air flow. The oxygen concentration is lower since the nitrogen has displace some of it and the region of intense energy release distribution has also decreased. However, the overall region of energy release has increased resulting in an increased in the flame height. As the water vapour is replaced with water mist, the flame temperature further reduces, becuase of vapourization of the fuel droplets.

6 Conclusion and Recommendations

The primary conclusion from this study is that numerical simulations can be used as an effective tool to address the effects of additives on the structure of jet diffusion flames. The current simulations predict the expected decrease in temperature and change in the flame structure. However, in order to predict actual extinction of flames some of the input parameters and simplifying assumptions used in the above simulations need to be changed. More detailed species dependent diffusion coefficients could cause some changes in the structure of the flame. However, a major improvement would be replacing the finite-width flame sheet model for the chemistry with a calibrated single-step Arrhenius reaction rate. Additional simulation with detailed multi-step chemical kinetics are also warranted. Such simulations could predict the time-dependent extinguishment of flames as well as be a valuable tool to address the chemical-kinetic effects of additives.

List of Figures

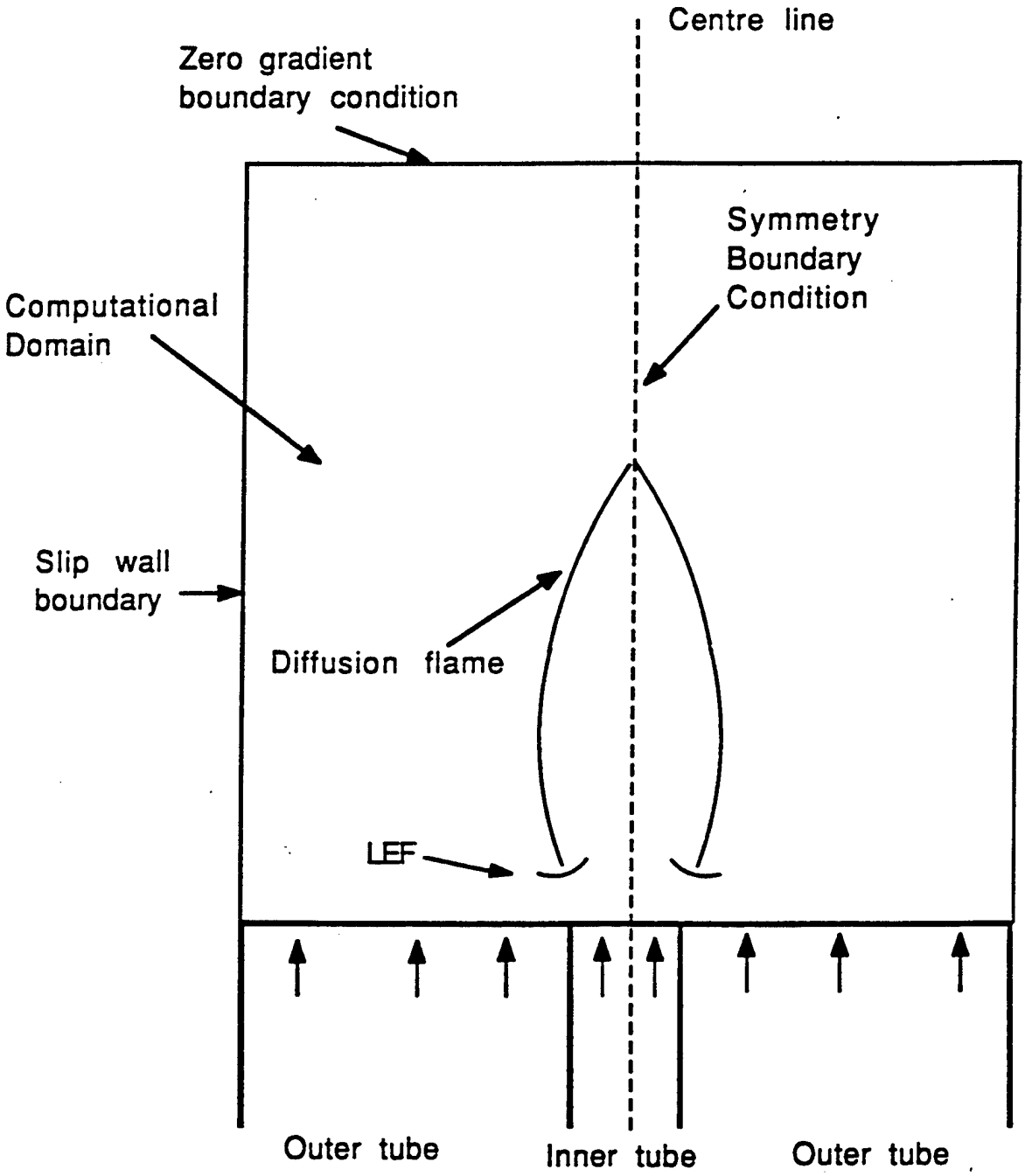
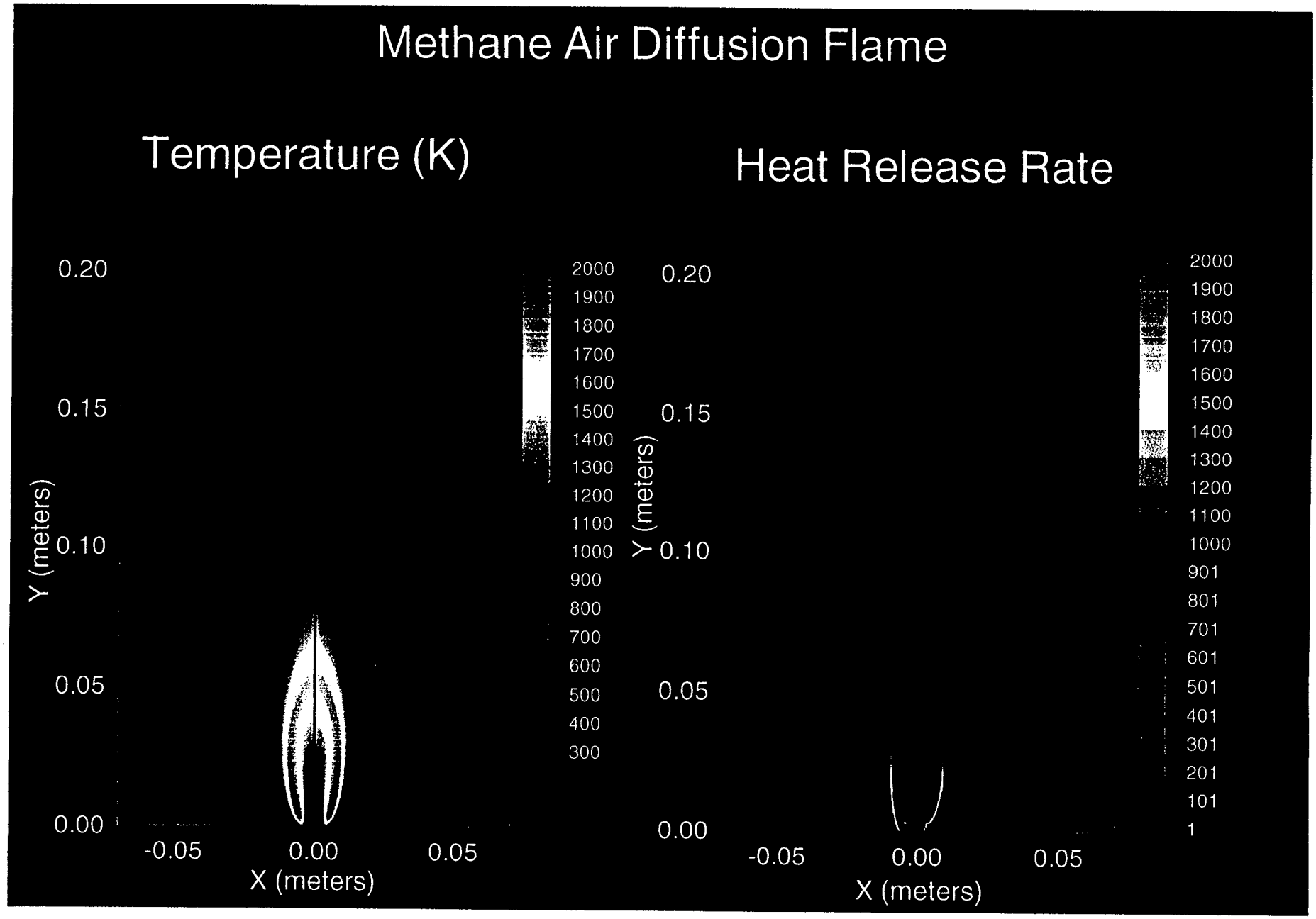
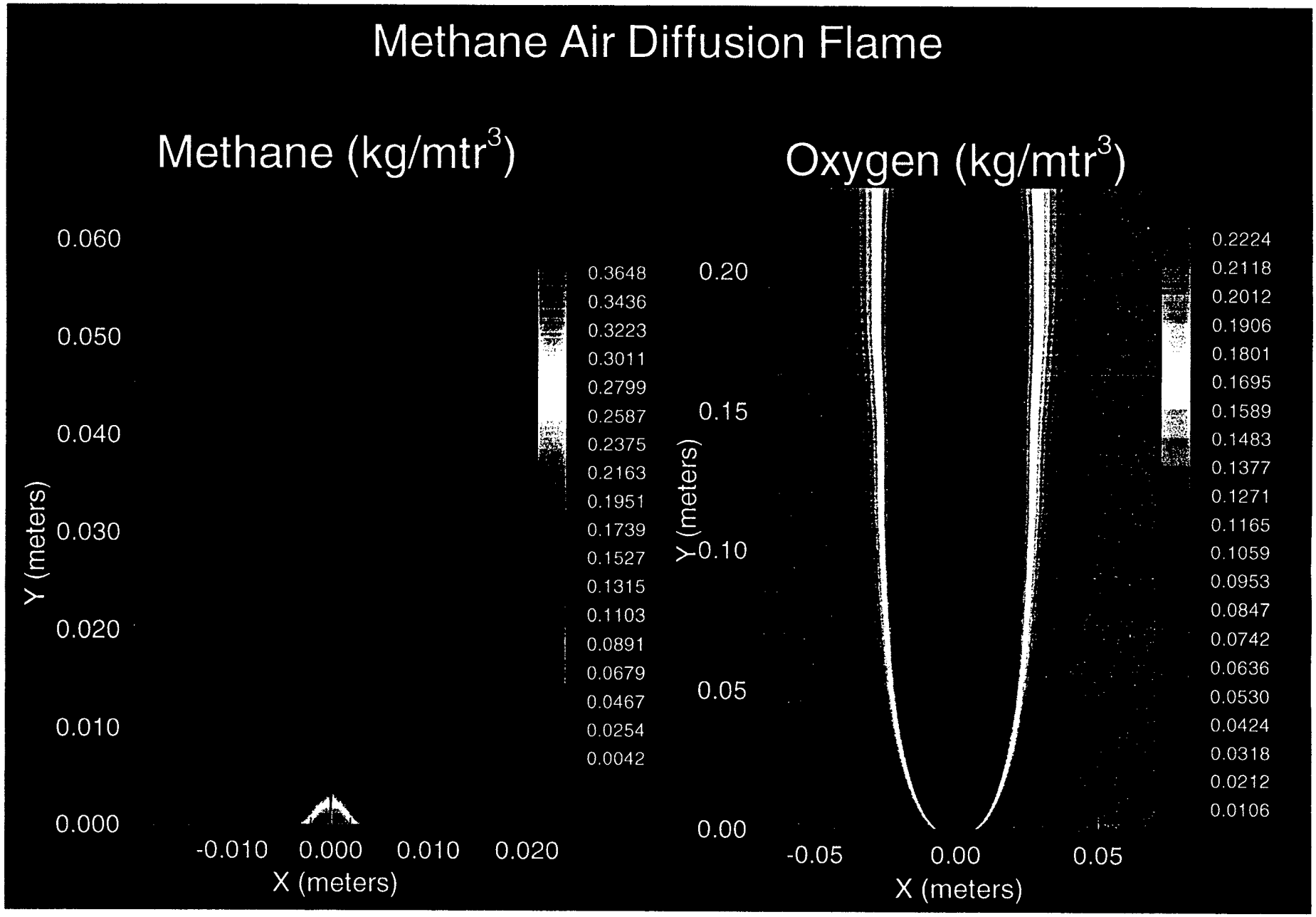
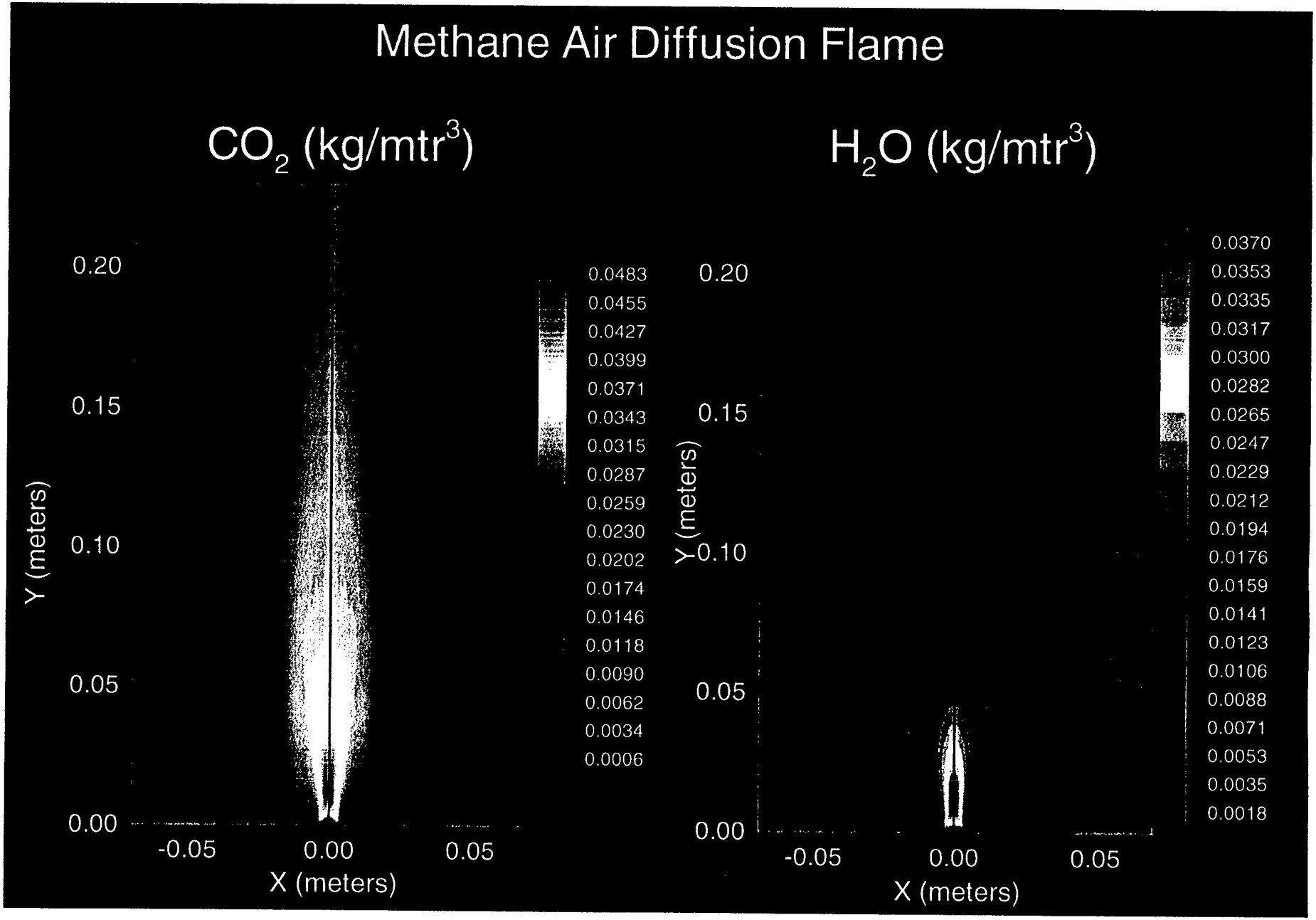
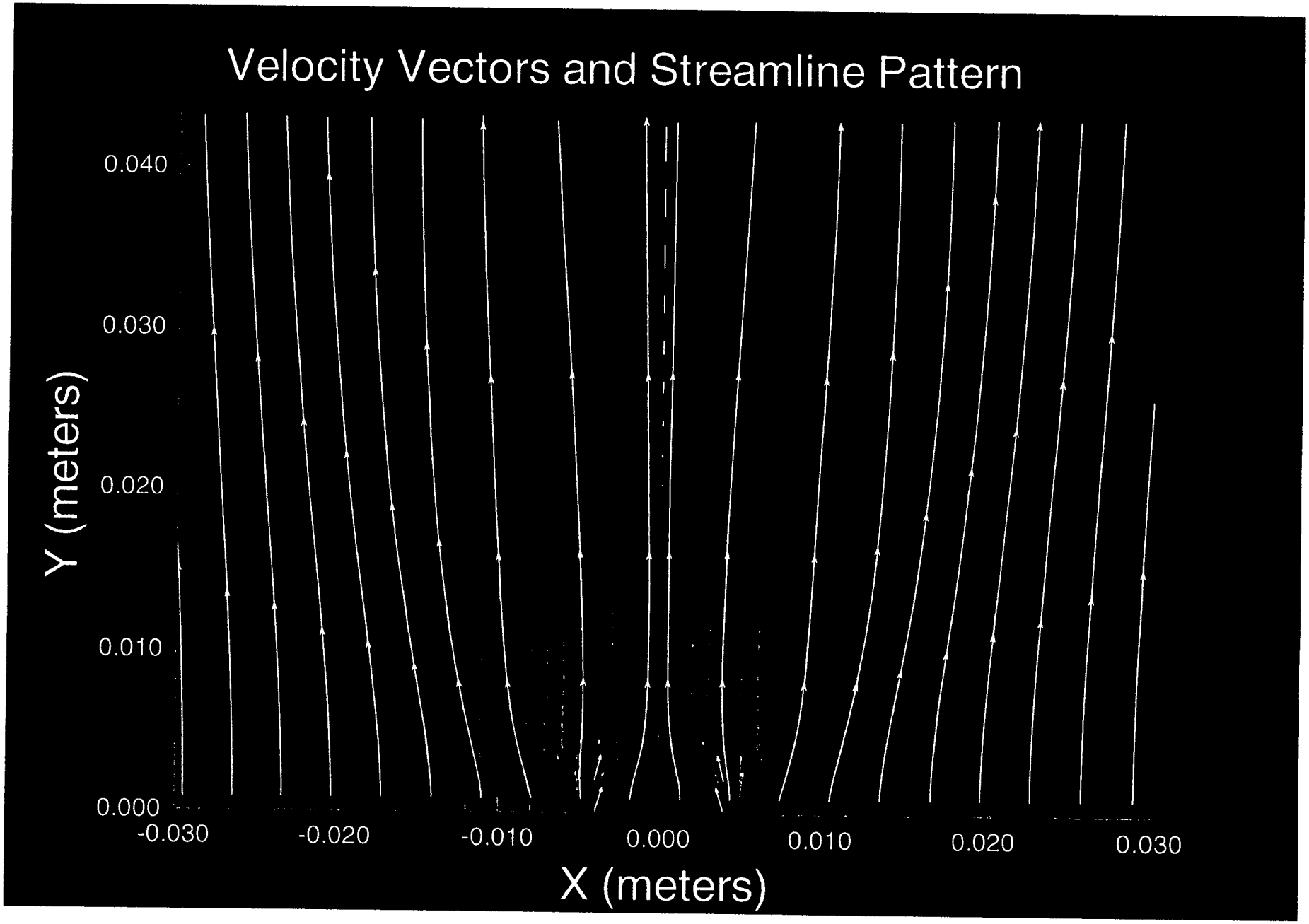


Figure 1.1. Schematic diagram of burner geometry and the associated computational domain.


Methane Air Diffusion Flame
Temperature (K)
Heat Release Rate
0.20
0.15
0.10
0.05
0.00
Y (meters)
-0.05
0.00
0.05
X (meters)
2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
901
801
701
601
501
401
301
201
101
1


Methane Air Diffusion Flame
Methane (kg/mtr3)
0.060
0.050
0.040
0.030
0.020
0.010
0.000
Y (meters)
0.3648
0.3436
0.3223
0.3011
0.2799
0.2587
0.2375
0.2163
0.1951
0.1739
0.1527
0.1315
0.1103
0.0891
0.0679
0.0467
0.0254
0.0042
-0.010
0.000
0.010
0.020
X (meters)
Oxygen (kg/mtr3)
0.20
0.15
0.10
0.05
0.00
Y (meters)
0.2224
0.2118
0.2012
0.1906
0.1801
0.1695
0.1589
0.1483
0.1377
0.1271
0.1165
0.1059
0.0953
0.0847
0.0742
0.0636
0.0530
0.0424
0.0318
0.0212
0.0106
-0.05
0.00
0.05
X (meters)

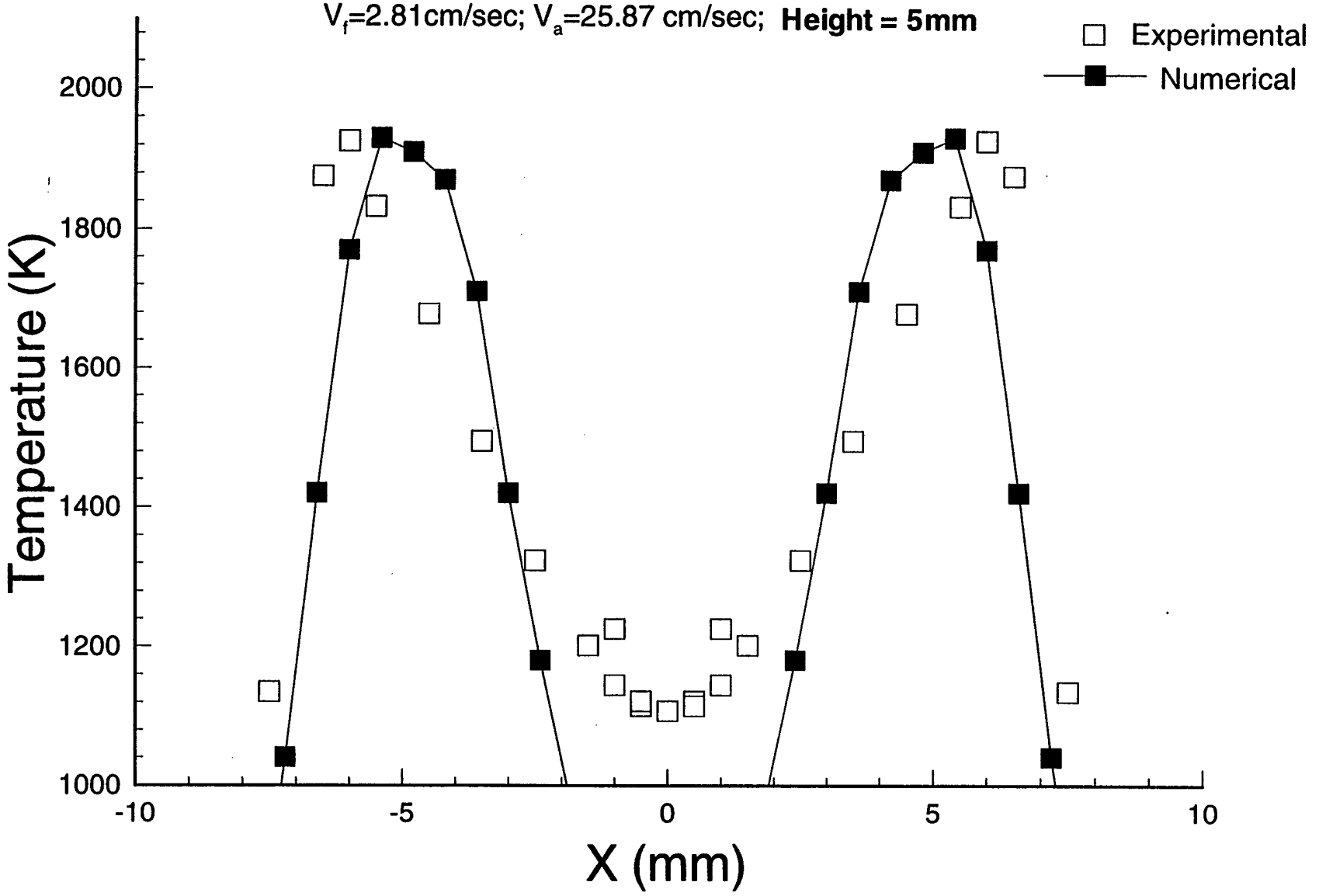

Methane Air Diffusion Flame
CO_2 (kg/mtr^3)
H_2O (kg/mtr^3)
Y (meters)
X (meters)
0.20
0.15
0.10
0.05
0.00
-0.05
0.00
0.05
0.0483
0.0455
0.0427
0.0399
0.0371
0.0343
0.0315
0.0287
0.0259
0.0230
0.0202
0.0174
0.0146
0.0118
0.0090
0.0062
0.0034
0.0006
0.0370
0.0353
0.0335
0.0317
0.0300
0.0282
0.0265
0.0247
0.0229
0.0212
0.0194
0.0176
0.0159
0.0141
0.0123
0.0106
0.0088
0.0071
0.0053
0.0035
0.0018


Velocity Vectors and Streamline Pattern
0.040
0.030
0.020
0.010
0.000
Y (meters)
-0.030
-0.020
-0.010
0.000
0.010
0.020
0.030
X (meters)

Planar Methane-Air Diffusion Flame

Comparison of Numerical and Experimental Results

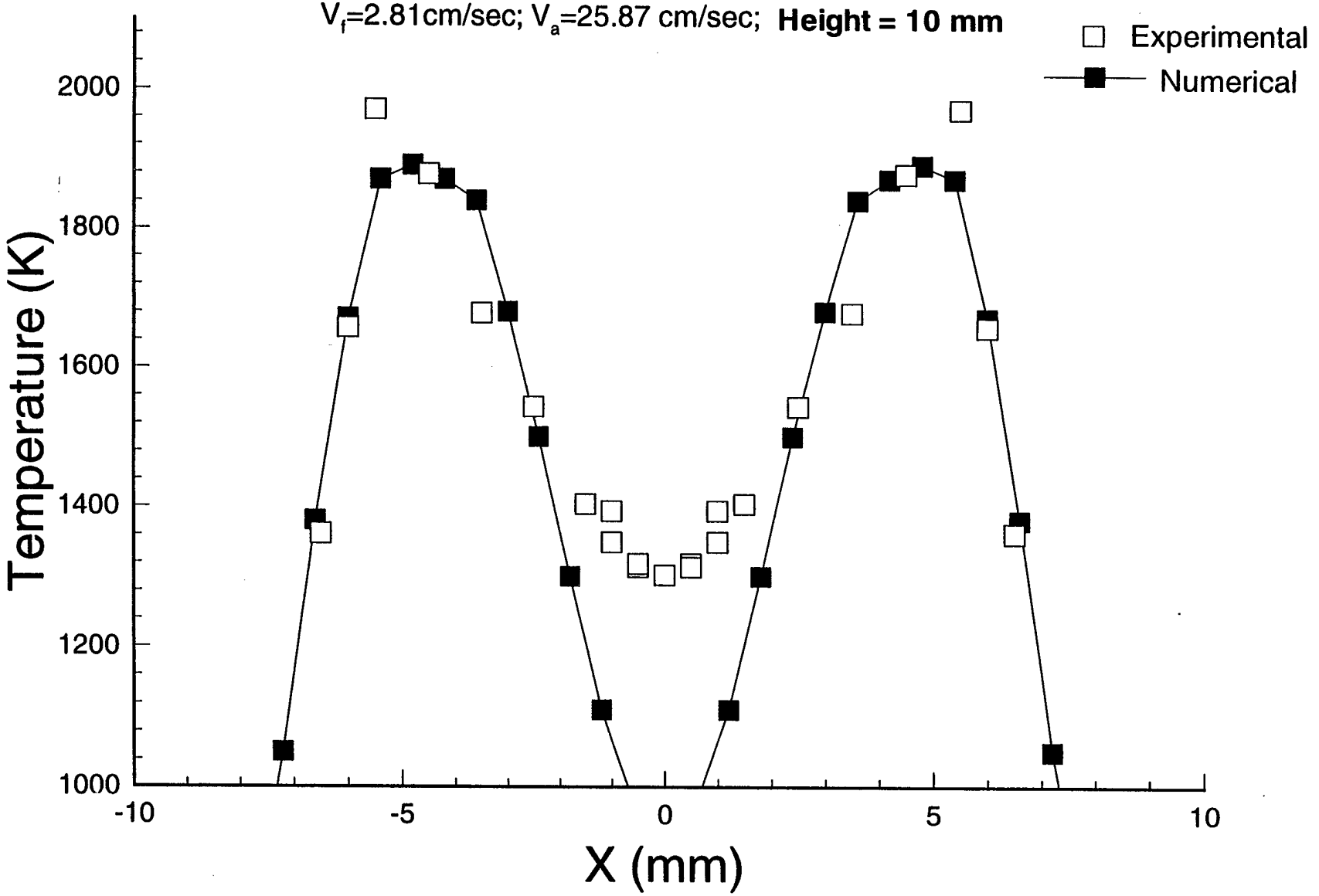
V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 5mm**

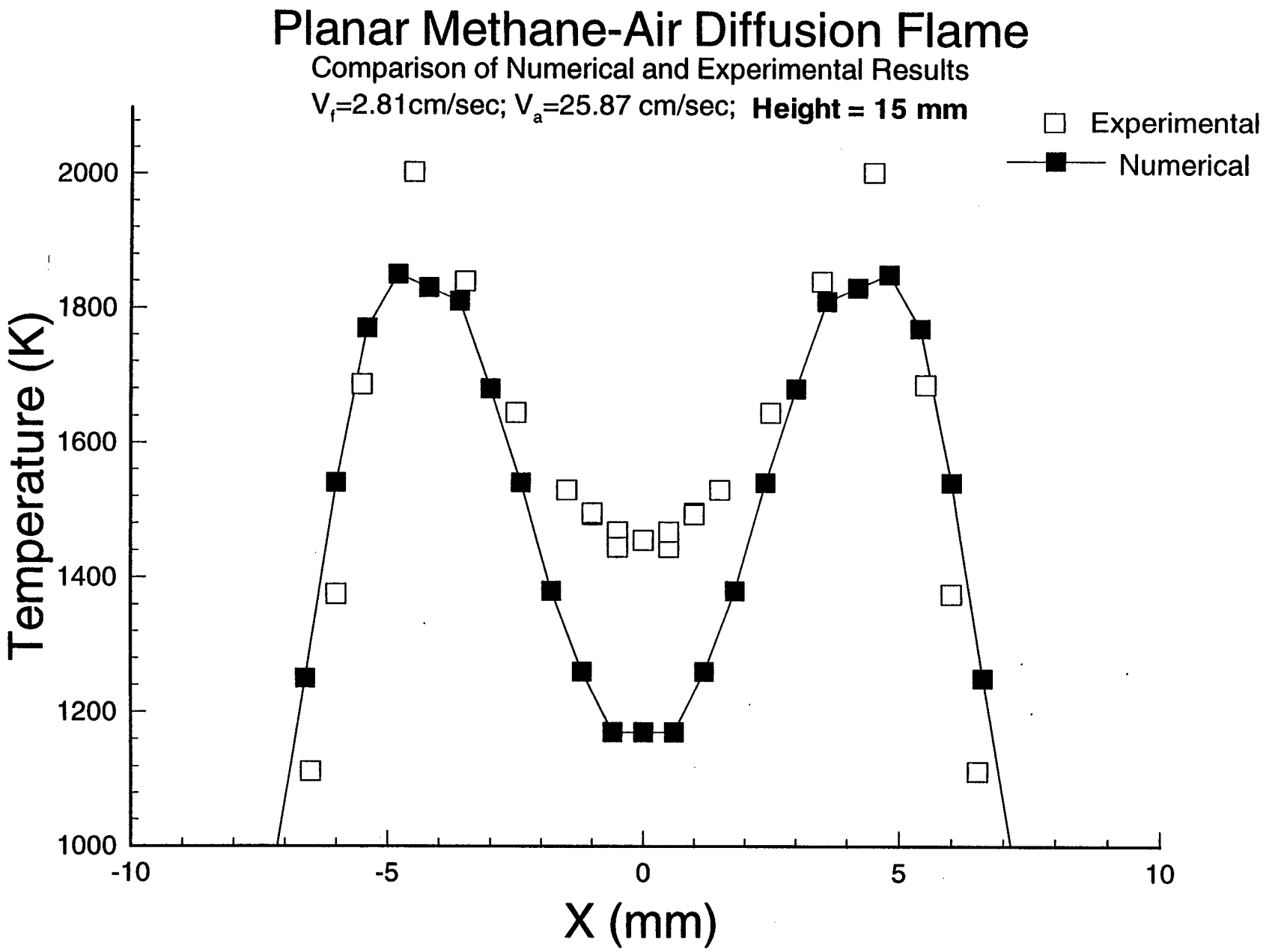


Planar Methane-Air Diffusion Flame

Comparison of Numerical and Experimental Results

V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 10 mm**

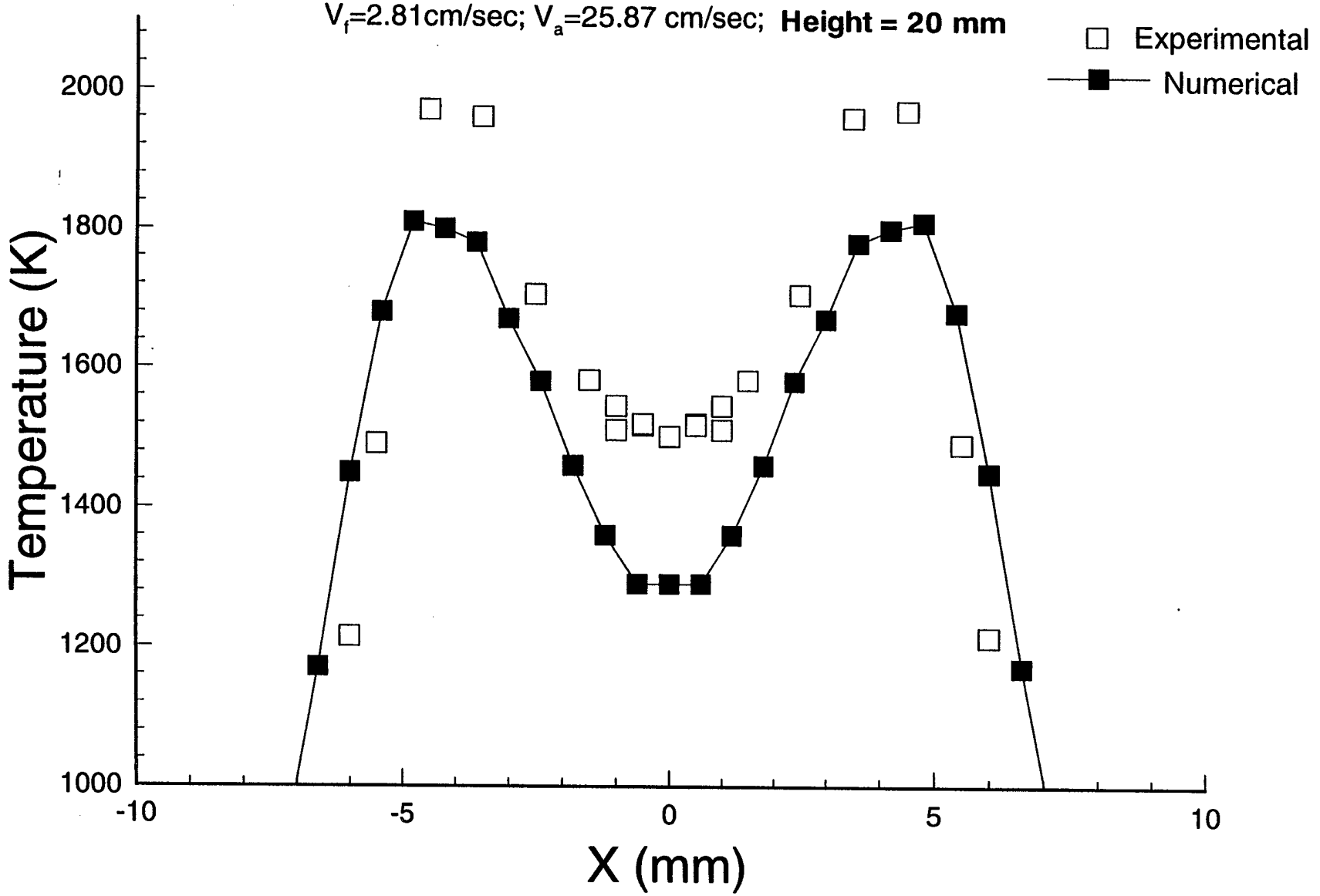


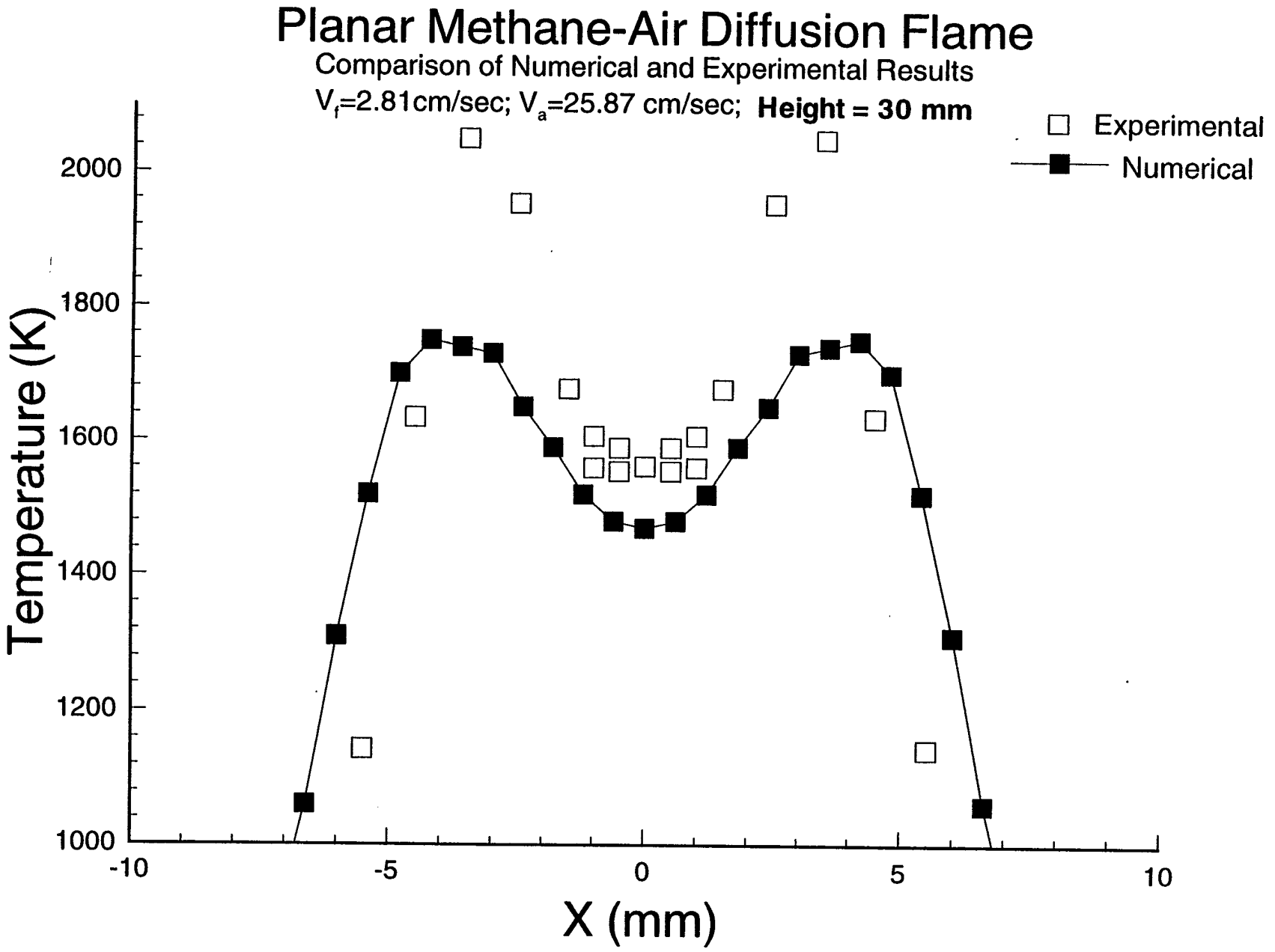

Planar Methane-Air Diffusion Flame
Comparison of Numerical and Experimental Results
V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 15 mm**
Experimental
Numerical
Temperature (K)
2000
1800
1600
1400
1200
1000
-10
-5
0
5
10
X (mm)

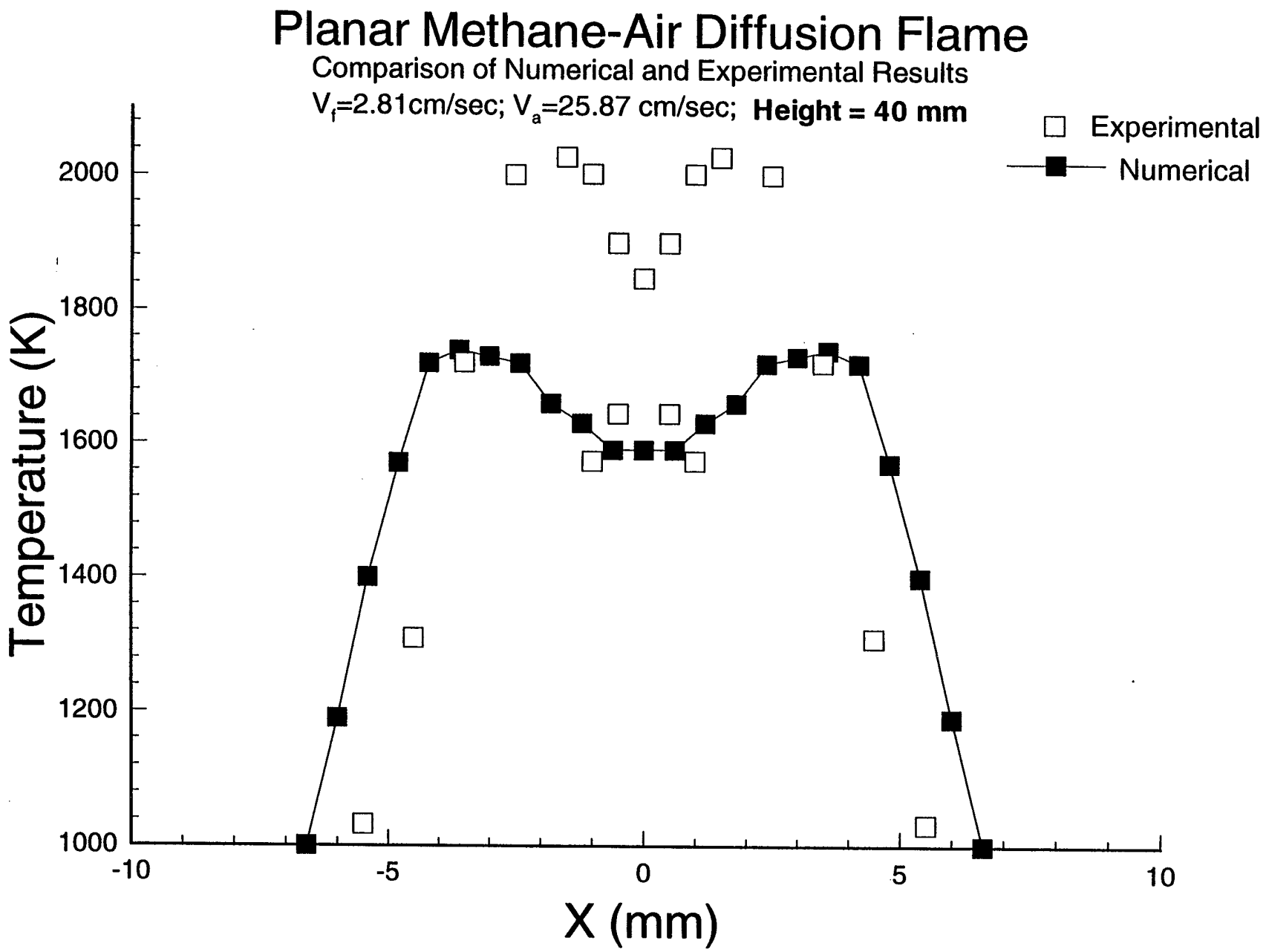
Planar Methane-Air Diffusion Flame

Comparison of Numerical and Experimental Results

V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 20 mm**



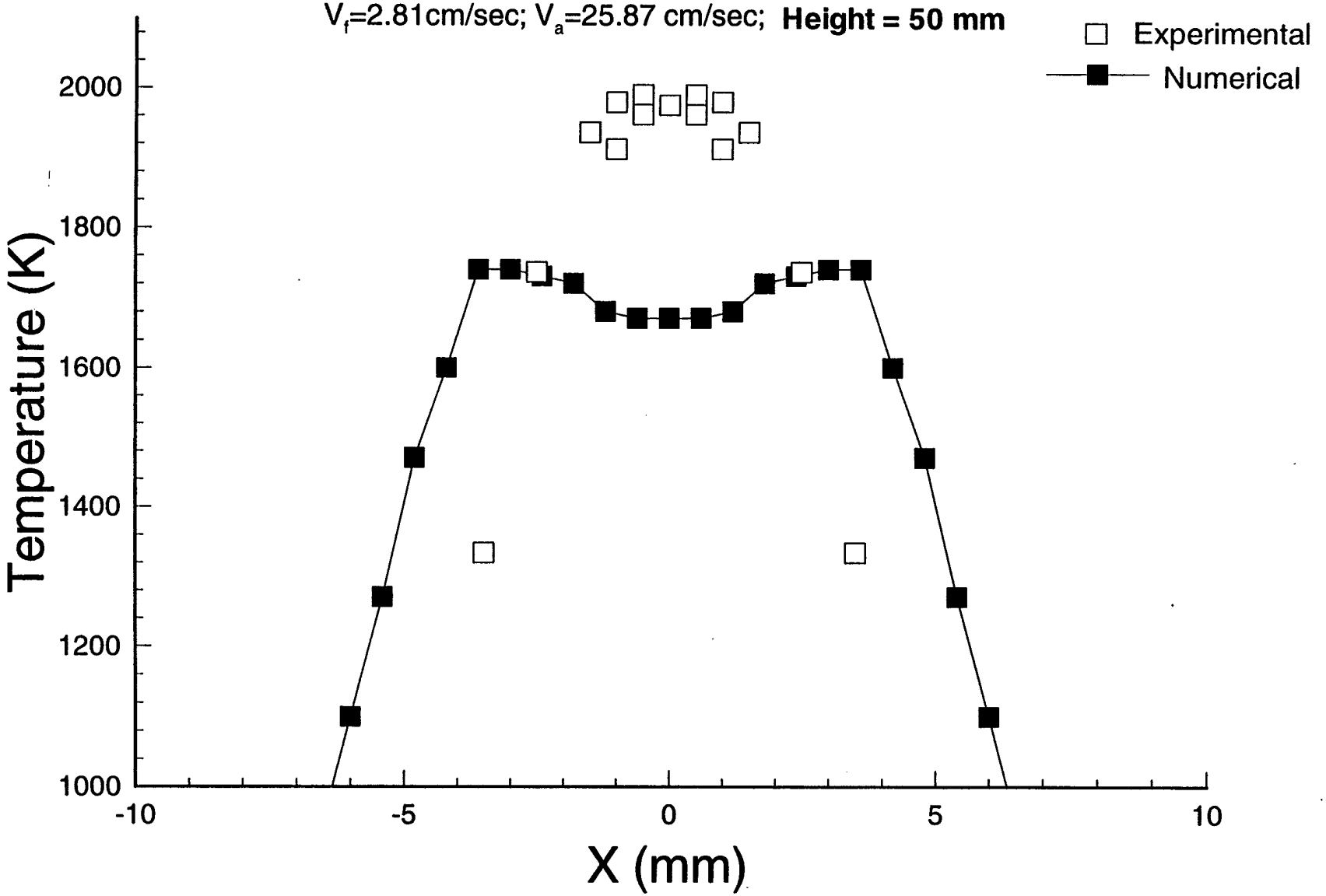

Planar Methane-Air Diffusion Flame
Comparison of Numerical and Experimental Results
V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 30 mm**
Experimental
Numerical
Temperature (K)
2000
1800
1600
1400
1200
1000
-10
-5
0
5
10
X (mm)


Planar Methane-Air Diffusion Flame
Comparison of Numerical and Experimental Results
V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 40 mm**
Experimental
Numerical
Temperature (K)
2000
1800
1600
1400
1200
1000
-10
-5
0
5
10
X (mm)

Planar Methane-Air Diffusion Flame

Comparison of Numerical and Experimental Results

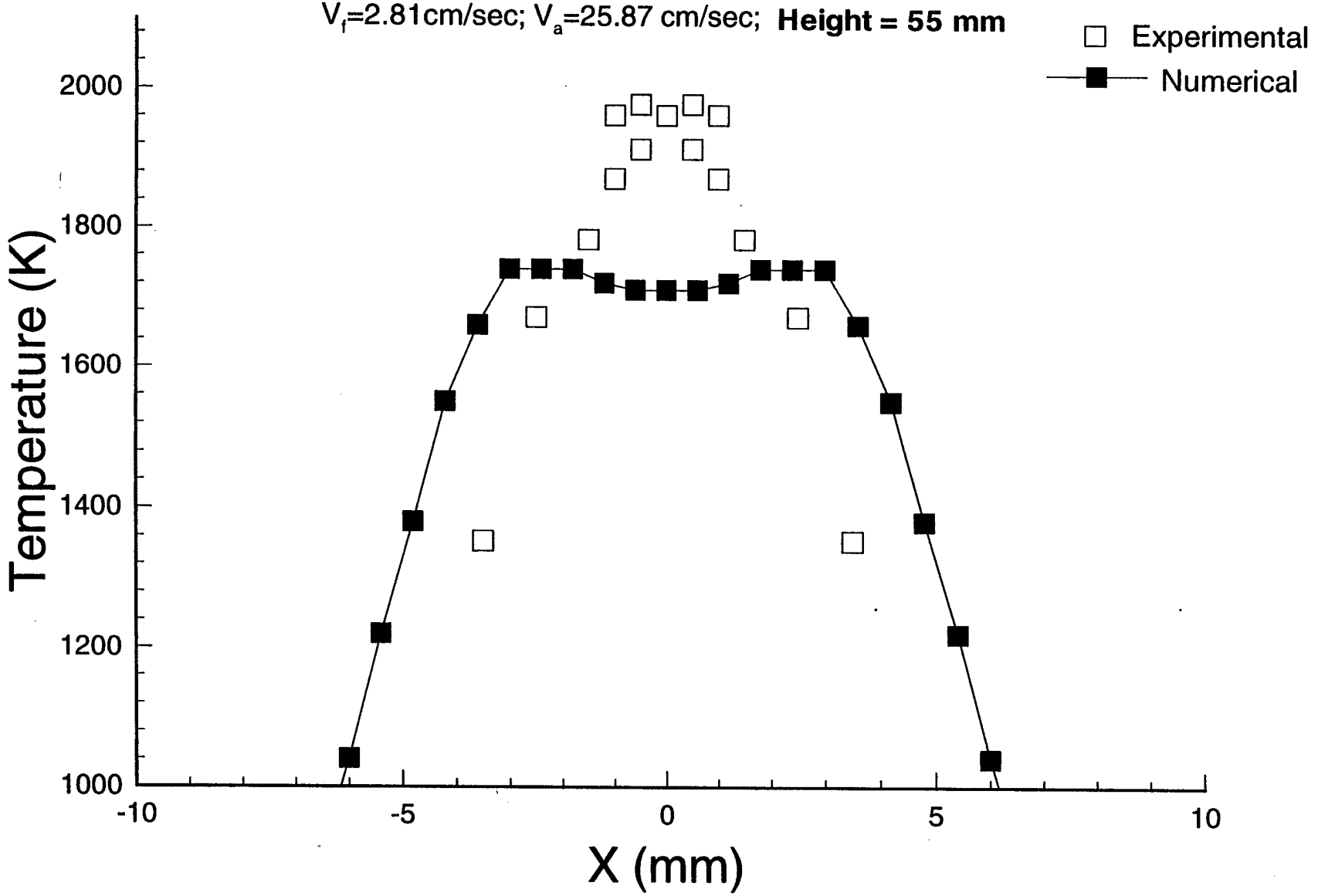
V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 50 mm**



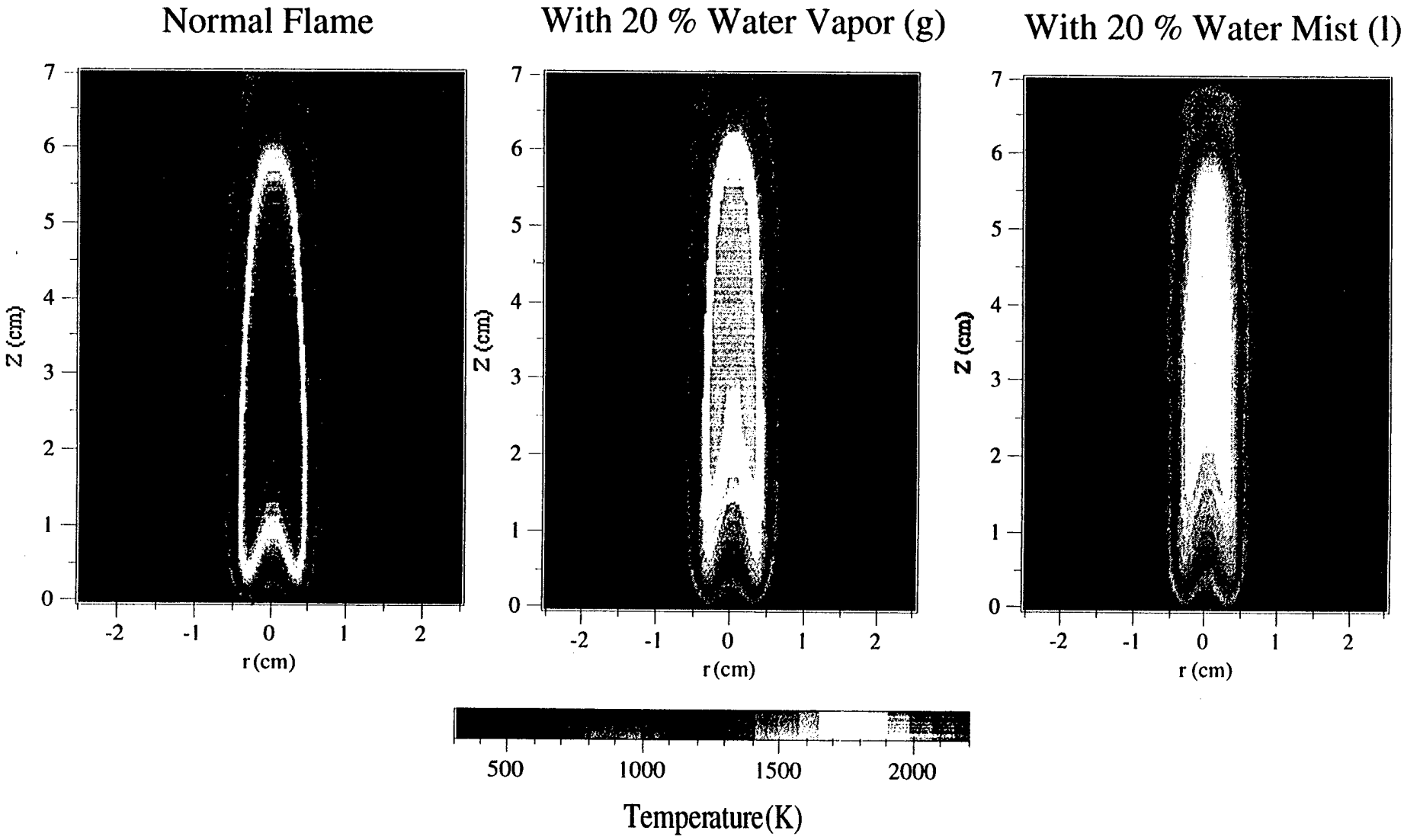
Planar Methane-Air Diffusion Flame

Comparison of Numerical and Experimental Results

V_f=2.81cm/sec; V_a=25.87 cm/sec; **Height = 55 mm**



Effects of Water Addition
Methane-Air Jet Diffusion Flame



Appendix II

An Eigenvalue Method for Computing the Burning Rates of RDX Propellants

AN EIGENVALUE METHOD FOR COMPUTING THE BURNING RATES OF RDX PROPELLANTS*

Kuldeep Prasad

Science Applications International Corporation
Applied Physics Operations
1710 Goodridge Drive
McLean, VA 22102

Richard A. Yetter
Department of Mechanical and Aerospace Engineering
Princeton University
Princeton, NJ 08544

Mitchell D. Smooke
Department of Mechanical Engineering
Yale University
New Haven, CT 06520

May 7, 1996

*Supported by a grant from ONR

An Eigenvalue Method for Computing the Burning Rates of RDX Propellants

Kuldeep Prasad, Richard A. Yetter & Mitchell D. Smooke

ABSTRACT

A mathematical model for a three-tiered system consisting of solid, liquid and gas is derived for studying the combustion of RDX propellants. The resulting nonlinear two-point boundary value problem is solved by Newton's method with adaptive gridding techniques. In this study the burning rate is computed as an eigenvalue, which can remove the uncertainty associated with employing evaporation and condensation rate laws in its evaluation. Results are presented for laser-assisted and self-deflagration of RDX monopropellants and are compared with experimental results. The burning rates are computed over a wide range of ambient pressures and compare well with experimental results from one to ninety atmospheres. The burning rate is found to be proportional to the pressure raised to the 0.76 power. Sensitivity of the burning rate to initial propellant temperature is calculated and found to be extremely low, in agreement with past theoretical predictions and experimental data. Results for laser-assisted combustion show a distinct primary and secondary flame separated by a dark zone, the length of which is dependent upon the incident laser flux intensity.

1 INTRODUCTION

Propellants are used for rockets, guns, gas generators and pressure generators. These applications require high-pressure combustion gases in an open or a closed chamber. Numerous chemicals have been used in making solid propellants with the goal of optimizing combustion characteristics for different purposes. There are two types of propellants which are distinguished by the condition in which their ingredients are connected. When the oxidizer and the fuel molecules are linked chemically, and are made of one material, the material is called a homogeneous propellant [1]. A typical

example of a homogeneous propellant is nitrocellulose. Since nitrocellulose is a fibrous material, it is difficult to form a specified propellant grain. Thus, liquid materials (nitroglycerin or trimethylolethane trinitrate) called plasticizers are mixed with the nitrocellulose to gelatinize it and to form a specified propellant grain. Propellants which are composed of nitrocellulose and nitroglycerin and a small amount of stabilizer are called double-base propellants and are typical homogeneous propellants. When the oxidizer and fuel ingredients are mixed and combined physically for a solid propellant, the pair of materials is called a heterogeneous propellant [2], [3], [4]. Typical crystalline particles used as an oxidizer include ammonium perchlorate, ammonium nitrate and potassium perchlorate. The fuels used in heterogeneous propellants have a hydrocarbon structure, such as polyurethane and polybutadiene, and act as binders to adhere the oxidizer particles together. Propellants have also been classified by their physical state, e.g., solid [5], [6] or liquid [7].

The cyclic nitramines cyclotrimethylenetrinitramine ($C_3H_6N_6O_6$), commonly called RDX, and cyclotetramethylenetetranitramine ($C_4H_8N_8O_8$), commonly called HMX are important ingredients in solid propellants. The addition of nitramine particles such as RDX and HMX to double-base homogeneous propellants increase the flame temperature and specific impulse. This type of propellant is called a nitramine-based composite modified double base (RDX-CMDB or HMX-CMDB) propellant. RDX has also been called hexogen (Soviet literature) and T-4. It is widely known that the use of these ingredients offers many advantages. Due to the high energy release of these compounds and the large amount of low molecular weight gaseous combustion products, high values of specific impulse for rocket propellants and impetus for gun propellants can be achieved. In contrast to propellants based on ammonium perchlorate (AP), nitramine propellants do not produce hydrochloric acid (HCl). Besides being corrosive, HCl provides nucleation sites for condensation of moisture droplets, thereby producing a visible contrail or secondary smoke.

Due to the lack of understanding of the various elementary chemical and physical processes involved during the combustion of RDX propellants, very little work has been

done in the area of detailed numerical modeling. Microscopic observations, mostly by Boggs [8] and Hanson-Parr and Parr [9], of quenched samples using a scanning electron microscope and of burning samples using high speed motion-picture photography have revealed many aspects of the deflagration process that can aid in the development of theories. Ermolin et al. [10] have theoretically investigated the detailed chemistry of the secondary luminous stage of an RDX flame which they have compared with corresponding experimental composition profiles. They do not include mass or energy diffusion, thereby reducing the problem to a strictly time-dependent system. Ben-Reuven et al. have incorporated the nitramine flame chemistry in theoretical models for studying the self-deflagration of RDX [11] and HMX [12]. The flame structure and liquid layer reactions of deflagrating RDX were expressed in terms of the energy, continuity, and species equations corresponding to RDX decomposing in liquid and gaseous phases and the NO_2/CH_2O reactions adjacent to the surface. A nine species gas phase reaction model was employed and the burning rate was computed using a Clausius-Clapeyron evaporation law. A simplified theoretical model was developed by Mitani and Williams [13] for describing the flame structure and deflagration velocity of steady, planar, adiabatic combustion of nitramines. The model involves exothermic decomposition in a liquid layer, equilibrium vaporization and exothermic combustion in the gas. Their selection of energetic and rate parameters (based on experimental results), resulted in reasonable agreement between predicted and measured burning rates and pressure and temperature sensitivities.

The problem of modeling the deflagration of a homogeneous solid propellant becomes especially difficult if it is necessary to account for two-phase flow that occurs in liquid melt layers as well as the gas phase region above the liquid melt layer. A two phase region has been called a dispersed phase, and the combustion of double-base propellants usually involves reactions occurring in the condensed phase, in the dispersed phase and in the gas phase. Li, Williams and Margolis [14] used asymptotic analysis to extend the Mitani-Williams model [13] to account for the presence of bubbles and droplets in a two-phase layer at the propellant surface. They identified two

zones in the two-phase region: one with high liquid volume fractions, which maintains evaporative equilibrium, and the other with low liquid volume fractions, which exhibits nonequilibrium vaporization. It was shown that by introducing reasonable estimates for simplified, two-phase, chemical and physical parameters of nitramines, the steady burning rates were found to be close to those obtained for models with a sharp liquid-gas interface. They again report good agreement with measured burning rates and pressure and temperature sensitivities using an overall chemical-kinetic parameter for describing gas phase reactions. These models [13], [14] do not predict a dark zone as observed in experimental results.

Melius [15], [16] developed a two layer solid-gas model for studying the ignition and self-deflagration of RDX. He constructed a gas-phase model consisting of 158 reactions and 38 species. An evaporation-condensation rate law was used to model the regressing planar interface separating the solid and the gas phase. In the solid phase there was a single irreversible reaction whose rate was pertinent to liquid phase RDX. In his formulation the solid phase reaction was considered to have no influence on the regression rate. This approach assumes a-priori that only a small amount of the RDX decomposes in the condensed phase. Thus, in the limit of all the RDX decomposing in the condensed phase, the rate of desorption of the decomposition products is assumed to be exactly the same as that for pure RDX evaporation (sublimation). If there is more than a little RDX decomposition in the condensed phase, the barrier for RDX escape from the surface will no longer be related only to the heat of transformation of the pure substance, but will involve molecular interactions between RDX and decomposition products. Under these circumstances, the net rate of evaporation of RDX and the desorption of products will be difficult to describe.

Fetherolf and Litzinger [17] have studied the CO_2 laser induced pyrolysis using both experimental studies and kinetic modelling of the gas phase processes. Gas phase species evolving from the surface were measured for laser assisted combustion of RDX at 0.5 and 1.0 atmospheres. CH_2O was found to be the most abundant species, followed by H_2O, NO_2, N_2O, N_2 and CO. They report a nonluminous primary flame

zone produced by the reaction of CH_2O and NO_2 directly above the surface followed immediately by a final luminous flame produced by the reaction of HCN, NO, and to a lesser extent N_2O. The gas phase processes were numerically modelled as a one-dimensional, premixed flame. The species measured at the gas-solid interface were provided as an input as well as an approximation to the temperature profile.

Hatch [18] has developed a steady-state one-dimensional combustion model to study the gas phase chemistry of a nitrate-ester flame and has recently applied this approach for studying the deflagration of HMX. Decomposition of the condensed phase was assumed to be a first-order Arrhenius rate law. A 77 reaction mechanism was employed to describe the gas phase combustion for HMX. The effects on combustion of different mechanisms of decomposition were determined and some of the more important gas phase reactions were identified.

The ultimate goal of research in the area of solid propellant combustion is to obtain a detailed understanding of the various physical and chemical processes involved during the burning of homogeneous and heterogeneous propellants. In the present study, we will focus on the development of a detailed model for studying the combustion of RDX homogeneous monopropellants. From the viewpoint of a detailed investigation, homogeneous propellants are attractive since their components are premixed and their flames are controlled by one-dimensional chemical processes, rather than by complex three-dimensional diffusion processes as is the case for composite propellants. As a result, the flames should be more amenable to analysis, both theoretically and experimentally. However, as will be shown, the flames of homogeneous propellants can involve multiple-stage chemical reactions which can complicate an understanding of their structure.

In the next section we develop an eigenvalue approach for computing the burning rates of solid rocket propellants. The physical system consists of a three layered model composed of solid, liquid and gas phases. Section 3 discusses the use of Newton's method and adaptive gridding techniques to solve the governing equations. In Section 4, we describe the gas and condensed phase reaction mechanisms. Finally, in Section 5,

we discuss results obtained from detailed simulations of both laser-assisted and unassisted combustion of RDX propellants. These results are compared with experimental results in the literature, particularly those from Hanson-Parr and Parr [9] and Zenin [19]. For laser-assisted deflagration, solutions are obtained for varying intensity of the incident laser flux, while for unassisted RDX deflagration, the burning rates and the sensitivity of the burning rate to initial propellant temperature are computed over a wide range of ambient pressures.

2 PROBLEM FORMULATION

Figure 1 shows a schematic diagram of the physical model we will employ for describing the steady-state burning of a homogeneous solid propellant. It consists of a three layered (solid, liquid and gas) model with sharp planar boundaries, with the liquid layer sandwiched between the other two phases. Details of the chemical and physical processes supporting the model for RDX combustion are given in a subsequent section. Because of the large characteristic time for the solid decomposition reaction as compared to that for the propagation of the thermal wave, it is assumed that the propellant components pass unaffected through a preheated zone consisting of pure RDX. Heat conducted into the solid from the reaction region raises its temperature to a point at which a phase transformation or liquefication occurs. In this liquid layer, sometimes referred to as the foam zone, temperatures are high enough for molecular degradation to take place, initiated by the rupture of the $N-NO_2$ bond. Simultaneous secondary reactions occur so that a mixture of RDX, NO_2, N_2O, NO, aldehydes and CN compounds (e.g., HCN) emerges from the surface and the net energy balance of the degradation is exothermic. Unless the liquid layer on the surface of the solid is very thin, bubbles may develop within it during the course of the exothermic, gas producing reaction, and a foam reaction zone may exist. This foam zone may be observed, in the burning of double-base propellants with sufficiently high contents of nitroglycerin. At low nitroglycerin fractions, the foam zone typically is less prominent than a fizz

reaction zone, a gas-phase region in which condensed particles that leave the surface of the homogeneous solid experience exothermic gasification. RDX monopropellant combustion is modeled with a solid region followed by a liquid reaction zone followed by a fizz zone and then, after a dark zone, by a luminous gas-phase flame. Under certain conditions the gas phase shows a clearly separated primary flame (fizz zone) and a secondary flame (luminous zone) separated by a dark zone which exhibits essentially isothermal chemical reactions. The fizz zone is known to involve NO_2 - aldehyde reactions and the luminous zone involves $NO - CO$ and $NO - H_2$ reactions. In general, the secondary flame is too far away to have any effect on the surface or even to induce a temperature gradient into the primary flame.

We consider the steady-state deflagration of a semi-infinite piece of pure RDX monopropellant. Microscopic observations of the burning surface [8], [20] have shown that the deflagration is not one-dimensional. In spite of these observations we employ a one-dimensional model. We do so with the knowledge that the model will not be capable of explaining the detailed structures observed (bubbles, ridges, active sites, needles) and with the hope that by implicitly averaging over these structures we will still explain properly the average temperature field, chemical compositions, chemical kinetics, heat flux and regression rate. As in most steady, planar flow problems, it is convenient to assume that the solid and gas are infinite in extent and to adopt a coordinate system in which the flow is steady. It is assumed that the propellant is mounted on a platform moving at a velocity equal but opposite to the regression velocity of the propellant. As a result, during steady-state RDX deflagration, the locations of the gas-liquid interface, the solid-liquid interface, and the primary and secondary flames are all fixed in the laboratory frame of reference. We further choose that the gas-liquid interface is located at the origin of the coordinate frame of reference. The choice of the coordinate frame is for convenience only, and does not in any way constrain the deflagration of the propellant. Thus in Figure 1, x is the coordinate normal to the surface; the interface is located at $x = 0$; the solid extends to $x = -\infty$, and the gas extends to $x = +\infty$. We are interested in computing the mass flux rate $\dot{m}$

$(gm/cm^2/sec)$ at which the solid must move in the $+x$ direction so that the interface will remain at $x = 0$.

The burning rate or the rate of decomposition of liquid RDX can be obtained as a difference between the rate of evaporation for liquid RDX and the rate of condensation for the gaseous products. The evaporation and condensation rates are generally complicated empirically derived expressions that involve the temperature and the various species concentrations at the interface. The evaporation and condensation rates are large numbers $O(10^2)$ with a very small difference $O(10^{-2})$. The result of employing this approach in a computational model is that numerical instability/convergence difficulties can occur. In the approach taken in this paper, the mass flux rate (burning rate) is obtained as an eigenvalue of the problem. Another potential problem with the use of evaporation and condensation rate laws is that their use is restricted to self-deflagration conditions. It would be necessary to re-calibrate these rate laws for laser-assisted conditions. This is due to the fact that these rate laws do not account for the temperature gradient that exists at the gas-liquid interface or the additional laser flux heating that might be added to the system.

Our goal is to predict theoretically the mass fractions of the species and the temperature as functions of the independent coordinate, x, together with the mass flux rate which is an eigenvalue of the problem. If in the gas phase we neglect viscous effects, body forces, radiative heat transfer and the diffusion of heat due to concentration gradients, the equations [21], [22] governing the structure of a steady, one-dimensional, gas phase, isobaric flame are

$$\dot{m} = \rho v A = K1 = constant, \tag{1}$$

$$\dot{m}\frac{dY_k}{dx} = -\frac{d}{dx}(\rho A Y_k V_k) + A\dot{\omega}_k W_k, \qquad k = 1, 2, \ldots, K, \tag{2}$$

$$\dot{m}c_p\frac{dT}{dx} = \frac{d}{dx}\left(\lambda A\frac{dT}{dx}\right) - A\sum_{k=1}^{K}\rho Y_k V_k c_{pk}\frac{dT}{dx} - A\sum_{k=1}^{K}\dot{\omega}_k h_k W_k, \tag{3}$$

$$\rho = \frac{p\bar{W}}{RT}. \tag{4}$$

In these equations, x denotes the independent spatial coordinate; $\dot{m}$, the mass flux rate; T, the temperature; Y_k, the mass fraction of the k^{th} species; p, the pressure; v, the velocity of the fluid mixture; ρ, the mass density; W_k, the molecular weight of the k^{th} species; $\bar{W}$, the mean molecular weight of the mixture; R, the universal gas constant; λ, the thermal conductivity of the mixture; c_p, the constant pressure heat capacity of the mixture; c_{pk}, the constant pressure heat capacity of the k^{th} species; h_k, the enthalpy of the k^{th} species; $\dot{\omega}_k$, the molar rate of production of the k^{th} species; and V_k, the diffusion velocity of the k^{th} species. A represents the local cross-sectional area of the stream tube encompassing the flame.

Liquid RDX can undergo decomposition reactions resulting in soluable and liquid phase species such as N_2O, CH_2O NO_2 and CN compounds. For the liquid phase, it is assumed that the diffusion velocities are negligible. The equations of motion therefore simplify into the form

$$\dot{m} = \rho v A = K1 = constant, \tag{5}$$

$$\dot{m}\frac{dY_k}{dx} = A\dot{\omega}_k W_k, \quad k = 1, 2, \ldots, K, \tag{6}$$

$$\dot{m}c_p\frac{dT}{dx} = \frac{d}{dx}\left(\lambda A\frac{dT}{dx}\right) - A\sum_{k=1}^{K}\dot{\omega}_k h_k W_k + \dot{Q}A, \tag{7}$$

$$\rho = constant = K2. \tag{8}$$

The term $\dot{Q}A$ represents the energy input due to laser flux heating.

The solid phase consists of a homogeneous propellant with no chemical reactions. The diffusion velocities in the condensed phase are again zero. The equations further simplify to

$$\dot{m} = \rho v A = K1 = constant, \tag{9}$$

$$Y_{RDX} = 1.0, \quad Y_k = 0, \quad k \neq RDX, \tag{10}$$

$$\dot{m}c_p\frac{dT}{dx} = \frac{d}{dx}\left(\lambda A\frac{dT}{dx}\right) + \dot{Q}A, \tag{11}$$

$$\rho = constant = K2. \tag{12}$$

Note that for the combined one-dimensional system under consideration, the mass flux rate is a constant in the gas, liquid and solid phases.

The net chemical production rate $\dot{\omega}_k$ of each species results from a competition between all the chemical reactions involving that species. We presume that each reaction proceeds according to the law of mass action and the forward rate coefficients are in the modified Arrhenius form. The net production rate $\dot{\omega}_k$ for the k^{th} species can be written in the form

$$\dot{\omega}_k = \sum_{j=1}^{m} \left(\nu_{jk}^P - \nu_{jk}^R\right) \left[k_j^f(T) \prod_{n=1}^{K} \left(\frac{\rho Y_n}{W_n}\right)^{\nu_{jn}^R} - k_j^r(T) \prod_{n=1}^{K} \left(\frac{\rho Y_n}{W_n}\right)^{\nu_{jn}^P}\right], \tag{13}$$

where $\nu_{jk}^R(\nu_{jk}^P)$ is the stoichiometric coefficient of species k appearing as a reactant (product) in reversible reaction j, $j = 1, 2, \ldots, m$. The function $k_j^f(k_j^r)$ is the rate constant for the forward (reverse) path of reaction j. We assume k_j^f has the following modified Arrhenius temperature dependence

$$k_j^f = A_j^f T^{\beta_j^f} e^{-(E_j^f/RT)}, \tag{14}$$

and similarly for k_j^r. A_j^f is the rate constant pre-exponential factor, β_j^f is the temperature exponent and E_j^f is the activation energy for the j^{th} forward reaction. The reverse rate constants k_j^r can be written in terms of the forward rate constants and the equilibrium constant K_j^c by

$$k_j^r = k_j^f / K_j^c. \tag{15}$$

The diffusion velocity appearing in the species and energy equations are divided into three parts

$$V_k = v_k + w_k + v_c, \qquad k = 1, 2, \ldots, K, \tag{16}$$

where v_k is the ordinary diffusion velocity due to mole fraction gradients, w_k is the thermal diffusion velocity, and v_c is a constant diffusion velocity (independent of species). We approximate v_k by the Curtiss-Hirschfelder approximation as follows,

$$v_k = -(1/X_k) D_k \frac{d}{dx} X_k, \quad k = 1, 2, \ldots, K, \tag{17}$$

where X_k is the mole fraction of the k^{th} species and D_k is related to the binary diffusion coefficients $\mathcal{D}_{jk}$ through the relation

$$D_k = \frac{(1 - Y_k)}{\sum_{j \neq k} X_j / \mathcal{D}_{jk}}. \tag{18}$$

Thermal diffusion is incorporated into the model only in the trace light component limit for low molecular weight species such as H and H_2. The thermal diffusion velocity is given by

$$w_k = (D_k t_{D_k}/X_k)(1/T)dT/dx, \tag{19}$$

where t_{D_k} is the thermal diffusion ratio of species k. The constant diffusion velocity v_c is introduced in order to satisfy the condition

$$\sum_{k=1}^{K} V_k Y_k = 0, \tag{20}$$

which must be satisfied if mass is guaranteed to be conserved. Upon making use of equation (20) we have

$$v_c = -\sum_{k=1}^{K} Y_k(w_k + v_k). \tag{21}$$

The binary diffusion coefficients, the thermal conductivity of the mixture and the chemical production rates are evaluated using vectorized and highly optimized transport and chemistry libraries [23], [24].

To complete the specification of the problem, boundary conditions must be imposed on each end of the computational domain $(-L_1, L_2)$ and at the interfaces between the three phases. The temperature and the species mass flux fractions are prescribed at the unreacted solid boundary. Thus we have

$$T(-L1) = T_u, \tag{22}$$

$$Y_k(-L1) = \varepsilon_k, \quad k = 1, \ldots, K, \tag{23}$$

where T_u is the cold reactant stream temperature and ε_k is the known incoming mass flux fraction of the k^{th} species. In the hot stream outflow boundary we have

$$\frac{dT}{dx}(L2) = 0, \tag{24}$$

$$\frac{dY_k}{dx}(L2) = 0, \quad k = 1, 2, \ldots, K. \tag{25}$$

We point out that in this problem, the mass flux rate $\dot{m}$ is not known; it is an eigenvalue to be determined. Calculation of the flux rate proceeds by introducing the trivial differential equation

$$\frac{d\dot{m}}{dx} = 0. \tag{26}$$

To maintain a well-posed problem, we must specify an additional boundary condition for the trivial differential equation (26). We choose to impose the boiling point temperature or the evaporation temperature at the gas - liquid interface, the location of which is fixed at the origin in the chosen coordinate system. This temperature can be obtained from experimental or theoretical data at the gas-liquid interface. Other possibilities for this boundary condition are discussed in Section 5.

In addition to specifying the boiling point temperature at the gas-liquid interface, one needs to specify specialized interface conditions that account for the transition from the liquid to the gas phase as well as any laser flux heating that might occur at the interface. Energy balance at this interface [22] is maintained by solving the difference form of the energy equation shown below

$$\dot{m}\frac{d}{dx}\left(\sum_{k=1}^{K} Y_k h_k\right) = \frac{d}{dx}\left(\lambda A \frac{dT}{dx}\right) - A\sum_{k=1}^{K} \rho Y_k V_k c_{pk}\frac{dT}{dx} + \dot{Q}A. \tag{27}$$

The term $\dot{Q}A$ represents the heat input due to laser flux heating, whereas the difference form of the convective term $\dot{m}\frac{d}{dx}\left(\sum_{k=1}^{K} Y_k h_k\right)$ takes account of the heat of vaporization associated with the phase transition from the liquid to the gas phase. Under self-deflagration conditions the heat input term $\dot{Q}A$ is set to zero in equation (27). Although there are significantly fewer species in the condensed phase as compared to the gas phase, by introducing each species over the entire computational domain (condensed phase and the gas phase), we obtain a very compact computational formulation. This approach, however, may require additional memory and CPU time for solution of the redundant species equations that are solved in the condensed phase.

Unlike the gas-liquid interface, whose location remains fixed at a specified grid point in the computational domain, the location of the solid-liquid interface is not known a-priori. The melting point temperature determines the transition from the

solid to the liquid phase. Since the location of this melt point is not known, the melt point may lie between two adjacent node points in the condensed phase. The algorithm uses liquid phase properties when the temperature is above the melt point and solid phase properties at temperatures below that point. Since we do not allow for chemical reactions in the solid phase, the species entering the liquid phase is pure RDX. Energy balance is maintained across the solid-liquid interface by solving equation (27) without the diffusion velocity term.

3 NUMERICAL MODEL

Solution of the governing equations (1)-(12), (22)-(27) proceeds with an adaptive non-linear boundary value method [25]. The solution procedure has been discussed in detail elsewhere, and we outline only the essential features here. Our goal is to obtain a discrete solution of the governing equations on the mesh $\mathcal{M}$

$$\mathcal{M} = [-L_1 = x_0 < x_1 < \ldots < x_m = L_2]. \tag{28}$$

With the continuous differential operators replaced by finite difference expressions, we convert the problem of finding an analytic solution of the governing equations to one of finding an approximation to this solution at each point of the mesh $\mathcal{M}$. We seek the solution U_h^* of the nonlinear system of difference equations

$$F(U_h; \alpha) = 0, \tag{29}$$

where α is an M-dimensional parameter vector. For an initial solution estimate U^0 that is sufficiently close to U_h^*, the system of equations can be solved by Newton's method. We write

$$J(U^n)(U^{n+1} - U^n) = -\lambda_n F(U^n), \quad n = 0, 1, \ldots, \tag{30}$$

where U^n denotes the n^{th} solution iterate, λ_n, the n^{th} damping parameter $(0 < \lambda < 1)$ and $J(U^n) = \partial F(U^n)/\partial U$ the Jacobian matrix. A system of linear block tridiagonal equations must be solved at each iteration for corrections to the previous solution

vector. In premixed flame problems, the cost of forming a numerical Jacobian and factoring the Jacobian matrix can be a significant part of the total cost of the calculation. In such problems, we apply a modified Newton method in which the Jacobian is re-evaluated only periodically [26], [27].

The solution of combustion problems, such as the RDX propellant system, requires that the computational mesh be determined adaptively. Many of the methods that have been used to determine adaptive grids for two-point boundary value problems can be interpreted in terms of equidistributing a positive weight function over a given interval. We say that a mesh $\mathcal{M}$ is equidistributed on the interval $[-L_1, L_2]$ with respect to the non-negative function W and the constant C if

$$\int_{x_j}^{x_{j+1}} W dx = C, \quad j = 0, 1, \ldots, m-1. \tag{31}$$

We determine the mesh by employing a weight function that equidistributes the difference in the components of the discrete solution and its gradient between adjacent mesh points. Upon denoting the vector of N dependent solution components by $U = [\tilde{U}_1, \tilde{U}_2, \ldots, \tilde{U}_N]^T$, we seek a mesh $\mathcal{M}$ such that

$$\int_{x_j}^{x_{j+1}} \left|\frac{d\tilde{U}_i}{dx}\right| dx \leq \delta \left| \underset{-L_1 \leq x \leq L_2}{\max \tilde{U}_i} - \underset{-L_1 \leq x \leq L_2}{\min \tilde{U}_i} \right| \begin{matrix} j = 0, 1, \ldots, m-1 \\ i = 1, 2, \ldots, N \end{matrix} \tag{32}$$

and

$$\int_{x_j}^{x_{j+1}} \left|\frac{d^2\tilde{U}_i}{dx^2}\right| dx \leq \gamma \left| \underset{-L_1 \leq x \leq L_2}{\max \frac{d\tilde{U}_i}{dx}} - \underset{-L_1 \leq x \leq L_2}{\min \frac{d\tilde{U}_i}{dx}} \right| \begin{matrix} j = 0, 1, \ldots, m-1 \\ i = 1, 2, \ldots, N \end{matrix} \tag{33}$$

where δ and γ are small numbers less than one and the maximum and minimum values of $\tilde{U}_i$ and $d\tilde{U}_i/dx$ are obtained from a converged numerical solution on a previously determined mesh.

4 CONDENSED AND GAS-PHASE CHEMISTRY MODELS

RDX is an orthorhombic crystal. It exists in solid form at room temperature with a density of 1.806 gm/cm^3 [8]. The melting temperature often quoted for RDX ranges

from $453 - 478K$ but this temperature does not represent equilibrium between the solid and liquid states. Although several investigators discuss RDX melting, experimental results indicate that the liquefication is not a true melting in the sense of a phase change from solid to liquid occurring at a rather precise temperature. Several experiments have been performed below the melting point to study the decomposition of solid RDX. Batten and Murdie [28] showed that the rate of decomposition in open ended glass ampules is greatly influenced by the reaction of active gaseous products with undecomposed RDX. They determined the activation energy for solid RDX decomposition as 63 $kcal/mole$ throughout the course of the reaction. The major decomposition products identified were N_2, N_2O, NO and H_2CO. Batten [29] has also shown that nonvolatile residue from solid RDX decomposition can cause initial liquefication of RDX. Although solid-phase decomposition is important to shelf-life, the associated high activation energy generally precludes it from playing a significant role in RDX deflagration. In the present work, we assume a single melting point temperature of 478 K.

Rogers and Smith [30] used several data-fit methods for the decomposition of RDX liquid at $491 - 515K$. For an assumed first-order reaction, they obtained Arrhenius parameters of $10^{18.8}s^{-1}$ and 48.2 $kcal/mole$. The major decomposition products from liquid RDX were NO, NO_2, N_2, CO and CO_2. Liquid RDX is known to evaporate at a temperature of roughly $573 - 620K$. The evaporation temperature is known to be a weak function of the ambient pressure [31], [32]. At their respective evaporation temperatures, liquid RDX as well as the other liquid decomposition products evaporate to form their counterpart gaseous species.

Fifer [33] has given a detailed review of the ignition and combustion chemistry of the nitrate-ester nitrocellulose (NC) and the nitramines HMX and RDX. Based on thermal decomposition experiments, it is assumed that the RDX surface products consist mainly of NO_2 plus a variety of aldehydes such as formaldehyde ($HCHO$) and glyoxal ($CHOCHO$), although it is probable that some of the NO_2 reacts within the condensed phase and that significant amounts of CO, CO_2 and H_2O are also produced.

In the fizz zone, the NO_2 is partially reduced to NO by reactions with the aldehydes to release about 400-500 cal/gm and to produce significant amounts of final product type molecules. The fizz zone products have been determined with sampling probes in the dark zone and by final product analysis for fizz burning samples (no secondary flame). In the secondary flame zone, the less reactive NO is reduced to N_2, with some leftover formaldehyde perhaps playing a role in initiating the reaction. The reaction releases another 500 cal/gm of energy and changes the ratios of the final-product-type molecules. The products are highly underoxidized (high CO, H_2) and, if burning takes place in air, considerably more energy is released in the flame zone.

An important experimental flame study on the decomposition and gas-phase speciation above deflagrating RDX was performed by Korobeinichev and co-workers [34]. Their study with mass spectrometer probe measurements identified HCN as a primary product of decomposition. As noted above, earlier studies (Cosgrove and Owen [35], [36], Ben-Reuven et al. [12]) had proposed mainly gas-phase products of CH_2O, CO, CO_2, H_2O, NO, NO_2, N_2 and N_2O. The experimental work of Fetherolf and Litzinger also observed HCN as a major product of the primary decomposition.

In recent T-jump/FTIR spectrospcopy experiments, which attempt to simulate the spatial chemistry at different depths into the condensed phase through rapid quenching of surface products from a thin layer of RDX reacting time dependently in a well controlled environment, Brill and co-workers [37]-[40] have inferred possible mechanisms occurring in the condensed-phase. Their results are generally consistent with the gaseous products that are liberated near the surface of deflagrating RDX. A simple representation of the condensed phase process has been given by the two competing branches

$$RDX \rightarrow 3CH_2O + 3N_2O \quad (L1)$$

$$RDX \rightarrow 3H_2CN + 3NO_2 \quad (L2)$$

This semiglobal scheme is also consistent with the decomposition process in the

gas-phase. In the T-jump experiments of Brill [38], N_2O and NO_2 have been observed to precede evolution of all other gas-phase species from the surface of RDX. Brill [41] identifies this finding as an indication that the products in the simplified scheme above are formed by different steps. The ratio of N_2O to NO_2 was found to decrease with increasing temperatures, being approximately equal for temperatures of 570-620 K, which is consistent with the NO_2 channel having a higher activation energy than the N_2O channel. The ratio of CH_2O to HCN was not observed to follow this trend. The NO_2 channel is endothermic by approximately 180 $kcal/mol$ whereas the N_2O channel is exothermic by approximately 40 $kcal/mol$. When reaction L2 is written to form $3HCN + 3HONO$, its endothermicity drops to approximately 25 $kcal/mol$. Because the rates of the competing channels are nearly equal at the RDX propellant surface temperature, the overall reaction can range from endothermic to thermally neutral. Thus, to sustain the reaction, additional exothermic steps have been speculated involving products from the condensed phase reactions. Products from these reactions, including CO, CO_2, H_2O, $HNCO$, and NO, have been measured in the T-jump experiments. These reactions are speculated to occur in the bulk phase, in bubbles, and on the near field gaseous side of the surface. Plausible secondary reactions considered by Brill include $CH_2O + NO_2$, $CH_2O + N_2O$, and $HCN + NO_2$. Using kinetic results from gas-phase studies of these systems, the most likely one to occur at lower temperatures in the condensed phase is the reaction between CH_2O and NO_2. One representation of this process is given by the overall step

$$CH_2O + NO_2 \rightarrow CO + NO + H_2O. \quad (L3)$$

The condensed phase reaction mechanism used in the present study is given in Table 1 [39], [42], [43]-[45]. It consists of the three overall steps identified above. The rate parameters for this scheme have been obtained from the work of Thynell et al. [46], who re-evaluated previous T-jump data with a detailed heat-transfer model of the T-jump filament and sample, a model of the control circuit, and the global decomposition and heat release mechanism described above. Predictions from this model were fitted

to experimental ignition times and experimental measurements of the NO_2/N_2O ratio to obtain the kinetic parameters. These results yielded a considerably higher activation energy for the overall reaction between CH_2O and NO_2.

In the future, inclusion of the reaction between CH_2O and N_2O may be necessary to account for N_2 formation, which has been measured above deflagrating RDX monopropellants. Nitrogen was not measureable with the FTIR analysis used in the previous T-jump experiments. Also, reactions L1 through L3 were postulated to be irreversible by Thynell et al.. Under conditions where these reactions play a significant role in its gasification process of RDX, i.e., dominate over pure vaporization, this assumption may no longer be valid. This is particularly true for methylene amidogen, H_2CN, which is an unstable radical with a heat of formation of 59 $kcal/mole$. At surface temperatures near 600 K, recombination of H_2CN and NO_2 to methylene nitramine H_2CNNO2 or reaction to HCN and $HONO$ are thermodynamically (and kinetically) favored. To investigate these trends, a more detailed condensed phase model will be required. (Note that neither H_2CN or H_2CNNO_2 have been observed in T-jump experiments).

In contrast to the present model, Melius [15] has previously employed the single step irreversible reaction L1 in his RDX self-deflagration studies with rate constant parameters of A = 4.66×10^{18} s^{-1} and E=47.8 $kcal/mol$. He employs this reaction throughout the condensed phase, whereas the present formulation invokes these reactions only in the liquid melt layer.

A system of 228 elementary reactions involving 48 species was used to describe the gas phase decomposition of RDX propellants. The reaction mechanism is listed in Table 2. This mechanism was developed starting with the RDX decomposition mechanism of Melius [15] and adding to it the species and reactions reviewed by Tsang and Herron [47] and Tsang [48]. The HCN isomer, HNC, has recently been shown to play a role in high temperature reacting HCN/NO_2 mixtures [49] and is also included in the present model. As denoted by Melius [15], the species $RDXR$ is the radical formed by removing a nitro group from RDX. The species $RDXRO$ denotes the ring-opened structure of $RDXR$.

The gas-phase model is based on sub-model development and analysis of reaction mechanisms for fuels HCN, CH_2O, CO, H_2, and NH_3 with oxidizers NO_2 and N_2O. Reaction rate and thermochemical constants are literature values that were selected from mechanistic studies based on comparisons between model predictions and experimental measurements from static and flow reactor studies, shock tube studies, and flame studies using various combinations of the above fuels and oxidizers and from critical reviews and independent isolated parameter measurements and estimations. The results from these studies were included without further modifications to the parameters of the sub-models.

The decomposition of and initial reactions with RDX and its secondary species are from Melius [15]. This mechanism is based on simple $N-N$ bond fission forming NO_2 and a cyclic radical species, which subsequently undergoes fragmentation to form two methylene nitramine molecules H_2CNNO_2, and one methylene amidogen radical, H_2CN. The methylene nitramine molecules thermally decompose to form additional methylene amidogen radicals and NO_2 or react with H_2O and OH through a catalyzed hydrogenation reaction to form CH_2O and N_2O. Further reaction of the amidogen radical subsequently leads to production of HCN. This mechanism differs from the recent work of Zhao et al. [50], who proposed that the majority (70%) of the RDX decomposes through concerted detrimerization to produce three methylene nitramine molecules. The H_2CNNO_2 molecules are then proposed to decompose by $HONO$ elimination (forming $HCN + HONO$) or through molecular rearrangement (forming $CH_2O + N_2O$). The latter step was found to be structurally unfavorable by Melius [51] while the $N-N$ bond fission of H_2CNNO_2 recommended by Melius [52] was not reported by Zhao et al. [50]. In the mechanism proposed by Zhao et al., the production of both NO_2 and H atoms from the initial and secondary reactions of RDX decomposition would thus be anticipated to be small.

Preliminary sensitivity analysis calculations of RDX decomposition using the Melius mechanism indicated that the rate controlling steps during the first stage primarily involved H_2CN and H_2CNNO_2. The most sensitive reactions in decreasing

order of importance were the thermal decomposition of H_2CN, the $N-N$ bond fission of H_2CNNO_2, reaction of H_2CNNO_2 with H_2O and OH, reaction of HCO with NO_2 and NO, and the reactions of H_2CN with NO_2 and NO. Early comparisons of model predicted species profiles with the data of Parr [9]. Hanson-Parr and Parr [53], Fetherolf and Litzinger [17], and Korobeinichev and co-workers [34], [54], [55] indicated insufficient quantities of NO_2 and N_2 at the end of the first stage and an abundance of CH_2O, furthermore suggesting insufficient radical production during the stage. In the present model, three modifications have been made to the Melius mechanism. First, a reaction allowing for the formation of N_2 during the initial stage of reaction was added to the mechanism. This reaction was between H_2CN and N_2O yielding H_2CNO and N_2. Second, the rate of thermal decomposition of H_2CN was increased relative to its reaction with NO_2, NO, and N_2O. This allowed for greater H-atom generation during the initial stage of reaction, leading to faster consumption of formaldehyde and larger NO_2 concentrations. Finally, a step between H_2CNNO_2 and NO_2 was added leading to $CH_2O + N_2O + NO_2$. However, the major step for consumption of H_2CNNO_2, other than through thermal decomposition, remained the water catalyzed reaction. The effect of the NO_2 reaction could also have been obtained by increasing the rate of the water catalyzed reaction. Clearly, additional work is needed on the initial decomposition of RDX, and in particular, on the reactions of methylene nitramine and methylene amidogen.

5 RESULTS AND DISCUSSION

In this section we describe numerical results that have been obtained for the case of self-deflagration and laser assisted combustion of RDX monopropellants. Using the model and the numerical procedures described in Sections 2 and 3, the governing equations describing the three-tiered RDX model are solved. The computed temperature and species profiles have been compared with experimental results that have been pub-

lished previously in the literature. For the case of self-deflagration, burning rates have been computed over a wide range of ambient pressure and have been compared with experimental results. A sensitivity analysis has been performed to determine the sensitivity of the burning rate to initial propellant temperature. Wherever possible, an attempt has been made to compare these results with experimental data.

Laser assisted combustion

A detailed experimental investigation of RDX flame structure was performed at the Naval Air Warfare Center [9]. Non-intrusive diagnostics were used to measure temperature and species profiles during neat RDX deflagration at one atmosphere. UV-Visible absorption was measured to obtain absolute concentration profiles of NO, NO_2, CN, NH, CH_2O and OH. Temperature and species concentrations were obtained by spectral fitting. Planar Laser-Induced Fluorescence (PLIF) of these same species was also measured to obtain 2D profiles in the flame with excellent spatial resolution. From these experiments, the flame structure was characterized by a "dark zone" close to the surface and a visible flame sheet above the dark zone.

Experimental studies often employ laser heating for ignition purposes. Another advantage of employing laser heating in experiments is that it increases the burning rates. The higher mass evolution rate results in a gas phase flame that is located further away from the propellant surface as compared to laser unassisted studies. Thus, the detailed structure in the gas phase can be resolved more easily experimentally. For a consistent comparison of model predictions with corresponding experimental results, a valid model for the laser flux heating of the propellant is of critical importance. In particular, the added laser flux energy significantly affects the energy balance at the gas-liquid interface, which results in larger mass flux rates (burning velocities) as mentioned above and also higher final flame temperatures. It is assumed that most of the gas phase species do not absorb at a laser wavelength of 10.6 microns (Hanson-Parr and Parr [9]) and that the laser flux energy is completely absorbed in the condensed

phase. Furthermore, the absorptivity of solid RDX at 10.6 microns is such that the $\frac{1}{e}$ depth is only $2\mu m$. This implies that the energy is absorbed in a distance quite small compared to the thermal profile in the solid. Employing Beer's law, the reduction in beam irradiance brought about by absorption in the condensed phase can be described as

$$I(x) = I(0)e^{-\alpha x}, \tag{34}$$

where $I(0)$ is the irradiance of the beam at the input to the gas-liquid interface, $\alpha = 1/(2\mu m)$ is the absorption coefficient, and x is the absolute distance measured from the gas-liquid interface. The energy input ($\dot{Q}$ appearing in equations (7, 11, 27)) at any node point in the condensed phase is obtained by subtracting the beam irradiance leaving the cell from that at the input to the cell.

The laser-assisted combustion experiments employ a laser flux profile that matches a Gaussian distribution $I(r)$ $cal/cm^2/sec$ given by

$$I(r) = 337 \exp^{\left(\frac{-r^2}{1.355}\right)} \quad \frac{cal}{cm^2 sec}, \tag{35}$$

where r (mm) is the radial distance measured from the centerline (point of maximum intensity) of the incident laser beam. In order to incorporate the effect of laser heating, we first compute the standard deviation $\sigma = 1.182$, a measure of the spread or dispersion of the values which the laser flux intensity can assume. The Gaussian profile was integrated and then averaged over a circular area of radius equal to σ resulting in averaged flux intensities of 62% of the peak value of 337 $cal/cm^2/sec$. Numerical simulations for laser assisted combustion have been performed with 70%, 80%, 90% and 100% of the peak value and have been compared with experimental results. The corresponding values for $I(0)$ used in equation (34) are 236, 270, 303 and 337 $cal/cm^2/sec$. The lower values used in the numerical simulations implicitly account for any losses in the experiments such as reflection at the surface as well as the gaussian distribution of the laser flux profile.

The laser-assisted experiments also show that the RDX flame is not a truly one-dimensional flame but a quasi-one-dimensional system. This is due to the fact that the

cross sectional area varies as a function of the height above the gas-liquid interface. Measurements of the variation in cross sectional area were obtained experimentally by generating streamlines from particle image velocimetry (PIV) data and measuring the relative change in area of these streamlines. It was shown that the area increases by a factor of about five between the surface and the region in the burnt gas where the flame stops expanding radially. A curve fit was developed for this area variation normalized to the region of the burnt gases very far from the gas-liquid interface. This area variation is expressed as

$$\begin{aligned} A(x) &= 0.2079, \qquad x < 0.2, \\ A(x) &= 0.148 + 0.305x - 0.0272x^2, \qquad 0.2 \leq x \leq 5.7; \\ A(x) &= 1.0, \qquad x > 5.7. \end{aligned} \tag{36}$$

where x is the distance measured above the gas-liquid interface in mm.

Using the numerical procedures described in Section 3, the case of laser-assisted deflagration for RDX propellants was studied. As discussed earlier these laser assisted calculations were performed with a laser flux intensity of 70%, 80%, 90% and 100% of the peak value (337 $cal/cm^2/sec$). The pressure was set to one atmosphere. The reactant stream temperature was $T_u = 298.5K$ and the computational length was given by $L = 8mm$. Starting with an initial uniformly spaced grid consisting of 40 points, grid points were adaptively placed in regions of sharp gradients. Laser-assisted combustion simulations typically were solved with 120-150 node points. The smallest grid spacing occurred at the gas-liquid interface, where adjacent node points were less than $1\mu m$ apart. The largest grid spacing occurred in regions where the temperature had reached its final equilibrium flame temperature. The CPU time for a calculation initiated from scratch takes roughly 45 minutes on an IBM RS6000 Model 590 machine. Subsequent calculations performed by varying the laser flux intensity took approximately 10 minutes.

Figure 2 shows the temperature profile in the solid, liquid and gas phases for laser-assisted combustion of RDX propellants. The gas-liquid interface is constrained

to be located at the origin of the coordinate frame ($x = 0.0$). The experimental temperature at which evaporation occurs is $T_{evap} = 610K$. Based on the structure of the thermal wave propagating into the condensed phase, we find that the temperature rises monotonically from the initial propellant temperature ($T_u = 298.15$) to the melting point temperature ($T_{melt} = 478K$). During this portion RDX exists in pure solid form, and it is assumed that there are no chemical reactions taking place in this regime. This region is followed by the melt layer in which the temperature further increases from $478K$ to the surface temperature. The thickness of the melt layer based on these calculations is approximately 50 μm. This compares reasonably well with results that have been published in the literature. In the liquid melt layer, RDX can undergo reactions in the condensed phase that decompose liquid RDX into CH_2O, CO, N_2O, NO, H_2CN, NO_2 and H_2O. Based on our calculations, we find approximately 43% decomposition of liquid RDX in the melt layer.

The liquid melt layer is followed by gas phase combustion of gases evaporating from the melt layer. The temperature profile in the gas phase shows three distinct regions termed in the literature as fizz zone, dark zone and flame zone. The fizz zone is the region immediately above the gas-liquid interface. In the fizz zone, the temperature steadily rises from the surface temperature to approximately $1500K$. The fizz zone ends at a height of approximately $0.5mm$ above the gas-liquid interface. It is followed by a dark zone which is characterized by nearly isothermal reactions. The temperature remains almost constant throughout the width of the dark zone. The length of the dark zone based on our calculations was found to be $1.5mm$. At a height of roughly $2.0mm$ above the gas-liquid interface, the temperature again starts to rise sharply to a value close to the adiabatic flame temperature for RDX propellants. This region is called the flame zone.

The symbols in Figure 2 represent temperature measurements from the experiments done by Hanson-Parr and Parr [9]. Temperature profiles were obtained by piecing together OH and NO rotational temperature data and thermocouple measurements. In general, there is good agreement between the experimental measurements

and the numerically computed values for the temperatures in the gas, liquid and solid phases. The experimentally obtained temperature profile in the condensed phase is not a direct measurement under laser-assisted conditions, but consists of scaled values obtained from direct thermocouple measurement for the non-laser-assisted combustion of RDX propellants. This treatment of the temperature data assumes that there are no endothermic or exothermic reactions in the condensed phase and that the thermal wave profile scales as the inverse ratio of the burn rates. The gas temperature near the surface is about 610K in good agreement with prior thermocouple measurements. The temperature rises sharply to about 1500K at 0.35 mm and then turns over to a much more gradual slope in the dark zone. At about 2 mm it becomes steeper again and finally levels off to 2600K at 3.0 mm, at the top edge of the CN flame sheet.

Figures 3 and 4 show experimental and numerically computed profiles for the NO and NO_2 mole fractions above the gas-liquid interface. The numerically computed values for NO mole fractions are slightly higher than those obtained experimentally, whereas those for NO_2 mole fraction underpredict the experimental values. The location of the peak values and the widths of the NO and NO_2 mole fraction profiles are predicted accurately. The NO_2 peaks very close to the surface at a value of approximately 0.17 mm. It then decays rapidly to zero at 1.5 mm. With the exception of a small region close to the surface, the NO mole fraction lies in the range between 0.2 and 0.24 mm in the dark zone out to 2.0 mm. It falls sharply to zero at 2.5 mm as NO is consumed in the CN/NH flame sheet.

The profiles for the CN mole fraction (Figure 5) and the OH mole fraction (Figure 6) also show good agreement between the numerical and the experimental results. The CN mole fraction profile is indicative of the location of the flame zone. Although the experimental predictions are slightly lower than the numerical values, the location of peak concentrations as well as the width of the CN profile again compare well with the experimental results. The experimental CN profile is slightly wider and peaks at 2.5 mm at a value of 660 ppm. Similar results are obtained for the OH mole fraction as shown in Figure 6. The OH profile peaks outside the CN/NH flame sheet with a

mole fraction of 0.055.

Figure 7 shows a comparison between numerical and experimental results for the NH mole fraction. For CO_2 laser-assisted deflagration, experiments indicate that the narrow NH profile peaks at 2.3 mm above the surface at a value of 100 ppm. Although the numerically predicted profiles for the NH mole fraction are higher than those predicted experimentally, the NH radicals obtained experimentally were based on line strength from Herzberg. Hanson-Parr and Parr [56] have recently suggested that the measured NH concentrations are much higher based on different line strengths [57]. This implies that the results obtained experimentally could more closely match the numerical results. Additional computed gas phase species profiles are illustrated in Figures 8-14. Although no formaldehyde was detected in the experiments, the computations show CH_2O to form early and to have been consumed prior to the dark zone.

As the laser flux intensity is increased, the numerical calculations suggest that the burning rate increases correspondingly, with very small changes in the final flame temperatures. The burning rates corresponding to 70%, 80%, 90% and 100% were 0.190, 0.213, 0.236 and 0.258 cm/sec respectively. Hanson-Parr and Parr [9] have reported a burning rate value of 0.18 cm/sec for the laser-assisted combustion experimentally studied. Note that the experimental value is not a directly measured value, but a value scaled from non-laser-supported measurements based on the CN species height above the liquid interface. This increase in the mass flux rate (as compared with those for self-deflagration) is responsible for increasing the length of the dark zone as shown in Figure 15. Experimental results also indicate that the length of the dark zone varies as a function of the laser flux intensity. Hanson-Parr and Parr [9] have measured the height of the secondary flame as a function of laser flux intensity. They report values of 1.0, 2.2 and 3.5 mm for laser flux intensities of 50, 100 and 150 $cal/cm^2/sec$ respectively, which is consistent with the computed findings reported here.

The energy necessary to sustain a steady deflagration may be liberated by condensed-phase reactions, gas-phase reactions, interface reactions, or any combination of these,

and the endothermic part of the surface gasification process may be an equilibrium process, a rate-controlled process or of intermediate character. Figure 16 provides profiles for heat release rates above the gas-liquid interface for the laser assisted conditions. This figure clearly shows the presence of two distinct heat release regimes located at the primary and secondary flames. The convective and diffusive terms in the energy equation are also shown in this figure.

Self-deflagration

Numerical calculations have also been performed for self-deflagrating RDX propellants. As compared to the calculations for laser-assisted deflagration, the laser flux intensity which appears as a source term in the energy equation was removed. The temperature profile for self-deflagration at one atmosphere ambient pressure is illustrated in Figure 17. The temperature profile shows that the two flame structure that was observed in the laser-assisted deflagration is not present. This finding is in accordance with experimental results [9]. When the laser flux heating is removed, the mass burning rate is reduced and, as a result, the primary and the secondary flames superimpose on each other producing a single flame. This flame structure is very similar to that of conventional premixed flames. Although the temperature profile does not show the two flame structure, the species profiles show the characteristics of a two-stage flame.

Calculations have also been performed for self-deflagration of RDX propellants over a wide range of ambient pressures ranging from one to ninety atmospheres. The dependence of the burning rate on ambient pressure is shown in Figure 18, and has been compared with experimental data from Atwood et al. [58], Hanson-Parr and Parr [9] and Zenin [31]. In these calculations the temperature of the gas-liquid interface was varied as observed from experimental data (Table 3). It is clear that the present model provides an excellent comparison between numerical and experimental values. Figure 18 also shows comparison with data from Zimmer-Galler [59] and that of Glaskova [60], [61]. Note that there is significantly more scatter in the experimental burning rate data at higher ambient pressures.

The pressure in the combustion chamber of a solid rocket motor is dependent on the burning rate of the propellants. A large thrust occurs when the burning rate is high, however, when a long burning time is required, a lower burning rate propellant may be suitable. The burning rate is often expressed in a form proportional to the n^{th} power of pressure,

$$r_b = ap^n \tag{37}$$

where r_b is the burning rate, p is the pressure, n is the pressure exponent of the burning rate and a is a constant. Figure 19 shows the pressure sensitivity of RDX monopropellants and comparison with experimental results [58]. The pressure sensitivity n was found to be approximately 0.76, and matches very well with experimental results and theoretical work of Mitani and Williams [13], Ben-Reuven et al. [12], Li et al. [14] and Price et al. [62]. Table 4 and 5 summarizes some of the properties at the gas-liquid interface as a function of the ambient pressure. As the ambient pressure increases, the flame comes closer to the gas-liquid interface resulting in higher temperature gradients in the gas and the condensed phases. Due to the higher burning rates associated with higher ambient pressures, we find that the thickness of the melt layer decreases as the ambient pressure increases. Although the surface temperature increases as the ambient pressure increases, the amount of decomposition in the liquid phase decreases from about 40% at one atmosphere to almost 0% at 90 atmospheres. This is due to a combination of smaller melt layer thickness and higher burning rates at higher ambient pressures. When the initial propellant temperature is increased from 298 K to 373 K we find that the temperature gradient in the gas and liquid phase reduces, but the thickness of the melt layer, the amount of liquid phase decomposition and the burn rate increase slightly.

As pointed out in Figures 18 and 19, the burning rate is strongly dependent on pressure. However, it is also dependent on the initial temperature of the propellant even when the pressure is kept constant. If one defines σ_p as the variation of the burning rate with changes in initial propellant temperature, at a constant pressure, σ_p

then can be written as

$$\sigma_p = \frac{1}{r}\frac{r_1 - r_0}{T_1 - T_0}, \tag{38}$$

where r_0 and r_1 are the burning rates at temperatures T_0 and T_1, respectively [5]. Figures 20 and 21 show the sensitivity of the burn rate to initial propellant temperature. The burning rate has been computed at initial propellant temperatures of 20^0C and 100^0C (Figure 20) over a wide range of ambient pressure. These results compare favourably with experimental results from Atwood et al. [58] that have also been plotted in the same figure. Figure 21 shows the experimental and numerically obtained values for σ_p over a wide range of pressure calculated using equation (38). The value of σ_p for conventional propellants ranges between 0.002/K and 0.008/K [5]. The values obtained for RDX monopropellant are approximately 0.0008/K, and compare favourably with experimental results. The value for σ_p predicted by Mitani and Williams [13], Ben-Reuven et al. [12], Price et al. [62] and Li et al. [14] were significantly higher.

The eigenvalue approach presented in this paper requires the modeler to prescribe the temperature at the gas-liquid interface. The prescription of the gas-liquid interface temperature as well as the location of the gas-liquid interface removes the translational invariance associated with the freely propagating problem. This temperature also decides the point at which liquid decomposition products are converted into the corresponding gas phase species. Mathematically this translates into provision of an additional boundary condition for the extra differential equation (26) that is introduced into the system. Physically this is equivalent to specifying the vapour phase equilibrium diagram. An alternative / equivalent approach could be to specify the location and temperature of the melting point where the transition of the solid to liquid takes place. This approach would still require knowledge of the boiling point temperature, although in this case the location of the boiling point does not have to be fixed in the coordinate frame of reference. From a formulation point of view, these two approaches are equivalent. In the former approach the solid-liquid interface may fall between two adjacent grid points, whereas in the latter approach the gas-liquid interface, which is

more critical in determining the burning rate, may fall between two adjacent nodes. The latter approach may therefore be more susceptible to numerical oscillations. Apart from these potential numerical problems in constraining the melting point as compared to the boiling point, experimental results have indicated that the liquefication process is not a true melting in the sense of a phase change from solid to liquid occurring at a rather precise temperature. In general we believe that the melting process for RDX monopropellants is much less understood as compared to the evaporation process.

The temperature at the gas-liquid interface can be obtained from phase equilibrium diagrams, or measured experimentally. Due to the sharp temperature gradients and intense heat released at the propellant surface, it might be difficult to measure this temperature accurately. Small errors in temperature measurements at the gas-liquid interface and the use of this erroneous temperature in the eigenvalue approach could potentially result in incorrect burning rates. In such cases one could equally well specify the adiabatic flame temperature of the system and then compute not only the mass flux rate ($\dot{m}$) but the interface temperature as well. In addition, the eigenvalue approach can be used for propellant systems, where an appreciable amount of propellant decomposes in the condensed phase. If the condensed phase reaction is highly exothermic or if a significant amount of decomposition occurs, we expect the burning rate to vary substantially as the gas-liquid interface temperature is changed. Use of the eigenvalue approach could still provide accurate results, with either a specified gas-liquid interface temperature or an adiabatic flame temperature.

The eigenvalue approach performs an energy and species balance at the interface and makes use of properties of the gas and the liquid phases. This approach should produce correct results under laser-assisted conditions as well as self-deflagration conditions. We believe that models based on evaporation and condensation rate laws would be unable to predict the increased burning rates obtained under laser assisted conditions when calibrated for the unassisted cases. This is due to the fact that these rate laws do not account for the temperature gradient that exists at the gas-liquid interface or the additional laser flux heating that might be added to the system. We have shown

earlier that the increased burning rates obtained under laser assisted conditions result in a dark zone (dark zone is absent at lower burning rates). It has also been shown that as the flux intensity increases, the burning rate increases and this controls the length of the dark zone.

The quasi-one-dimensional nature of the laser assisted experiments is found to play an important role in the stabilization of the secondary flame. Numerical calculations were performed for laser assisted combustion with no assumed area variation. (Constant stream tube area $A(x) = 1.0$.) These calculations resulted in primary flame profiles and burning rates similar to those obtained when area profiles shown in equation (36) were used. However the secondary flame was pushed completely out of the computational domain, resulting in very large dark zones. This is due to the fact that the assumption of constant stream tube area results in larger gas velocities at the secondary flame as compared to those obtained with a divergent stream tube. Thus the area variation shown in equation (36) plays an important role in determining the location and structure of the secondary flame. In addition, by varying the laser intensity variable length dark zones could be achieved.

6 CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

Research in the field of solid rocket propellants over the past few decades has focused on obtaining a detailed understanding of the various physical and chemical processes involved during the burning of homogeneous and heterogeneous propellants. In this paper, a mathematical model for a three-tiered system consisting of solid, liquid and gas was derived for studying the combustion of RDX propellants. The resulting non-linear two-point boundary value problem was solved by Newton's method and adaptive gridding techniques. In this study, the burning rate was computed as an eigenvalue, which removes the uncertainty associated with employing evaporation and condensation rate laws in its evaluation. Results were presented for laser-assisted and self-deflagration of RDX monopropellants and were compared with experiments. The burning rates

were computed over a wide range of ambient pressures and were shown to compare favourably with experimental results. Results for laser-assisted combustion showed a distinct primary and secondary flame separated by a dark zone.

7 ACKNOWLEDGEMENT

The authors wish to acknowledge support for this work from the Office of Naval Research and to acknowledge Dr. R. S. Miller as Technical Monitor. The authors would also like to thank Drs. Tim Parr and Donna-Hanson Parr at the Naval Weapons Center for providing experimental results.

Table 1: RDX Liquid Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a	Ref.
1L	$RDX \rightarrow 3CH_2O + 3N_2O$	$.600 \times 10^{14}$	0.00	36000.0	[37]
2L	$RDX \rightarrow 3NO_2 + 3H_2CN$	$.160 \times 10^{18}$	0.00	45000.0	[37]
3L	$CH_2O + NO_2 \rightarrow CO + NO + H_2O$	$.802 \times 10^{03}$	2.77	13730.0	[42]

Table 2: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
1G	$H_2 + M \rightleftharpoons H + H + M^a$	$.457 \times 10^{20}$	−1.40	104000.0
2G	$O + H_2O \rightleftharpoons OH + OH$	$.295 \times 10^{07}$	2.02	13400.0
3G	$O + H_2 \rightleftharpoons H + OH$	$.508 \times 10^{05}$	2.67	6290.0
4G	$O + O + M \rightleftharpoons O_2 + M^a$	$.617 \times 10^{16}$	−0.50	0.0
5G	$H + O_2 \rightleftharpoons O + OH$	$.352 \times 10^{17}$	−0.70	17070.0
6G	$H + O_2 + M \rightleftharpoons HO_2 + M^a$	$.452 \times 10^{14}$	0.00	0.0
	Fall-off Parameters	$.675 \times 10^{-06}$	1.42	0.0
7G	$H + O + M \rightleftharpoons OH + M^a$	$.472 \times 10^{19}$	−1.00	0.0
8G	$OH + H_2 \rightleftharpoons H_2O + H$	$.216 \times 10^{09}$	1.51	3430.0
9G	$OH + H + M \rightleftharpoons H_2O + M^a$	$.221 \times 10^{23}$	−2.00	0.0
10G	$HO_2 + O \rightleftharpoons O_2 + OH$	$.175 \times 10^{14}$	0.00	−397.0
11G	$HO_2 + H \rightleftharpoons H_2 + O_2$	$.662 \times 10^{14}$	0.00	2130.0
12G	$HO_2 + H \rightleftharpoons OH + OH$	$.169 \times 10^{15}$	0.00	874.0
13G	$HO_2 + OH \rightleftharpoons H_2O + O_2$	$.190 \times 10^{17}$	−1.00	0.0
14G	$HO_2 + HO_2 \rightleftharpoons H_2O_2 + O_2$	$.420 \times 10^{15}$	0.00	11980.0
15G	$H_2O_2 + M \rightleftharpoons OH + OH + M$	$.300 \times 10^{15}$	0.00	48460.0
	Fall-off Parameters	$.250 \times 10^{-02}$	0.00	2960.0
16G	$H_2O_2 + O \rightleftharpoons OH + HO_2$	$.964 \times 10^{07}$	2.00	3970.0
17G	$H_2O_2 + H \rightleftharpoons H_2O + OH$	$.100 \times 10^{14}$	0.00	3590.0
18G	$H_2O_2 + H \rightleftharpoons HO_2 + H_2$	$.482 \times 10^{14}$	0.00	7950.0
19G	$H_2O_2 + OH \rightleftharpoons H_2O + HO_2$	$.580 \times 10^{15}$	0.00	9557.0
20G	$CH_2O + M \rightleftharpoons HCO + H + M$	$.163 \times 10^{34}$	−4.10	92550.0
21G	$CH_2O + M \rightleftharpoons H_2 + CO + M$	$.825 \times 10^{16}$	0.00	69540.0
22G	$CH_2O + O_2 \rightleftharpoons HCO + HO_2$	$.205 \times 10^{14}$	0.00	38920.0
23G	$CH_2O + O \rightleftharpoons HCO + OH$	$.181 \times 10^{14}$	0.00	3078.0
24G	$CH_2O + H \rightleftharpoons HCO + H_2$	$.794 \times 10^{08}$	1.66	2000.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
25G	$CH_2O + OH \rightleftharpoons HCO + H_2O$	$.343 \times 10^{10}$	1.18	−447.0
26G	$CH_2O + HO_2 \rightleftharpoons HCO + H_2O_2$	$.199 \times 10^{13}$	0.00	11660.0
27G	$HCO + M \rightleftharpoons H + CO + M^b$	$.186 \times 10^{18}$	−1.00	17000.0
28G	$HCO + O_2 \rightleftharpoons CO + HO_2$	$.758 \times 10^{13}$	0.00	410.0
29G	$HCO + O \rightleftharpoons CO + OH$	$.300 \times 10^{14}$	0.00	0.0
30G	$HCO + O \rightleftharpoons CO_2 + H$	$.300 \times 10^{14}$	0.00	0.0
31G	$HCO + H \rightleftharpoons CO + H_2$	$.723 \times 10^{14}$	0.00	0.0
32G	$HCO + OH \rightleftharpoons CO + H_2O$	$.300 \times 10^{14}$	0.00	0.0
33G	$HCO + HO_2 \rightleftharpoons CO_2 + OH + H$	$.300 \times 10^{14}$	0.00	0.0
34G	$CO + O + M \rightleftharpoons CO_2 + M^a$	$.251 \times 10^{14}$	0.00	−4540.0
35G	$CO + O_2 \rightleftharpoons CO_2 + O$	$.253 \times 10^{13}$	0.00	47700.0
36G	$CO + OH \rightleftharpoons CO_2 + H$	$.150 \times 10^{08}$	1.30	−765.0
37G	$CO + HO_2 \rightleftharpoons CO_2 + OH$	$.603 \times 10^{14}$	0.00	23000.0
38G	$N + H_2 \rightleftharpoons H + NH$	$.160 \times 10^{15}$	0.00	25140.0
39G	$N + O_2 \rightleftharpoons NO + O$	$.640 \times 10^{10}$	1.00	6280.0
40G	$N + OH \rightleftharpoons NO + H$	$.380 \times 10^{14}$	0.00	0.0
41G	$N + HO_2 \rightleftharpoons NH + O_2$	$.100 \times 10^{14}$	0.00	2000.0
42G	$N + HO_2 \rightleftharpoons NO + OH$	$.100 \times 10^{14}$	0.00	2000.0
43G	$N + CO_2 \rightleftharpoons NO + CO$	$.190 \times 10^{12}$	0.00	3400.0
44G	$N + NO \rightleftharpoons N_2 + O$	$.327 \times 10^{13}$	0.30	0.0
45G	$N + NO_2 \rightleftharpoons NO + NO$	$.400 \times 10^{13}$	0.00	0.0
46G	$N + NO_2 \rightleftharpoons N_2O + O$	$.500 \times 10^{13}$	0.00	0.0
47G	$N + NO_2 \rightleftharpoons N_2 + O_2$	$.100 \times 10^{13}$	0.00	0.0
48G	$N + HNO \rightleftharpoons NH + NO$	$.100 \times 10^{14}$	0.00	2000.0
49G	$N + HNO \rightleftharpoons N_2O + H$	$.500 \times 10^{11}$	0.50	3000.0
50G	$N + N_2O \rightleftharpoons N_2 + NO$	$.100 \times 10^{14}$	0.00	19870.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
51G	$NO + M \rightleftharpoons N + O + M^c$	$.964 \times 10^{15}$	0.00	148400.0
52G	$NO + H_2 \rightleftharpoons HNO + H$	$.139 \times 10^{14}$	0.00	56530.0
53G	$NO + O + M \rightleftharpoons NO_2 + M$	$.130 \times 10^{16}$	−0.75	0.0
	Fall-off Parameters	$.275 \times 10^{-09}$	2.12	−1551.0
54G	$NO + H + M \rightleftharpoons HNO + M$	$.152 \times 10^{16}$	−0.41	0.0
	Fall-off Parameters	$.170 \times 10^{-04}$	0.91	−735.2
55G	$NO + OH + M \rightleftharpoons HONO + M^d$	$.199 \times 10^{13}$	−0.05	−721.0
	Fall-off Parameters	$.392 \times 10^{-11}$	2.46	−653.0
56G	$NO + HCO \rightleftharpoons HNO + CO$	$.140 \times 10^{14}$	0.00	0.0
57G	$NO_2 + O \rightleftharpoons O_2 + NO$	$.100 \times 10^{14}$	0.00	600.0
58G	$NO_2 + O + M \rightleftharpoons NO_3 + M$	$.133 \times 10^{14}$	0.00	0.0
	Fall-off Parameters	$.893 \times 10^{-15}$	4.08	−2467.0
59G	$NO_2 + H \rightleftharpoons NO + OH$	$.132 \times 10^{15}$	0.00	361.6
60G	$NO_2 + OH + M \rightleftharpoons HNO_3 + M$	$.241 \times 10^{14}$	0.00	0.0
	Fall-off Parameters	$.375 \times 10^{-19}$	5.49	−2350.0
61G	$HO_2 + NO \rightleftharpoons NO_2 + OH$	$.211 \times 10^{13}$	0.00	−479.0
62G	$NO_2 + CH_2O \rightleftharpoons HONO + HCO$	$.783 \times 10^{03}$	2.77	13730.0
63G	$NO_2 + HCO \rightleftharpoons CO + HONO$	$.124 \times 10^{24}$	−3.29	2354.0
64G	$NO_2 + HCO \rightleftharpoons H + CO_2 + NO$	$.839 \times 10^{16}$	−0.75	1927.0
65G	$NO_2 + CO \rightleftharpoons CO_2 + NO$	$.903 \times 10^{14}$	0.00	33780.0
66G	$NO_2 + NO_2 \rightleftharpoons NO_3 + NO$	$.964 \times 10^{10}$	0.73	20920.0
67G	$NO_2 + NO_2 \rightleftharpoons 2NO + O_2$	$.163 \times 10^{13}$	0.00	26120.0
68G	$NH + M \rightleftharpoons N + H + M$	$.265 \times 10^{15}$	0.00	75510.0
69G	$NH + O_2 \rightleftharpoons HNO + O$	$.389 \times 10^{14}$	0.00	17890.0
70G	$NH + O_2 \rightleftharpoons NO + OH$	$.760 \times 10^{11}$	0.00	1530.0
71G	$NH + O \rightleftharpoons NO + H$	$.550 \times 10^{14}$	0.00	0.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
72G	$NH + O \rightleftharpoons N + OH$	$.372 \times 10^{14}$	0.00	0.0
73G	$NH + OH \rightleftharpoons HNO + H$	$.200 \times 10^{14}$	0.00	0.0
74G	$NH + OH \rightleftharpoons N + H_2O$	$.500 \times 10^{12}$	0.50	2000.0
75G	$NH + N \rightleftharpoons N_2 + H$	$.300 \times 10^{14}$	0.00	0.0
76G	$NH + NO \rightleftharpoons N_2O + H$	$.294 \times 10^{15}$	−0.40	0.0
77G	$NH + NO \rightleftharpoons N_2 + OH$	$.216 \times 10^{14}$	−0.23	0.0
78G	$NH + NO_2 \rightleftharpoons NO + HNO$	$.100 \times 10^{12}$	0.50	4000.0
79G	$NH + NO_2 \rightleftharpoons N_2O + OH$	$.100 \times 10^{14}$	0.00	0.0
80G	$NH + NH \rightleftharpoons N_2 + H + H$	$.510 \times 10^{14}$	0.00	0.0
81G	$NH_2 + O_2 \rightleftharpoons HNO + OH$	$.178 \times 10^{13}$	0.00	14900.0
82G	$NH_2 + O \rightleftharpoons HNO + H$	$.663 \times 10^{15}$	−0.50	0.0
83G	$NH_2 + O \rightleftharpoons NH + OH$	$.675 \times 10^{13}$	0.00	0.0
84G	$NH_2 + H \rightleftharpoons NH + H_2$	$.692 \times 10^{14}$	0.00	3650.0
85G	$NH_2 + OH \rightleftharpoons NH + H_2O$	$.400 \times 10^{07}$	2.00	1000.0
86G	$NH_2 + N \rightleftharpoons N_2 + 2H$	$.720 \times 10^{14}$	0.00	0.0
87G	$NH_2 + NO \rightleftharpoons NNH + OH$	$.280 \times 10^{14}$	−0.55	0.0
88G	$NH_2 + NO \rightleftharpoons N_2 + H_2O$	$.882 \times 10^{16}$	−1.25	0.0
89G	$NH_2 + NO \rightleftharpoons N_2O + H_2$	$.500 \times 10^{14}$	0.00	24640.0
90G	$NH_2 + NO \rightleftharpoons HNO + NH$	$.100 \times 10^{14}$	0.00	40000.0
91G	$NH_2 + NO_2 \rightleftharpoons N_2O + H_2O$	$.328 \times 10^{19}$	−2.20	0.0
92G	$NH_3 + M \rightleftharpoons NH_2 + H + M$	$.220 \times 10^{17}$	0.00	93470.0
93G	$NH_3 + O \rightleftharpoons NH_2 + OH$	$.940 \times 10^{07}$	1.94	6460.0
94G	$NH_3 + H \rightleftharpoons NH_2 + H_2$	$.640 \times 10^{06}$	2.39	10170.0
95G	$NH_3 + OH \rightleftharpoons NH_2 + H_2O$	$.204 \times 10^{07}$	2.04	566.0
96G	$NH_3 + HO_2 \rightleftharpoons NH_2 + H_2O_2$	$.300 \times 10^{12}$	0.00	22000.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
97G	$NH_2 + HO_2 \rightleftharpoons NH_3 + O_2$	$.100 \times 10^{14}$	0.00	0.0
98G	$NH_2 + NH_2 \rightleftharpoons NH_3 + NH$	$.500 \times 10^{14}$	0.00	10000.0
99G	$NNH + M \rightleftharpoons N_2 + H + M$	$.100 \times 10^{15}$	0.00	3000.0
100G	$NNH + H \rightleftharpoons N_2 + H_2$	$.100 \times 10^{15}$	0.00	0.0
101G	$NNH + NO \rightleftharpoons N_2 + HNO$	$.500 \times 10^{14}$	0.00	0.0
102G	$NNH + O \rightleftharpoons N_2O + H$	$.100 \times 10^{15}$	0.00	0.0
103G	$NNH + OH \rightleftharpoons N_2 + H_2O$	$.500 \times 10^{14}$	0.00	0.0
104G	$NNH + NH \rightleftharpoons N_2 + NH_2$	$.500 \times 10^{14}$	0.00	0.0
105G	$NNH + NH_2 \rightleftharpoons N_2 + NH_3$	$.500 \times 10^{14}$	0.00	0.0
106G	$HNO + O \rightleftharpoons OH + NO$	$.181 \times 10^{14}$	0.00	0.0
107G	$HNO + OH \rightleftharpoons H_2O + NO$	$.100 \times 10^{14}$	0.00	993.5
108G	$HNO + HCO \rightleftharpoons CH_2O + NO$	$.602 \times 10^{12}$	0.00	1987.0
109G	$HNO + NO \rightleftharpoons N_2O + OH$	$.200 \times 10^{13}$	0.00	26000.0
110G	$HNO + NO_2 \rightleftharpoons HONO + NO$	$.602 \times 10^{12}$	0.00	1987.0
111G	$HNO + HNO \rightleftharpoons H_2O + N_2O$	$.851 \times 10^{09}$	0.00	3080.0
112G	$HNO + O_2 \rightleftharpoons NO + HO_2$	$.100 \times 10^{14}$	0.00	25000.0
113G	$HNO + NH_2 \rightleftharpoons NO + NH_3$	$.200 \times 10^{14}$	0.00	1000.0
114G	$HONO + O \rightleftharpoons OH + NO_2$	$.120 \times 10^{14}$	0.00	5961.0
115G	$HONO + H \rightleftharpoons H_2 + NO_2$	$.120 \times 10^{14}$	0.00	7352.0
116G	$HONO + OH \rightleftharpoons H_2O + NO_2$	$.126 \times 10^{11}$	1.00	135.1
117G	$HCN + M \rightleftharpoons H + CN + M$	$.830 \times 10^{18}$	−0.93	123800.0
	Fall-off Parameters	$.232 \times 10^{-08}$	1.67	−1100.0
118G	$HCN + O \rightleftharpoons CN + OH$	$.270 \times 10^{10}$	1.58	29200.0
119G	$HCN + O \rightleftharpoons NH + CO$	$.345 \times 10^{04}$	2.64	4980.0
120G	$HCN + O \rightleftharpoons NCO + H$	$.138 \times 10^{05}$	2.64	4980.0
121G	$HCN + OH \rightleftharpoons H_2O + CN$	$.390 \times 10^{07}$	1.83	10290.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
122G	$HCN + OH \rightleftharpoons H + HOCN$	$.585 \times 10^{05}$	2.40	12500.0
123G	$HCN + OH \rightleftharpoons H + HNCO$	$.198 \times 10^{-02}$	4.00	1000.0
124G	$HCN + OH \rightleftharpoons NH_2 + CO$	$.783 \times 10^{-03}$	4.00	4000.0
125G	$HCN \rightleftharpoons HNC$	$.206 \times 10^{15}$	−1.11	43710.0
126G	$HNC + O \rightleftharpoons NH + CO$	$.289 \times 10^{13}$	0.00	0.0
127G	$HNC + O \rightleftharpoons H + NCO$	$.160 \times 10^{02}$	3.08	−224.0
128G	$HNC + OH \rightleftharpoons HNCO + H$	$.280 \times 10^{14}$	0.00	3700.0
129G	$HNC + OH \rightleftharpoons CN + H_2O$	$.150 \times 10^{13}$	0.00	7680.0
130G	$HNC + NO_2 \rightleftharpoons HNCO + NO$	$.100 \times 10^{13}$	0.00	32000.0
131G	$HNC + CN \rightleftharpoons C_2N_2 + H$	$.100 \times 10^{14}$	0.00	0.0
132G	$N_2O + M \rightleftharpoons N_2 + O + M^e$	$.791 \times 10^{11}$	0.00	56040.0
	Fall-off Parameters	$.900 \times 10^{-04}$	0.00	−1510.0
133G	$N_2O + O \rightleftharpoons O_2 + N_2$	$.100 \times 10^{15}$	0.00	28000.0
134G	$N_2O + O \rightleftharpoons 2NO$	$.100 \times 10^{15}$	0.00	28000.0
135G	$N_2O + H \rightleftharpoons N_2 + OH$	$.223 \times 10^{15}$	0.00	16750.0
136G	$N_2O + OH \rightleftharpoons HO_2 + N_2$	$.200 \times 10^{13}$	0.00	40000.0
137G	$N_2O + CO \rightleftharpoons N_2 + CO_2$	$.319 \times 10^{12}$	0.00	20330.0
138G	$CN + H_2 \rightleftharpoons H + HCN$	$.550 \times 10^{03}$	3.18	−223.0
139G	$CN + O_2 \rightleftharpoons NCO + O$	$.750 \times 10^{13}$	0.00	−389.0
140G	$CN + O \rightleftharpoons CO + N$	$.180 \times 10^{14}$	0.00	0.0
141G	$CN + OH \rightleftharpoons NCO + H$	$.400 \times 10^{14}$	0.00	0.0
142G	$CN + OH \rightleftharpoons NH + CO$	.000	0.00	0.0
143G	$CN + OH \rightleftharpoons HNCO$	.000	0.00	0.0
144G	$CN + CH_2O \rightleftharpoons HCN + HCO$	$.422 \times 10^{14}$	0.00	0.0
145G	$CN + HCO \rightleftharpoons HCN + CO$	$.602 \times 10^{14}$	0.00	0.0
146G	$CN + NO \rightleftharpoons NCO + N$	$.964 \times 10^{14}$	0.00	42120.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
147G	$CN + CO_2 \rightleftharpoons CO + NCO$	$.367 \times 10^{07}$	2.16	26900.0
148G	$CN + NO_2 \rightleftharpoons NCO + NO$	$.159 \times 10^{14}$	0.00	−1133.0
149G	$CN + HNO \rightleftharpoons HCN + NO$	$.181 \times 10^{14}$	0.00	0.0
150G	$CN + HONO \rightleftharpoons HCN + NO_2$	$.120 \times 10^{14}$	0.00	0.0
151G	$CN + HCN \rightleftharpoons H + C_2N_2$	$.121 \times 10^{08}$	1.71	1530.0
152G	$CN + N_2O \rightleftharpoons NCN + NO$	$.385 \times 10^{04}$	2.60	3696.0
153G	$CN + CN + M \rightleftharpoons C_2N_2 + M$	$.566 \times 10^{13}$	0.00	0.0
	Fall-off Parameters	$.165 \times 10^{-12}$	2.61	0.0
154G	$C_2N_2 + O \rightleftharpoons NCO + CN$	$.457 \times 10^{13}$	0.00	8880.0
155G	$C_2N_2 + OH \rightleftharpoons HOCN + CN$	$.186 \times 10^{12}$	0.00	2900.0
156G	$NCN + H \rightleftharpoons HCN + N$	$.100 \times 10^{15}$	0.00	0.0
157G	$NCN + O \rightleftharpoons CN + NO$	$.100 \times 10^{15}$	0.00	0.0
158G	$NCN + OH \rightleftharpoons HCN + NO$	$.500 \times 10^{14}$	0.00	0.0
159G	$NCN + O_2 \rightleftharpoons NO + NCO$	$.100 \times 10^{15}$	0.00	0.0
160G	$NCO + M \rightleftharpoons N + CO + M^f$	$.310 \times 10^{17}$	−0.50	48300.0
161G	$NCO + H_2 \rightleftharpoons HNCO + H$	$.760 \times 10^{03}$	3.00	4000.0
162G	$NCO + O \rightleftharpoons CO + NO$	$.200 \times 10^{14}$	0.00	0.0
163G	$NCO + H \rightleftharpoons NH + CO$	$.500 \times 10^{14}$	0.00	0.0
164G	$NCO + OH \rightleftharpoons NO + CO + H$	$.100 \times 10^{14}$	0.00	0.0
165G	$NCO + OH \rightleftharpoons NO + HCO$	$.500 \times 10^{13}$	0.00	15000.0
166G	$NCO + O_2 \rightleftharpoons NO + CO_2$	$.200 \times 10^{13}$	0.00	20000.0
167G	$NCO + CH_2O \rightleftharpoons HNCO + HCO$	$.602 \times 10^{13}$	0.00	0.0
168G	$NCO + HCO \rightleftharpoons HNCO + CO$	$.361 \times 10^{14}$	0.00	0.0
169G	$NCO + NO \rightleftharpoons N_2O + CO$	$.620 \times 10^{18}$	−1.73	763.0
170G	$NCO + NO \rightleftharpoons CO_2 + N_2$	$.780 \times 10^{18}$	−1.73	763.0
171G	$NCO + NO_2 \rightleftharpoons CO + 2NO$	$.139 \times 10^{14}$	0.00	0.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
172G	$NCO + NO_2 \rightleftharpoons CO_2 + N_2O$	$.417 \times 10^{13}$	0.00	0.0
173G	$NCO + HNO \rightleftharpoons HNCO + NO$	$.181 \times 10^{14}$	0.00	0.0
174G	$NCO + HONO \rightleftharpoons HNCO + NO_2$	$.361 \times 10^{13}$	0.00	0.0
175G	$NCO + N_2O \rightleftharpoons N_2 + NO + CO$	$.903 \times 10^{14}$	0.00	27820.0
176G	$NCO + CN \rightleftharpoons NCN + CO$	$.181 \times 10^{14}$	0.00	0.0
177G	$NCO + NCO \rightleftharpoons N_2 + 2CO$	$.100 \times 10^{14}$	0.00	0.0
178G	$NCO + N \rightleftharpoons N_2 + CO$	$.200 \times 10^{14}$	0.00	0.0
179G	$CNO + O \rightleftharpoons CO + NO$	$.100 \times 10^{14}$	0.00	0.0
180G	$CNO + NO_2 \rightleftharpoons CO + 2NO$	$.100 \times 10^{14}$	0.00	0.0
181G	$CNO + N_2O \rightleftharpoons N_2 + CO + NO$	$.100 \times 10^{13}$	0.00	15000.0
182G	$HNCO + M \rightleftharpoons NH + CO + M$	$.600 \times 10^{14}$	0.00	99800.0
	Fall-off Parameters	$.276 \times 10^{-14}$	3.10	-2100.0
183G	$HNCO + O \rightleftharpoons CO_2 + NH$	$.964 \times 10^{08}$	1.41	8524.0
184G	$HNCO + O \rightleftharpoons OH + NCO$	$.667 \times 10^{-03}$	4.55	1780.0
185G	$HNCO + O \rightleftharpoons HNO + CO$	$.158 \times 10^{09}$	1.57	44300.0
186G	$HNCO + H \rightleftharpoons NH_2 + CO$	$.220 \times 10^{08}$	1.70	3800.0
187G	$HNCO + OH \rightleftharpoons H_2O + NCO$	$.638 \times 10^{06}$	2.00	2563.0
188G	$HNCO + CN \rightleftharpoons HCN + NCO$	$.151 \times 10^{14}$	0.00	0.0
189G	$HNCO + HO_2 \rightleftharpoons NCO + H_2O_2$	$.300 \times 10^{12}$	0.00	29000.0
190G	$HNCO + O_2 \rightleftharpoons HNO + CO_2$	$.100 \times 10^{13}$	0.00	35000.0
191G	$HNCO + NH_2 \rightleftharpoons NH_3 + NCO$	$.500 \times 10^{13}$	0.00	6200.0
192G	$HNCO + NH \rightleftharpoons NH_2 + NCO$	$.300 \times 10^{14}$	0.00	23700.0
193G	$HCNO + O \rightleftharpoons HCO + NO$	$.100 \times 10^{13}$	0.00	9000.0
194G	$HCNO + OH \rightleftharpoons HCO + HNO$	$.100 \times 10^{14}$	0.00	5000.0
195G	$HCNO + OH \rightleftharpoons CNO + H_2O$	$.100 \times 10^{13}$	0.00	2000.0
196G	$HCNO + CN \rightleftharpoons HCN + CNO$	$.100 \times 10^{13}$	0.00	2000.0
197G	$HOCN + H \rightleftharpoons HNCO + H$	$.200 \times 10^{08}$	2.00	2000.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
198G	$HOCN + OH \rightleftharpoons NCO + H_2O$	$.640 \times 10^{06}$	2.00	2560.0
199G	$HOCN + O \rightleftharpoons NCO + OH$	$.150 \times 10^{05}$	2.64	4000.0
200G	$H_2CN + M \rightleftharpoons HCN + H + M$	$.100 \times 10^{18}$	0.00	30000.0
201G	$H_2CN + CH_2O \rightleftharpoons H_2CNH + HCO$	$.100 \times 10^{12}$	0.00	14000.0
202G	$H_2CN + NO \rightleftharpoons HCN + HNO$	$.100 \times 10^{12}$	0.00	3000.0
203G	$H_2CN + NO_2 \rightleftharpoons HCN + HONO$	$.100 \times 10^{12}$	0.00	1000.0
204G	$H_2CN + NO_2 \rightleftharpoons H_2CNO + NO$	$.100 \times 10^{12}$	0.00	3000.0
205G	$H_2CN + N_2O \rightleftharpoons H_2CNO + N_2$	$.100 \times 10^{12}$	0.00	3000.0
206G	$H_2CN + HNO \rightleftharpoons H_2CNH + NO$	$.100 \times 10^{12}$	0.00	4000.0
207G	$H_2CN + HONO \rightleftharpoons H_2CNH + NO_2$	$.100 \times 10^{12}$	0.00	12000.0
208G	$H_2CNH + OH \rightleftharpoons H_2CN + H_2O$	$.100 \times 10^{14}$	0.00	0.0
209G	$H_2CNH + CN \rightleftharpoons H_2CN + HCN$	$.100 \times 10^{14}$	0.00	0.0
210G	$H_2CNO + M \rightleftharpoons HCNO + H$	$.100 \times 10^{17}$	0.00	50000.0
211G	$H_2CNO + OH \rightleftharpoons HCNO + H_2O$	$.100 \times 10^{14}$	0.00	0.0
212G	$H_2CNO + NO \rightleftharpoons HCNO + HNO$	$.100 \times 10^{13}$	0.00	25000.0
213G	$H_2CNO + NO_2 \rightleftharpoons HCNO + HONO$	$.100 \times 10^{13}$	0.00	2000.0
214G	$H_2CNO + NO_2 \rightleftharpoons CH_2O + NO + NO$	$.100 \times 10^{13}$	0.00	0.0
215G	$H_2CNO + HNO \rightleftharpoons H_2CN + HONO$	$.100 \times 10^{13}$	0.00	2000.0
216G	$H_2CNNO + M \rightleftharpoons H_2CN + NO + M$	$.100 \times 10^{17}$	0.00	2000.0
	Fall-off Parameters	.130	0.00	−13000.0
217G	$H_2CNNO_2 + M \rightleftharpoons H_2CN + NO_2 + M$	$.100 \times 10^{17}$	0.00	31000.0
	Fall-off Parameters	.130	0.00	5000.0
218G	$H_2CNNO_2 + NO_2 \rightleftharpoons CH_2O + N_2O + NO_2$	$.100 \times 10^{12}$	0.00	2000.0
219G	$H_2CNNO_2 + N_2O \rightleftharpoons CH_2O + N_2O + N_2O$	$.100 \times 10^{12}$	0.00	2000.0
220G	$H_2CNNO_2 + H_2O \rightleftharpoons CH_2O + N_2O + H_2O$	$.100 \times 10^{12}$	0.00	2000.0
221G	$H_2CNNO_2 + H \rightleftharpoons H_2CN + HONO$	$.100 \times 10^{13}$	0.00	5000.0
222G	$H_2CNNO_2 + OH \rightleftharpoons HCN + NO_2 + H_2O$	$.100 \times 10^{14}$	0.00	3000.0

Table 2-cont.: Detailed RDX Gas Phase Reaction Mechanism

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

No.	Reaction	A	n	E_a
223G	$H_2CNNO_2 + OH \rightleftharpoons CH_2O + N_2O + OH$	$.100 \times 10^{14}$	0.00	0.0
224G	$RDX + M \rightleftharpoons RDXR + NO_2 + M$	$.200 \times 10^{17}$	0.00	45000.0
	Fall-off Parameters	.127	0.00	17000.0
225G	$RDX + H \rightleftharpoons RDXR + HONO$	$.100 \times 10^{14}$	0.00	5000.0
226G	$RDX + OH \rightarrow 2H_2CNNO_2 + HCN + NO_2 + H_2O$	$.100 \times 10^{14}$	0.00	5000.0
227G	$RDXR + M \rightarrow RDXRO + M$	$.100 \times 10^{17}$	0.00	23000.0
	Fall-off Parameters	.130	0.00	5000.0
228G	$RDXRO + M \rightarrow 2H_2CNNO_2 + H_2CN + M$	$.100 \times 10^{17}$	0.00	23000.0
	Fall-off Parameters	.130	0.00	5000.0

[a] Efficiency factors for the collision partners of this pressure dependent reaction are: $\varepsilon_{H_2} = 2.5$, $\varepsilon_{H_2O} = 12.0$, $\varepsilon_{CO} = 1.9$ and $\varepsilon_{CO_2} = 3.8$. All other species have efficiences equal to unity.

[b] Efficiency factors for the collision partners of this pressure dependent reaction are: $\varepsilon_{H_2} = 1.89$, $\varepsilon_{H_2O} = 12.0$, $\varepsilon_{CO} = 1.9$ and $\varepsilon_{CO_2} = 3.8$. All other species have efficiences equal to unity.

[c] Efficiency factors for the collision partners of this pressure dependent reaction are: $\varepsilon_{N_2} = 1.5$ and $\varepsilon_{CO_2} = 10.0$. All other species have efficiences equal to unity.

[d] Efficiency factors for the collision partners of this pressure dependent reaction are: $\varepsilon_{H_2O} = 5.0$. All other species have efficiences equal to unity.

[e] Efficiency factors for the collision partners of this pressure dependent reaction are: $\varepsilon_{H_2O} = 7.5$, $\varepsilon_{NO} = 2.0$, $\varepsilon_{CO} = 2.0$, $\varepsilon_{CO_2} = 3.0$ and $\varepsilon_{HCN} = 3.0$. All other species have efficiences equal to unity.

[f] Efficiency factors for the collision partners of this pressure dependent reaction are: $\varepsilon_{N_2} = 1.5$. All other species have efficiences equal to unity.

Table 3: Burning Rate vs. Pressure

Pressure atm	Surface Temperature ^oK	Burn Rate Parr [9],[58] $r_b$cm/sec	Burn Rate Zenin [19] $r_b$cm/sec	Burn Rate Numerical $r_b$cm/sec
1.0	590.0[a]	0.0400	0.0500	0.0533
5.0	630.0[b]	0.1464	0.1700	0.1596
10.0	650.0[a]	0.2560		0.2809
20.0	690.0[a]	0.4476	0.5000	0.4735
50.0	730.0[b]	0.9369		1.0100
90.0	790.0[a]	1.5047	1.8000	1.6054

[a] Refer Zenin [19]. [b] Interpolated value.

Table 4: Gas Liquid Interface Conditions as a function of Ambient Pressure : $T_u = 298K$

Pressure (atm)	1.0	5.0	10.0	20.0	50.0	90.0
Gas Phase Heat Conduction (K/cm)	1.66E5	3.85E5	6.38E5	9.70E5	1.96E6	3.09E6
Condensed Phase Heat Conduction (K/cm)	3.84E4	1.00E5	1.66E5	2.77E5	5.24E5	7.62E5
Thickness of Melt Layer (μm)	70	33	22	12	7	5
Liquid Phase Decomposition (%)	40.0	17.5	7.8	3.5	0.3	0.0

Table 5: Gas Liquid Interface Conditions as a function of Ambient Pressure : $T_u = 373K$

Pressure (atm)	1.0	5.0	10.0	20.0	50.0	90.0
Gas Phase Heat Conduction (K/cm)	1.54E5	3.74E5	5.94E5	9.50E5	1.93E6	3.06E6
Condensed Phase Heat Conduction (K/cm)	3.62E4	9.20E4	1.49E5	2.54E5	4.74E5	6.93E5
Thickness of Melt Layer (μm)	72	35	26	19	11	5
Liquid Phase Decomposition (%)	43.0	20.5	10.7	5.3	0.6	0.1

References

[1] Glassman, I. and Sawyer, R. F. (1970). The Performance of Chemical Propellants, Circa Publication, Inc., New York.

[2] Summerfield, M., Sutherland, G., Webb, M., Taback, H., and Hall, K.P. (1960). Burning Mechanism of Ammonium Perchlorate Propellants, In Summerfield M. (Ed.), ARS Progress in Astronautics and Rocketry: Solid Propellant Rocket Research, Academic Press Inc., New York, **1**, 141.

[3] Hermance, C.E. (1966). A Model of Composite Propellant Combustion Including Surface Heterogeneity and Heat Generation, AIAA Journal, **4**, 1629.

[4] Beckstead, M.W., Derr, R.L., and Price, C.F. (1970). A Model of Composite Solid-Propellant Combustion Based on Multiple Flames, AIAA Journal, **8**, 2200.

[5] Sutton, G.P. (1986). Rocket Propulsion Elements: An Introduction to the Engineering of Rockets, Fifth Edition John Wiley & Sons, New York, 262.

[6] Kubota, N. (1984). Survey of Rocket Propellants and Their Combustion Characteristics, In Kuo, K.K. and Summerfield, M. (Ed.), Progress in Astronautics and Aeronautics: Fundamentals of Solid-Propellant Combustion, AIAA, Inc., New York, **90**, 1.

[7] Summerfield, M. (1959). The Liquid Propellant Rocket Engine, Princeton Univ. press, Princeton, NJ.

[8] Boggs, T.L. (1984). The Thermal Behavior of Cyclotrimethylenetrinitramine (RDX) and Cyclotetramethylenetetranitramine (HMX), In Kuo, K.K. and Summerfield, M. (Ed.) Progress in Astronautics and Aeronautics: Fundamentals of Solid-Propellant Combustion, AIAA, Inc., New York, **90**, 121.

[9] Parr, D. H. and Parr, T. (1994). RDX Flame Structure, Proceedings of the Twenty-Fifth Symposium (International) on Combustion, The Combustion Institute.

[10] Ermolin N.E., Korbeinichev, O. P., Kuibida, L.V. and Fomin, V. M. (1989). Flame Structure, Combustion Explosion and Shock Waves, **24**, 400.

[11] Ben-Reuven, M. Caveny, L.H., Vichnevetsky, R., and Summerfield, M. (1976). Flame Zone and Subsurface Reaction Model for Deflagrating RDX, Proceedings of the Sixteenth Symposium (International) on Combustion, The Combustion Institute.

[12] Ben-Reuven, M. and Caveny, L.H. (1981). Nitramine Flame Chemistry and Deflagration Interpreted in Terms of a Flame Model, AIAA J., **19**, 10, 1276.

[13] Mitani, T. and Williams, F.A. (1988). A Model for the Deflagration of Nitramines, Proceedings of the Twenty-First Symposium (International) on Combustion, The Combustion Institute.

[14] Li, S.C. and Williams, F.A. and Margolis, S.B. (1990). Effects of Two-Phase Flow in a Model for Nitramine Deflagration, Combustion and Flame, **80**, 329.

[15] Melius, C.F. (1990). Thermochemical Modeling: II. Application to Ignition and Combustion of Energetic Materials, In Bulusu, S.N. (Ed.), Chemistry and Physics of Energetic Materials, Kluwer Academic Publishers, The Netherlands, 51.

[16] Melius, C.F. (1988). The Gas-Phase Chemistry of Nitramine Combustion, Proceedings of the Twenty-Fifth JANNAF Combustion Meeting.

[17] Fetherolf, B.L. and Litzinger, T.A. (1993). Chemical Structure of the Gas-Phase Above Deflagrating RDX: Comparison of Experimental Measurements and Model Predictions, Proceedings of the Thirteeth JANNAF Combustion Meeting.

[18] Hatch, R. L. (1987). Chemical Kinetics Modeling of HMX Combustion, Proceedings of the Twenty-Fourth JANNAF Combustion Meeting.

[19] Zenin, A. (1995). Journal of Power and Propulsion, submitted.

[20] Blomshield, F.S. (1989). Nitramine Composite Solid Propellant Modelling, NWC TP 6992.

[21] Williams, F.A. (1985). Combustion Theory, Addison-Wesley Publishing Co., Inc., New York.

[22] Kuo, K.K. (1986). Fundamentals of Combustion, Wiley, NY.

[23] Kee, R. J., Miller, J. A., and Jefferson, T.H. (1980). CHEMKIN: A general purpose, Transportable, Fortran Chemical Kinetics Code Package, Sandia National Laboratories Report, SAND80-8003.

[24] Giovangigli, V., and Darabiha, N. (1987). Vector computers and complex chemistry combustion, Proceedings of the Conference on Mathematical Modeling in Combustion, Lyon, France, NATO ASI Series.

[25] Smooke, M.D. (1982). Solution of Burner-Stabilized Premixed Laminar Flames by Boundary Value Methods, Journal of Computational Physics, **48**, 72.

[26] Giovangigli, V., and Smooke, M.D. (1992). Application of Continuation Methods to Plane Premixed Laminar Flames, Combustion Science and Technology, **87**, 241.

[27] Smooke, M.D. (1983). Error Estimates for the Modified Newton Method with Applications to the Solution of Nonlinear, Two-Point Boundary-Value Problems, Journal of Optimization Theory and Applications, **39**, 489.

[28] Batten, J. J., and Murdie, D.C. (1970). The Thermal Decomposition of RDX at Temeratures Below the Melting Point, Australian Journal of Chemistry, **23**, 737.

[29] Batten, J.J. (1971). The Thermal Decomposition of RDX at Temperatures Below the Melting Point : Elucidation of the Mechanism, Australian Journal of Chemistry, **24**, 945.

[30] Rogers, R. N. and Smith, L.C. (1967). Estimation of Preexponential Factor from Thermal Decomposition Curve of an Unweighed Sample, Analytical Chemistry, **39**, 1024.

[31] Zenin, A. A. (1966). Structure of Temperature Distribution in Steady State Burning of a Ballistic Powder, Combustion, Explosion and Shock Waves, **2**, 67.

[32] Zenin, A. A. and Novozhilov, B. V. (1973). Single Value Dependence of the Surface Temperature of Ballistite on the Burning Rate, Combustion Explosion and Shock Waves, **9**, 246.

[33] Fifer, R. A. (1984). Chemistry of Nitrate Ester and Nitramine Propellants, Kuo, K. K. and Summerfield, M. (Ed.) Progress in Astronautics and Aeronautics : Fundamentals of Solid-Propellant Combustion, AIAA, **90**, 177.

[34] Korobeinichev, O.P., Kuibida, L.V., Orlov, V.N., Tereschenko, A.G., Kutsenogii, K.P., Mavliev, R.A., Ermolin, N.E., Fomin, V.M., and Emel'yanov, I.D. (1985). Mass Spectrometric Probe Study of the Flame Structure and Kinetics of Chemical Reactions in Flames, Mass-Spektrom. Khim. Kinet., 73.

[35] Cosgrove, J.D. and Owen, A.J. (1974). The Thermal Decomposition of 1, 3, 5 Trinitrohexahydro-1, 3, 5 triazine (RDX) Part I: The Products and Physical Parameters, Combustion and Flame, **22**, 13.

[36] Cosgrove, J.D. and Owen, A.J. (1974). The Thermal Decomposition of 1, 3, 5 Trinitrohexahydro-1, 3, 5 triazine (RDX) Part I: The Effects of the Products, Combustion and Flame, **22**, 19.

[37] Oyumi, Y. and Brill, T.B. (1985). Thermal Decomposition of Energetic Materials 3. A High-Rate, In- Situ, FTIR Study of the Thermolysis of RDX and HMX with Pressure and Heating Rate as Variables, Combustion and Flame, **62**, 213.

[38] Brill, T.B., Brush, P.J., Patil, D.G., and Chen, J.K. (1992). Chemical Pathways at the Burning Surface, Proceedings of the Twenty-Fourth Symposium (International) on Combustion, The Combustion Institute.

[39] Brill, T.B. and Brush, P.J. (1992). Condensed Phase Chemistry of Explosives and Propellants at High Temperature: HMX, RDX, and BAMO, Phil. Trans. Royal Soc. of London A, **339**, 377.

[40] Brill, T.B. (1992). Connecting the Chemical Composition of a Material to its Combustion Characteristics, Prog. Energy Combustion Sci., **18**, 91.

[41] Brill, T.B. (1995). Multiphase Chemistry Considerations at the Surface of Burning Nitramine Monopropellants," Journal of Propulsion and Power, submitted.

[42] Lin, C.Y., Wang, H.T.,Lin, M.C. and Melius, C.F. (1990). A Shock Tube Study of the $CH_2O + NO_2$ Reaction at High Temperatures, International Journal of Chemical Kinetics, **22**, 455.

[43] Behrens, Jr., R. and Bulusu, S. (1992). Thermal Decomposition of Energetic Materials 3. Temporal Behaviors of the Rates of Formation of the Gaseous Pyrolysis Products of Condensed Phase Decomposition of 1,3,5-Trinitrohexahydro-s-triazine, J. Phys. Chem., **96**, 8877.

[44] Behrens, Jr., R. and Bulusu, S. (1992). Thermal Decomposition of Energetic Materials 4. Deuteriom Isotope Effects and Isotopic Scrambling (H/D, 13C/18O, 14N/15N) in Condensed-Phase Decomposition of 1,3,5-Trinitrohexahydro-s-triazine, J. Phys. Chem., **96**, 8891.

[45] Behrens, Jr., R., Land, T.A., and Bulusu, S. (1993). Thermal Decomposition Mechanisms of 1- Nitroso-3,5-Dinitro-s-Triazine (ONDNTA), Proceedings of the Thirteeth JANNAF Combustion Meeting.

[46] Thynell, S.T., Gongwer, P.E. and Brill, T.B. (1995). Modelling of Thermal Response of Filament used in T-Jump Experiment, submitted.

[47] Tsang, W. and Herron, J.T. (1991). Chemical Kinetic Data Base for Propellant Combustion: I Reactions involving NO, NO2, HNO, HONO, HCN and N2O, J. Phys. Chem. Ref. Data, **20**, 609.

[48] Tsang, W. (1992). Chemical Kinetic Data Base for Propellant Combustion II. Reactions Involving CN, NCO, and HNCO, J.Phys. Chem. Ref. Data, **21**, 753.

[49] Thaxton, A.G., Diau, E.X.G., Lin, M.C., Lin, C-Y, and Melius, C.F. (1994). Thermal Oxidation of HCN by NO2 at High Temperatures, International Journal of Chemical Kinetics, submitted.

[50] Zhao, X., Hintsa, E.J., and Lee, Y.T. (1988). Infrared Multiphoton Dissociation of RDX in a Molecular Beam, J. Chem. Phys., **88**, 801.

[51] Melius, C.F. (1987). Theoretical Studies of the Chemical Reactions Involved in the Ignition of Nitramines, Proceedings of the Twenty-Fourth JANNAF Combustion Meeting.

[52] Melius, C.F. (1990). Thermochemical Modeling: I. Application to Decomposition of Energetic Materials, In Bulusu, S. N. (Ed.) Chemistry and Physics of Energetic Materials, Kluwer Academic Publishers, The Netherlands, 21.

[53] Hanson-Parr, D. and Parr, T. (1993). RDX Flame Structure and Chemistry, Proceedings of the Thirteeth JANNAF Combustion Meeting.

[54] Ermolin, N.E., Korobeinichev, O.P., Kuibida, L.V., and Fomin, V.M. (1986). Study of the Kinetics and Mechanism of Chemical Reactions in Hexogen Flames, Fizika Goreniya i Vzryva, **22**, 54.

[55] Ermolin, N.E., Korobeinichev, O.P., Kuibida, L.V., and Fomin, V.M. (1988). Processes in Hexogene Flames, Fizika Goreniya i Vzryva, **24**, 21.

[56] Parr D. H. and Parr, T. (1995). Private Communication.

[57] A.Y. Chang and R.K. Hanson (1989). J.Quant. Spectrosc. Radiat. Transfer, **42**, 207.

[58] Atwood, A., Price, C., Curran, P., and Wiknich, J. (1995). Laboratory Ignition and Combustion Data and the Determination of Global Kinetics and Energetic Parameters of RDX, Proceedings of the Thirty-Second JANNAF Combustion Meeting, submitted.

[59] Zimmer-Galler, R. (1968). Correlations between Deflagration Characteristics and Surface Properties of Nitramine-Based Propellants, AIAA Journal, **6**, 2107.

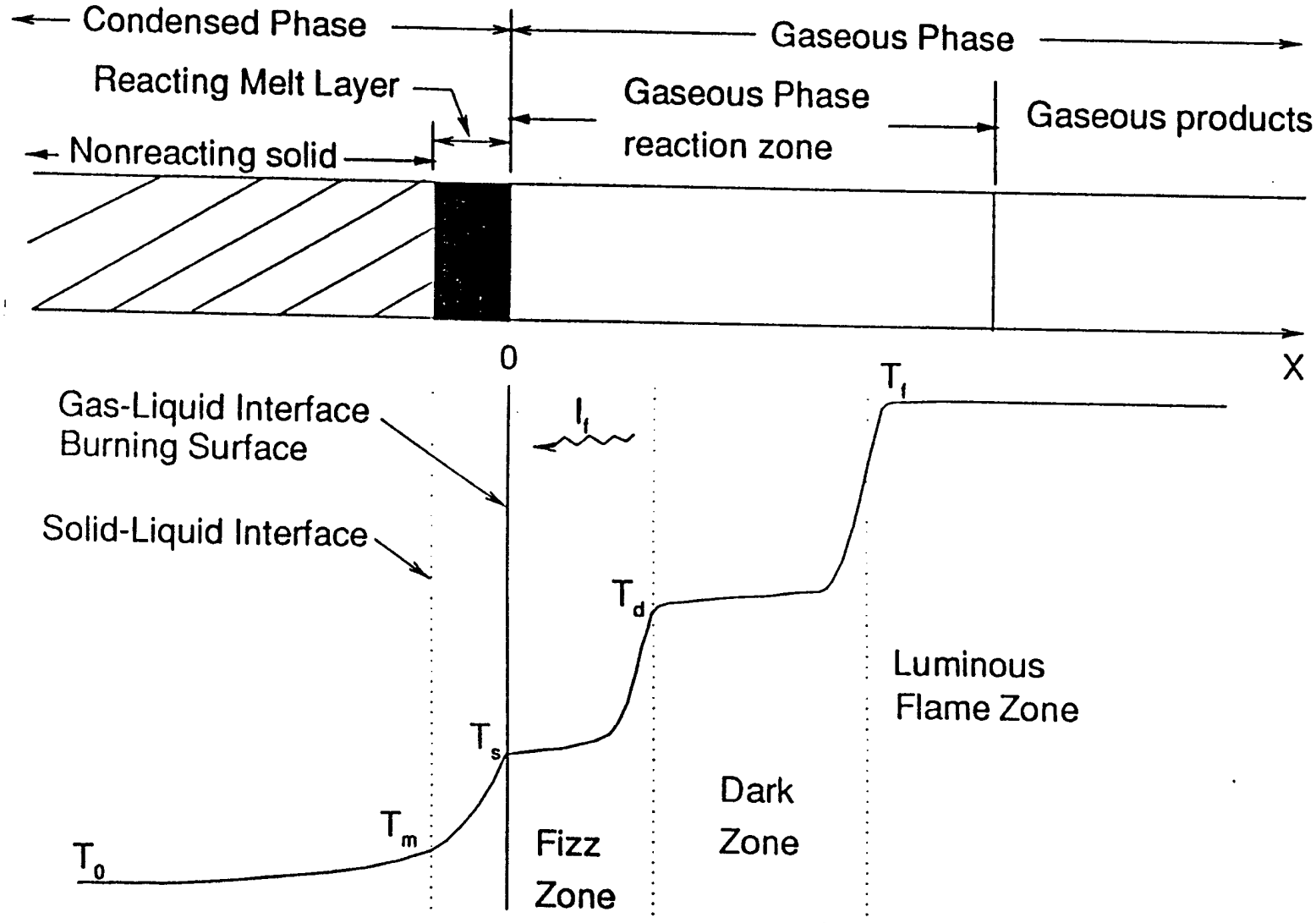
[60] Glaskova, A.P. (1974). The Effect of Catalysts on the Combustion of Explosives, Fizika Goreniya i Vzryva, **10**, 323.

[61] Glaskova, A.P. (1974). The Effect of Catalysts on the Combustion of Explosives, Fizika Goreniya i Vzryva, **10**, 323.

[62] Price, C.F., Boggs, T.L. and Derr, R.L. (1979). The steady state combustion behavior of ammonium perchlorate and HMX, AIAA Paper 79-0164.

List of Figures

List of Tables

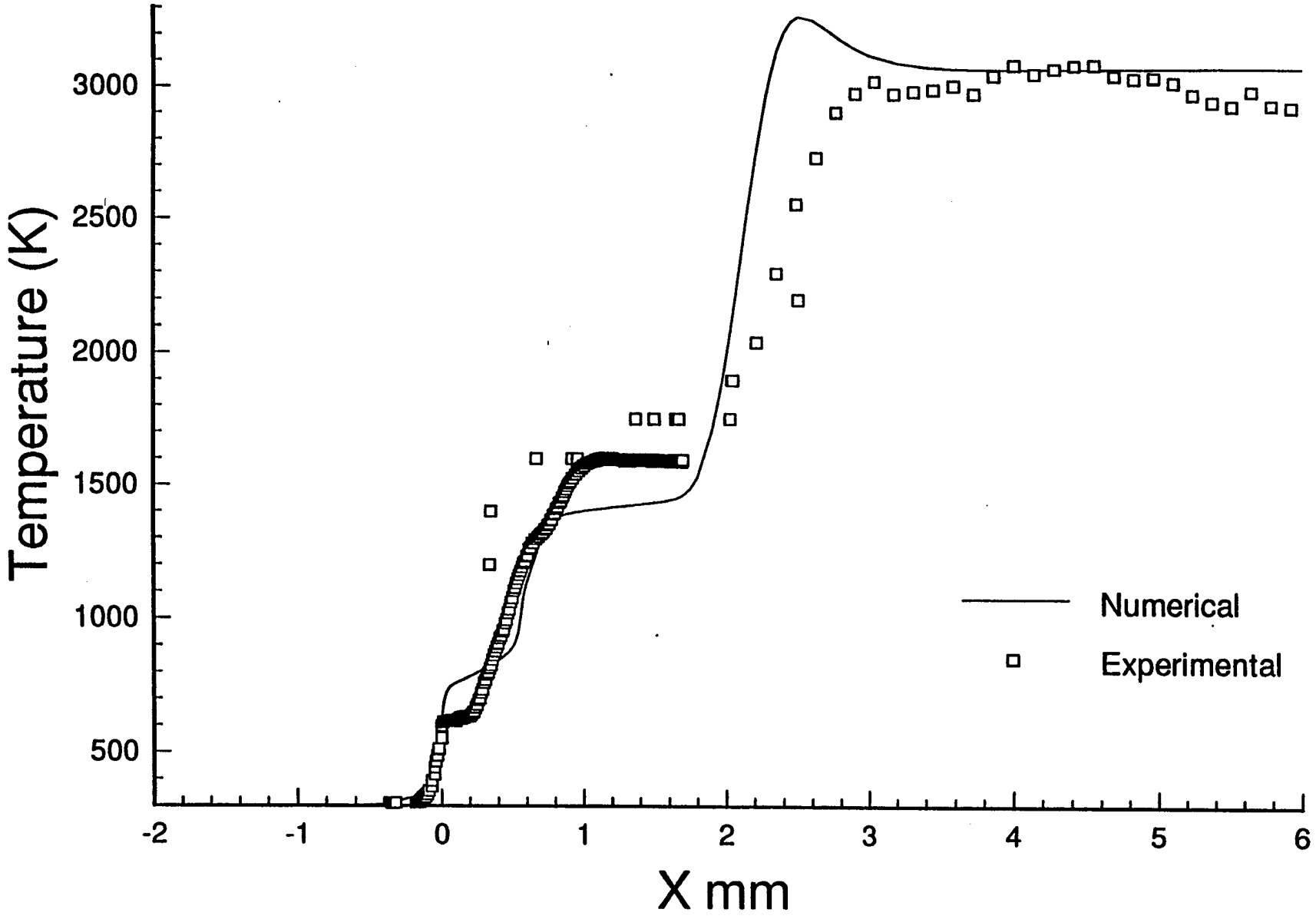


Combustion of RDX monopropellant

Schematic description of the flame (combustion wave) structure

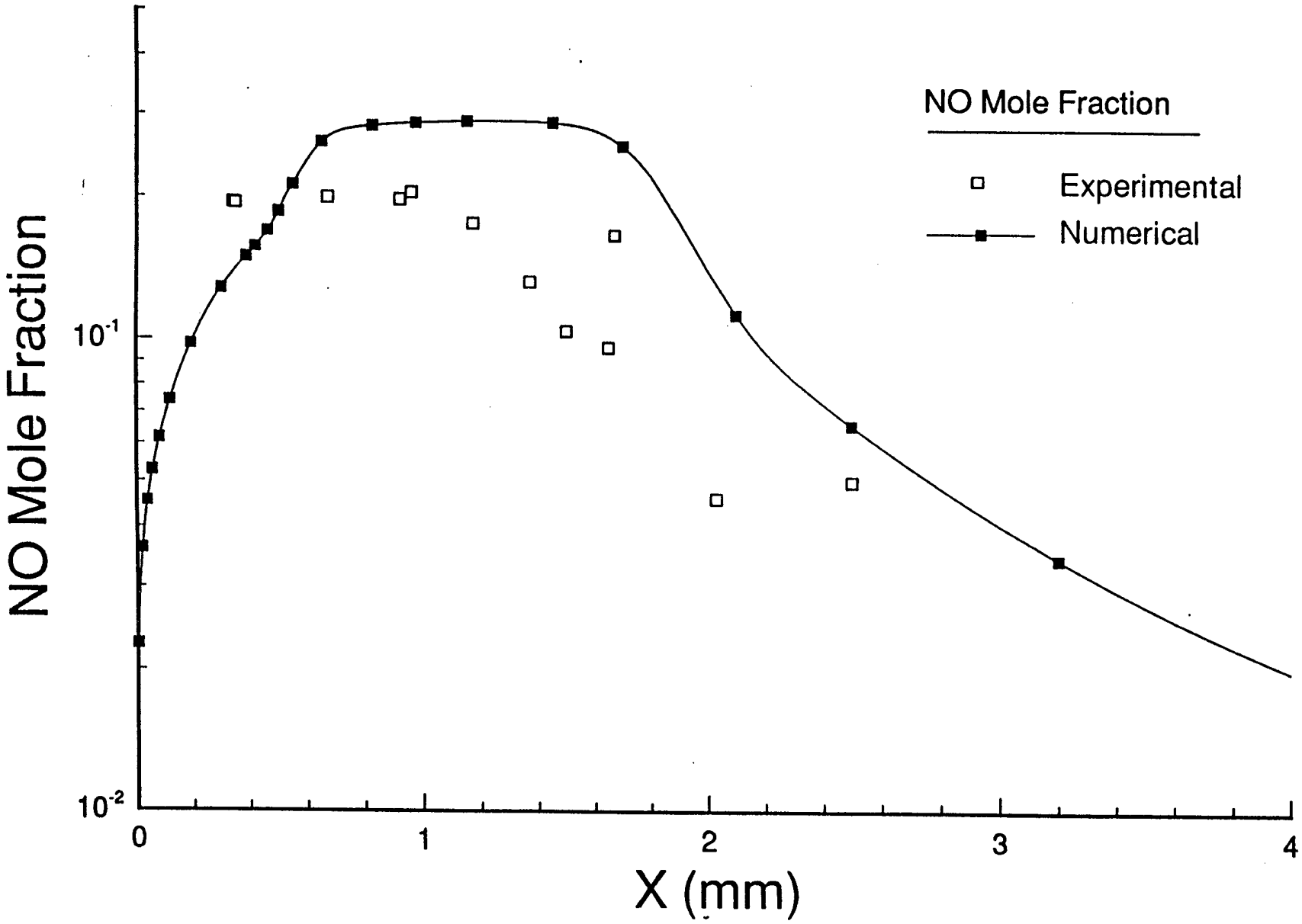
Laser Assisted Deflagration of RDX Propellant

Comparison of Numerical and Experimental Results : Pressure = 1 atm



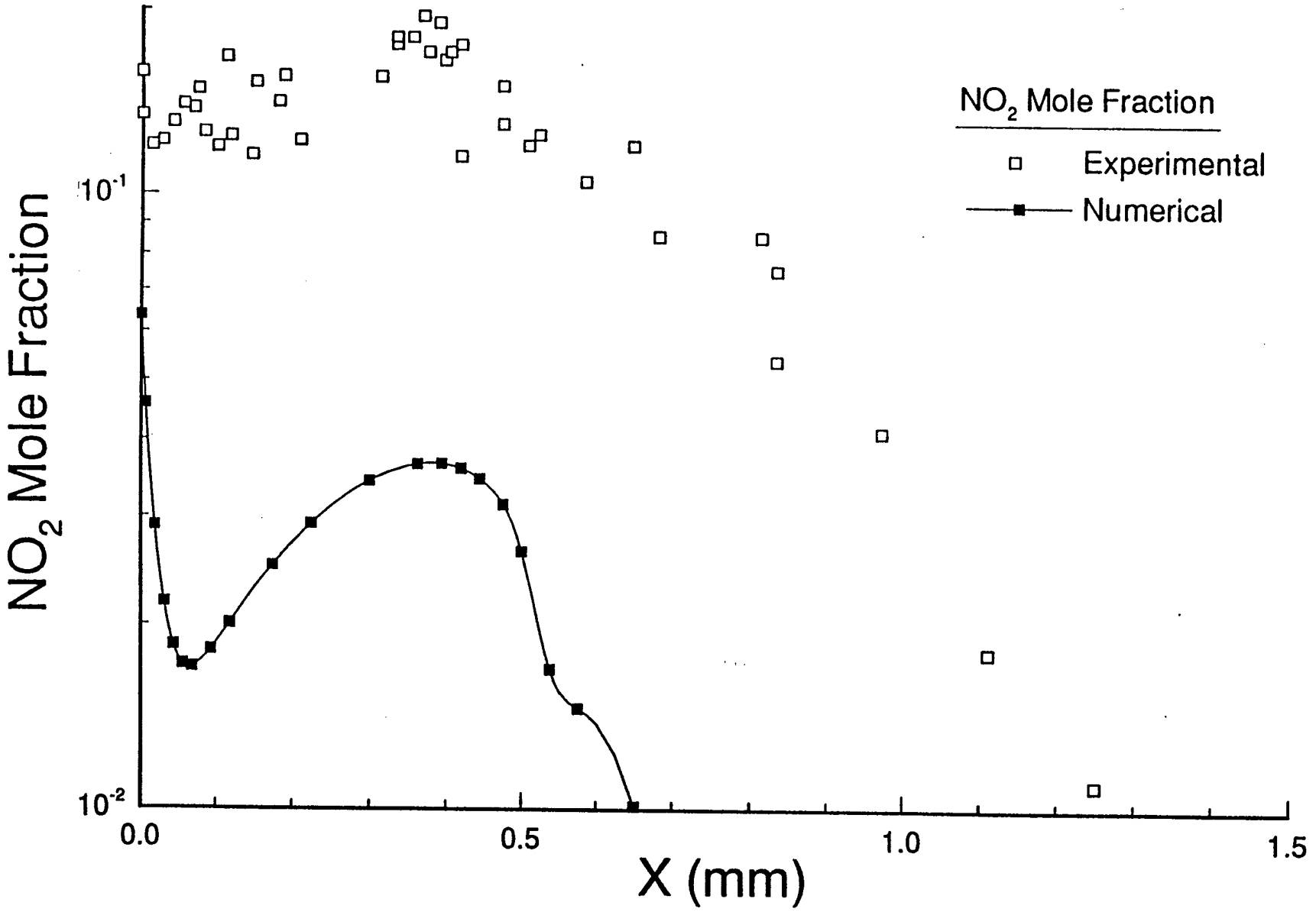
Laser Assisted Deflagration of RDX Propellant

Comparison of Numerical and Experimental Results : Pressure = 1 atm



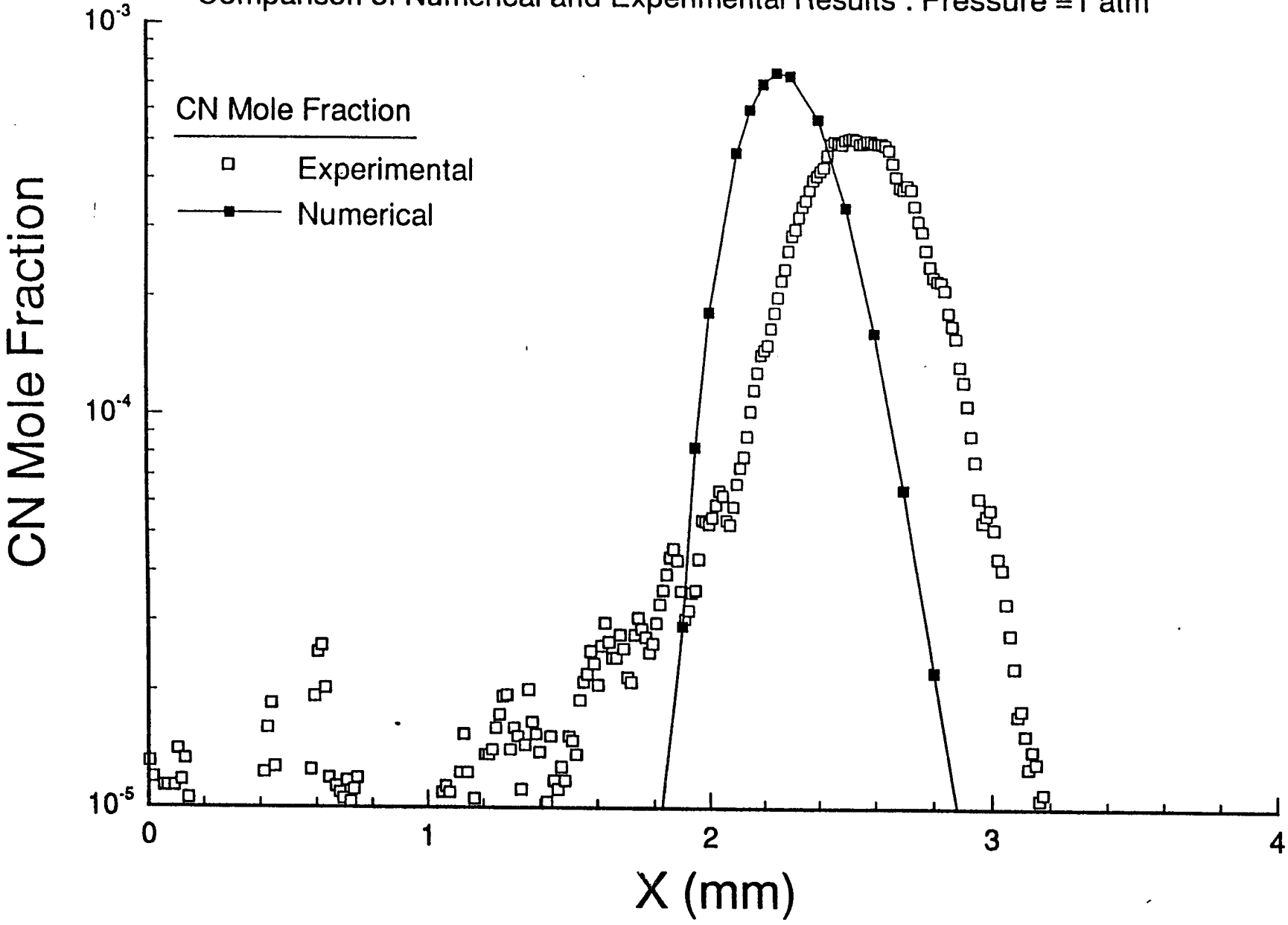
Laser Assisted Deflagration of RDX Propellant

Comparison of Numerical and Experimental Results : Pressure = 1 atm



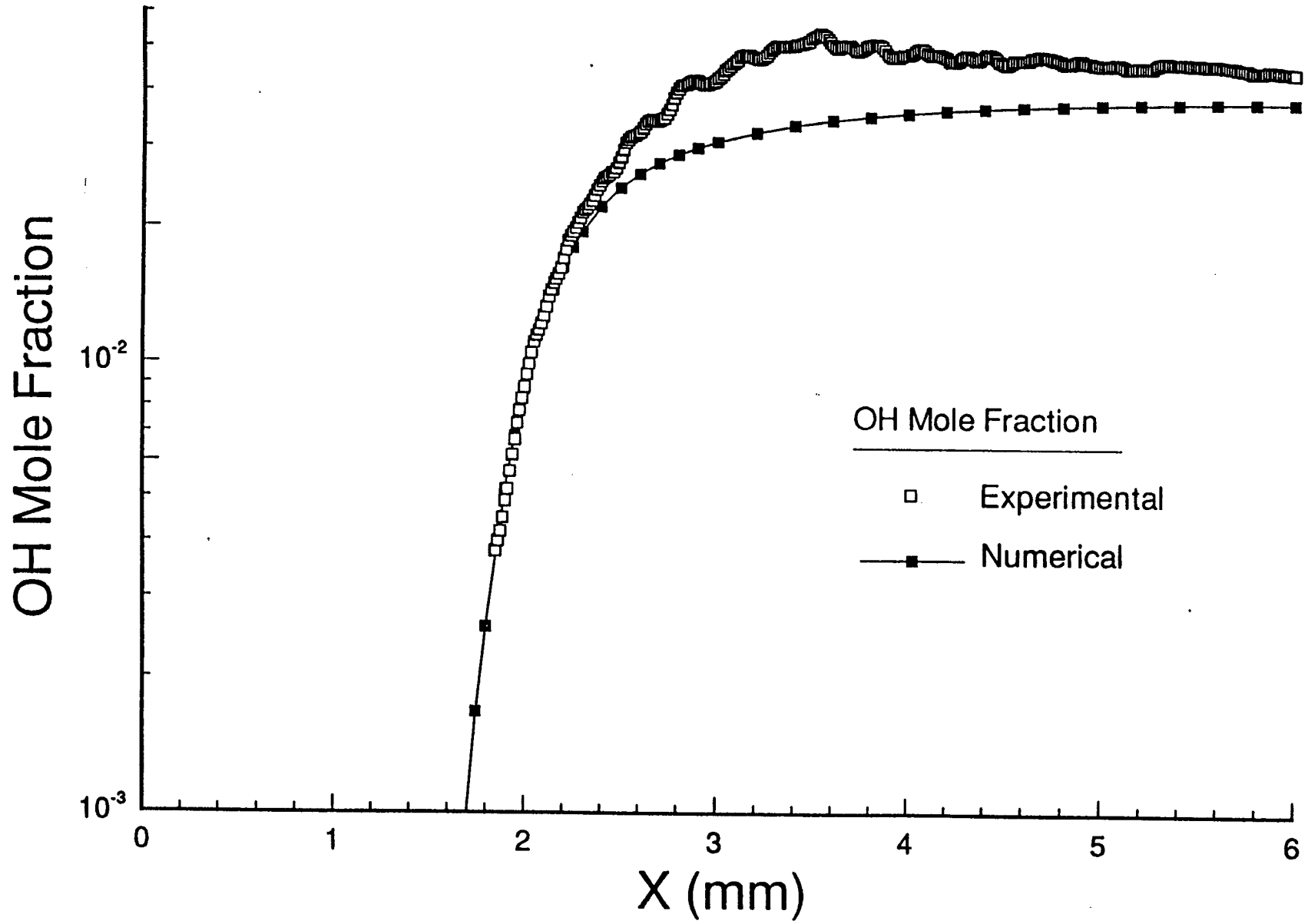
Laser Assisted Deflagration of RDX Propellant

Comparison of Numerical and Experimental Results : Pressure =1 atm



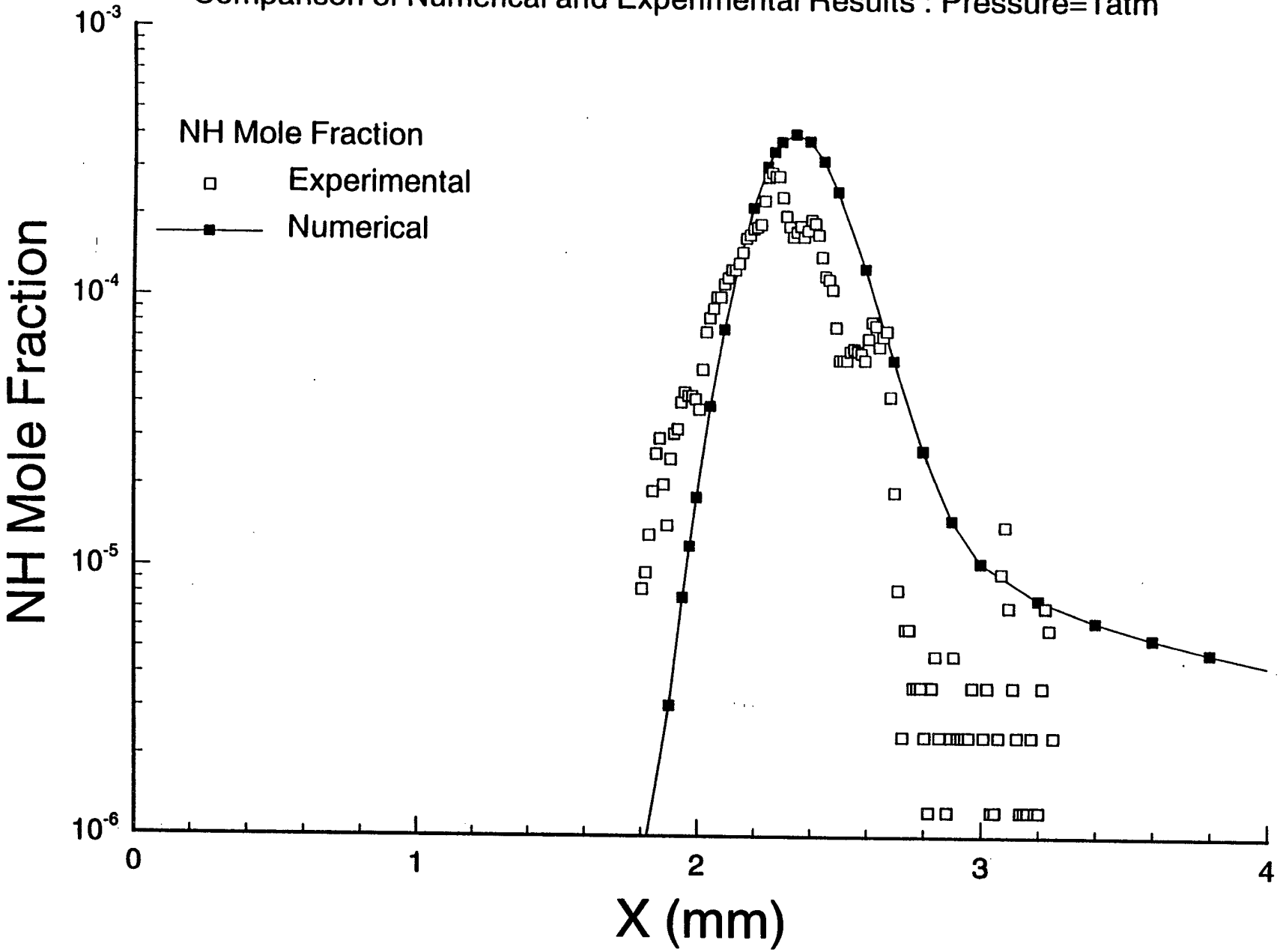
Laser Assisted Deflagration of RDX Propellant

Comparison of Numerical and Experimental Results : Pressure = 1 atm



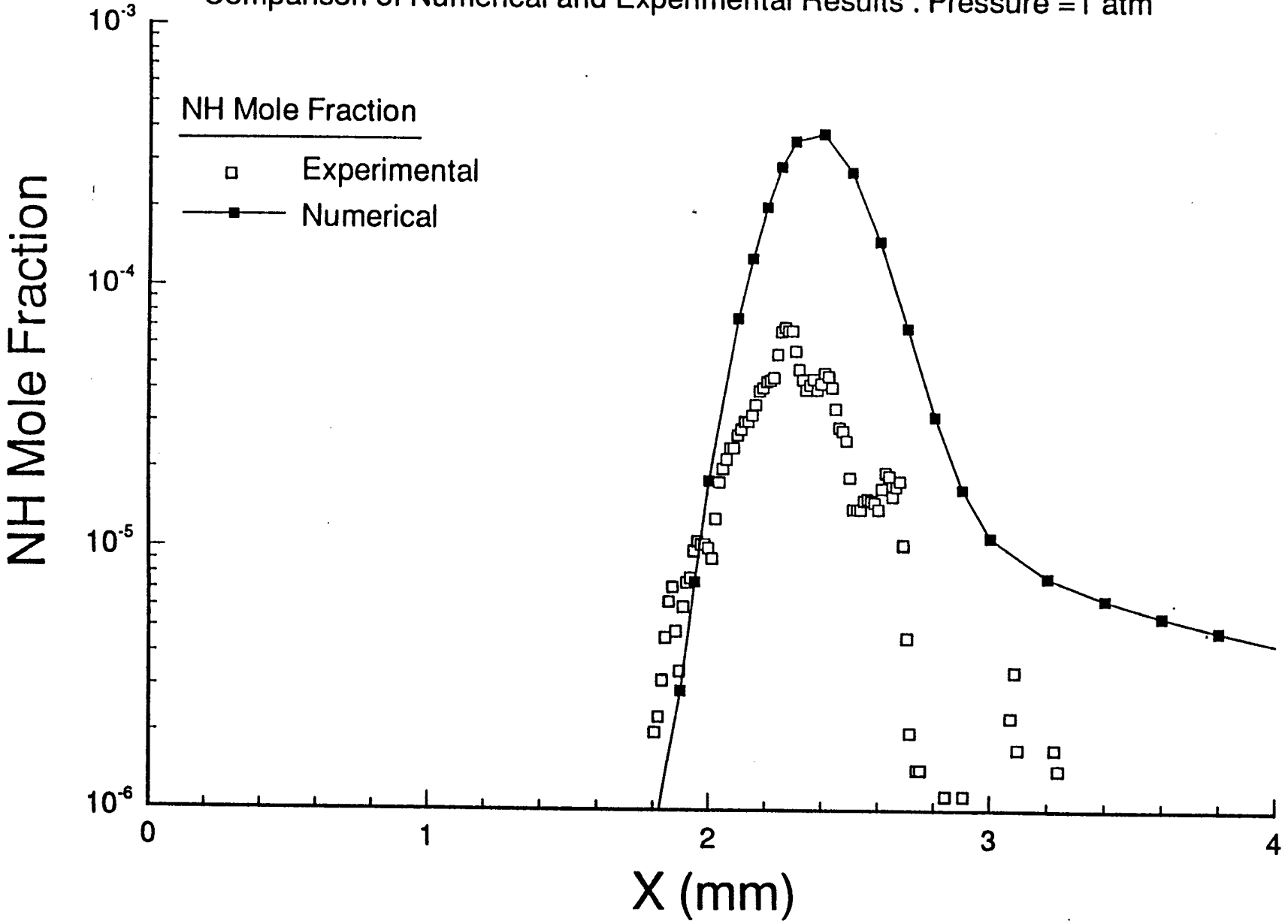
Laser Assisted Deflagration of RDX Propellant

Comparison of Numerical and Experimental Results : Pressure=1atm



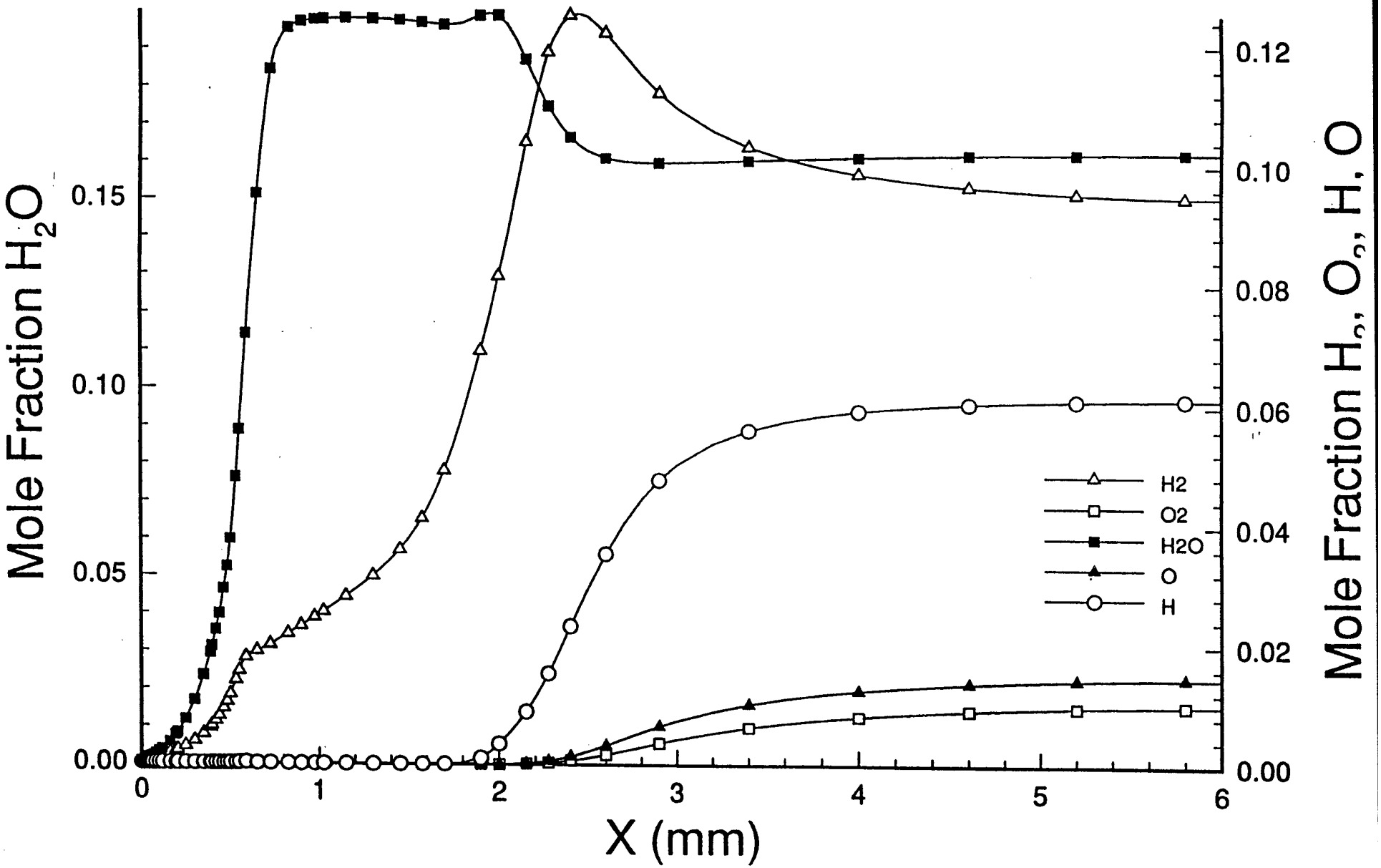
Laser Assisted Deflagration of RDX Propellant

Comparison of Numerical and Experimental Results : Pressure =1 atm

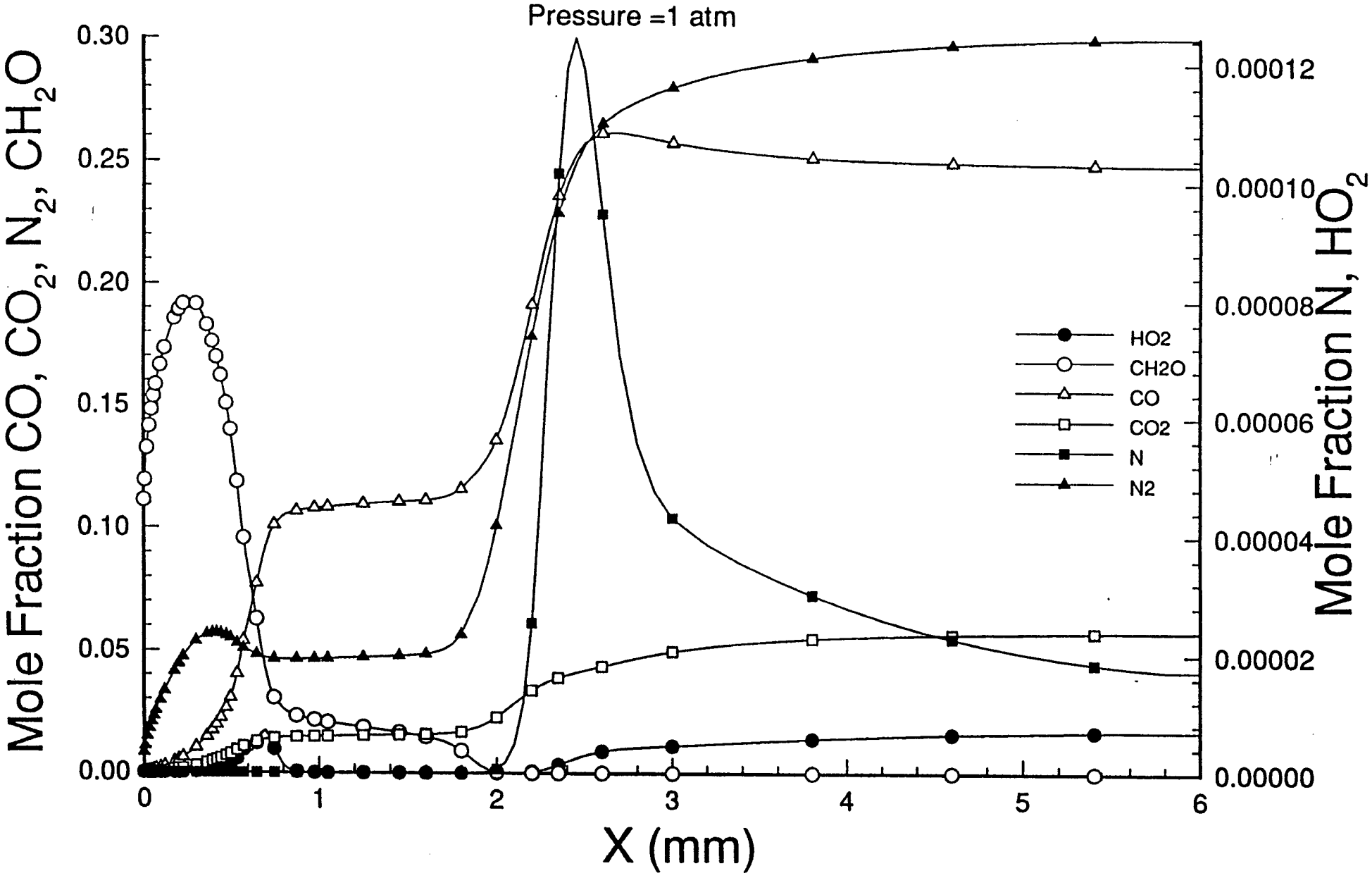


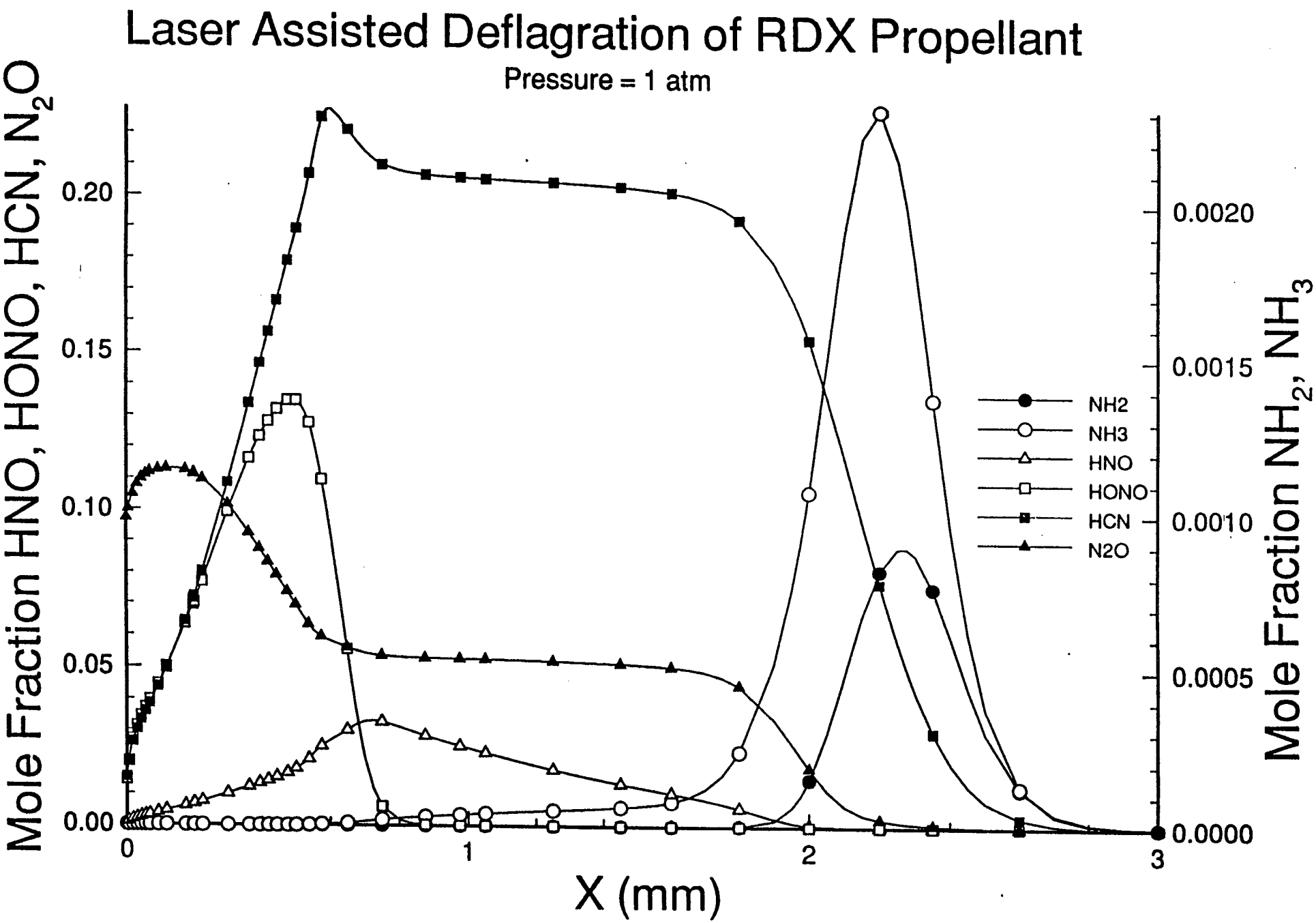
Laser Assisted Deflagration of RDX Propellant

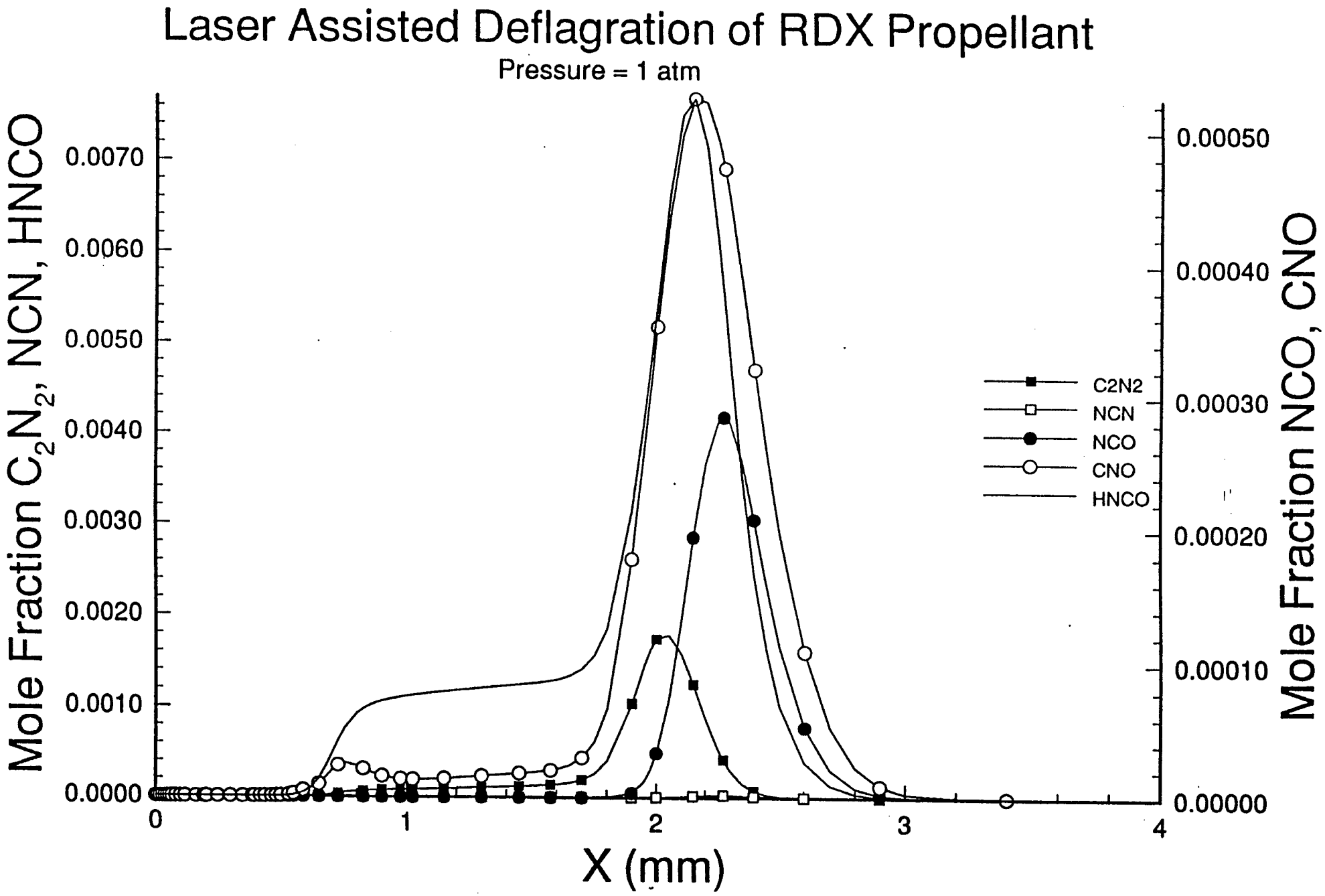
Pressure = 1 atm



Laser Assisted Deflagration of RDX Propellant

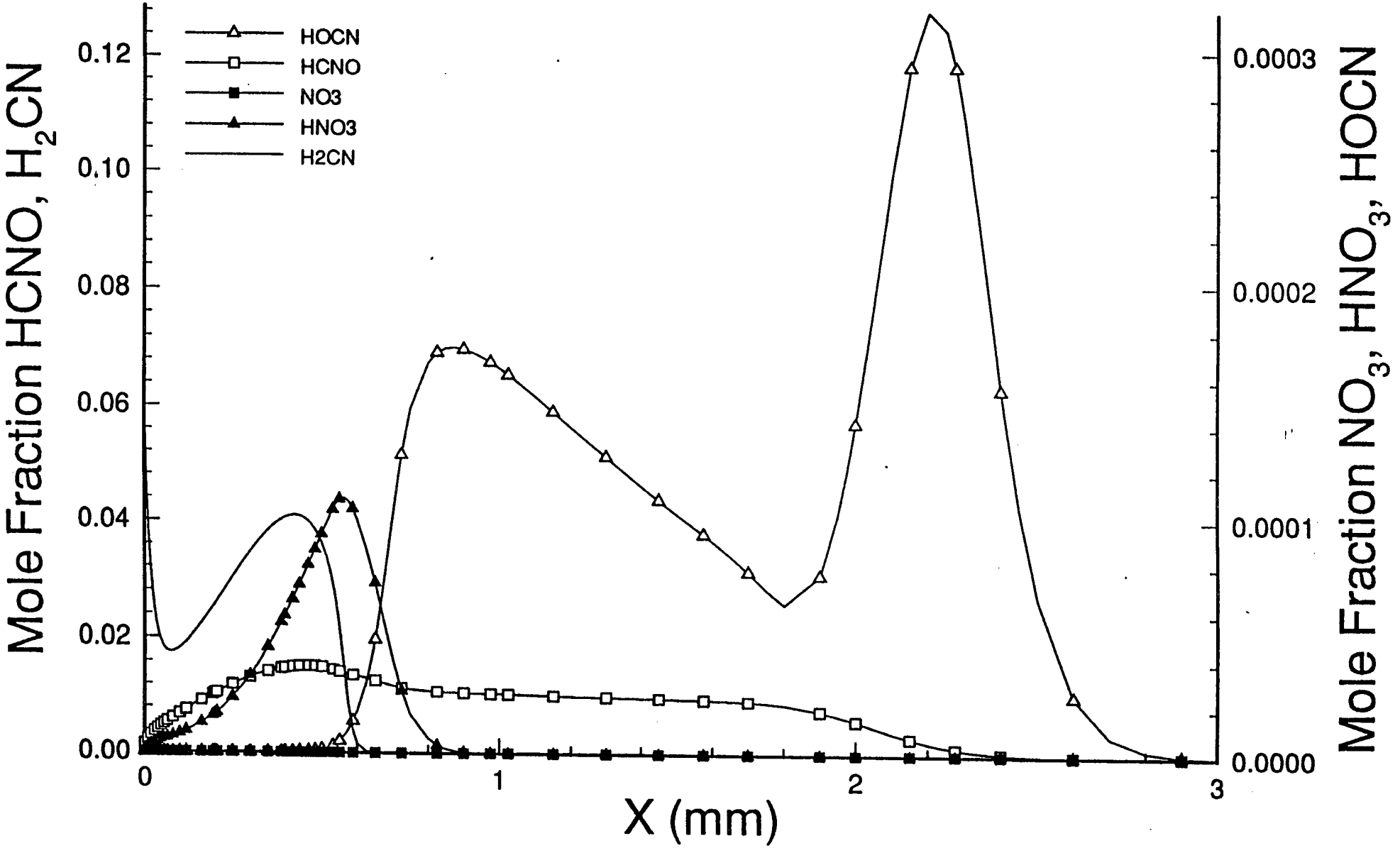


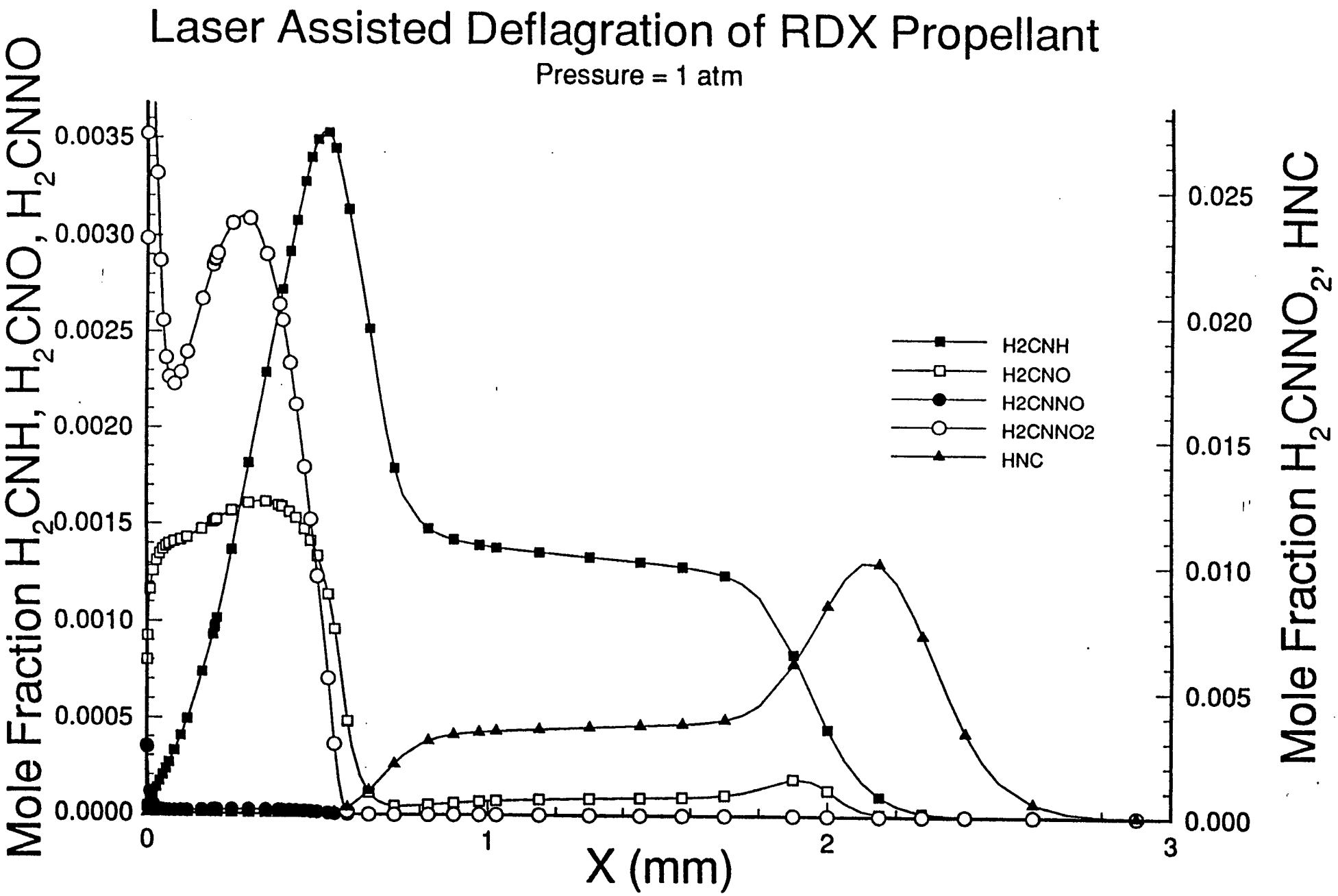

Laser Assisted Deflagration of RDX Propellant
Pressure = 1 atm
Mole Fraction HNO, HONO, HCN, N2O
Mole Fraction NH2, NH3
X (mm)
0.00
0.05
0.10
0.15
0.20
0.0000
0.0005
0.0010
0.0015
0.0020
0
1
2
3
NH2
NH3
HNO
HONO
HCN
N2O


Laser Assisted Deflagration of RDX Propellant
Pressure = 1 atm
Mole Fraction C_2N_2, NCN, HNCO
Mole Fraction NCO, CNO
X (mm)
0.0000
0.0010
0.0020
0.0030
0.0040
0.0050
0.0060
0.0070
0.00000
0.00010
0.00020
0.00030
0.00040
0.00050
0
1
2
3
4
C2N2
NCN
NCO
CNO
HNCO

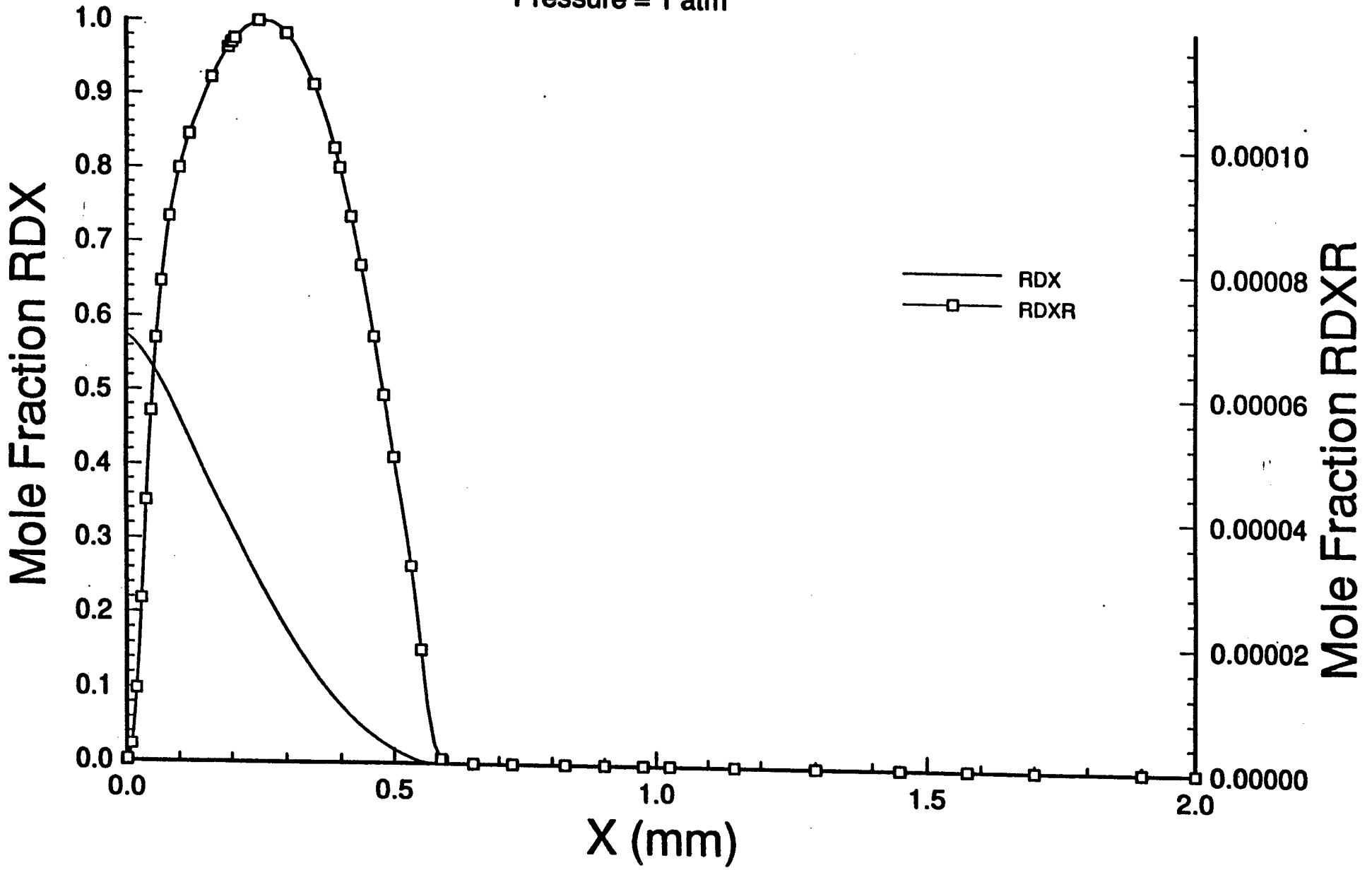
Laser Assisted Deflagration of RDX Propellant

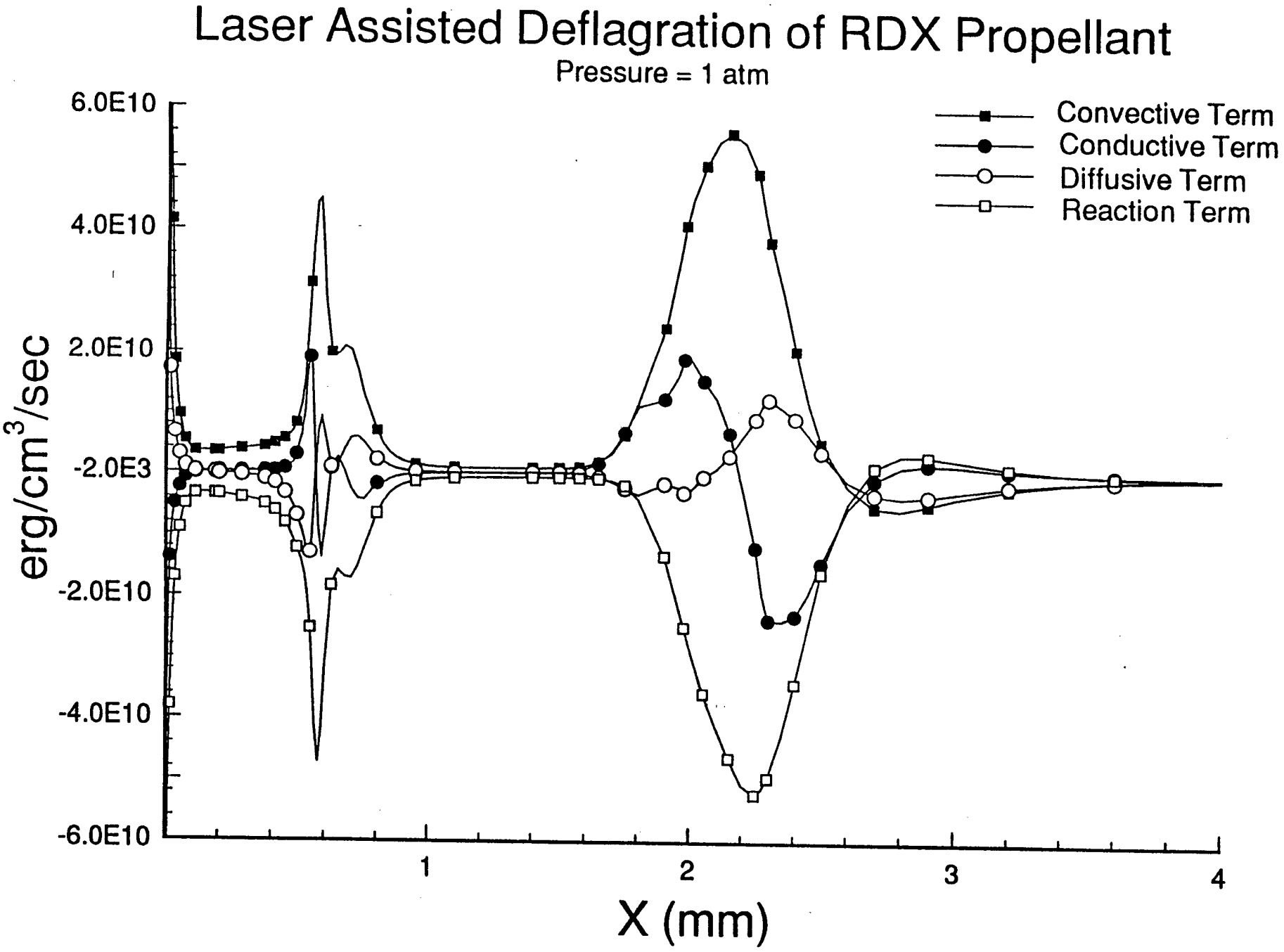
Pressure = 1 atm




Laser Assisted Deflagration of RDX Propellant
Pressure = 1 atm
Mole Fraction H2CNH, H2CNO, H2CNNO
Mole Fraction H2CNNO2, HNC
X (mm)
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0.025
0.020
0.015
0.010
0.005
0.000
0
1
2
3
H2CNH
H2CNO
H2CNNO
H2CNNO2
HNC

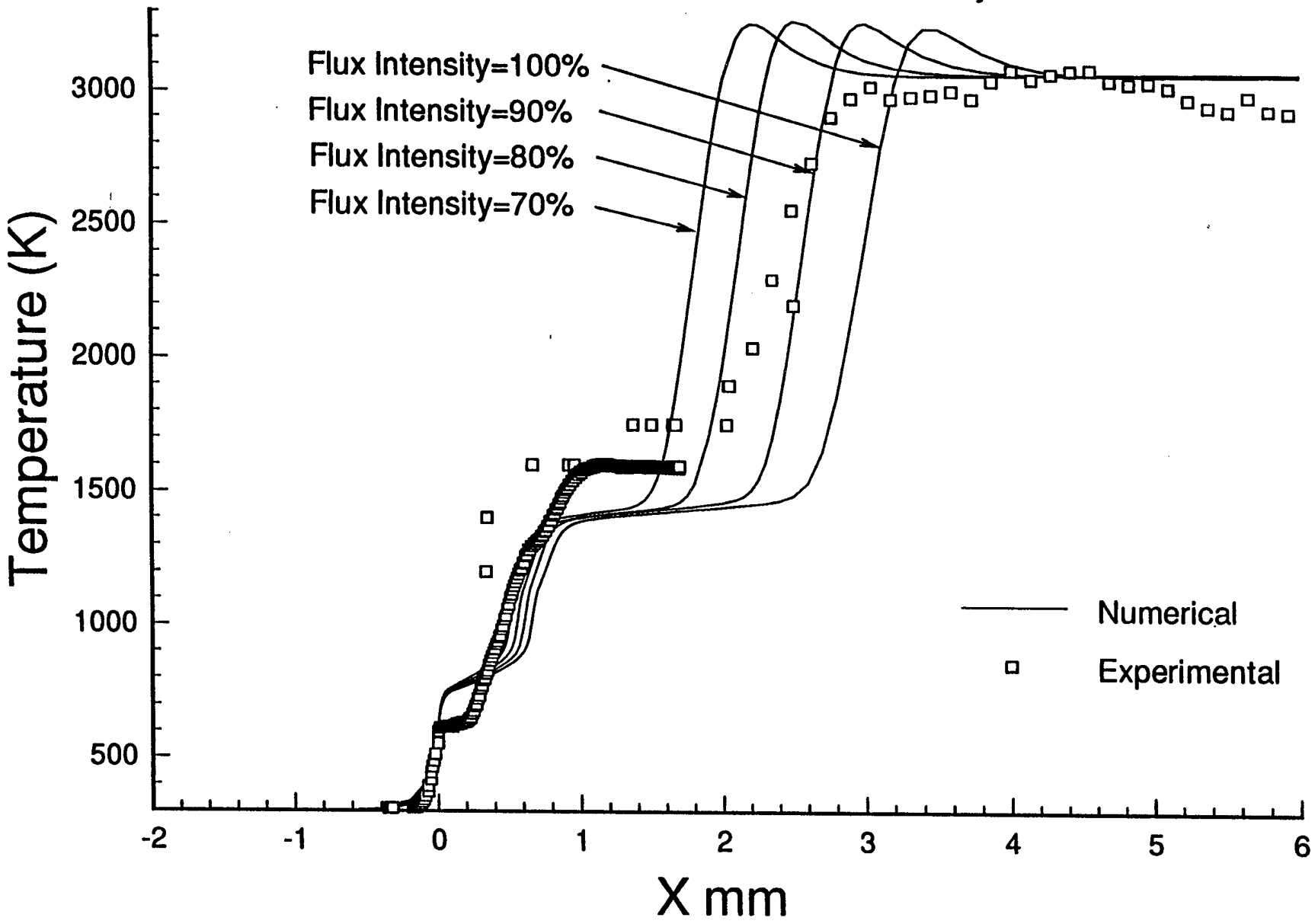
Laser Assisted Deflagration of RDX Propellant

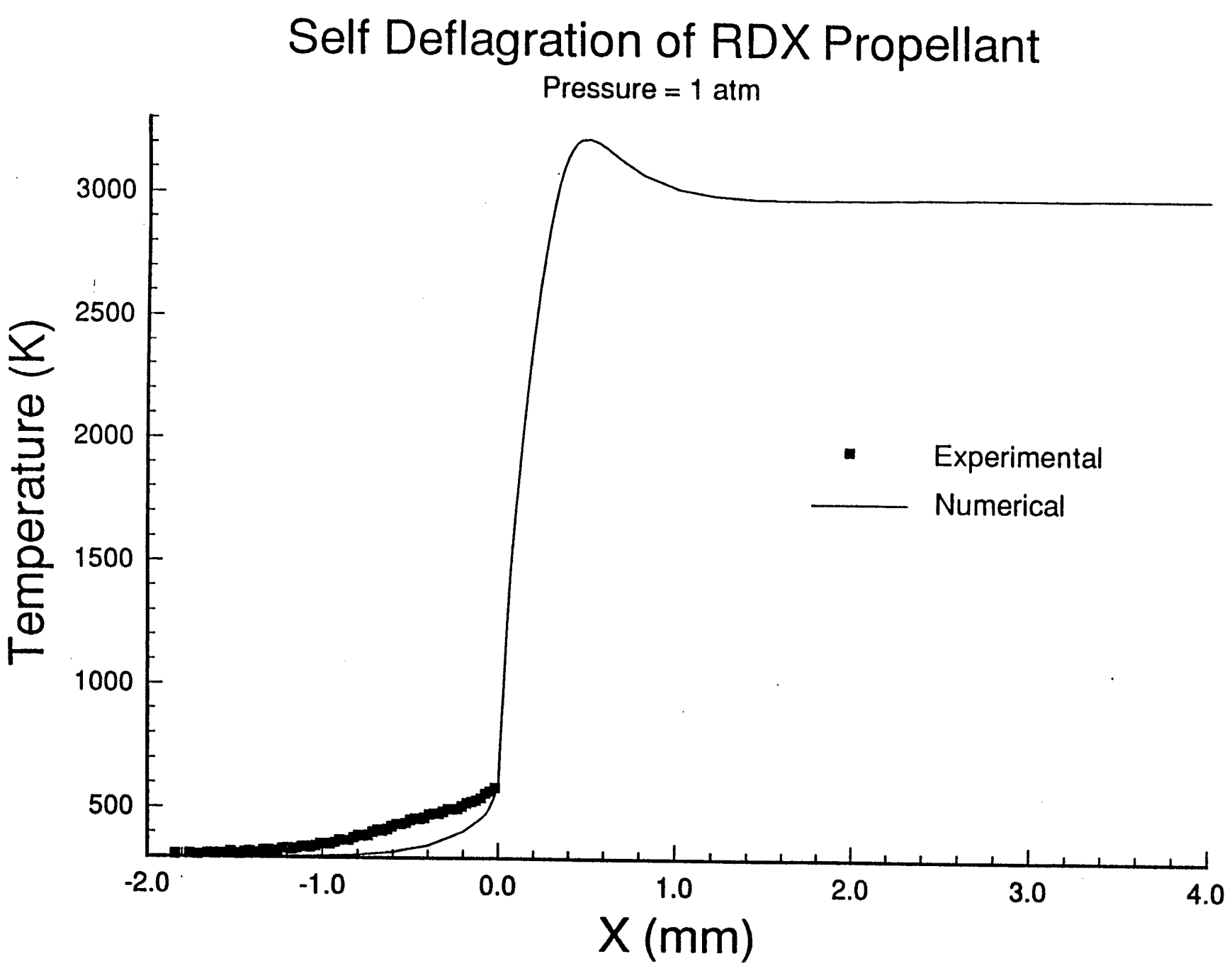


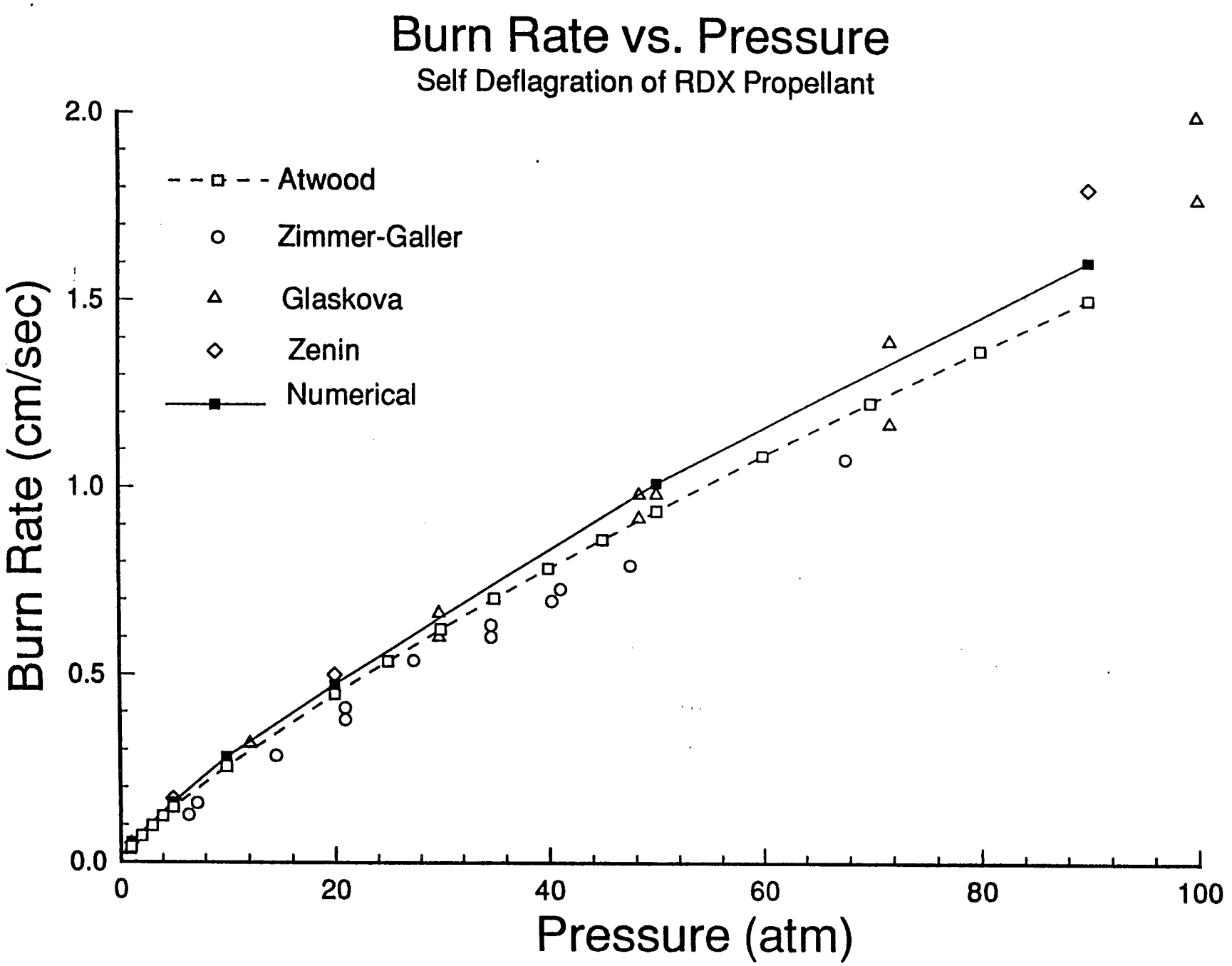

Laser Assisted Deflagration of RDX Propellant
Pressure = 1 atm
Convective Term
Conductive Term
Diffusive Term
Reaction Term
6.0E10
4.0E10
2.0E10
-2.0E3
-2.0E10
-4.0E10
-6.0E10
erg/cm3/sec
1
2
3
4
X (mm)

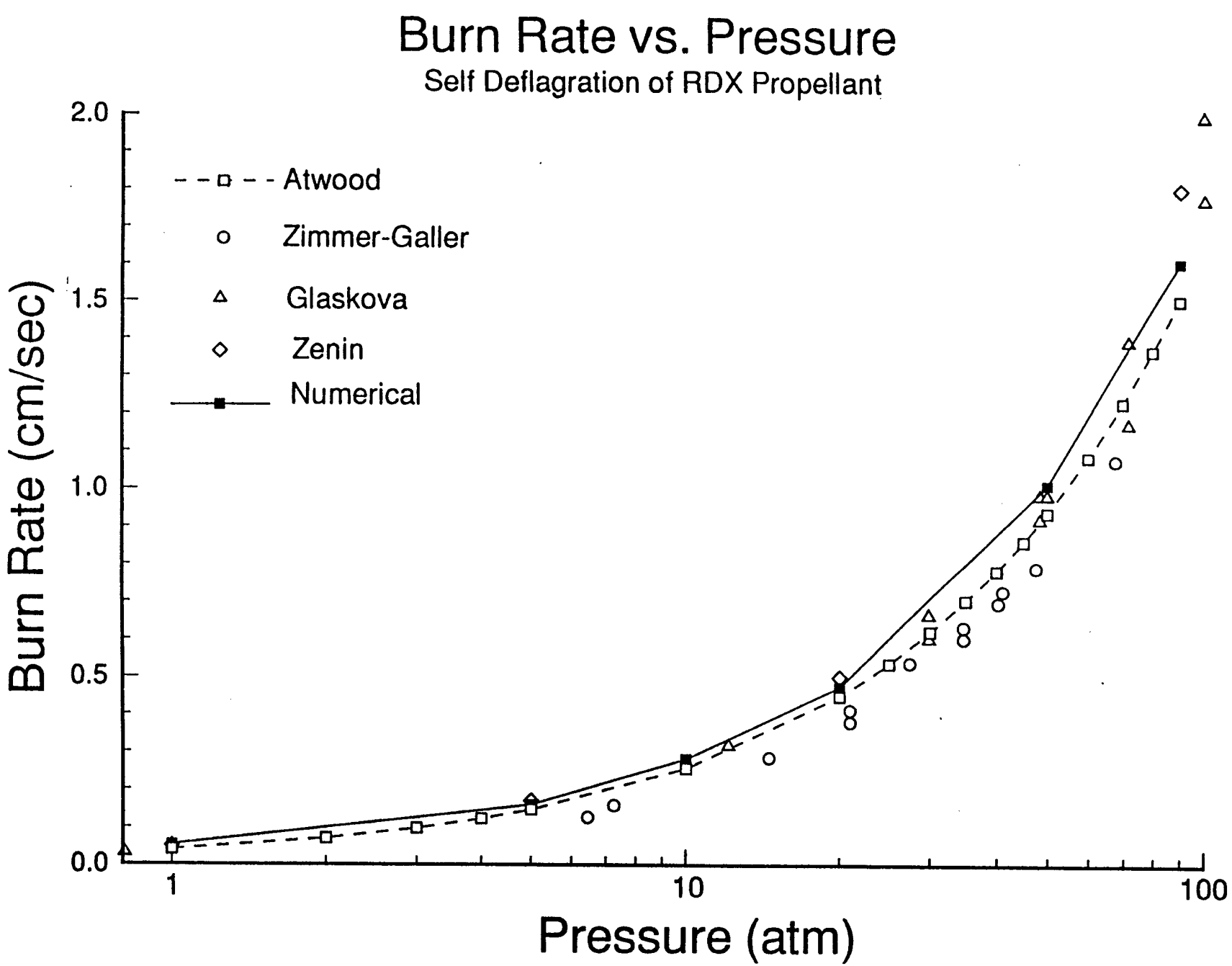
Laser Assisted Deflagration of RDX Propellant

Length of Dark Zone vs. Laser Flux Intensity

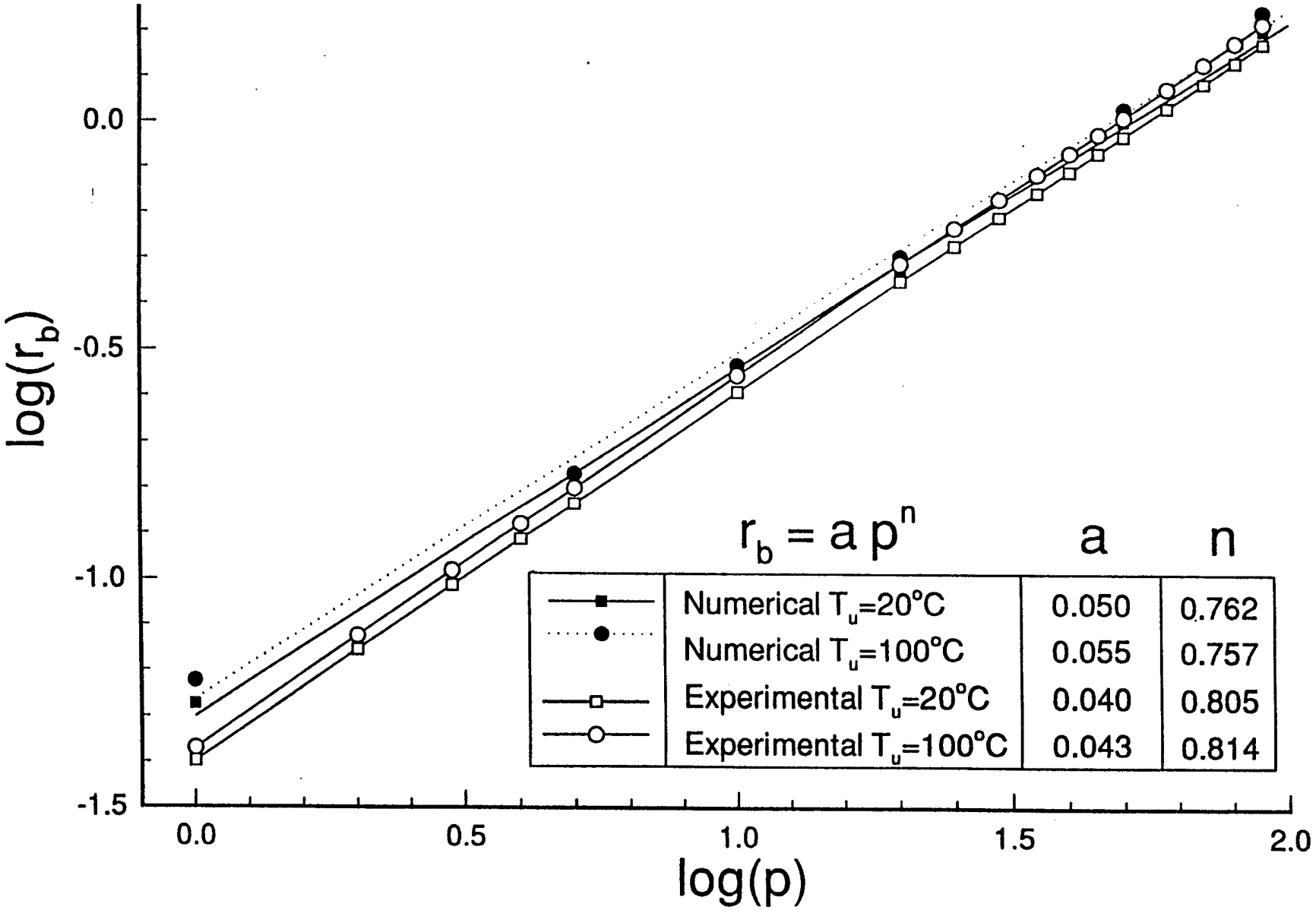



Self Deflagration of RDX Propellant
Pressure = 1 atm
Temperature (K)
3000
2500
2000
1500
1000
500
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
X (mm)
Experimental
Numerical


Burn Rate vs. Pressure
Self Deflagration of RDX Propellant
Atwood
Zimmer-Galler
Glaskova
Zenin
Numerical
Burn Rate (cm/sec)
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80
100
Pressure (atm)

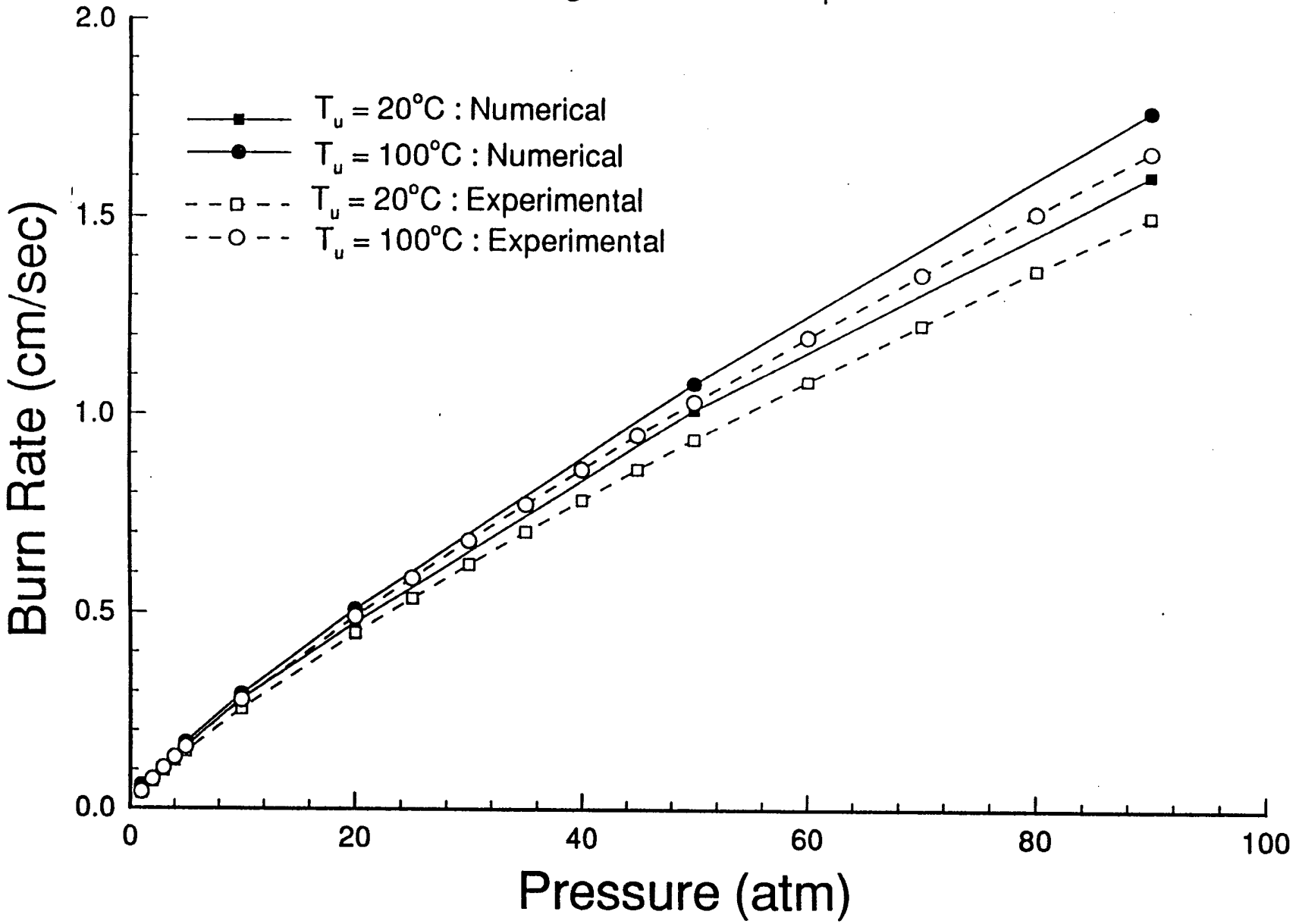

Burn Rate vs. Pressure
Self Deflagration of RDX Propellant
Atwood
Zimmer-Galler
Glaskova
Zenin
Numerical
Burn Rate (cm/sec)
2.0
1.5
1.0
0.5
0.0
1
10
100
Pressure (atm)

Pressure Sensitivity of RDX monopropellants



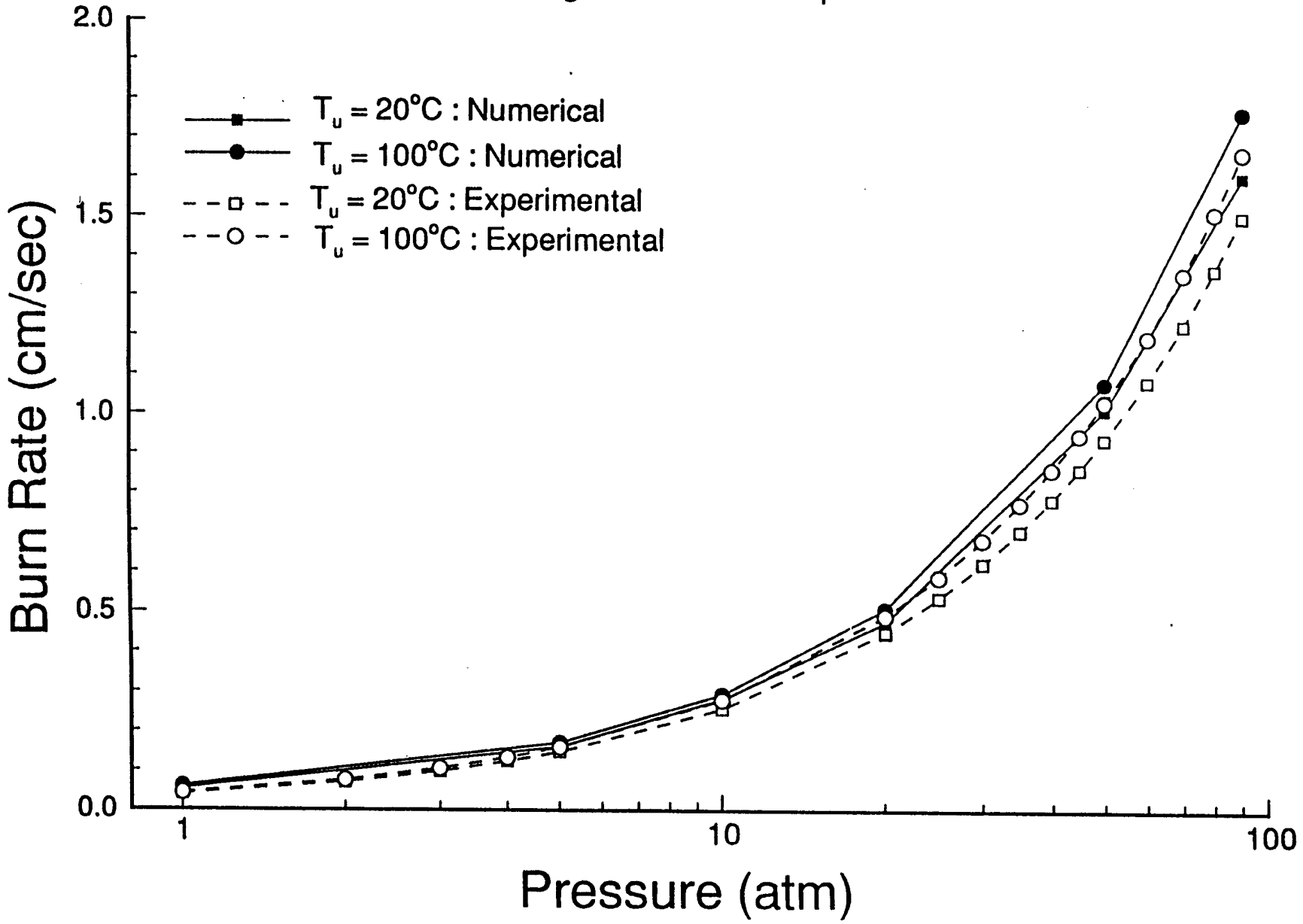
Sensitivity of Burn Rate to Initial Propellant Temperature

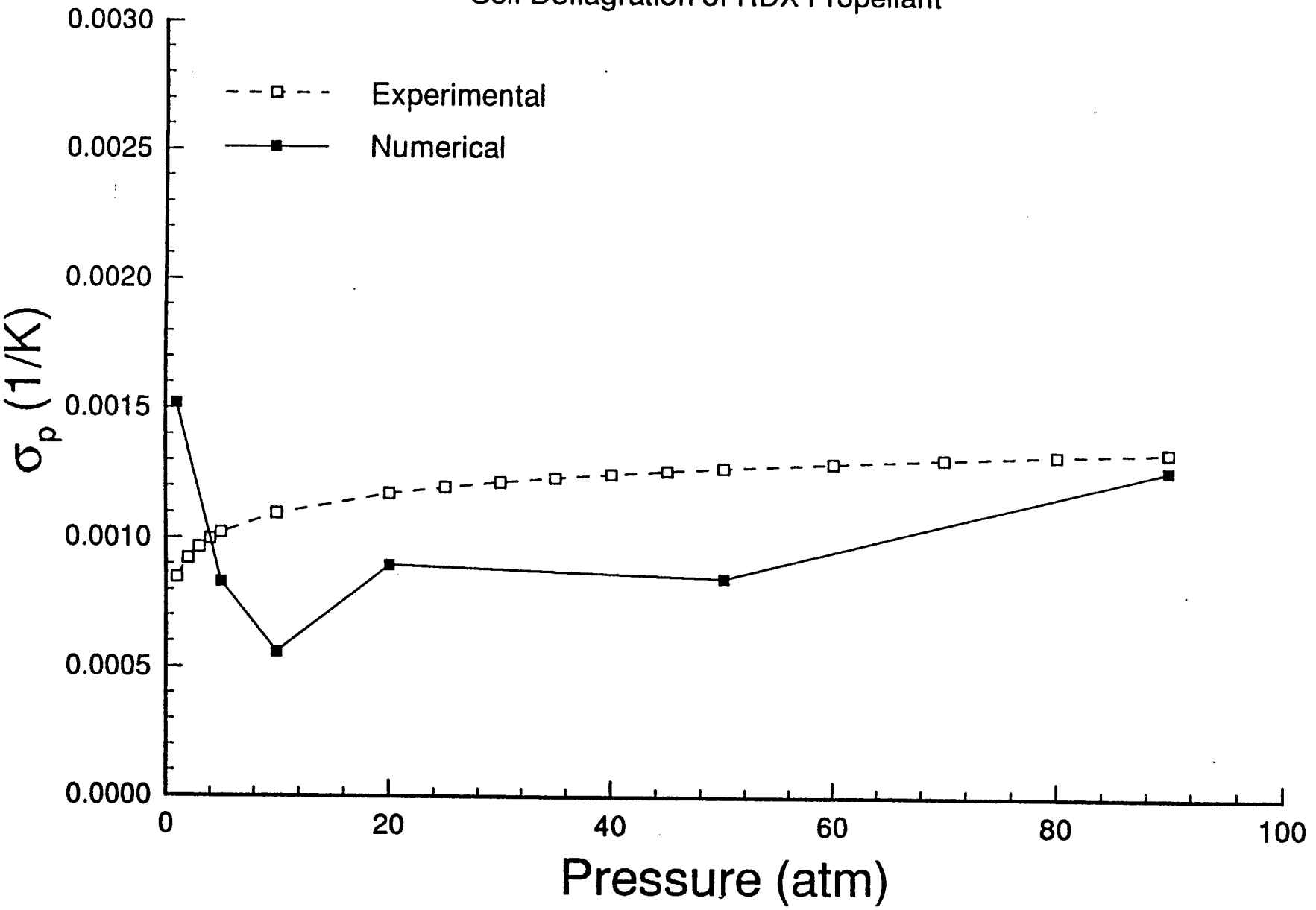
Self Deflagration of RDX Propellant



Sensitivity of Burn Rate to Initial Propellant Temperature

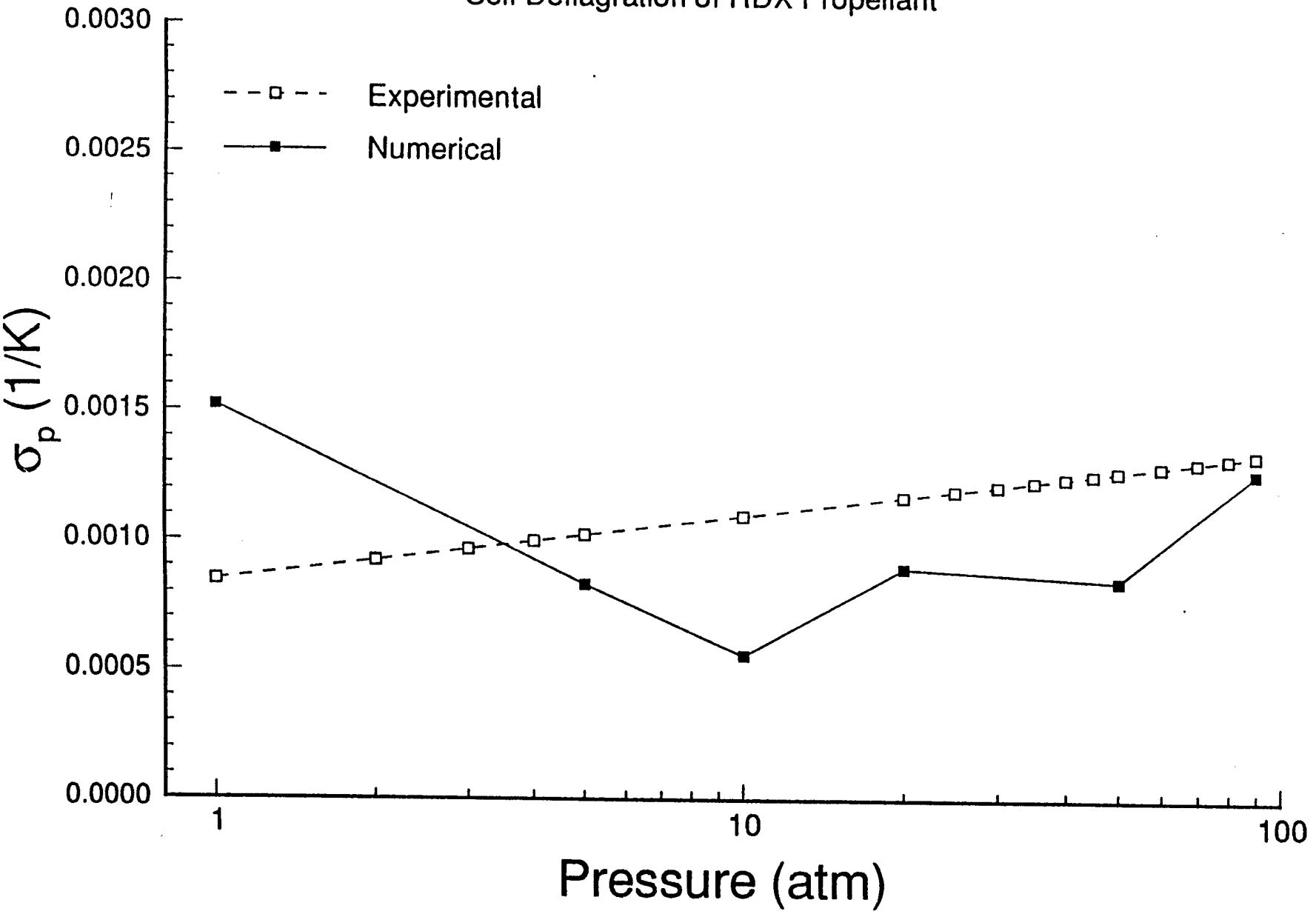
Self Deflagration of RDX Propellant




Sensitivity (σ_p) to Initial Propellant Temperature
Self Deflagration of RDX Propellant
Experimental
Numerical
σ_p (1/K)
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0
20
40
60
80
100
Pressure (atm)

Sensitivity (σ_p) to Initial Propellant Temperature

Self Deflagration of RDX Propellant



Appendix JJ

An Eigenvalue Method for Computing the Burning Rates of HMX Propellants

AN EIGENVALUE METHOD FOR COMPUTING THE BURNING RATES OF HMX PROPELLANTS

Kuldeep Prasad*
Science Applications International Corporation
Applied Physics Operation
1710 Goodridge Drive
McLean, VA 22102

Richard A. Yetter
Princeton University
Princeton, NJ 08544

Mitchell D. Smooke
Yale University
New Haven, CT 06520

Tim P. Parr & Donna Hanson-Parr
Naval Air Warfare Center Weapons Division
China Lake, CA 93555

Submitted for the 26th International Symposium on Combustion
Oral Presentation, Colloquia: Laminar Flames
Word Count: Text 3950, Six figures 1200, Two tables 400
Total Word Count: 5550

May 7, 1996

***Corresponding Author: e-mail:prasad@cp.nrl.navy.mil, Phone: 202-404-6222, FAX: 202-767-4798**

An Eigenvalue Method for Computing the Burning Rates of HMX Propellants

ABSTRACT

A mathematical model for a three-tiered system consisting of solid, liquid and gas is derived for studying the combustion of HMX propellants. The resulting nonlinear two-point boundary value problem is solved by Newton's method with adaptive gridding techniques. In this study the burning rate is computed as an eigenvalue, which removes the uncertainty associated with employing evaporation and condensation rate laws in its evaluation. Results are presented for laser-assisted and self-deflagration of HMX monopropellants and are compared with experimental results. The burning rates are computed over a wide range of ambient pressures and compare well with experimental results from one to ninety atmospheres. The burning rate is found to be proportional to the pressure raised to the 0.82 power. Sensitivity of the burning rate to initial propellant temperature is calculated and found to be extremely low, in agreement with past theoretical predictions and experimental data. Results for laser-assisted combustion show a distinct primary and secondary flame separated by a dark zone, the length of which is dependent upon the incident laser flux intensity.

1 INTRODUCTION

Numerous chemicals have been used in making solid propellants with the goal of optimizing combustion characteristics for different purposes. The cyclic nitramines cyclotrimethylenetrinitramine ($C_3H_6N_6O_6$), commonly called RDX, and cyclotetramethylenetetranitramine ($C_4H_8N_8O_8$), commonly called HMX are important ingredients in solid propellants. The addition of nitramine particles such as RDX and HMX to double-base homogeneous propellants increases the flame temperature and energy release. It is widely known that the use of these ingredients offers many advantages. Due to the high energy release of these compounds and the large amount of low molecular weight gaseous combustion products, high values of specific impulse for rocket propellants and impetus for gun propellants can be achieved.

Due to the lack of understanding of the various elementary chemical and physical processes involved during the combustion of HMX propellants, little work has been done in the area of detailed numerical modeling. Microscopic observations, mostly by Boggs [1] and Hanson-Parr and Parr [2], of quenched samples using a scanning electron microscope and of burning samples using high speed motion-picture photography have revealed many aspects of the deflagration process that can aid in the development of theories. Ben-Reuven et al. have incorporated basic nitramine flame chemistry into theoretical models for studying the self-deflagration of HMX [3]. A simplified theoretical model was also developed by Mitani and Williams [4] for describing the flame structure and deflagration of nitramines, which included exothermic decomposition in a liquid layer, equilibrium vaporization, and exothermic combustion in the gas.

The problem of modeling the deflagration of a homogeneous solid propellant becomes especially difficult if it is necessary to account for two-phase flow that occurs in liquid melt layers as well as the gas-phase region above the liquid melt layer. The combustion of double-base propellants usually involves reactions occurring in the condensed-phase, in the dispersed-phase and in the gas-phase. Li, Williams and Margolis [5] used asymptotic analysis to extend the Mitani-Williams model [4] to account

for the presence of bubbles and droplets in a two-phase layer at the propellant surface. They again report good agreement with measured burning rates and pressure and temperature sensitivities using an overall chemical-kinetic parameter for describing gas-phase reactions. These models [4], [5] do not predict a dark zone as observed in experimental results.

Melius [6], [7] developed a two layer solid-gas model for studying the ignition and self-deflagration of RDX. He constructed a gas-phase model with an evaporation-condensation rate law that was used to model the regressing planar interface separating the solid and the gas-phase. In the solid-phase there was a single irreversible reaction whose rate was pertinent to liquid phase RDX. Fetherolf and Litzinger [8] have studied the CO_2 laser induced pyrolysis of HMX using both experimental studies and kinetic modelling of the gas-phase processes. Hatch [9] has developed a steady-state one-dimensional combustion model to study the gas-phase chemistry of a nitrate-ester flame and has recently applied this approach for studying the deflagration of HMX. The effects on combustion of different mechanisms of decomposition were determined and some of the more important gas-phase reactions were identified.

More recently, Prasad, Yetter and Smooke [10] have developed a detailed three layered model for studying RDX self and laser assisted deflagrations using detailed finite rate chemistry in the gas-phase and a three step liquid-phase decomposition mechanism. This model accurately predicts the burn rates over a wide range of ambient pressures and initial propellant temperatures under self deflagration conditions. In the present study, we apply the approach used in [10] to develop a detailed model for studying the combustion of HMX homogeneous monopropellants.

2 PROBLEM FORMULATION

Figure 1 shows a schematic diagram of the physical model we will employ for describing the steady-state burning of a homogeneous solid propellant. It consists of a three layered (solid, liquid and gas) model with sharp planar boundaries, with the liquid

layer sandwiched between the other two phases. HMX monopropellant combustion is modeled with a solid region followed by a liquid reaction zone followed by a fizz zone and then, after a dark zone, by a luminous gas-phase flame. Unlike RDX, there are four polymorphic forms of HMX [1]. The βHMX is the commonly encountered stable form at room temperature, whereas δHMX is most relevant to the thermal decomposition of HMX. The $\beta \rightarrow \delta$ phase transformation is again modeled as a sharp planar interface within the solid-phase to account for the different enthalpies and specific heats associated with the two polymorphs.

We consider the steady-state deflagration of a semi-infinite piece of pure HMX monopropellant. It is assumed that the propellant is mounted on a platform moving at a velocity equal but opposite to the regression velocity of the propellant. As a result, during steady-state HMX deflagration, the locations of the gas-liquid interface, the solid-liquid interface, and the primary and secondary flames are all fixed in the laboratory frame of reference. We further choose that the gas-liquid interface is located at the origin of the coordinate frame of reference. Thus in Figure 1, x is the coordinate normal to the surface; the interface is located at $x = 0$; the solid extends to $x = -\infty$, and the gas extends to $x = +\infty$. We are interested in computing the mass flux rate $\dot{m}$ $(gm/cm^2/sec)$ with which the solid must move in the $+x$ direction so that the interface will remain at $x = 0$.

The burning rate or the rate of decomposition of liquid HMX can be obtained as a difference between the rate of evaporation for liquid HMX and the rate of condensation for the gaseous products. The evaporation and condensation rates are generally complicated empirically derived expressions that involve the temperature and the various species concentrations at the interface. The evaporation and condensation rates are large numbers $O(10^2)$ with a very small difference $O(10^{-2})$. The result of employing this approach in a computational model is that numerical instability/convergence difficulties can occur. In the approach taken in this paper, the mass flux rate (burning rate) is obtained as an eigenvalue of the problem [10].

Our goal is to predict theoretically the mass fractions of the species and the

temperature as functions of the independent coordinate, x, together with the mass flux rate. If in the gas-phase we neglect viscous effects, body forces, radiative heat transfer and the diffusion of heat due to concentration gradients, the equations [11], [12] governing the structure of a steady, one-dimensional, gas-phase, isobaric flame are

$$\dot{m} = \rho v A = K1 = constant, \quad \rho = \frac{p\bar{W}}{RT}, \tag{1}$$

$$\dot{m}\frac{dY_k}{dx} = -\frac{d}{dx}(\rho A Y_k V_k) + A\dot{\omega}_k W_k, \qquad k = 1, 2, \ldots, K, \tag{2}$$

$$\dot{m}c_p\frac{dT}{dx} = \frac{d}{dx}\left(\lambda A\frac{dT}{dx}\right) - A\sum_{k=1}^{K}\rho Y_k V_k c_{pk}\frac{dT}{dx} - A\sum_{k=1}^{K}\dot{\omega}_k h_k W_k, \tag{3}$$

In these equations, x denotes the independent spatial coordinate; $\dot{m}$, the mass flux rate; T, the temperature; Y_k, the mass fraction of the k^{th} species; p, the pressure; v, the velocity of the fluid mixture; ρ, the mass density; W_k, the molecular weight of the k^{th} species; $\bar{W}$, the mean molecular weight of the mixture; R, the universal gas constant; λ, the thermal conductivity of the mixture; c_p, the constant pressure heat capacity of the mixture; c_{pk}, the constant pressure heat capacity of the k^{th} species; h_k, the enthalpy of the k^{th} species; $\dot{\omega}_k$, the molar rate of production of the k^{th} species; and V_k, the diffusion velocity of the k^{th} species. A represents the local cross-sectional area of the stream tube encompassing the flame.

In the liquid layer, sometimes referred to as the foam zone, temperatures are high enough for molecular degradation to take place. Simultaneous secondary reactions occur so that a mixture of soluable and liquid-phase species emerge from the surface and the net energy balance of the degradation is exothermic. For the liquid-phase, it is assumed that the diffusion velocities are negligible. The equations of motion therefore simplify into the form

$$\dot{m} = \rho v A = K1 = constant, \quad \rho = constant = K2, \tag{4}$$

$$\dot{m}\frac{dY_k}{dx} = A\dot{\omega}_k W_k, \qquad k = 1, 2, \ldots, K, \tag{5}$$

$$\dot{m}c_p\frac{dT}{dx} = \frac{d}{dx}\left(\lambda A\frac{dT}{dx}\right) - A\sum_{k=1}^{K}\dot{\omega}_k h_k W_k + \dot{Q}A, \tag{6}$$

The term $\dot{Q}A$ represents the energy input due to laser flux heating.

The solid-phase (both β and δ HMX) consists of a homogeneous propellant with no chemical reactions. The diffusion velocities in the condensed phase are again zero. The equations further simplify to

$$\dot{m} = \rho v A = K1 = constant, \quad \rho = constant = K2, \tag{7}$$

$$Y_{HMX} = 1.0, \quad Y_k = 0, \quad k \neq HMX, \tag{8}$$

$$\dot{m} c_p \frac{dT}{dx} = \frac{d}{dx}\left(\lambda A \frac{dT}{dx}\right) + \dot{Q}A, \tag{9}$$

Note here that the mass flux rate is a constant in the gas, liquid and solid phases.

To complete the specification of the problem, boundary conditions must be imposed on each end of the computational domain $(-L_1, L_2)$ and at the interfaces. The temperature and the species mass flux fractions are prescribed at the unreacted solid boundary. In the hot stream outflow boundary the gradient of temperature and species mass fractions are assumed to be negligibly small.

Calculation of the mass flux rate proceeds by introducing the trivial differential equation

$$\frac{d\dot{m}}{dx} = 0, \tag{10}$$

and an additional boundary condition. We choose to impose the boiling point temperature or the evaporation temperature at the gas - liquid interface, the location of which is fixed at the origin in the chosen coordinate system.

In addition, one needs to specify specialized interface conditions that account for the transition from the liquid to the gas-phase as well as any laser flux heating that might occur at the interface. Energy balance at this interface [12] is maintained by solving the difference form of the energy equation shown below

$$\dot{m}\frac{d}{dx}\left(\sum_{k=1}^{K} Y_k h_k\right) = \frac{d}{dx}\left(\lambda A \frac{dT}{dx}\right) - A\sum_{k=1}^{K} \rho Y_k V_k c_{pk} \frac{dT}{dx} + \dot{Q}A. \tag{11}$$

The term $\dot{Q}A$ represents the heat input due to laser flux heating, whereas the difference form of the convective term $\dot{m}\frac{d}{dx}\left(\sum_{k=1}^{K} Y_k h_k\right)$ takes account of the heat of vaporization associated with the phase transition from the liquid to the gas-phase.

Unlike the gas-liquid interface, whose location remains fixed at a specified grid point, the locations of the δHMX-liquid (melt point) and the βHMX-δHMX (phase transformation point) interfaces are not known a-priori. The temperature decides the location of these interfaces and the algorithm uses thermodynamic properties relevant to that phase for evaluating the interface conditions. Energy balance is maintained across the δHMX-liquid and the βHMX-δHMX interfaces by solving equation (11) without the diffusion velocity term.

The governing equations (1)-(11), are solved by Newton's method with adaptive gridding techniques [13]. The various diffusion coefficients in the gas-phase are evaluated using vectorized and highly optimized transport and chemistry libraries. The specific heats and thermal diffusivities for βHMX, δHMX and liquid HMX were obtained from the DSC work of Shoemaker [14].

The liquid-phase reaction mechanism is based on the work of Brill [15] and Thynell et al. [16]. The mechanism consists of three steps (Table 1) in which HMX thermally decomposes through two global channels to form either CH_2O and N_2O or H_2CN and NO_2. The third global step describes low temperature secondary chemistry of CH_2O and NO_2 and is derived from the gas-phase studies of Lin et al. [17].

The gas-phase reaction mechanism consists of 48 chemical species and 228 reactions and was formulated from the original RDX mechanism of Melius [6] and the RDX mechanism in [10]. Table 1 lists the modifications made to the RDX mechanism for HMX. The remainder reactions of the gas-phase mechanism can be found in [10]. The gas-phase kinetic parameters for thermal decomposition of HMX were obtained from [18]. Rate parameters for reactions of the species, HMXR, the radical formed by removing a nitro group from HMX and the species, HMXRO, the ring open-structure of the HMXR were taken directly from the RDX mechanism.

3 RESULTS AND DISCUSSION

In this section we present results for the case of self-deflagration and laser assisted combustion of HMX monopropellants.

Laser assisted combustion

Experimental studies often employ laser heating for ignition purposes and for obtaining higher burning rates [2]. In particular, the added laser flux energy significantly affects the energy balance at the gas-liquid interface, which results in larger mass flux rates (burning velocities) and also higher final flame temperatures. It is assumed that gas-phase species do not absorb at a laser wavelength of 10.6 microns [2] and that the laser flux energy is completely absorbed in the condensed-phase. Furthermore, the absorptivity of solid HMX at 10.6 microns is such that the $\frac{1}{e}$ depth is only $1\mu m$. Employing Beer's law, the reduction in beam irradiance brought about by absorption in the condensed-phase can be described as

$$I(x) = I(0)e^{-\alpha x}, \tag{12}$$

where $I(0)$ is the irradiance of the beam at the input to the gas-liquid interface, $\alpha = 1/(1\mu m)$ is the absorption coefficient, and x is the absolute distance measured from the gas-liquid interface. The energy input ($\dot{Q}$) at any node point in the condensed-phase is obtained by subtracting the beam irradiance leaving the cell from that at the input to the cell.

The laser-assisted experiments also show that the HMX flame is not a truly one-dimensional flame but a quasi-one-dimensional system. Measurements of the variation in cross sectional area were obtained experimentally by generating streamlines from particle image velocimetry (PIV) data and measuring the relative change in area of these streamlines. It was shown that the area increases by a factor of about five between the surface and the region in the burnt gas where the flame stops expanding radially. A curve fit was developed for this area variation normalized to the region of the burnt

gases very far from the gas-liquid interface.

Numerical simulations for laser assisted combustion have been performed at 1 atm with 70%, 80%, 90% and 100% of the peak value (337 $cal/cm^2/sec$), and have been compared with experimental results. The lower values used in the numerical simulations implicitly account for losses in the experiments such as reflection at the surface as well as the Gaussian distribution of the laser flux profile. The reactant stream temperature was $T_u = 298.5K$ and the computational length was given by $L = 8mm$.

Figure 2 shows the temperature profile in the solid, liquid and gas-phases. The experimental temperature at which evaporation occurs is $T_{evap} = 660K$. Based on the structure of the thermal wave propagating into the condensed-phase, we find that the temperature rises monotonically from the initial propellant temperature ($T_u = 298.15$) to the melting point temperature ($T_{melt} = 555K$). During this portion HMX exists in solid forms (β or δ crystals), and it is assumed that there are no chemical reactions taking place in this regime. This region is followed by the melt layer in which the temperature further increases from $555K$ to the surface temperature. The thickness of the melt layer based on these calculations is approximately 40 μm. This compares reasonably well with results published in the literature. We also find approximately 53% decomposition of liquid HMX in the melt layer. This is in contrast to 43% found for laser assisted RDX deflagration.

The liquid melt layer is followed by gas-phase combustion of gases evaporating from the melt layer. The temperature profile in the gas-phase shows three distinct regions termed in the literature as fizz, dark and flame zones. The fizz zone temperature steadily rises from the surface temperature to $1500K$. The temperature of the dark zone which follows the fizz zone remains almost constant throughout its width (2 mm), while that of the flame zone rises from the end of the dark zone to the adiabatic flame temperature.

The symbols in Figure 2 represent temperature measurements from the experiments of Hanson-Parr and Parr [2]. Temperature profiles were obtained by piecing

together OH and NO rotational temperature data in the gas phase and thermocouple measurements in the condensed phase. In general, there is good agreement between experimental and numerically computed temperatures in the gas, liquid and solid-phases.

As the laser flux intensity is increased, the calculations suggest that the burning rate increases correspondingly, with small changes in the final flame temperatures. The burning rates corresponding to 70%, 80%, 90% and 100% were 0.135, 0.152, 0.165 and 0.182 cm/sec respectively. Hanson-Parr and Parr [2] have reported a burning rate value of 0.18 cm/sec for the laser-assisted combustion experimentally studied. This increase in the mass flux rate (as compared with those for self-deflagration) is responsible for increasing the length of the dark zone as shown in Figure 2. Experimental results also indicate that the length of the dark zone varies as a function of the laser flux intensity.

The profiles for the NH, OH and CN mole fraction are shown in Figures 3 and 4, and show good agreement between the numerical and the experimental results. The CN profile is indicative of the location of the flame zone. Although the experimental measurements are slightly lower than the numerical values, the location of peak concentrations as well as the width of the CN profile again compare well with the experimental results. The OH profile peaks outside the CN/NH flame sheet. For CO_2 laser-assisted deflagration, experiments indicate that the narrow NH profile peaks at 3.6 mm above the surface at a value of 100 ppm.

Self-deflagration

Numerical calculations have also been performed for self-deflagrating HMX propellants. When the laser flux heating is removed, the mass burning rate is reduced and, as a result, the primary and the secondary flames superimpose on each other producing a single flame. The temperature profile for self-deflagration at 1 atm (Figure 5) shows that the two-zone flame structure is no longer present. This finding is in accordance with experimental results [2].

Calculations have also been performed for self-deflagration of HMX propellants over a wide range of ambient pressures ranging from one to ninety atmospheres. The dependence of the burning rate on ambient pressure is shown in Figure 6, for both

model and experiment [2], [19], [20]. The pressure sensitivity n was found to be approximately 0.82, and matches very well with experimental results and theoretical work [3], [4], [5] [21]. Gas-Liquid interface conditions for a self deflagrating HMX propellant as a function of ambient pressure are shown in Table 2. As the ambient pressure increases, the flame comes closer to the gas-liquid interface resulting in higher temperature gradients in the gas and the condensed-phases. Due to the higher burning rates associated with higher ambient pressures, we find that the thickness of the melt layer and the δHMX layer decreases as the ambient pressure increases. Although the surface temperature increases as the ambient pressure increases, the amount of decomposition in the liquid-phase decreases from about 59% at one atmosphere to almost 0% at 90 atmospheres (Table 2). This is due to a combination of smaller melt layer thickness and higher burning rates at higher ambient pressures.

The burning rate has been computed at initial propellant temperatures (T_u) ranging from 173 K to 423 K over a wide range of ambient pressure. These results compare favourably with experimental results from Atwood et al. [19] (Burning rate comparison for $T_u = 373K$ is also shown in Figure 6). When the initial propellant temperature is increased from 298 K to 373 K, the temperature gradient in the gas and liquid-phase reduces, but the thickness of the melt layer, the amount of liquid-phase decomposition and the burn rate increase slightly. The average value of temperature sensitivity (σ_p) obtained for HMX monopropellant is approximately 0.0008/K, and compares favourably with experimental results. The value for σ_p predicted in [3], [4], [5] and [21] were significantly higher.

The eigenvalue approach presented in this paper requires the modeler to prescribe the temperature at the gas-liquid interface. Physically this is equivalent to specifying the vapour phase equilibrium diagram. An alternative / equivalent approach could be to specify the location and temperature of the melting point where the transition of the solid to liquid takes place. This approach would still require knowledge of the boiling point temperature, although in this case the location of the boiling point does not have to be fixed in the coordinate frame of reference. One could equally well specify the

adiabatic flame temperature of the system and then compute not only the mass flux rate ($\dot{m}$) but the interface temperature as well.

4 CONCLUSIONS

Research in the field of solid rocket propellants over the past few decades has focused on obtaining a detailed understanding of the various physical and chemical processes involved during the burning of homogeneous and heterogeneous propellants. In this paper, a mathematical model for a three-tiered system consisting of solid, liquid and gas was derived for studying the combustion of HMX propellants. The resulting non-linear two-point boundary value problem was solved by Newton's method and adaptive gridding techniques. In this study, the burning rate was computed as an eigenvalue, which removes the uncertainty associated with employing evaporation and condensation rate laws in its evaluation. Self-deflagration burning rates were computed over a wide range of ambient pressures and were shown to compare favourably with experimental results. Results for laser-assisted combustion showed a distinct primary and secondary flame separated by a dark zone.

5 ACKNOWLEDGEMENT

This work was supported by the Office of Naval Research.

Table 1: HMX Chemistry Mechanisms

Units are $cm^3, mole, sec, cal, K$, $k = AT^n exp(-E_a/RT)$

	HMX Liquid Phase Reaction Mechanism			
No.	Reaction	A	n	E_a
1L	$HMX \rightarrow 4CH_2O + 4N_2O$	$.301 \times 10^{15}$	0.00	41200.0
2L	$HMX \rightarrow 4NO_2 + 4H_2CN$	$.802 \times 10^{18}$	0.00	50200.0
3L	$CH_2O + NO_2 \rightarrow CO + NO + H_2O$	$.802 \times 10^{03}$	2.77	13730.0
	HMX Gas Phase Reactions			
No.	Reaction	A	n	E_a
224G*	$HMX + M \rightleftharpoons HMXR + NO_2 + M$	$.250 \times 10^{17}$	0.00	46200.0
	Fall-off Parameters	.125	0.00	17200.0
225G	$HMX + H \rightleftharpoons HMXR + HONO$	$.100 \times 10^{14}$	0.00	5000.0
226G	$HMX + OH \rightarrow 3H_2CNNO_2 + HCN + NO_2 + H_2O$	$.100 \times 10^{14}$	0.00	5000.0
227G	$HMXR + M \rightarrow HMXRO + M$	$.100 \times 10^{17}$	0.00	23000.0
	Fall-off Parameters	.130	0.00	5000.0
228G	$HMXRO + M \rightarrow 3H_2CNNO_2 + H_2CN + M$	$.100 \times 10^{17}$	0.00	23000.0
	Fall-off Parameters	.130	0.00	5000.0

* Numbers correspond to those in Reference [10].

Table 2: Gas Liquid Interface Conditions as a function of Ambient Pressure : $T_u = 298K$

Pressure (atm)	1.0	5.0	20.0	50.0	90.0
Gas Phase Heat Conduction (K/cm)	1.51E5	5.80E5	8.70E5	1.90E6	2.95E6
Condensed Phase Heat Conduction (K/cm)	2.50E4	1.10E5	2.33E5	4.74E5	7.45E5
Gas-liquid Interface Temperature (K)*	630.0	670.0	730.0	750.0	800.0
Thickness of Melt Layer (μm)	65	25	14	7	4
Liquid Phase Decomposition (%)	59.0	49.0	11.0	0.2	0.0
Thickness of δHMX layer (μm)	195.0	53.0	23.0	14.0	6.0
Temperature Sensitivity σ_p (1/K)	6.1E-4	8.2E-4	1.2E-3	1.4E-3	1.2E-3
Burn Rate (cm/sec)	0.035	0.153	0.401	0.896	1.509

*Refer Zenin [22]

References

[1] Boggs, T.L., in *Fundamentals of Solid-Propellant Combustion* (K.K. Kuo, and M. Summerfield, Ed.), New York, 1984, Vol. 90, p. 121.

[2] Parr, D. H., and Parr, T., *Twenty-Fifth Symposium (International) on Combustion*, The Combustion Institute, 1994, pp. 1281-1287.

[3] Ben-Reuven, M., and Caveny, L.H., *AIAA J.*, Vol. 19, 10, p. 1276 (1981).

[4] Mitani, T., and Williams, F.A., *Twenty-First Symposium (International) on Combustion*, The Combustion Institute, (1988).

[5] Li, S.C., Williams, F.A., and Margolis, S.B., *Combust. Flame*, 80:329-342, (1990).

[6] Melius, C.F., in *Chemistry and Physics of Energetic Materials*, (S.N. Bulusu, Ed.), The Netherlands, 1990, p. 51.

[7] Melius, C.F., *Proceedings of the Twenty-Fifth JANNAF Combustion Meeting*, 1988.

[8] Fetherolf, B.L., and Litzinger, T.A., *Proceedings of the Thirteeth JANNAF Combustion Meeting*, 1993.

[9] Hatch, R. L., *Proceedings of the Twenty-Fourth JANNAF Combustion Meeting*, 1987.

[10] Prasad, K., Yetter, R.A., and Smooke, M., *Comb. Sci. & Tech.*, submitted, 1995.

[11] Williams, F.A., *Combustion Theory*, 1985, Addison-Wesley Publishing Co., Inc., New York.

[12] Kuo, K.K., *Fundamentals of Combustion*, 1986, Wiley, NY.

[13] Smooke, M.D., *Jour. of Computational Phy.*, Vol. 48, p. 72, (1982).

[14] Shoemaker, R. L., Stark, J. A., and Taylor, R. E., *High Temperature-High Pressures*, Vol. 17, p. 129-435, 1985.

[15] Brill, T.B., *Journal of Prop. Power*, Vol. 11, p. 740, 1995.

[16] Thynell, S.T., Gongwer, P.E., and Brill, T.B., submitted, *Comb. Flame*, 1995.

[17] Lin, C.Y., Wang, H.T., Lin, M.C. and Melius, C.F., *Int. Journal of Chemical Kinetics*, Vol. 22, p. 455, 1990.

[18] Shaw, R., and Walker, F.E., *J. Phys. Chem.*, Vol. 81, p. 2572-2576, 1977.

[19] Atwood, A., Price, C., Curran, P., and Wiknich, J., *Proceedings of the Thirty-Second JANNAF Combustion Meeting*, (1995).

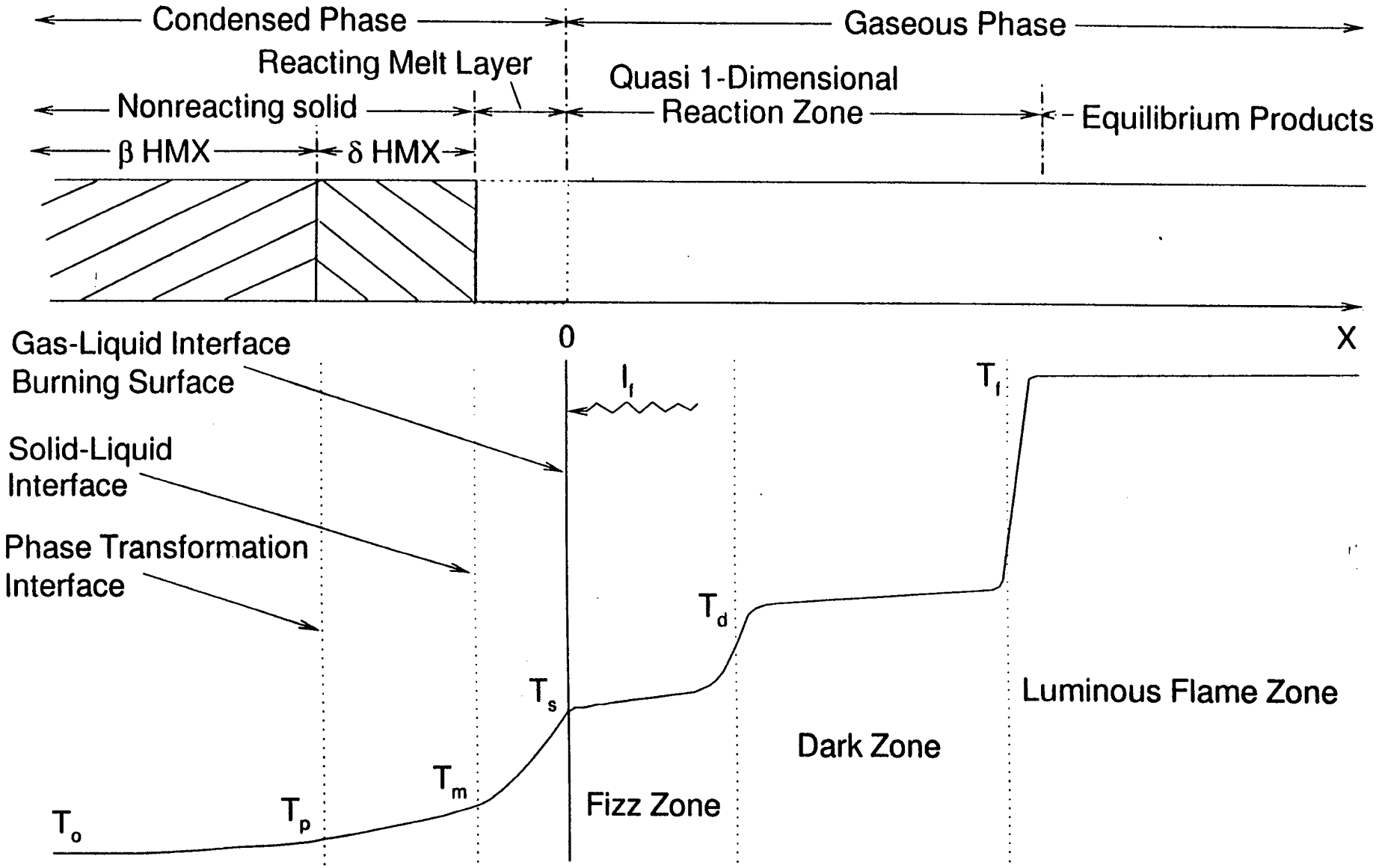
[20] Zenin, A. A., *Combustion Explosion and Shock Waves*, Vol. 2, p. 67, 1966.

[21] Price, C.F., Boggs, T.L., and Derr, R.L., *AIAA Paper* 79-0164, 1979.

[22] Zenin, A., *Jour. Power Propulsion*, Vol. 11, p. 752, 1995.

List of Figures

List of Tables

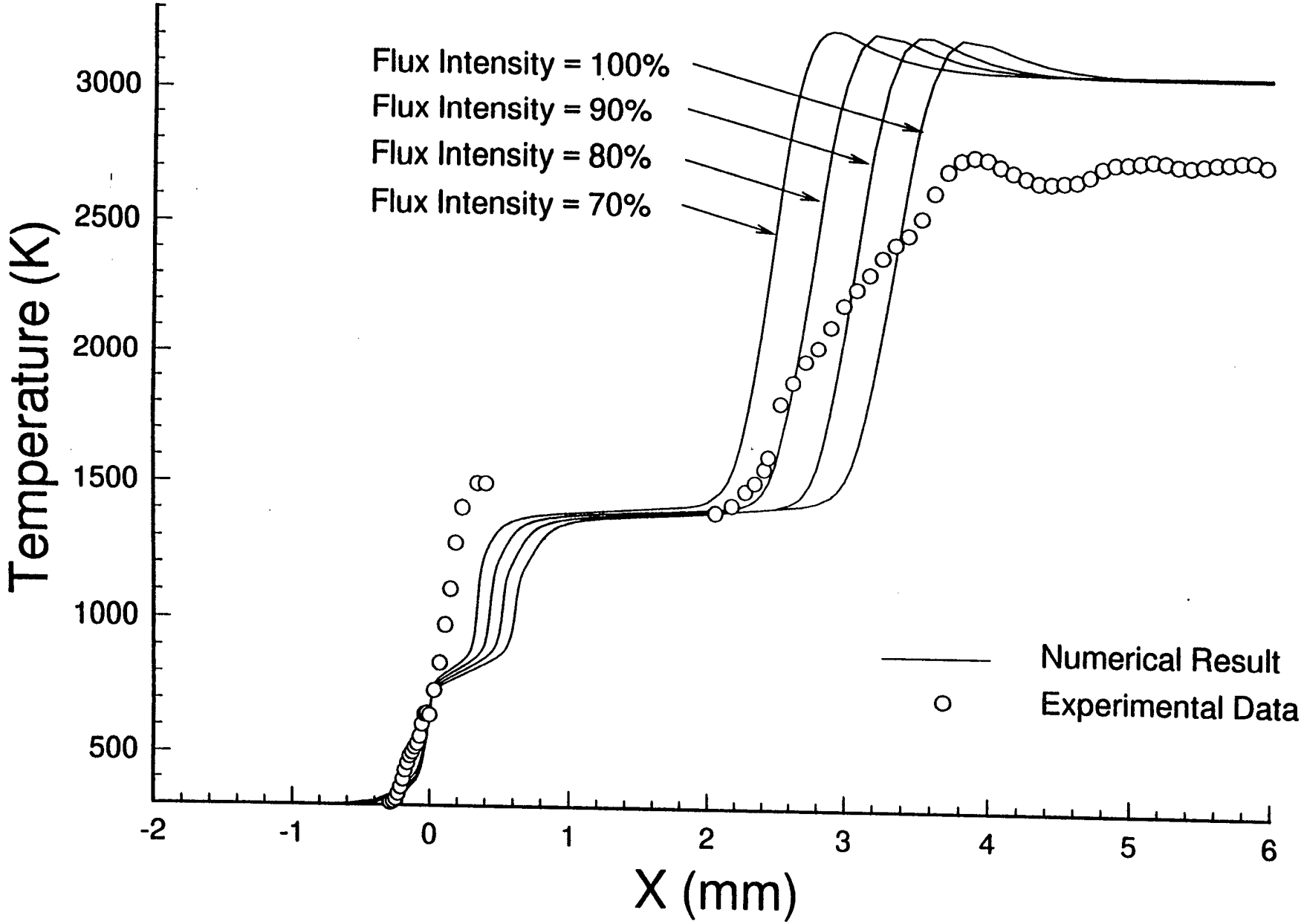


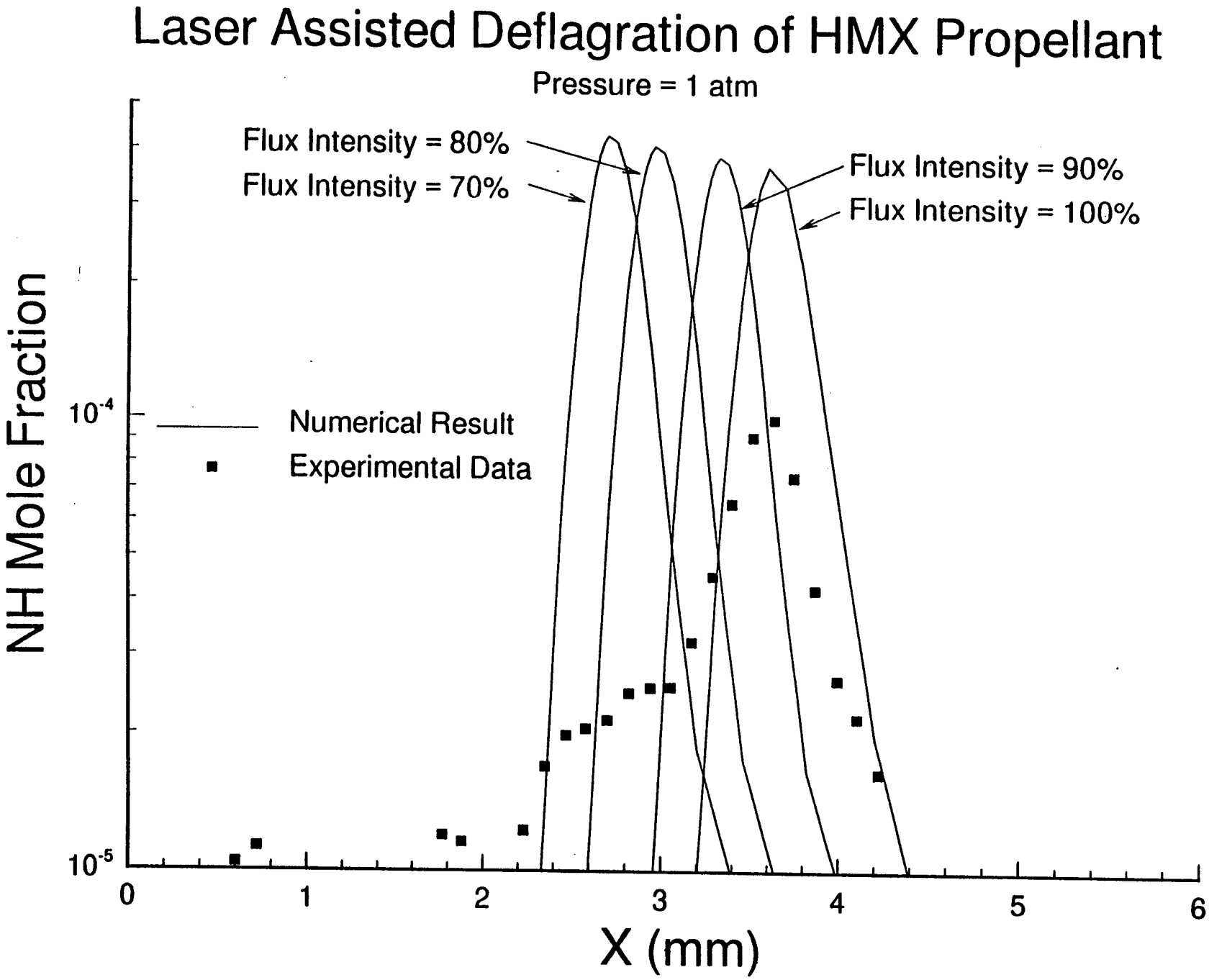
Combustion of HMX monopropellant

Schematic description of the flame (combustion wave) structure

Laser Assisted Deflagration of HMX Propellant

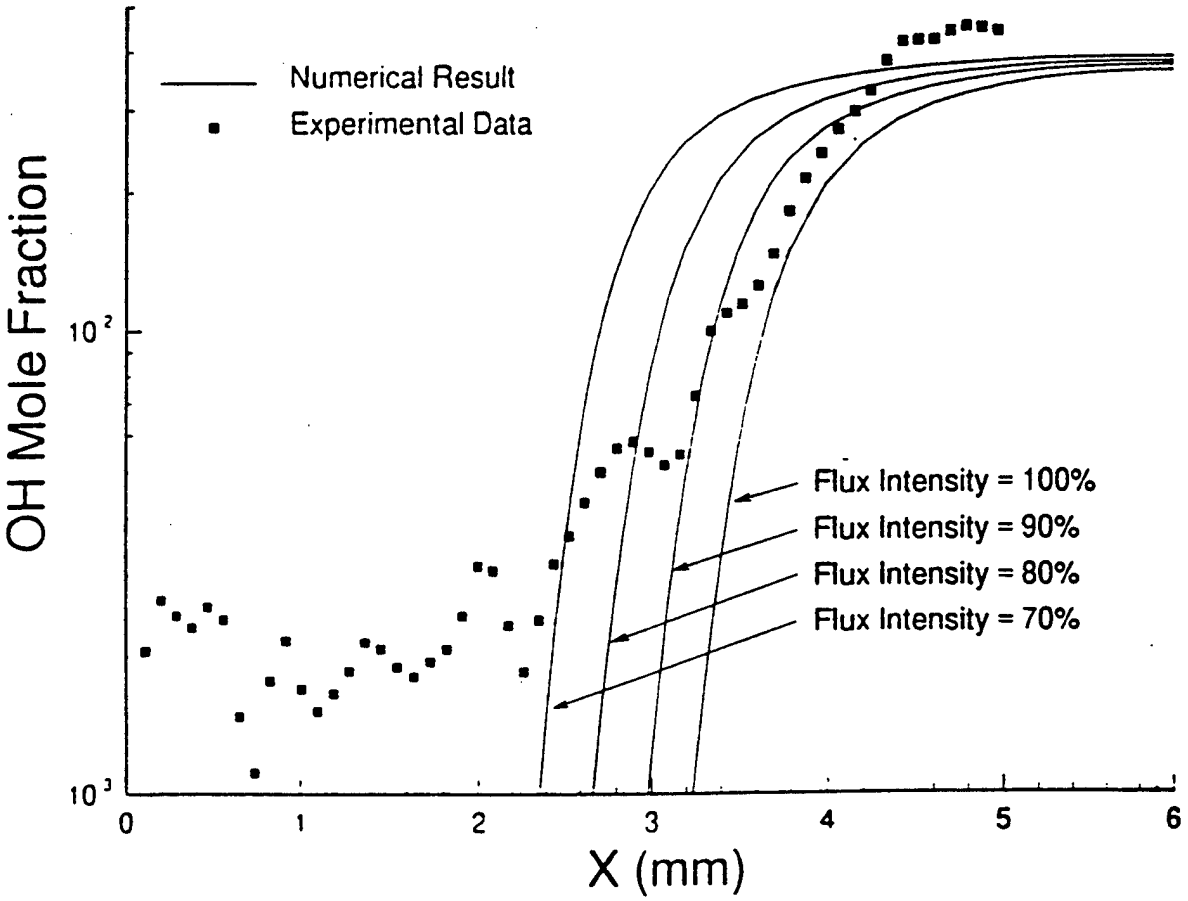
Pressure = 1 atm

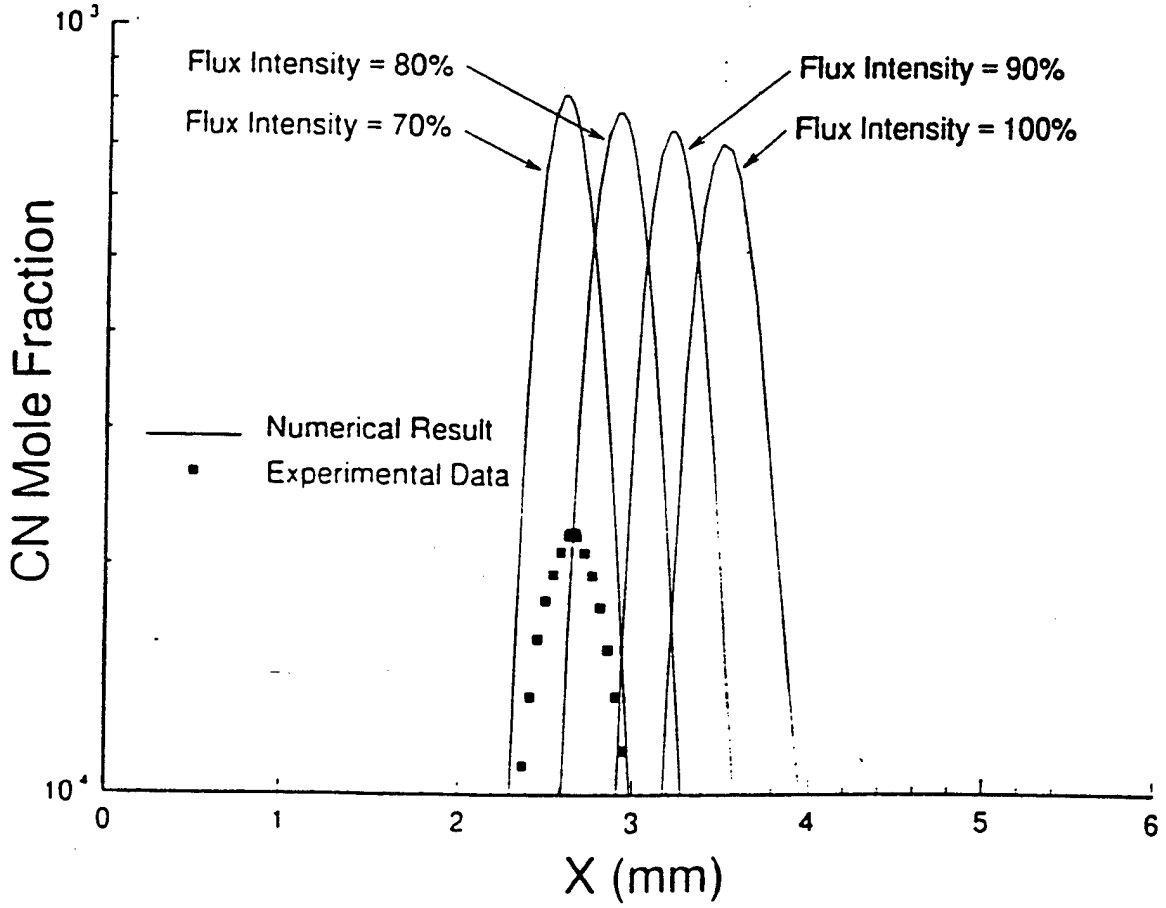


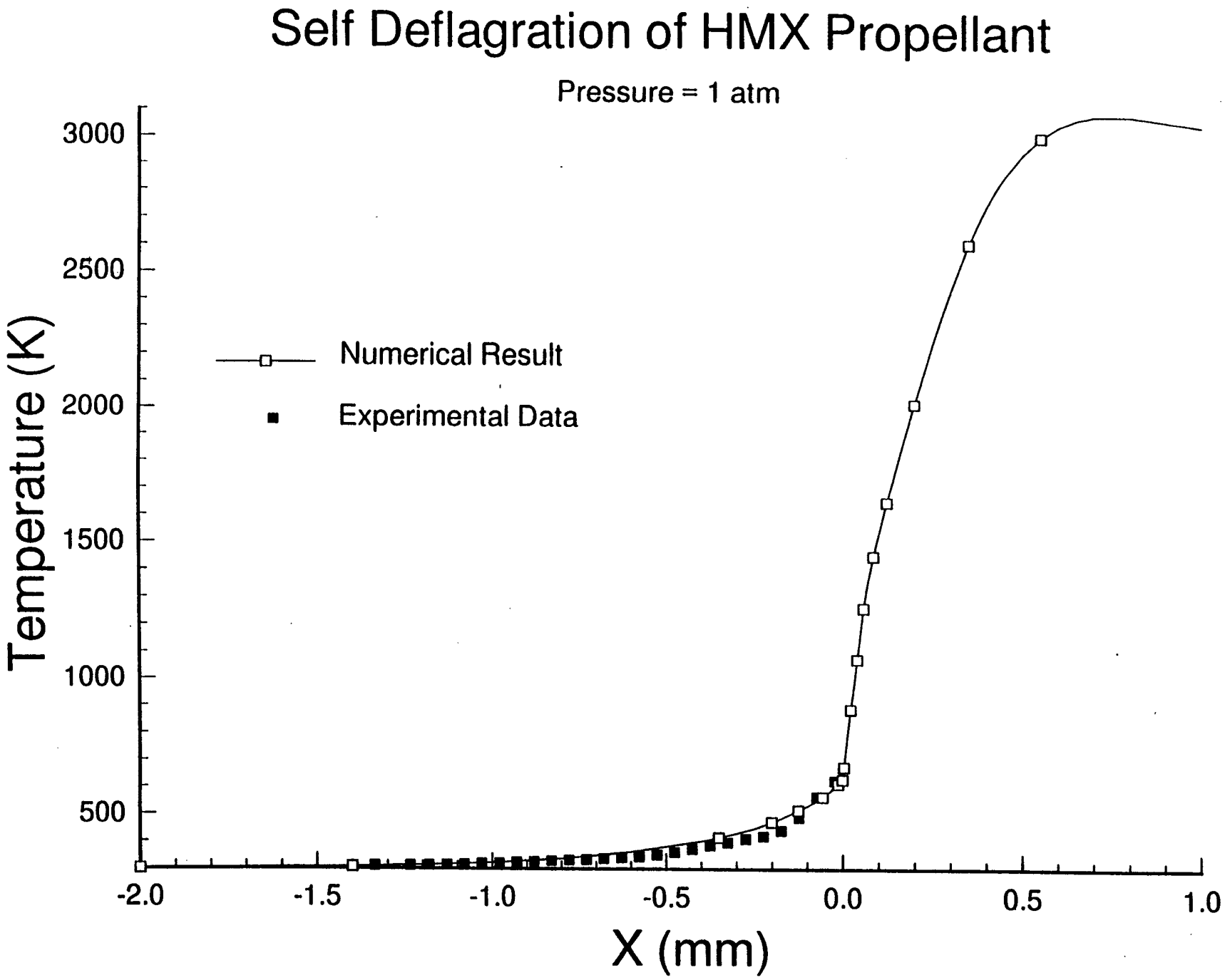

Laser Assisted Deflagration of HMX Propellant
Pressure = 1 atm
Flux Intensity = 80%
Flux Intensity = 70%
Flux Intensity = 90%
Flux Intensity = 100%
Numerical Result
Experimental Data
NH Mole Fraction
10-4
10-5
0
1
2
3
4
5
6
X (mm)

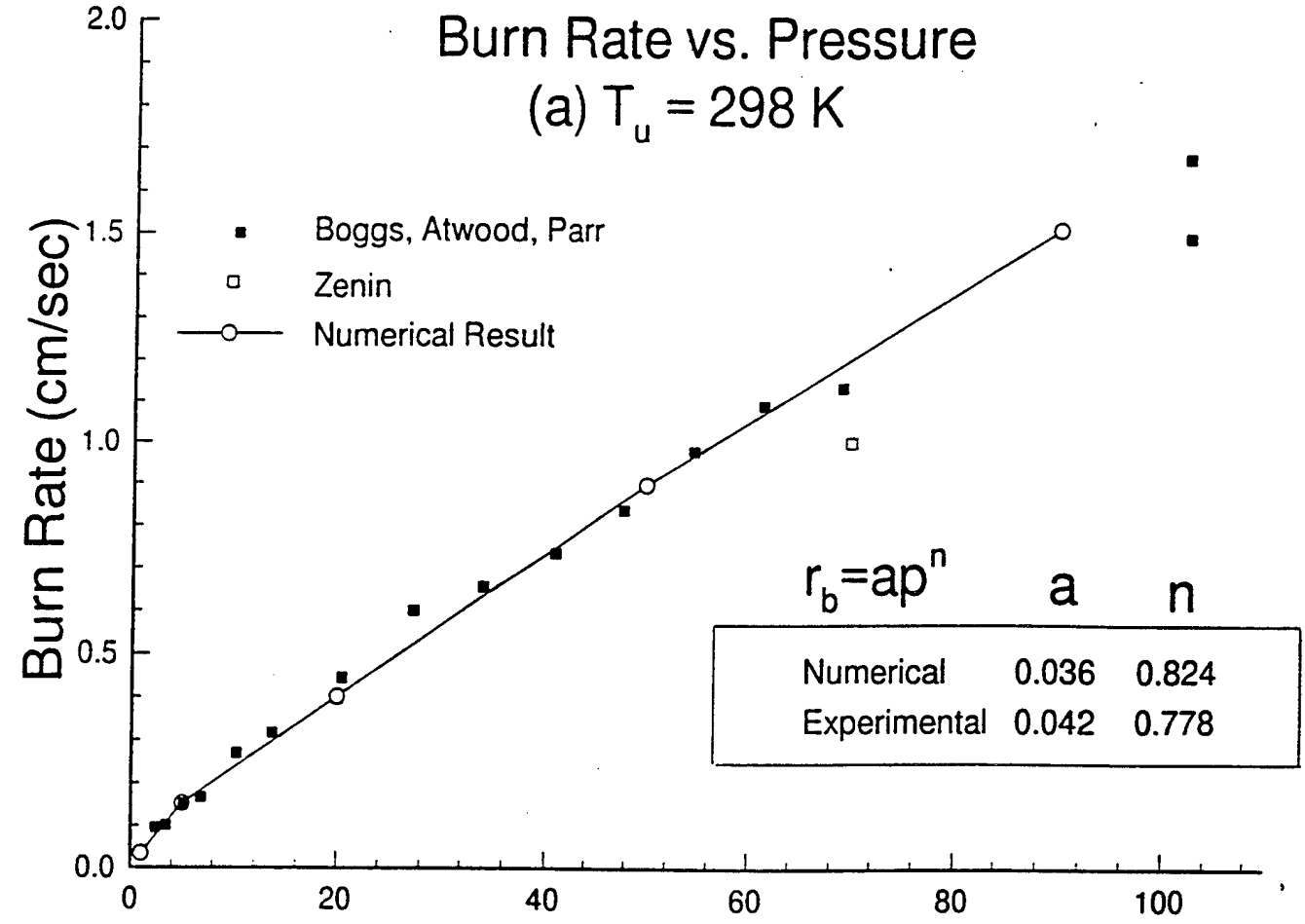
Laser Assisted Deflagration of HMX Propellant

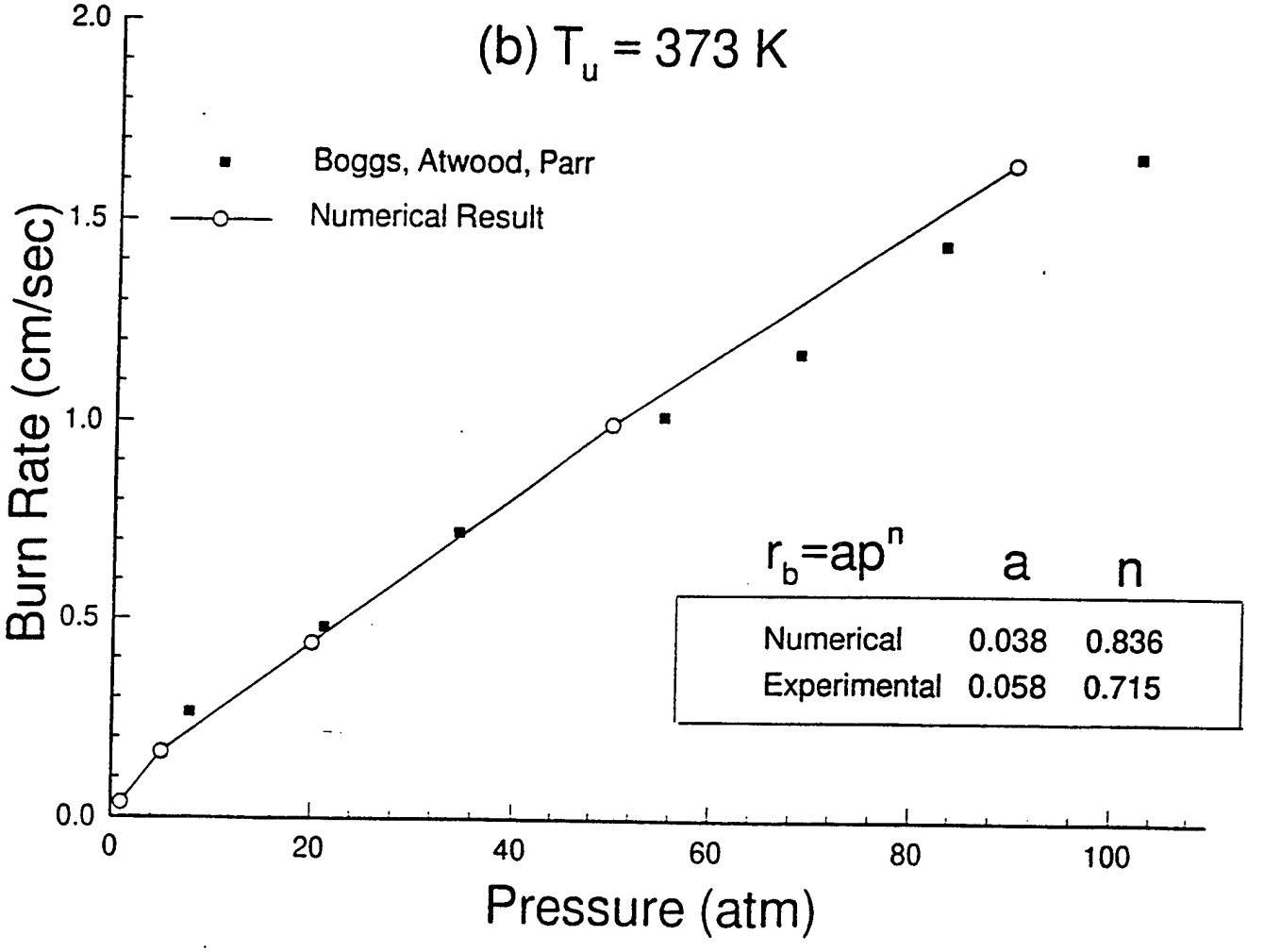
Pressure = 1 atm






Self Deflagration of HMX Propellant
Pressure = 1 atm
Temperature (K)
3000
2500
2000
1500
1000
500
Numerical Result
Experimental Data
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
X (mm)


Burn Rate vs. Pressure
(a) T_u = 298 K
Boggs, Atwood, Parr
Zenin
Numerical Result
Burn Rate (cm/sec)
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80
100
$r_b=ap^n$ a n
Numerical 0.036 0.824
Experimental 0.042 0.778


(b) T_u = 373 K
Boggs, Atwood, Parr
Numerical Result
Burn Rate (cm/sec)
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80
100
Pressure (atm)
$r_b=ap^n$ a n
Numerical 0.038 0.836
Experimental 0.058 0.715

Appendix KK

Numerical Analysis of a High-Velocity Oxygen-Fuel Thermal Spray System

JTTEE5 5:175-184

Numerical Analysis of a High-Velocity Oxygen-Fuel Thermal Spray System

X. Yang and S. Eidelman

The fluid and particle flow field characteristics of a high-velocity oxygen-fuel (HVOF) thermal spray (TS) system are analyzed using a two-phase flow model and simulated using computational fluid dynamics (CFD) techniques. The model consists of a conservation equation and constitutive relations for both gas and particle phases. Compressible, turbulent flow is modeled by a k-ε turbulent model. A Lagrangian formulation is used to model particle trajectory, and heat and momentum transfer. The fluid velocity fluctuations resulting from gas turbulence are simulated by a stochastic model and the particle motion in the turbulent flow is calculated in a Lagrangian Stochastic-Deterministic (LSD) method. Details of gas flow field, particle temperature and particle velocity histories, and particle temperature and velocity profiles in the system are presented. For the validation of the numerical analysis, the computed results are compared with available experimental measurement. Excellent agreement between simulations and measurements is obtained for both gas and particle flow fields. A parametric study is also conducted for different particle sizes and different nozzle barrel lengths. The flow phenomena for different flow parameters are analyzed and explained as the result of gas dynamics and heat and momentum transfer between the two phases. The developed methodology provides a means to analyze, design, and optimize the TS process. The numerical analysis presents a first comprehensive, fundamental quantitative analysis for the HVOF TS system.

1. Introduction

THE HIGH-VELOCITY oxygen-fuel (HVOF) thermal spray (TS) system has been used in the aerospace industry for a variety of surface coating applications for many years. Aerospace engineers use TS coatings to insulate parts from heat, reduce turbine blade wear, and protect against oxidation and corrosion (Ref 1). Although the process is widely used, TS system design has been primarily empirical and the understanding of the system mechanism is based on engineering intuition and analysis of operational data. The schematic diagram of the system, including combustion chamber, nozzle, barrel, particle injection, gas and particle flow field, is shown in Fig. 1. High-pressure and high-temperature combustion gases resulting from the combustion of oxygen and fuel expand through the converging and diverging de Laval nozzle and the barrel to supersonic speed (local Mach number $M \approx 2$). Particles are then injected into the barrel at the exit of the nozzle. Particles are turbulently mixed, heated, and accelerated in the barrel and jet at high speed and temperature to the substrate to be coated. From the fluid dynamics point of view, the system is very complex and involves two-phase (gas-particle) flow with turbulence, heat transfer, chemical reactions, and supersonic/subsonic flow transitions. In an engineering application, the microstructure and physical properties of the plated surface are determined by the physical and chemical conditions of the particles that impinge on the surface. The physical and chemical condition of the particle impinging on the substrate in turn is dependent on a large number of parameters such as gun design (length of barrel in particular), the gas jet formation, the position of the particle relative to the substrate, the particle size, shape, material, injection method, and so forth. In recent years, some publications have been dedicated to analyzing the TS process (Ref 2-4). However, these studies are either based on some highly simplified model or address only the gas and particle flow in the jet, without considering the key processes of particle injection and gas/particle flow inside barrel.

Keywords high-velocity oxygen-fuel, Lagrangian formulation, particle trajectory, particle temperature, supersonic, thermal spray, two-phase flow, turbulent flow

X. Yang and **S. Eidelman,** Science Applications International Corporation, McLean, VA, USA.

To advance the technology of the TS system and improve the quality and efficiency of the TS coating, we believe that a detailed understanding of the complete system is needed. The understanding will include gas dynamics, particle injection, the interaction between the injected particle and gas flow, particle conditions before they impinge on the substrate, and the relation of the particle condition to the final coating quality. In this paper, we present a study of the HVOF thermal spray system using computational fluid dynamics. We first formulate the compressible two-phase flow using a Eulerian (for gas) and Lagrangian (for particle) formulation. The dynamics of the flow are governed by conservation equations of each phase, and the particle phase is coupled with gas phase by momentum and heat transfer. We solve the system of equations numerically for the gas and particle flow field. The objectives of the study are (1) to formulate and solve the HVOF TS flow field and compare the simulation with existing experimental results and (2) to conduct a parametric study using a validated model to explore the characteristics of the two-phase flow for TS engineering applications.

2. Mathematical Model and Numerical Solution

The mathematical model for the two-phase flow consists of a conservation-governing equation and constitutive relations that provide closure models. The basic formulation adopted here follows the gas and low-loading particle flow dynamics model presented by Crowe et al. (Ref 5) and Chen and Crowe (Ref 6). For

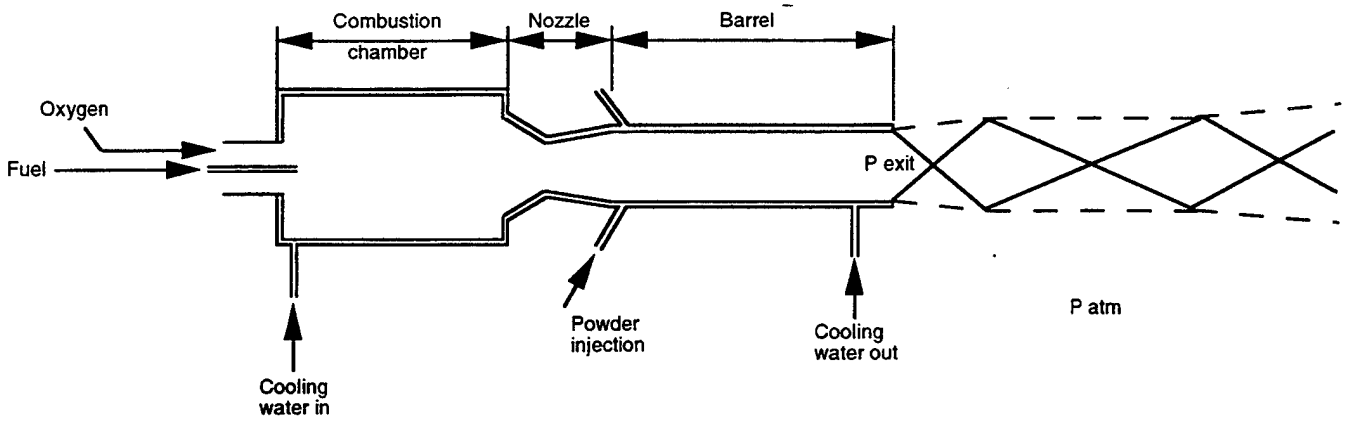


Fig. 1 Schematic of the high-pressure HVOF gun and flow field

the gas flow, it is natural to use the Eulerian approach for analyzing the continuum. A direct numerical simulation of the two-phase turbulent flow system requires solving the instantaneous compressible Navier-Stokes equation for the gas flow followed by simulating the motion of a large number of particles in the resolved instantaneous gas field. This method needs large capacity and very fast computers. It is not practical at the present time for engineering applications. At present, it is more appropriate to apply an averaged Navier-Stokes equation with turbulence models. In this study, the gas model is the compressible Navier-Stokes equation in Favre mass-averaged variables for the conservation of mass, momentum, and energy. To model flow turbulence, the Reynolds stress tensor is closed using a k-ε turbulence model (Ref 7). Here, k is turbulent kinetic energy and ε is the rate of dissipation of kinetic energy. A recently developed (Ref 8) three-dimensional, multiblock, upwind, fully implicit, finite volume code was modified to solve the governing equations. The code incorporates the high-order, upwind, flux-difference splitting process of Roe (Ref 9) with a monotonic upstream-centered scheme for conservation laws (MUSCL) integration scheme to obtain good shock-capturing, high-accuracy solutions in general coordinates for the three-dimensional geometry computational domain. For a three-dimensional flow, seven variable-governing equations are discretized and integrated in time to provide a steady-state solution. To deal with the well-known stiffness of the equations, an implicit time integration is used. The details of the code and numerical solution techniques can be found in Ref 7.

To describe particle motion, we use the Lagrangian formulation. The Lagrangian formulation allows a description of the particle/wall interactions that are very difficult to analyze numerically using a Eulerian approach. In addition, the Lagrangian formulation is also logical for the low-volume, collisionless particle flow that is typical of TS systems. To simplify the analysis, the following assumptions are made:

- The particles do not undergo a phase change.
- The particles are solid spheres and have a constant material density.
- The volume occupied by the particles is negligible.
- The interaction between particles can be ignored.
- The only force acting on a particle is drag force, and the only heat transfer between the two phases is convection.
- The weight of the solid particles and their buoyancy force are negligible compared to the drag force.
- The particles have a constant specific heat and are assumed to have a uniform temperature distribution inside each particle.

Furthermore, because the loading of the particle defined as total mass flow rate/total gas mass flow rate is very low (<4%), one can assume that the presence of particles will have a minimal effect on gas velocity and temperature field. This means that the momentum and heat exchanges from particle to gas are too small to change gas velocity and temperature distribution. As a result of the assumptions presented above, the two-phase problem can then be decoupled: The gas flow field can be simulated first, followed by the particle flow analysis.

The simulation of the particle flow field consists of calculating particle trajectories and temperature histories in the gun barrel and in the jet, after particles are injected into the gun barrel from the particle injection port. The interaction between particle and wall is also included in the simulation.

Particle motion in gas turbulent flow is predicted by means of the Lagrangian stochastic model (LSD) (Ref 10, 11). There are two elements in the LSD model: the description of the turbulent field and integration of the particle motion equations.

The numerical solution of gas fluid equations provides the fields of mean velocity components as well as turbulent kinetic energy k and dissipation rate of turbulent kinetic energy, ε. From k and ε, the scales in time and space of large turbulence eddies can be evaluated. The fluid instantaneous velocity is obtained by adding the fluctuating velocity resulting from large turbulent eddy to the mean velocity. The root mean fluctuating velocity can be calculated from turbulent kinetic energy k, which is defined as $\vec{V'}\vec{V'}/2$, where $\vec{V'}$ is fluctuating velocity defined as $\vec{V'} = (u'\vec{i} + v'\vec{j} + w'\vec{k})$. If the turbulent field is assumed to be isotropic and to possess a Gaussian distribution of fluctuating velocity, then the turbulent kinetic energy k is equal to $\frac{1}{2}(u'^2 + v'^2 + w'^2) = \frac{3}{2}u'^2$. The standard deviation of the fluctuating velocity distribution is $\sigma' = (3k/2)^{1/2}$.

Using this standard deviation and the assumption of a Gaussian and isotropic distribution of fluctuating velocities, we can

calculate turbulent velocity fluctuation. The instantaneous total fluid velocity is then:

$$\vec{V}_g = \vec{U}_g + \vec{V} \quad \text{(Eq 1)}$$

where $\vec{U}_g$ is the mean flow velocity.

The equation of motion for a particle in the gas flow (Ref 5) is

$$m_p \frac{d\vec{V}_p}{dt} = C_D \rho_g \left(\vec{V}_g - \vec{V}_p\right) \left| \vec{V}_g - \vec{V}_p \right| \frac{A_p}{2} + \vec{F} \quad \text{(Eq 2)}$$

where ρ_g is gas intensity, m_p is particle mass, $\vec{V}_p$ and $\vec{V}_g$ are particle and gas velocity vectors, respectively. A_p is the particle area. The term $\vec{F}$ denotes external force like gravity or centrifugal or Coriolis force in a cylindrical coordinate system. Other forces such as added mass effect and Basset history force are not presently considered because they are of the order of the gas/particle density ratio, which for most applications of interest is of the order of 10^{-3} (Ref 12, 13).

The drag coefficient, C_D, which appears in Eq 2, depends on a few parameters, for example, the Reynolds number, the Mach number, the surface roughness of the particle, the flow stream turbulence level, the rotation of the particle, and so forth. However, it depends primarily on the Reynolds number. For the present solution, the simple form proposed in Ref 14 is adopted:

$$C_D = \left(\frac{24}{Re_p}\right)\left(1 + 1.015\, Re_p^{0.687}\right) Re_p < 1000$$

$$C_D = 0.44 \; Rep > 1000 \quad \text{(Eq 3)}$$

Here the particle Reynolds number is defined by

$$Re_p = \rho_g \frac{\left(\left|\vec{V}_g - \vec{V}_p\right|\right)}{\mu_g} d_p \quad \text{(Eq 4)}$$

where U_g is gas viscosity.

Integrating Eq 1, assuming the gas velocity is constant over the time of integration and defining $f = C_d$ Re/24, yields

$$\vec{V}_p = \vec{V}_g - \left(\vec{V}_g - \vec{V}_p^o\right) \exp\left(\frac{-\Delta t}{\tau}\right) \quad \text{(Eq 5)}$$

where $\vec{V}_p^o$ is the initial particle velocity, Δt is the time interval, and τ is

$$\tau = \frac{\rho_s d_p^2}{18 \mu_g f} \quad \text{(Eq 6)}$$

where ρ_s is particle material density.

After calculating the particle velocity, the particle position at time Δt is determined as:

$$\vec{X} = \vec{X}_o + \left(\vec{V}_p + \vec{V}_p^o\right) \frac{\Delta t}{2} \quad \text{(Eq 7)}$$

The model assumes that the particle interacts with a sequence of turbulent eddies of given size and lifetime. After the particle traverses the eddy or the eddy dissipates, we assume that the particle enters a new eddy with a randomly sampled new fluctuation intensity and hence, a given size and dissipation time. As a result, the interaction time (the Δt in Eq 5) is the minimum of turbulence dissipation time and the time required for the particle to traverse a given eddy.

The characteristic size of the randomly sampled turbulent flow field (eddy size) is proportional to turbulent length scale given by $l_e = ck^{3/2}/\varepsilon$. Here ε is the energy dissipation rate that was calculated from the k-ε model in gas turbulent flow calculation. c is equal to 0.3 according to arguments in Ref 10. The eddy lifetime is estimated as $\Delta t_e = l_e/|\vec{V}|$.

The requirement that the particle must remain within the eddy during the interaction time interval leads to the second part of the requirement that

$$\Delta t_r \le \frac{l_e}{\left|\vec{V}_g - \vec{V}_p\right|}$$

Therefore, the interaction time between particle and gas (the eddy) will be the minimum of Δt_e and Δt_r; that is, $\Delta t = \text{Min}(\Delta t_e, \Delta t_r)$. After one interaction time, a new velocity fluctuation is randomly sampled and the process is repeated.

The governing equation to calculate particle temperatures is:

$$m_p C_p \frac{dT_p}{dt} = \dot{q} \quad \text{(Eq 8)}$$

where C_p is the specific heat of the particle and $\dot{q}$ is the heat transfer rate to the particle. The rate of heat transfer excluding radiation heat transfer is defined as:

$$\dot{q} = \text{Nu} \pi k_g d_p \left(T_g - T_p\right) \quad \text{(Eq 9)}$$

where k_g is the thermal conductivity of the gas and Nu is the Nusselt number and is a function of the Reynolds number and Prandtl number. The relation used here is:

$$\text{Nu} = 2 + 0.6 \text{Re}^{0.5} \text{Pr}^{0.33} \quad \text{(Eq 10)}$$

Integrating the particle heat transfer equation over a small time interval, Δt, and assuming a gas temperature constant over the integration time, Δt, the resulting particle temperature is

$$T_p = T_g - \left(T_g - T_p^o\right) \exp\left(-\frac{\Delta t}{\sigma}\right) \quad \text{(Eq 11)}$$

where σ is defined as

$$\sigma = \frac{\rho_s d_p^2 C_p}{6 \text{Nu} \cdot k_g} \quad \text{(Eq 12)}$$

3. Results

3.1 *Gas Flow Simulations*

To validate the current two-phase flow model and numerical solution procedure, we selected the TAFA JP-5000 HVOF TS system (Ref 15) as the test case because extensive experimental data are available for this system. We first resolve the gas flow field in the barrel. The chemical state of the gas flow in the barrel

can be considered to be in equilibrium or frozen composition. Experimental measurements (Ref 16) indicate that the flow is closer to the frozen state than to equilibrium. This is because the gas residence time is shorter than the chemical reaction time. Thus, we assumed in our simulations that the gas has a frozen chemical composition for a given fuel/oxygen ratio. The state of gas in a frozen condition can be calculated using the NASA Gordon-McBride program (Ref 17). We simulated the JP-5000 gun flow for the conditions given in Ref 16, for which the experimental measurements were made. We considered the gas flow for the TAFA JP-5000 gun (Hobart Tafa Technologies Inc., Concord, NH) with a 200 mm (8 in.) barrel with the following conditions in the combustion chamber:

- Molar mass of combustion products: M = 25.84 kg/kg mole
- Stagnation temperature: T^o = 3100 K
- Chamber pressure: p^o = 586.1 kPa
- Isentropic coefficient: α = 1.12

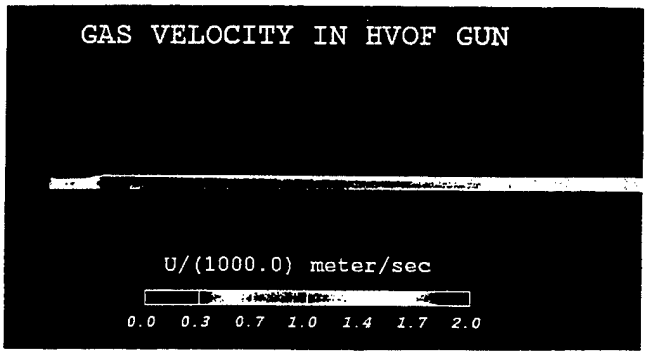


Fig. 2 Velocity contours inside the nozzle and barrel of the gun

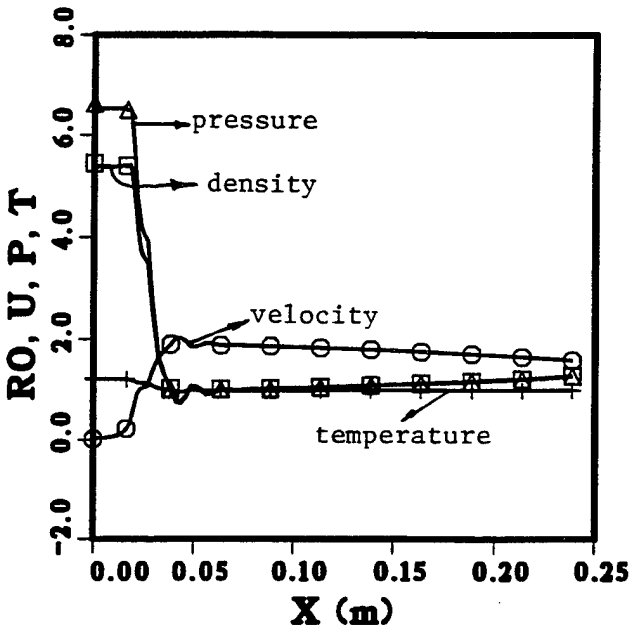


Fig. 3 Properties along the nozzle and barrel centerline. Density (□, ρ/ρ_{Ref}, ρ_{Ref} = 0.121 kg/m^3); velocity (○, u/a_{Ref}, a_{Ref} = 966 m/s); pressure (Δ, p/p_{Ref}, p_{Ref} = 101,325 Pa); temperature (+, T/T_{Ref}, T_{Ref} = 2600 K)

Given these conditions and the gun/barrel geometry, we simulated the steady-state flow regime for the gun de Laval nozzle and barrel, using barrel lengths of 8 and 4 in. The results for the 8 in. length are shown as velocity contour plots in Fig. 2, which shows flow acceleration in the de Laval nozzle from very low subsonic speeds on the order of less than 1 m/s in the combustion chamber to high supersonic speeds on the order of 2000 m/s in the barrel. We also notice that a significant boundary layer develops in the barrel. This leads to a decay of flow velocity along the barrel because of partial constriction of the channel. The nondimensional pressure, temperature, and velocity distribution along the nozzle and barrel centerline is given in Fig. 3. Here the following trends can be noted:

- The pressure and density drop rapidly in the de Laval nozzle, reaching a minimum and then steadily increasing along the barrel length with the reduction of flow velocity.
- The velocity increases to 2000 m/s at the nozzle exit and then gradually decreases to 1700 m/s at the barrel exit as a result of the boundary layer growth.
- The temperature decreases from 3100 K in the chamber to about 2500 K at the barrel exit.

Figure 4 compares the simulation results in the form of velocity distribution at the barrel exit with the experimentally measured values given in Ref 16 and shows an excellent agreement between the computational and experimental results, indicating that our numerical methodology can accurately predict the flow regimes of the TS guns. The maximum deviation between numerical and experimental results shown in Fig. 4 is less than 10%. Figure 5 shows the comparison between computational prediction and experimental measurements of radial distribution of gas temperature at barrel exit. Again, good agreement between experimental and numerical results is observed.

The external flow field is calculated with the flow condition at the exit of the gun barrel as the input boundary condition for jet simulation in the ambient air. Because the pressure at the barrel exit is greater than the ambient pressure, the nozzle flow is categorized to be underexpanded. The flow expands supersonically into the ambient air. The density contour plot for the jet is shown in Fig. 6. The jet flow forms a supersonic core region, generates a so-called diamond shock pattern of expansion and compression waves, forms a free shear layer between jet and external air, and mixes and entrains the ambient air into the jet to slow down the jet from supersonic flow to subsonic flow. Figure 6 shows the jet cross-section growth as a result of mixing with the external air. Currently, the only quantitative experimental measurements of such jets are given in Ref 16, which presents measurements of velocity at the jet centerline. Figure 7 compares the velocity data obtained from our simulation with the experimental measurements given in Ref 16 for the same conditions. We can see from this comparison that experimental observations validate our simulations, and the deviation between the numerical and experimental data is less than 10%.

3.2 *Particle Flow Simulations*

As discussed in the last section, the loading (defined as particle mass flow rate/gas mass flow rate) is low (<4%). The pres-

ence of particle flow will not influence the gas flow field. After calculating the gas flow field, we integrate the equation of particle motion (Eq 2) and the equation of heat transfer (Eq 8) to obtain particle trajectories and temperature histories for different particle diameters and injection speeds. Inconel 718 particles are used in the study. The following particle flow parameters are selected for the test cases:

- Particle material density: $\rho_s = 9000$ kg/m^3
- Particle diameters: $D_p = 10\,\mu m, 20\,\mu m, 40\,\mu m, 60\,\mu m$
- Specific heat: $C_s = 462$ J/kg·K
- Injection angle: $\alpha = 12° \pm 5°$
- Injection speed: $u_j = 20$ m/s

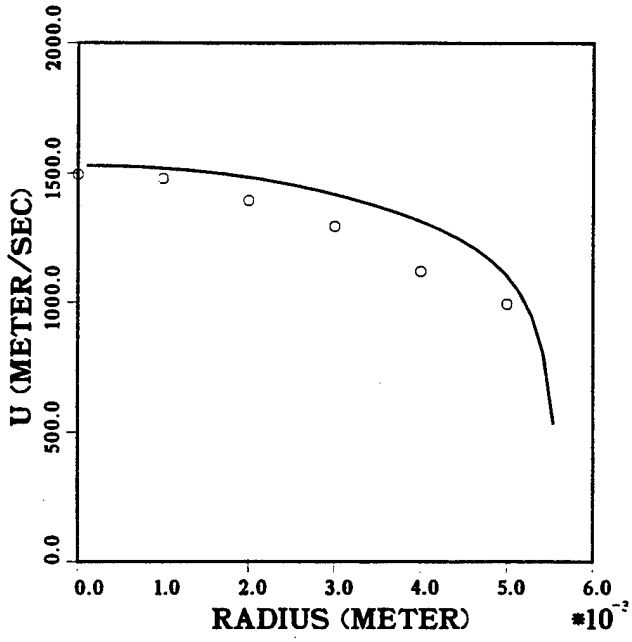


Fig. 4 Comparison between computational prediction and experimental measurement of radial distribution of gas velocity (O, experiment, Ref 16; ——, calculation), at barrel exit

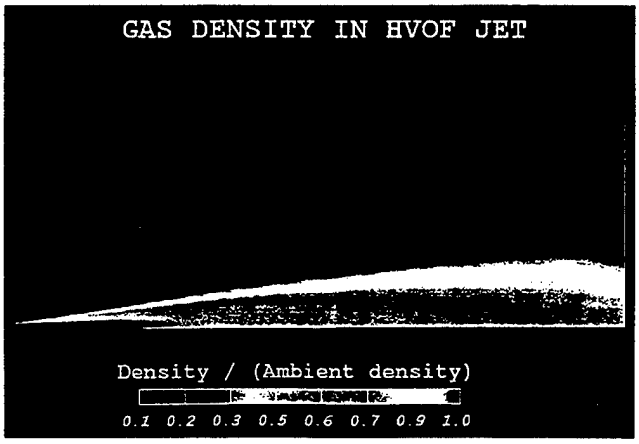


Fig. 6 Density contours in the area of the barrel and the jet for the JP-5000 gun

A large number (>1000) of particle trajectories and temperature histories are traced through the flow field (from nozzle injection to plating distance) in order to obtain mean particle flow properties.

The first set of simulations was done for particles injected into the barrel at a speed of 20 m/s for an 8 in. barrel gun. Be-

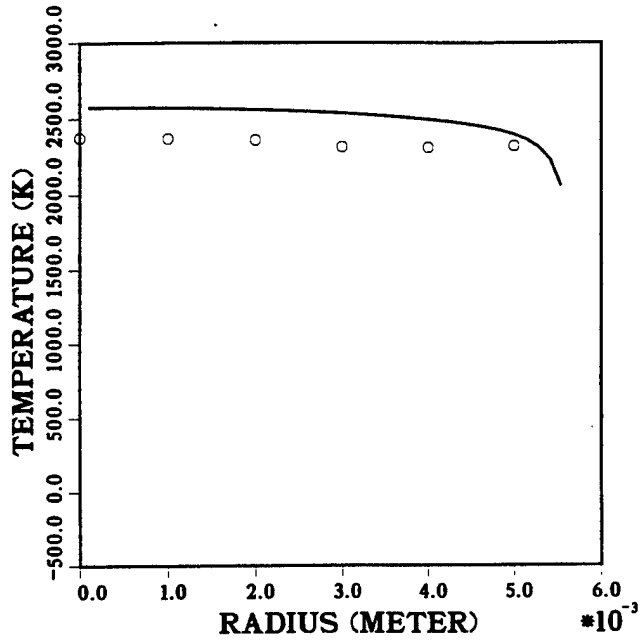


Fig. 5 Comparison between computational prediction and experimental measurements of the radial distribution of the gas temperature at barrel exit (O, experiment; ——, calculation)

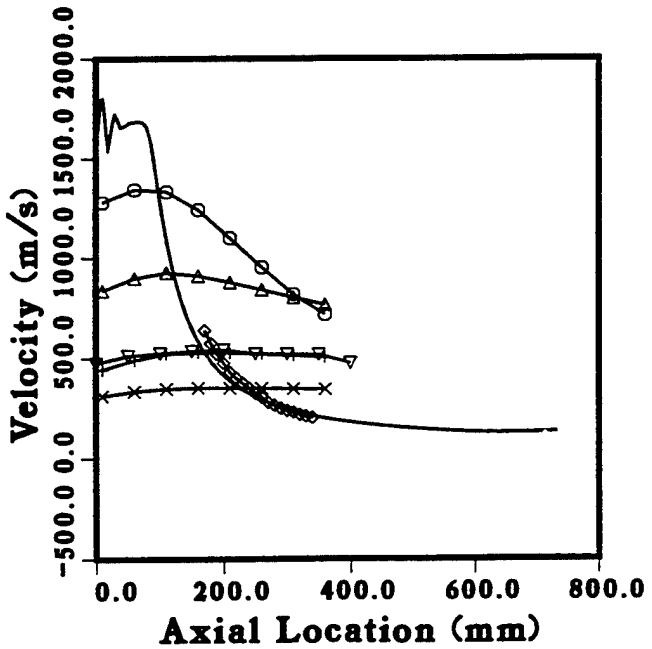


Fig. 7 Comparison between computational prediction and experimental measurement of gas velocity and particle velocity distribution at the jet centerline for the case of 8 in. gun barrel. (◊, experiment, gas velocity (Ref 16); ——, calculation; O, 10 μm; Δ, 20 μm; +, 40 μm; ×, 60 μm; ∇, 35 μm experiment particle velocity) (Ref 18)

cause particle size typically ranges from 10 to 60 μm in TS applications, Fig. 7 shows averaged particle velocities for 10, 20, 40, and 60 μm particles as a function of axial location from the exit of the barrel. The simulations were done for Inconel 718 alloy injected at the rate of 10 lb/h into an 8 in. barrel. Gas velocity is shown in the same figure. The 10 μm particles exit the barrel at a high speed of 1250 m/s and continue to accelerate to speeds as high as 1350 m/s at 100 mm from the barrel exit. However, the particles are then decelerated by gas due to a drastic decrease in the gas velocity. Particles end up at about 700 m/s when they reach the substrate at a normal standoff distance of 360 mm from the barrel. The 20 and 40 μm particles exit from the barrel at speeds of about 820 and 470 m/s, respectively. Due to the greater inertia of these particles, they are accelerated less when gas speed is greater than the respective particle speeds. However, these particles also decelerated less when the gas speed decreases to below the corresponding particle speed. Thus, in Fig. 7 at a distance of 360 mm from the exit, the 20 μm particle will have higher velocity than the 10 μm particle. The 60 μm particle exits from the gun at a speed of about 300 m/s and retains that velocity in the jet flow region. At the standoff distance, all four particles have higher velocity than the gas velocity. There is a large variation in particle velocity as a function of particle size. At the typical plating location, the difference is about 400 m/s for 10 and 60 μm particles. This difference translates into an order of magnitude difference in kinetic energy per unit mass of the impinging particle and obviously will lead to different plating conditions. In general, the understanding is that the higher the velocity, the better the coating. It is believed that the greater velocity simply packs the particle more tightly. These velocity differences for different size particles at the same injection speed (20 m/s) can be explained by the particle inertia and hence particle acceleration/deceleration mechanism. The significant influence of particle size on particle dynamics can be quantitatively seen from Eq 5, which shows that particle acceleration and deceleration relaxation time is dependent on the particle diameter-squared.

To compare with experimental results, we also draw average particle velocity along the centerline measured in Ref 19 in Fig. 7.

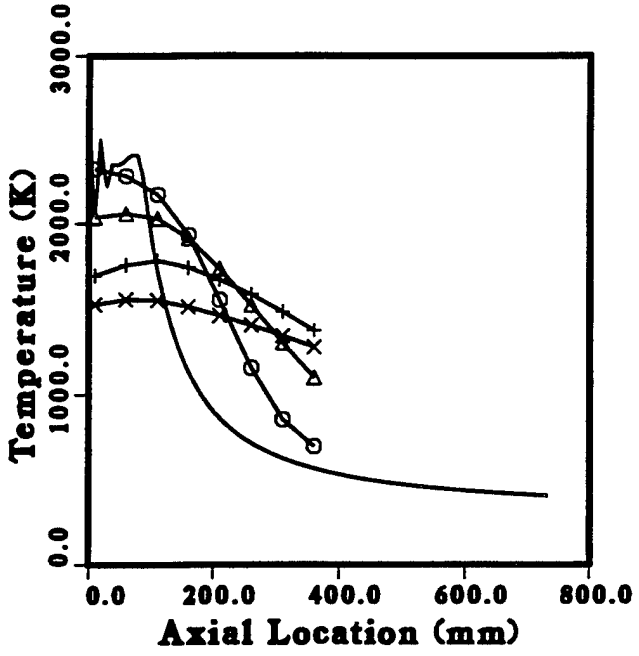


Fig. 8 Gas and particle temperature distribution at the jet centerline for the case of 8 in. gun barrel (——, gas; O, 10 μm; Δ, 20 μm; +, 40 μm; ×, 60 μm)

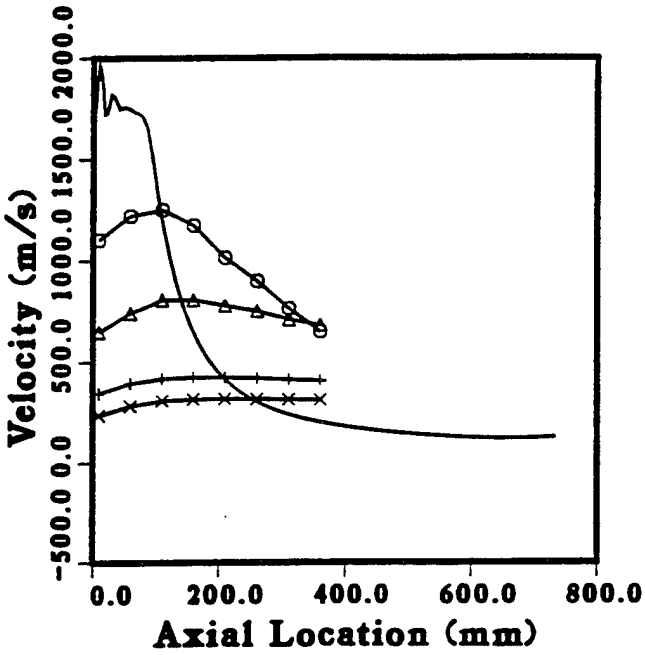


Fig. 9 Comparison between computational prediction and experimental measurement of gas velocity and particle velocity distribution at the jet centerline for the case of 4 in. gun barrel (——, calculation: O, 10 μm; Δ, 20 μm; +, 40 μm; ×, 60 μm)

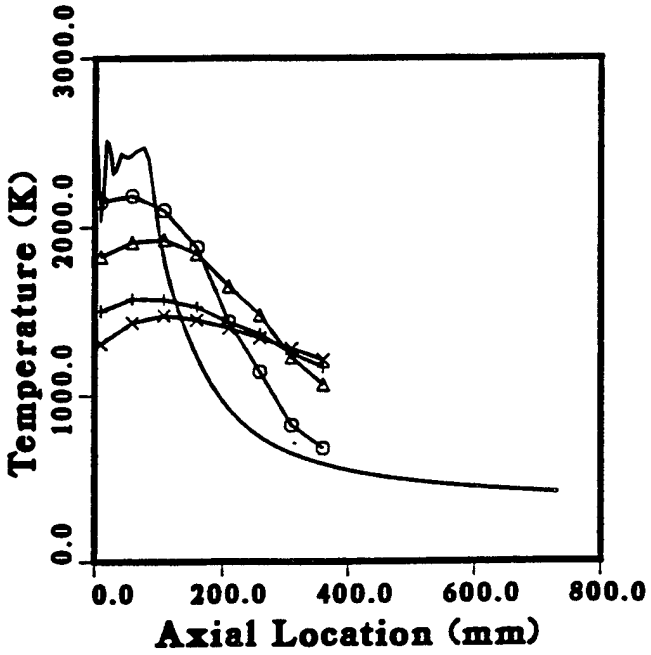


Fig. 10 Gas and particle temperature distribution at the jet centerline for the case of 4 in. gun barrel (——, gas; O, 10 μm; Δ, 20 μm; +, 40 μm; ×, 60 μm)

The number averaged particle size is 35 μm. As we can see, good agreement between experimental measurement and numerical prediction is obtained.

Figure 8 shows particle and gas temperature as a function of axial location. Again, we observe that particle temperatures (except in 10 μm particles, which reach thermal equilibrium at exit of nozzle) rise the first part of the jet, because gas temperature is higher than the particle temperatures. Because the gas temperature decreases sharply in the jet due to gas jet expansion, it drops below particle temperature at the distance of 100 mm from the exit. Particle temperatures are higher than gas temperature at the plating distance. The 10 μm particles almost reach thermal equilibrium with the gas when they reach the exit of the gun. The temperature of 10 μm particles is as high as 2400 K. The 20 μm particles also reach 2000 K at the gun exit. These particle temperatures are well above the melting point of Inconel 718 particles (about 1650 K). Thus, under these conditions, there is a good chance that particles smaller than 20 μm will be deposited

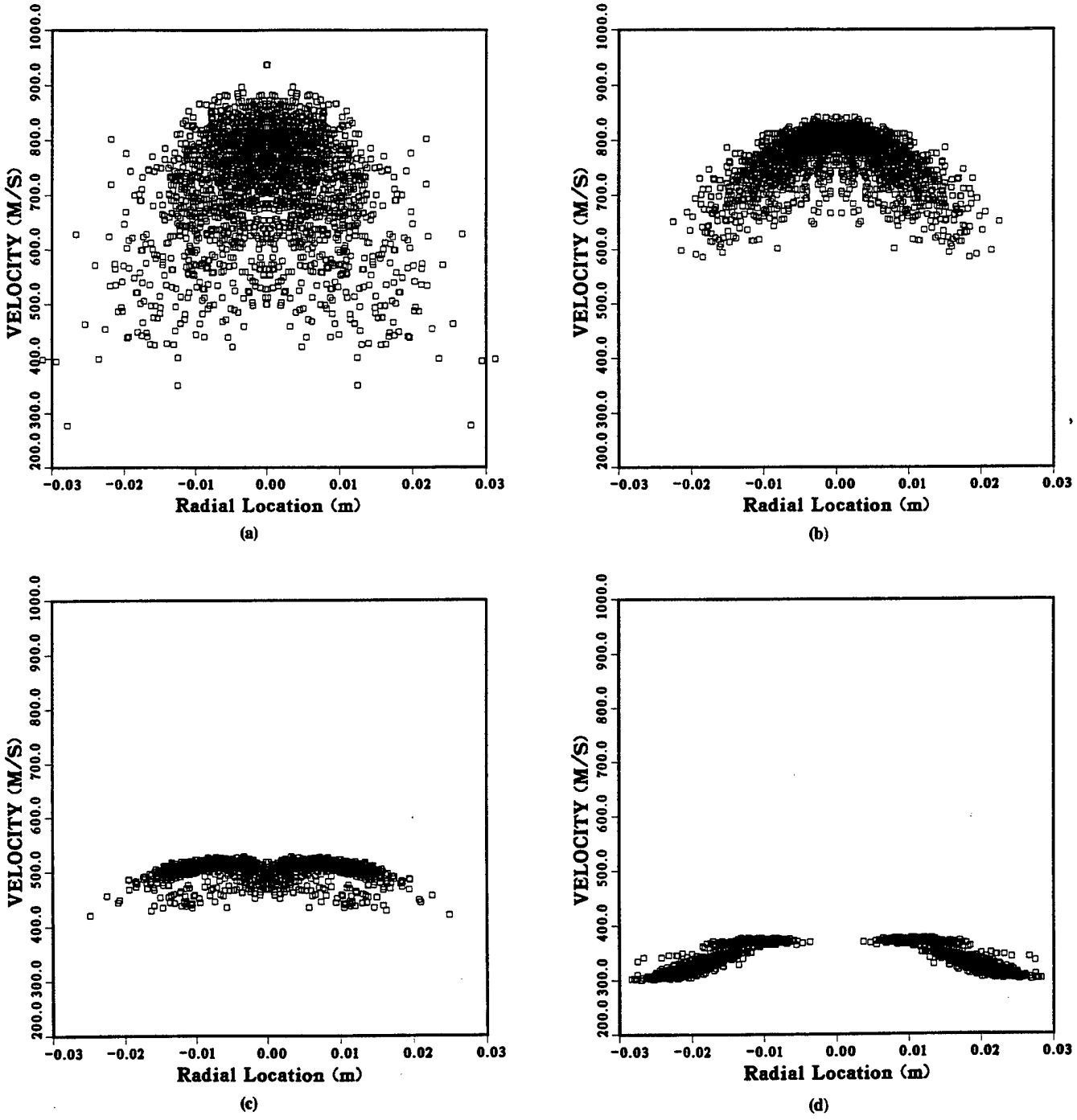


Fig. 11 Radial particle velocity distribution at standoff distance of 36 mm. (a) 10 μm particles. (b) 20 μm particles. (c) 40 μm particles. (d) 60 μm particles

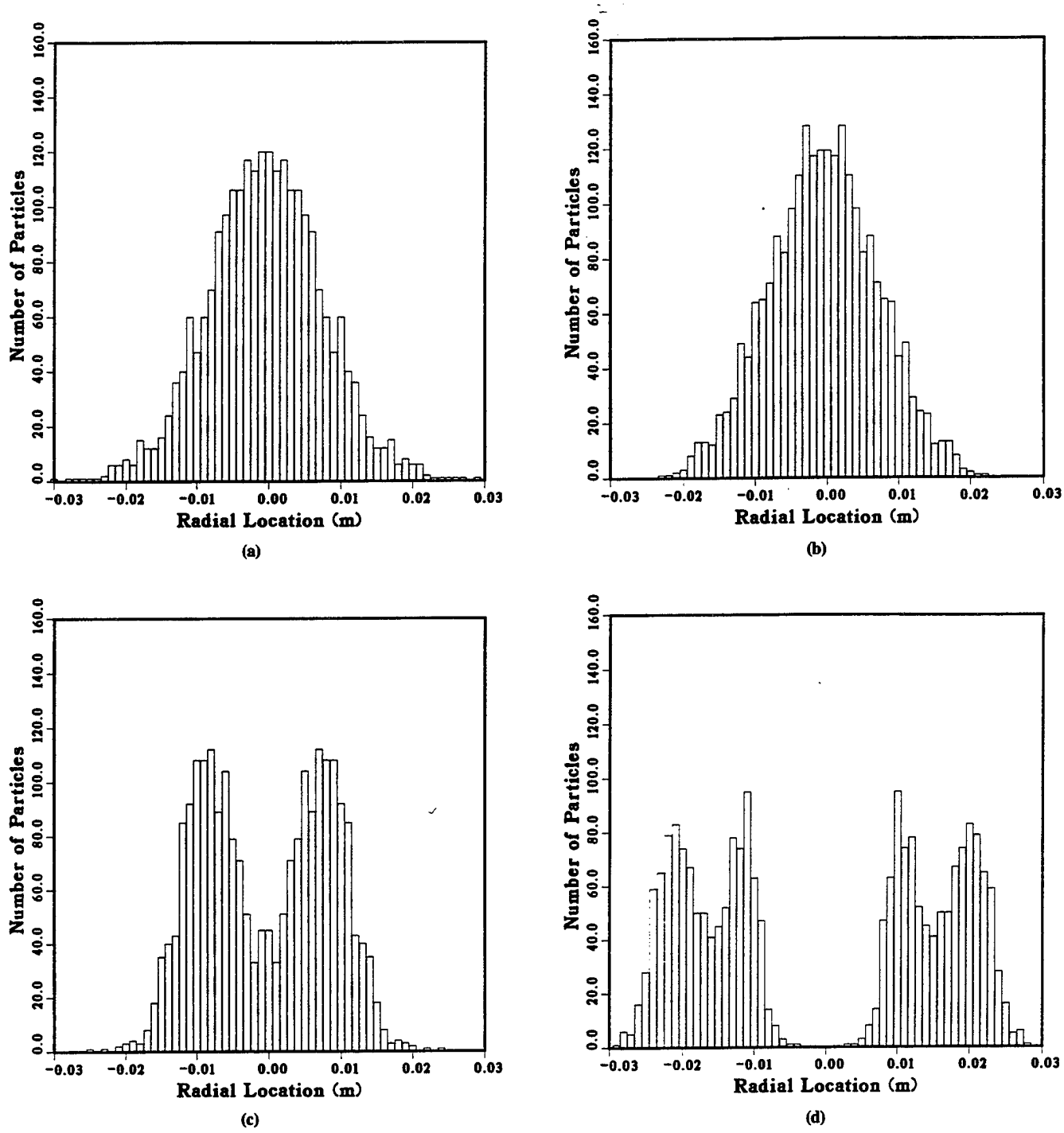


Fig. 12 Particle concentration distribution at cross section 36 cm from the barrel exit. (a) 10 μm. (b) 20 μm. (c) 40 μm. (d) 60 μm

at the barrel surface. The temperature of 60 μm particles may be too low to obtain good coating quality, because they may never reach the melting point during the entire process. At the plating distance, the average temperatures for 10, 20, 40, and 60 μm particles are 700, 1100, 1400, and 1300 K, respectively. At the 350 mm plating distance, 40 μm particles have the highest temperature of about 1400 K, whereas the temperature for 10 μm particles dropped to 700 K due to fast cooling of the small particle in the gas. This temperature difference will significantly affect plating quality. This trend again can be demonstrated through Eq 11, which shows that particle heating is proportional to the square of the particle diameter. However, the temperature differences are somewhat smaller between the different-sized particles because of the longer residence time of the slower moving, larger particles.

As we observe from Fig. 7 and 8, larger particles are more difficult to heat and to accelerate. However, they are also more difficult to cool and decelerate. They maintain their velocity and temperature longer and have greater kinetic energy at impact on the substrate. As observed, large differences in velocity and temperature exist for different particle sizes, which lead to the conclusion that optimal coating conditions can be obtained for a

much narrower range of particle sizes than 10 to 60 μm and that a large amount of coating material is wasted.

To study the effect of the gun barrel length on particle velocity and temperature, a second set of simulations was done for a 4 in. barrel gun. The resulting centerline gas velocity averaged particle velocities are shown in Fig. 9. Gas temperature and average particle temperature as a function of axial location from the gun exit is plotted in Fig. 10. In general, we observe the same trend of variation of particle temperature and velocity as the particles travel through the jet. Careful comparison of Fig. 7 and 9 shows that at the gun exits (axial location 0), the particle velocity of each particle size for the 8 in. barrel is higher than the particle velocity of the same size for the 4 in. barrel. This is expected, because the longer the barrel, the longer the time to accelerate the particles. At the plating distance, the particle velocity of each particle size for the 8 in. barrel case is about 50 to 100 m/s higher than the velocity of the same size particle in the 4 in. barrel case. Comparing the particle temperatures of Fig. 8 and 10, we observe the same general trend for particle temperature variations in both the 8 and 4 in. barrel cases. The variations in the particle velocity and temperature as a function of barrel length may well be used to achieve the optimal condition for different coating materials and requirements.

Simulation of particle interaction with flow microturbulence allows us to predict particle parameter distribution in the cross sections of the jet. Figure 11 displays particle velocity distribution as a function of radius at the standoff distance for particle sizes 10, 20, 40, and 60 μm for the case of an 8 in. gun barrel. This figure shows large variations in particle velocity for small particles at the gun exit, because they are more affected by the local turbulence and velocity variation in the cross section. In general, particles exit from the barrel and spread out in a radial direction as they travel downstream. However, we find that the spreading is highly restricted. At a standoff distance of 36 cm, the particle jet spread or the deposition spot is well defined within less than 30 mm diameter. For 10 μm particles, the velocity varies from 70 to 400 m/s; most of the particles have a velocity higher than 600 m/s. For 20 μm particles, the velocity variation is between 600 and 850 m/s. For 40 μm particles, the velocity variation is kept between 400 and 500 m/s. Finally, for 60 μm particles, the velocity variation is in the range of 300 to 375 m/s. In general, "uniform" or "top hat" distribution of particle velocity is desirable for uniform coating. In Fig. 12, the radial cross section of particle density distribution at the standard plating distance (360 mm) is given in the form of particle numbers at the radial locations. The particle distributions are presented for 10, 20, 40, and 60 μm particle sizes. In this figure, we observe that the particle density distribution is a strong function of the particle size. While 10 and 20 μm particles have a typical Gaussian distribution, 40 and 60 μm particles will produce coating spots or "rings" with most of the material concentrating at the edges. Our preliminary simulation results (Ref 19) show changing the particle injection velocity will allow modification of the particle density distribution; however, it will be very difficult to affect all classes of particles at once in the necessary direction. Thus, improving the distribution for the large particles might worsen it for the small particles.

4. Conclusion

A two-phase flow numerical model is developed and applied to study an HVOF thermal spraying system. The gas flow field and particle trajectories and temperature histories are calculated and presented. Validation of numerical simulation results with experimental data has shown the CFD methodology accurately predicts gas and particle flow fields in the HVOF TS system. A parametric study is conducted for different particle sizes and gun barrel lengths. The quantitative results obtained by this analysis offer a comprehensive, fundamental analysis of the HVOF thermal spray system. It appears that for the JP-5000 system, particle injection velocity should be carefully controlled in order to produce a high-quality coating. We also have shown that particle trajectories and parameters are very strong functions of the particle radius. The developed methodology allows the researcher or engineer to design optimal injection conditions for different particle and flow regimes. This approach can significantly reduce plating development time for new plating systems and improve the quality of the traditional platings.

Acknowledgments

The authors would like to thank Dr. A. Tsao and Dr. E. Oran for their encouragement and interest in this project. The authors would also like to express appreciation to Dr. J. Morrison and Dr. T.B. Gatski of NASA Langley Research Center for their contribution to this work. The work reported here was partially supported by ARPA.

References

1. A.S. Brown, Spraying for Strength, *Aerosp. Am.*, Jan 1992, p 52-56
2. M. Thorpe and H. Richter, A Pragmatic Analysis and Comparison of the HVOF Process, *Thermal Spray Coatings: Research, Design, and Applications*, C.C. Berndt and T.F. Bernecki, Ed., ASM International, 1993, p 157-167
3. E. Oberkampf and M. Talpallikar, Analysis of a High Velocity Oxygen-Fuel (HVOF) Thermal Spray Torch, Thermal Spray Industrial Applications, C. Berndt and S. Sampath, Ed., ASM International, 1994, p 381-386
4. E. Kadyrov, Y. Evdokimenko, V. Kisel, V. Kadyrov, and F. Worzala, Interaction of Particles with Carrier Gas in HVOF Spraying Systems, *J. Therm. Spray Technol.*, Vol 3, 1994, p 389-397
5. C. Crowe and D. Stock, A Computer Solution for Two-Dimensional Fluid-Particle Flows, *Int. J. Numer. Methods Eng.*, Vol 19, 1976, p 185-196
6. P.P. Chen and C.T. Crowe, On the Monte-Carlo Method for Modeling Particle Dispersion in Turbulence, *Symposium Gas-Solid Flows*, American Society of Mechanical Engineers, 1984, p 37-42
7. H. Speziale, R. Abid, and E. Anderson, "A Critical Evaluation of Two-Equation Models for Near Wall Turbulence," Paper 90-1491, American Institute of Aeronautics & Astronautics, 1990
8. J. Morrison, "Flux Difference Split Scheme for Turbulent Transport Equations," Paper 90-5251, American Institute of Aeronautics & Astronautics, 1990
9. P.L. Roe, Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes, *J. Compt. Phys.*, Vol 43, 1981, p 357-372
10. D. Milojevic, Lagrangian Stochastic-Deterministic (LSD) Prediction of Particle Dispersion in Turbulence, *Part. and Part. System Characterization*, Vol 7, 1990, p 181-190
11. M. Sommerfeld, Modeling of Particle/Wall Collision in Confined Gas-Particle Flows, *Int. J. Multiphase Flow*, Vol 18, 1992, p 905-926

12. C.T. Crowe, M.P. Sharma, and D.E. Stock, The Particle-Source-In Cell (PSI-CELL) Model for Gas-Droplet Flows. *ASME J. Fluids Eng.*, Vol 99 (No. 2), 1977, p 325-332

13. P.J. Thomas, On the Influence of the Basset History Force on the Motion of a Particle through a Fluid. *Phys. Fluids A*, Vol 4, 1992, p 2090-2093

14. R. Clift, J.R. Grace, and M.E. Weber, *Bubbles, Drops and Particles*, Academic Press, 1978

15. I. Gary, Higher Velocity Thermal Spray Processes Produce Better Aircraft Engine Coatings, JAE 28th Ann. Aerospace/Airline Plating & Metal Finishing Forum & Exposition (San Diego, CA), April 1992

16. W.D. Swank, J.R. Fincke, D.C. Haggard, and G. Irons, HVOF Gas Flow Field Characteristics, *Thermal Spray Industrial Applications*, C. Berndt and S. Sampath, Ed., ASM International, 1994, p 313-318

17. S. Gordon and B. McBride, "Computer Program for Calculation of Complex Chemical Equilibrium Compositions, Rocket Performance, Incident and Reflected Shocks, and Chapman-Jouguet Detonations," SP-273, National Aeronautics and Space Administration, May 1976, new version, 1989

18. W.D. Swank, J.R. Fincke, and D.C. Haggard, HVOF Particle Flow Field Characteristics, *Thermal Spray Industrial Applications*, C. Berndt and S. Sampath, Ed., ASM International, 1994, p 319-324

19. S. Eidelman and X. Yang, Optimization of the Thermal Spray Guns and Coating Processes Using Numerical Simulation, Annual Meeting and Exhibition (Anaheim, CA), Feb 1996, TMS

Appendix LL

Use of Thermal Spray Methods for Coating Nanoscale Materials

USE OF THERMAL SPRAY METHODS FOR COATING NANOSCALE MATERIALS

Shmuel Eidelman and Xiaolong Yang
Science Applications International Corporation
1710 Goodridge Drive
McLean, Virginia

Abstract—*Thermal Spray (TS) systems are a low cost versatile technology for significantly improving the properties of base materials by coating them with high performance materials. TS coating methods are simple and can be used for coating large surfaces at atmospheric pressures with a range of materials from metal alloys and cermets to ceramics and polymers. TS guns allow delivery of the plating materials to the substrates at high velocities (on the order of 1 to 1.5 km/sec) and high temperatures. The total pressure of the particles impinging on the surface can reach 10 GPa for some of the TS systems. Lower sintering temperatures and higher ductility of nanoscale materials open a range of attractive and unique possibilities for high rate deposition of nanostructured coatings. However, the current TS systems are very inefficient in their use of energy and materials, as well as in quality control. In addition, barrel erosion and barrel coating lead to additional equipment maintenance costs and problems with coating consistency. We use a recently developed and validated three dimensional simulation capability to model the TS systems' gas and coating powder flow for the TS process analysis, and to illustrate the roots of the TS systems' inefficiencies. The same capability can be used to design optimized TS systems, and to optimize and control the coating process. Examples are given for TS process designs with improved performance and system efficiency. The use of numerical simulation will be especially crucial for plating with nanoscale powders. We will discuss specific equipment and process solutions that will make nanoscale powers coating a viable industrial process.*

INTRODUCTION

Nanostructured materials have demonstrated a set of unique properties that make them attractive for their industrial applications. However, most of the nanomaterials are produced in powder form and require additional processing steps for creation of useful products such as coating, bulk single phase, or composite material sintering. Most of the nanoscale materials properties were studied for small samples produced by high pressure and high temperature sintering methods. The high pressure/high temperature sintering of the nanoscale powders has a set of distinct attributes. When some of these attributes make sintering nanostructured materials easier (lower sintering temperature), the small particle size

leads to high pressure requirements for the low porosity green body compaction. This limits the utility of the high pressure nanopowder sintering to relatively small parts. Using Thermal Spray processing to coat nanopowders offers a high rate deposition method that can provide both the effective pressure and temperature required to sinter high density nanostructured materials. For example, an agglomerate of WC/Co nanoparticles having 60% porosity accelerated to 1500 m/sec velocity impinging on a wall, generates pressure equivalent to 6.8 Gpa; this is sufficient to sinter high density WC/Co coatings. We will discuss the distinct characteristics of Thermal Spray deposition of nanoscale materials. We will address the issues of nanoscale materials coating by numerically simulating TS processes, which allows us to examine possible regimes for nanostructured coating deposition.

The High Velocity Oxygen-Fuel (HVOF) Thermal Spray (TS) system is used in the chemical, tooling, automotive and aerospace industry. HVOF TS systems use the velocity and temperature of a supersonic jet to accelerate and heat the metal or ceramic particles; that subsequently impinges on the surface of a substrate, and produces coatings at a high deposition rate. The quality, consistency, and deposition rate of the coatings depends on a set of particle, gas flow, and substrate parameters. In Figure 1, we show the schematic diagram of an HVOF system consisting of a combustion chamber with a fuel/oxidizer injection system; a de Laval nozzle, a barrel, particle injection ports, and a water cooling system. The diagram corresponds to the TAFA JP-5000 gun used in our analysis. High pressure and temperature combustion gases, resulting from the combustion of oxygen and fuel, expand through the converging and diverging de Laval nozzle and the barrel, to supersonic speed. This supersonic flow expands into atmospheric air, creating a supersonic jet that mixes with the air and quickly decays from supersonic to subsonic velocities. In the TAFA JP-5000 gun, the solid particles of metal or ceramic powders are injected into the barrel at the exit of the nozzle from two small inlets located on the barrel's wall. Particles are turbulently mixed, heated and accelerated in the barrel and jet, to high speed and temperature, and move towards the substrate where they are deposited in a coating layer. From the fluid dynamics point of view, the system is very complex and involves two phase (gas-particle) flow with turbulence, heat transfer, chemical reactions and supersonic/subsonic flow transitions. In an engineering application, the microstructure and physical properties of the plated surface are determined by the physical and chemical conditions of the particles that impinge on the surface. The physical and chemical conditions of the particle impinging on the substrate in turn are dependent on a large number of parameters such as gun design (length of barrel in particular), the gas jet formation, the position of the particle relative to the substrate, and the particle parameter such as size, shape, material and how it is injected into the system.

Here we use methods of Computational Fluid Dynamics to conduct a comprehensive analysis of the HVOF TS system. Direct simulations of the TS processes will allow process optimization, new process and coating tool design, and control design for the TS system. Below we describe the methodology used for the HVOF system analysis, present some validation results, and provide a process optimization footpath for the nanomaterials coating that is the result of our detailed TS system analysis.

The analysis presented below is based on a comprehensive model of the HVOF TS system that numerically simulates the gas dynamics of the gas expansion and flow in the barrel and in the free jet; particle injection and the interaction between the injected particle and the

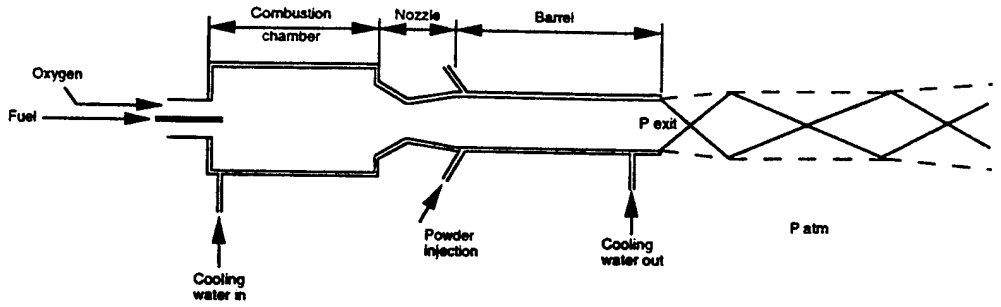


Figure 1. Schematic of the high pressure HVOF gun and flow field.

turbulent gas flow inside and outside the barrel, and; particle interaction with the barrel walls. We will show that our model allows accurate prediction of the particle conditions before they impinge on the substrate. Here we only give a short description of the mathematical model and numerical solution method. The details of the mathematical model, numerical solution method, and validation results are presented elsewhere (9). Our paper will focus on using this methodology to analyze the TS system operation requirements for nanomaterials coatings.

MATHEMATICAL MODEL AND NUMERICAL SOLUTION

The mathematical model for the two-phase flow consists of conservation-governing equations and constitutive relations that provide closure models. For the gas flow we use the Eulerian approach for analyzing the continuum. We use compressible Navier-Stokes equations in Favre mass-averaged variables for the conservation equations of mass, momentum, and energy describing the gas flow. To model flow turbulence, the Reynolds stress tensor is closed using a k-ε turbulence model (1). Here, k is turbulent kinetic energy and ε is the rate of dissipation of kinetic energy. A recently developed (2) three-dimensional, multiblock, upwind, fully implicit, finite volume code was modified to solve the governing equations. The code incorporates the high-order, upwind, flux-difference splitting processor of Roe (3) with a MUSCL integration scheme to obtain good shock-capturing, high-accuracy solutions in general coordinates for the 3-D geometry computational domain. For a 3-D flow, seven variable governing equations are discretized and integrated in time to provide a steady-state solution. To deal with the well-known stiffness of the equations, an implicit time integration is used. The details of the code and numerical solution techniques can be found in (1).

To describe particle motion, we use the Lagrangian formulation. The Lagrangian formulation allows a description of the particle/wall interactions that is very difficult to analyze numerically using the Eulerian approach. In addition, Lagrangian formulation is also logical for the low-volume, collisionless particle flow that is typical of TS systems. To simplify the analysis, the following assumptions are made: 1. the particles do not undergo a phase change; 2. the particles are solid spheres and have a constant material density; 3. the

volume occupied by the particles is negligible; 4. the interaction between particles can be ignored; 5. the only force acting on a particle is drag force and the only heat transfer between the two phases is convection; 6. the weight of the solid particles and their buoyancy force are negligible compared to the drag force; and 7. the particles have a constant specific heat and are assumed to have a uniform temperature distribution inside each particle.

Further, because the loading of the particle defined as total mass flow rate/total gas mass flow rate is very low (<4%), one can assume that the presence of particles will have a minimal effect on gas velocity and temperature field. This means that the momentum and heat exchanges from particle to gas are too small to change gas velocity and temperature distribution. As a result of the assumptions presented above, the two-phase problem can then be decoupled: the gas flow field can be simulated first, followed by the particle flow analysis. The simulation of the particle flow field consists of calculating particle trajectories and temperature histories in the gun barrel and in the jet, after particles are injected into the gun barrel from the particle injection port. The interaction between particle and wall is also included in the simulation. Particle motion in gas turbulent flow is predicted by means of the Lagrangian stochastic model (LSD) (4, 5).

MODEL VALIDATION

To validate the two-phase flow model and numerical solution procedure, we selected the TAFA JP-5000 HVOF TS system (6) as the test case, since extensive experimental data are available for this system. First, we resolved the gas flow field in the barrel. We assumed in our simulations that the gas has a frozen chemical composition for a given fuel/oxygen ratio, which can be calculated using the NASA Gordon-McBride program (7). We simulated the JP-5000 gun flow for the conditions given in (8), for which the experimental measurements were made. We simulated the steady-state flow regime for the gun's de Laval nozzle and barrel, using barrel lengths of 8 inches, and the expansion of the hot gas into atmospheric air. In Figure 2a, experimental data from gas velocity distribution on the centerline of the jet is compared with our numerical data. We observe that the experimental and simulation results compare well, with the difference between them being less than 10%.

In the TS systems, the loading (defined as particle mass flow rate/gas mass flow rate) is low (<4%). The presence of particle flow will not influence the gas flow field. After calculating the gas flow field, we integrate the equation of particle motion (9) and the equation of heat transfer (9) to obtain particle trajectories and temperature histories for different particle diameters and injection speeds. Following the experimental results presented in (8), we used Inconel 718 particles in our numerical study. A large number (> 1000) of particle trajectories and temperature histories are traced through the flow field (from nozzle injection to plating distance) in order to obtain mean particle flow properties. In Figure 2b, simulation results for the 40 μm particle axial velocity distribution are compared with the experimental results given in (8) for the same gas stream conditions. The experimental and computational results compare very well with less than 5% deviation, which is well within the accuracy limits of the experiments.

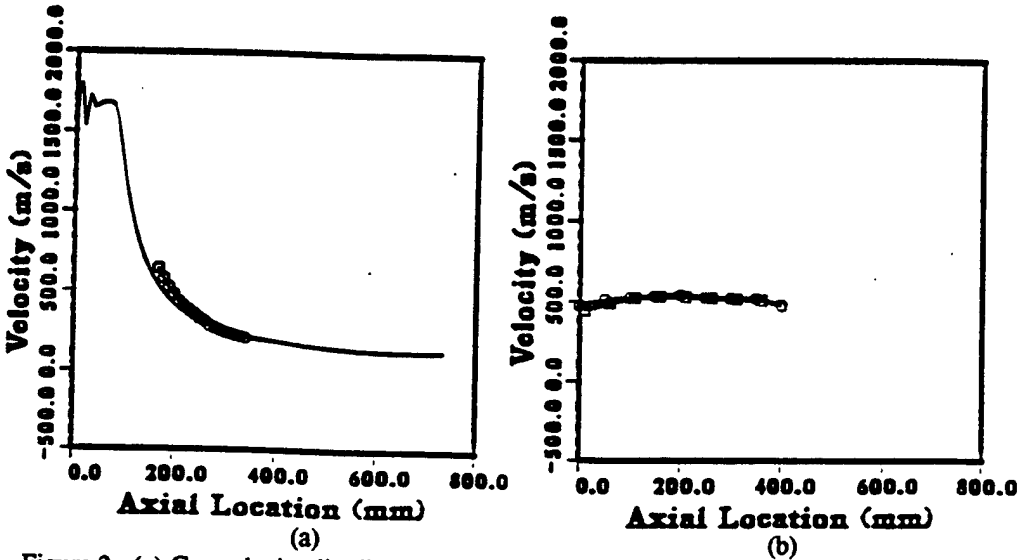


Figure 2. (a) Gas velocity distribution at the jet centerline (o experimental, — calculation). (b) particle velocity distributions at the jet centerline (o experimental, □ - calculation).

NANOSCALE MATERIALS COATINGS

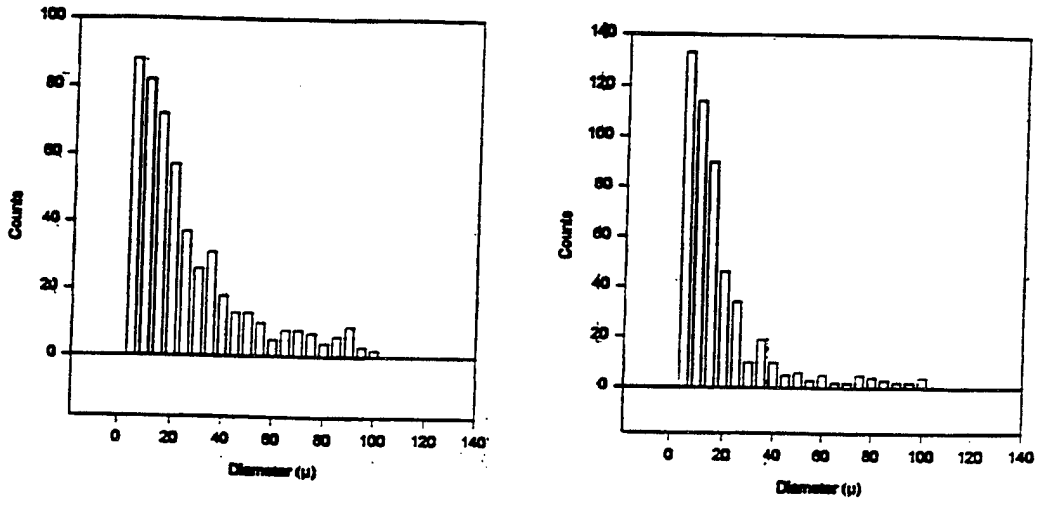
There is an immediate problem in coating nanomaterials. Most "as is" nanoscale powders cannot be coated, because very small particles will closely follow the stream lines of the carrying gas. Thus, when the TS gas jet is impinging on a substrate surface, very small particles will be slowed down and diverted by the flow in the stagnation region. Assuming that particles are accelerated (decelerated) by viscous drag, we estimate that a 1 μm Inconel 718 particle in 1500 m/sec gas flow will have a relaxation zone (the distance that it takes to accelerate a particle from a quiescent state to 86% of the gas velocity) of 1.3 cm and a 0.5 μm particle will have a relaxation zone of only 0.35 cm. These distances are on the order of the stagnation zone thickness for typical jet/substrate interaction. Thus even if you could accelerate small particles to velocities of the order of 1500 m/sec, they will lose most of their kinetic energy in the stagnation zone. Another practical difficulty is feeding (injecting) the small particles into the gas stream. Particles smaller than 10 μm are extremely difficult to feed into the gas flow and can result in plugged particle feed lines because of particle agglomeration.

One solution to this problem was proposed by P. Strutt and B. Kear (10). According to proposed methodology, nanosize particles can be processed to form controlled size agglomerates by spray drying of nanopowder/binder suspension. The nanosize particles in the agglomerates are retained by the vander Waals force as well as by the binder. During the coating process in the high temperature jet, the binder (which must be chosen so as not to contaminate the coating) evaporates and agglomerated particles can be coated. Recently this

concept was successfully demonstrated by P. Strutt (10) for a WC/Co system, where a fully dense single phase high quality WC/Co deposit was obtained in a Browning gun HVOF system. In this case an additional step in powder preparation was used. The nanoscale WC/Co powder, in the form of hollow shells, was processed first to break the shells. The fragments were mixed with a binder and spray dried to produce 10 μm to 50 μm agglomerates, which were then used for TS coatings.

Another method for coating nanostructured WC/Co particles is to inject 30 μm to 50 μm hollow spheres directly into the TS guns barrel. The large spheres will feed well due to their shape and size; however, they will quickly disintegrate into micron size fragments in the supersonic gas stream of the TS jet. In Figures 3a and 3b, size distribution of the WC/Co fragments is given for the axial location at a distance of 200 mm and 300 mm from the gun's barrel. Here we can observe that at 200 mm, most of the particles are smaller than 15 μm (23 μm average size). Particles continue to brake in the jet and are smaller at 300 mm than at 200 mm. At 300 mm distance most of the particles are smaller than 10 μm and the average size is 16 μm. In our preliminary experiments we obtained good quality WC/Co using thermally sprayed nanostructured WC/Co powders in the form of hollow spheres agglomerates.

Using particle agglomeration, we can prepare different size agglomerates with varying average density. What is the optimal strategy in preparing the agglomerated nanopowders for spray coatings? In Figures 4a and 4b, simulation results for a number of particle sizes for the case of WC/Co agglomerates with average particle density of 4000 kg/m^3 are shown in the form of velocity and temperature distributions at the jet axis of symmetry.



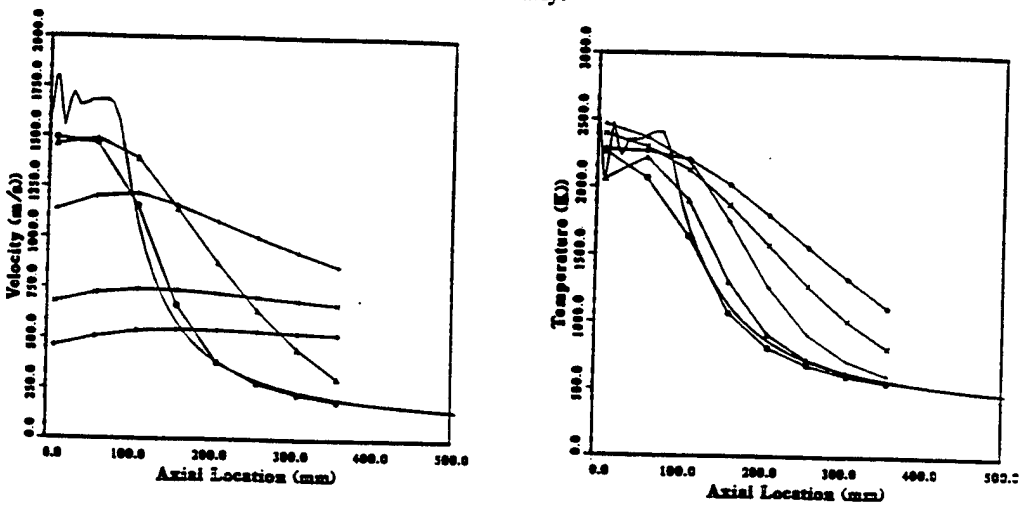
(a) (b)

Figure 3. WC/Co particle size distribution: (a) 200 mm from the TS guns barrel; (b) 300 mm from the guns barrel.

In Figure 4 we observe the following trends:

1. Large particles do not accelerate to high velocities. 60 μm particles will have a maximum velocity of 1/4 to 1/3 of the maximum velocity of the gas jet. Because the agglomerates are porous, and require higher pressures to produce dense coatings, 60 μm TS coatings will lead to porous coatings.
2. Particles smaller than 20 μm can be accelerated to velocities higher than 1000 m/sec; however, particles smaller than 10 μm decelerate rapidly in the expanding jet. This can lead to the condition where a very small change in the distance between the gun's barrel and the substrate produces a large change in the speed of particles impinging on the surface and therefore also changes the plating conditions. Particles smaller than 5 μm will be impossible to coat because they closely follow the gas stream lines and will decelerate in the stagnation region.
3. As a result of relatively low heat capacity and low bulk density of the WC/Co agglomerates, all classes of particles shown in Figure 4b will probably overheat in the gun barrel and rapidly cool-down in the jet. Even larger 60 μm particles will have a 50 K/cm rate of temperature reduction at the 200 mm distance from the barrel. Thus, coating these particles will require precise definition of the plating window.

Numerical simulation of the TS process allows a detailed examination of the range of parameters that affect plating of the nanoscale powders, and aid in the understanding of plating conditions. Our example was only focused on the velocity and particle temperature as function of particle size and thermodynamic parameters. Other important aspects such as substrate temperature as a function of geometry and distance form the barrel; TS gas jet conditions as a function of combustion regime; effects of the turbulent flow regimes on the size of the plating spot; gun barrel and nozzle design; and robotics arm trajectory assignment for uniform coatings, can be also addressed numerically.



(a) (b)

Figure 4. Gas and particle velocity (a) and temperature (b) distribution at the jet centerline for WC/Co agglomerates (— gas, o 5 μm, Δ 10 μm, + 20 μm, x 40 μm, and □ 60 μm).

CONCLUSIONS

The use of a Thermal Spray system to deposit high quality nanostructured coatings will allow to produce functional material coatings over large surfaces at low cost. However, to capitalize on the advantages of nanostructured materials, the coating regimes should be significantly different from the micronscale particles, and a number of challenges should be overcome. Numerical simulations of the HVOF thermal spraying system deposition process, offer a powerful tool for plating process development for this new class of coatings. We have presented a comprehensive, fundamental analysis of the HVOF thermal spray system that addresses a full spectrum of parameters that define TS system performance and allow consideration of the flow regimes and conditions that are outside the conventional operation regimes. We have demonstrated that use of the numerical simulation methodology permits the detailed analysis of a range of plating conditions that can be encountered when coating nanostructured WC/Co. According to our analysis, successful plating of high density WC/Co requires preparing 10 μm to 20 μm WC/Co agglomerates that should be stable and not disintegrate in supersonic flow. Also it will require lower temperature regimes to prevent grain growth, particle disintegration due to cobalt melting, and substrate overheat. Presented here results are given for simulation capability demonstration; in an applied study, the simulations should be linked with an experimental study for results validation.

Thermal Spray (TS) systems are low cost versatile technology for significantly improving the properties of the base materials by coating them with the high performance materials. The current TS systems are very inefficient in their use of energy and materials, as well as in quality control. In addition, barrel erosion and barrel coating lead to additional equipment maintenance cost and problems with coating consistency. The methodology illustrated above allows us to study the roots of TS system inefficiencies. The same capability can be used to design optimized TS systems, and for coating process optimization and control. Our approach can significantly reduce plating development time for new plating systems and improve the quality of traditional platings.

ACKNOWLEDGMENT

The work reported here was partially supported by DARPA.

The authors would like to thank Dr. A. Tsao and Dr. E. Oran for their encouragement and interest in this project.

REFERENCES

1. H. Speziale, R. Abid, and E. Anderson, "A Critical Evaluation of Two-Equation Models for Near Wall Turbulence," AIAA Paper 90-1491, (1990).

2. J. Morrison, "Flux Difference Split Scheme for Turbulent Transport Equations," AIAA Paper 90-5251, (1990).

3. P. L Roe, "Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes," J. of Compt. Phys., 43 , 357 (1981).

4. D. Milojevic, "Lagrangian Stochastic-Deterministic (LSD) Prediction of Particle Dispersion in Turbulence," Part. and Part. System Characterization, 7, 181 (1990).

5. M. Sommerfeld, "Modeling of Particle/wall Collision in Confined Gas-particle Flows." Int. J. Multiphase Flow, 18, 905 (1992).

6. I. Gary, "Higher Velocity Thermal Spray Processes Produce Better Aircraft Engine Coatings," (JAE 28th Ann., Aerospace/Airline Plating & Metal Finishing Forum & Exposition: San Diego, CA, April 1992).

7. S. Gordon and B. McBride, "Computer Program for Calculation of Complex Chemical Equilibrium Compositions, Rocket Performance, Incident and Reflected Shocks, and Chapman-Jouguet Detonations," (NASA SP-273, May 1976, new version, 1989).

8. W. D. Swank et al., "HVOF Gas Flow Field Characteristics," (Proc. of the 7th National Thermal Spray Conf., C. Berndt and S. Sampath, Eds.), 313 (June 1994).

9. Yang and S. Eidelman, "Numerical Analysis of the High Velocity Oxygen-Fuel (HVOF) Thermal Spray (TS) System," Accepted for Publication in Journal of Thermal Spray Technology, (1996).

10. P. Strutt and B. H. Kear, Private communications, (1996).

Appendix MM

A Comprehensive Model of Plasma Etch Reactors

A Comprehensive Model of Plasma Etch Reactors

Presentation at the

Applied Computational Mathematics Program Review

Section 3: Materials

to

Dr. Anna Tsao

DARPA/DSO

Presented by

Ellis Hyman

June 13, 1996

Summary of Presentation

- SAIC Modeling/University/Industry Team and Program Goals
- Overview of Major Etch Model Elements
- Frequency Domain Solver
- Status of Low Pressure (PMT) Reactor Model/PMT Simulation
- Status of High Pressure CDE (GaSonics) Reactor Model/GaSonics Simulations
- Investigation of Effects of Impurities on CDE Reactor/Surface Effects
- Investigation of Effect of Nozzle Placement on PMT Reactor Performance
- Technology Transfer Activities

SAIC®
An Employee-Owned Company

Comprehensive Model of Etch Reactors

Performing Organizations

- Science Applications International Corporation
- Cornell University
- GaSonics International, Inc.
- Plasma & Materials Technologies, Inc. (PMT)

Contracting Agent:

- Lab for Computational Physics and Fluid Dynamics Naval Research Laboratory, Washington, D.C.

Contract Objectives

- Develop a detailed computational model of Plasma Etch Reactors valid over a broad range of operating parameters (pressure, input power, flow rate, and geometry).
- Identify important physical mechanisms that govern the performance of Plasma Etch Reactors and validate model assumptions and algorithms with experimental data.
- Construct simplified codes for use by industry for design and control of Etch Reactor performance.

Plasma Etch Reactor Code (PERC) Status

MAJOR MODULES

• Plasma and Neutral Chemistry	Chlorine System for PMT / Oxygen, Nitrogen for GaSonics.
• Direct Simulation Monte Carlo (DSMC)	Parallelized Code Completed. New Nozzle Designs Simulated.
• Boltzmann Electron Kinetic Energy Distribution	Code Completed. Extensively Tested for Continuum System.
• RF Energy Deposition	New Algorithm Developed. In testing for PMT Reactor.
• Continuum Plasma and Neutral Flow	Modules for Conductivity, Diffusion, Viscosity Completed.

Overview of Computational Program

- Calculation of Electromagnetic Field including Plasma Response
- Electron Energy Distribution - (High Pressure) Local - Boltzmann code,

  (Low Pressure) Along Magnetic Field Lines - drift particle orbits, Monte Carlo collisions
- Ion Motion - (High Pressure) No Direct Role, (Low Pressure) Ambipolar Diffusion
- Neutral Motion - (High Pressure) Navier-Stokes, (Low Pressure) DSMC
- Chemistry - Plasma and Neutral
- Surface Effects - Substrate and Wall

Separation of Time Scales in Plasma Reactors

- On the fast time scale (RF cycle) the RF field pattern and the electron energy distribution change very slowly. The RF field structure is well approximated by linear theory. The field is linear but 'dressed' by the spatial plasma profile.

$$\text{Quiver velocity} = (qE/m\omega) << v_{th}$$

- In any one cycle of the RF field an electron gets a small kick in energy, $\Delta\varepsilon$

$$\Delta\varepsilon/\varepsilon << 1$$

- In the high pressure reactor, the electron gains and loses energy locally;
- In the low pressure reactor, the electrons are constrained by the static B field and bounce back and forth repeatedly through the region of high RF E field, where they get their energy kick.
- In both cases the energization process is diffusive in velocity space.

SAIC®
An Employee-Owned Company

Frequency Domain Electromagnetic Solver for Plasma Reactors

- Straightforward time integration for the long time (steady state) behavior of the reactor process is impractical due to wide range in process time scales.
- Frequency Domain (FD) solutions for the microwave modes can provide the self consistent electric field in the presence of plasma and a given background gas.
- The chemistry/diffusion/convection/heat transport can then be time integrated on their characteristic (and much longer) time scales.
- An FD solver that includes the plasma response terms in Maxwell's equations generates an operator that is non-positive-definite and non-diagonally dominant.
- Standard techniques for second order PDE's do not provide a converging solution.

SAIC®
An Employee-Owned Company

PMT Reactor Code Status

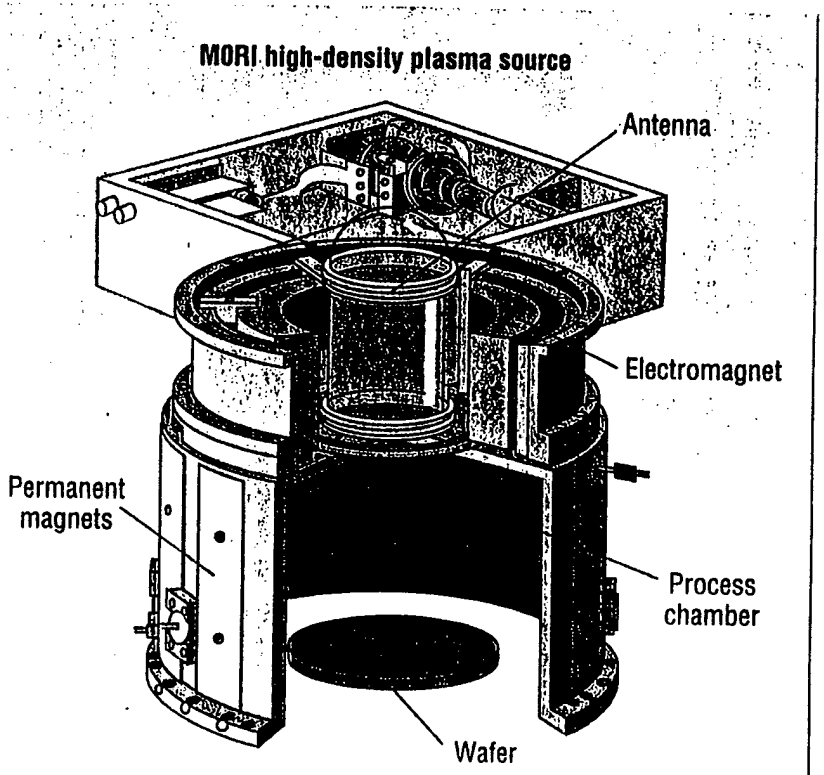
- Calculate RF E Field in Plasma using Frequency Domain Algorithm
- Calculate Typical (Low Energy) Electron Orbits in B Field using Particle Code [Electrons Constrained by Static B Field; Energized by RF E Field]
  - no collisions, no RF field
  - collisions, no RF field
  - collisions, RF field
- Perform DSMC Calculation with Chemistry for Ions, along B Field, and Neutrals

<u>Next Steps</u>

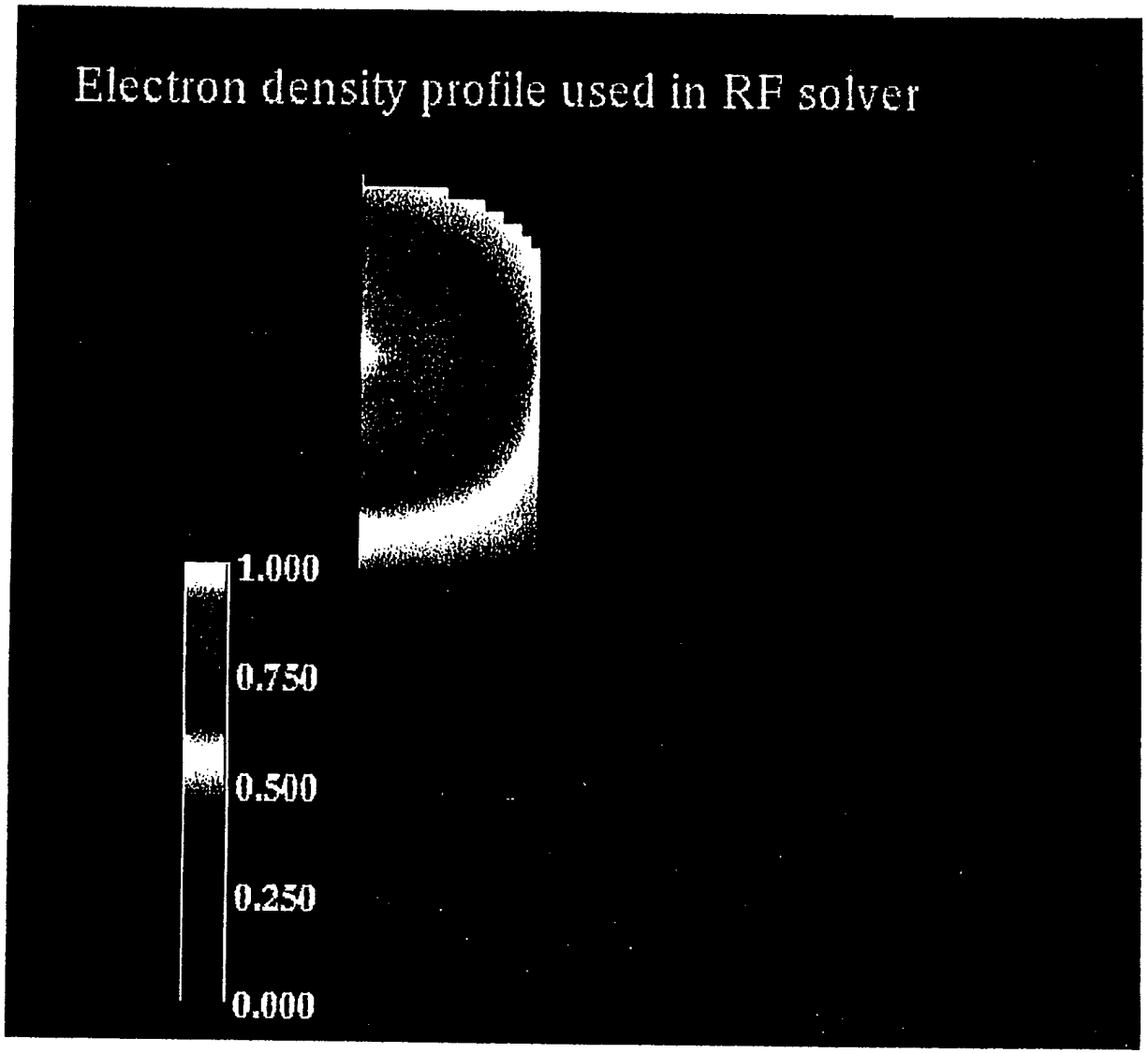
- Use Particle Code for Electron Motion: Include Effects of Ion Distribution (Ambipolar Potential), Static B Field, RF E Field, and Collisions
- Integrate into DSMC Code

SAIC®
An Employee-Owned Company

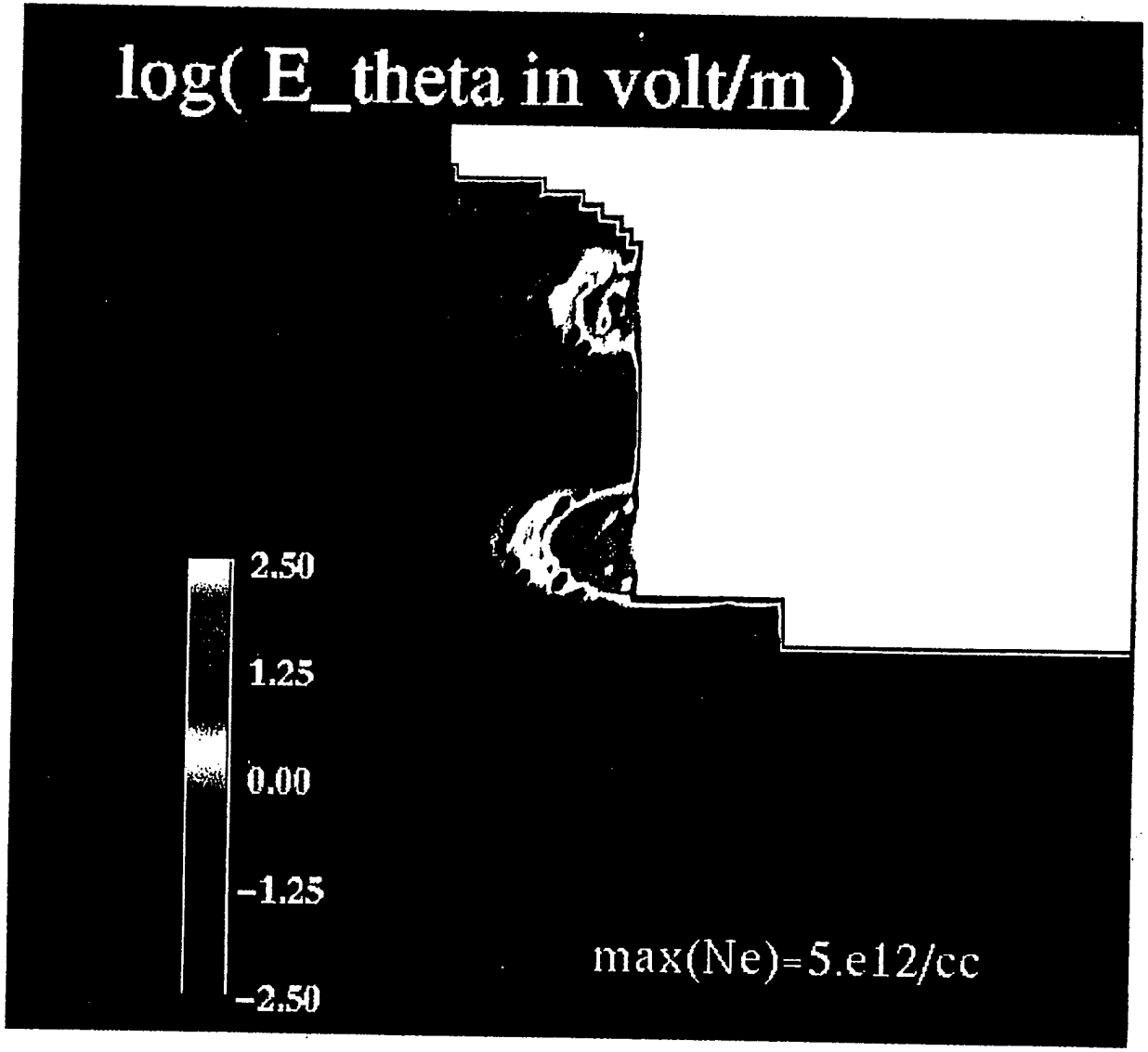
PMT Helicon Reactor



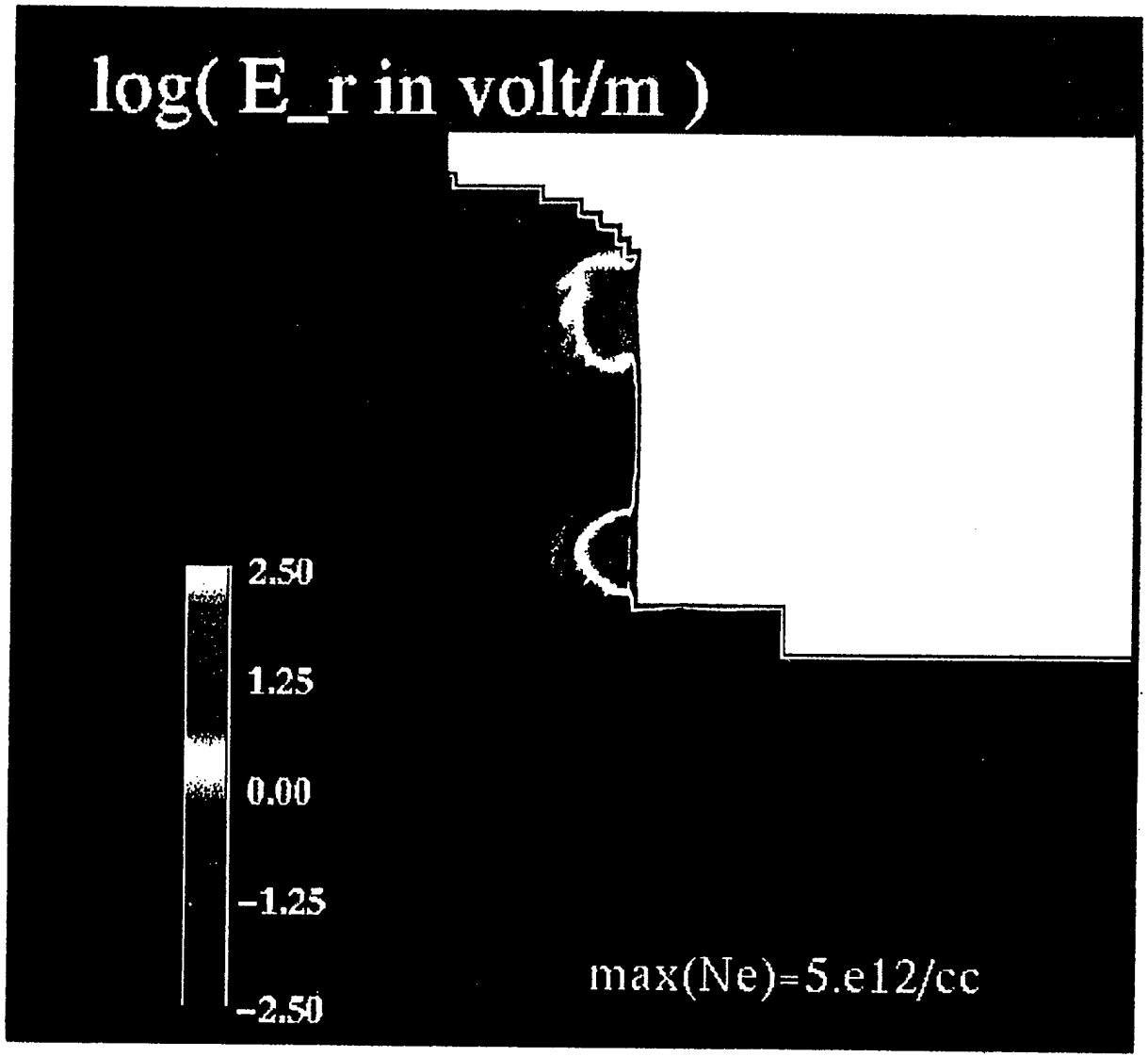
SAIC®
An Employee-Owned Company



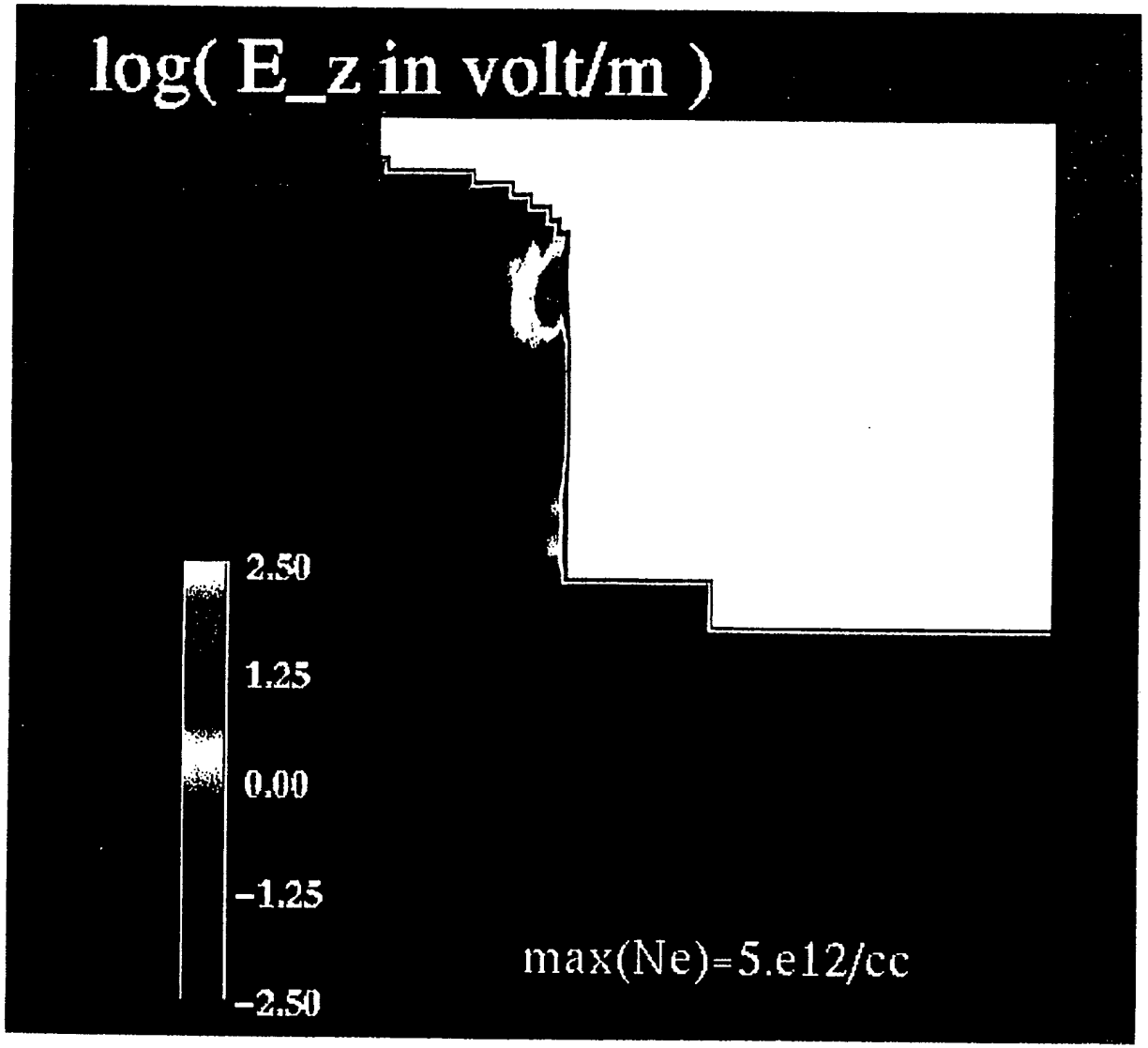
SAIC®
An Employee-Owned Company



SAIC®
An Employee-Owned Company

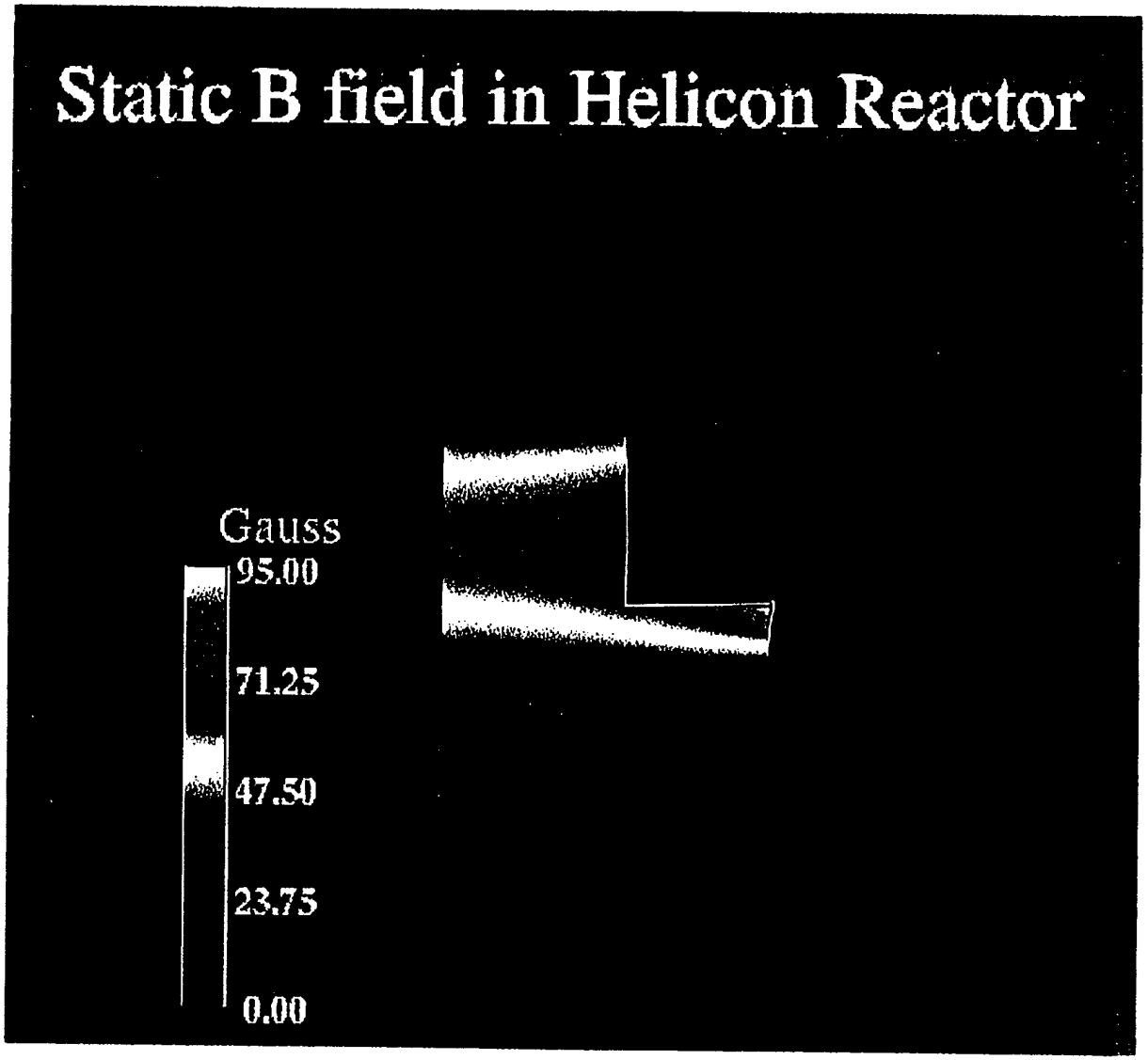


SAIC®
An Employee-Owned Company

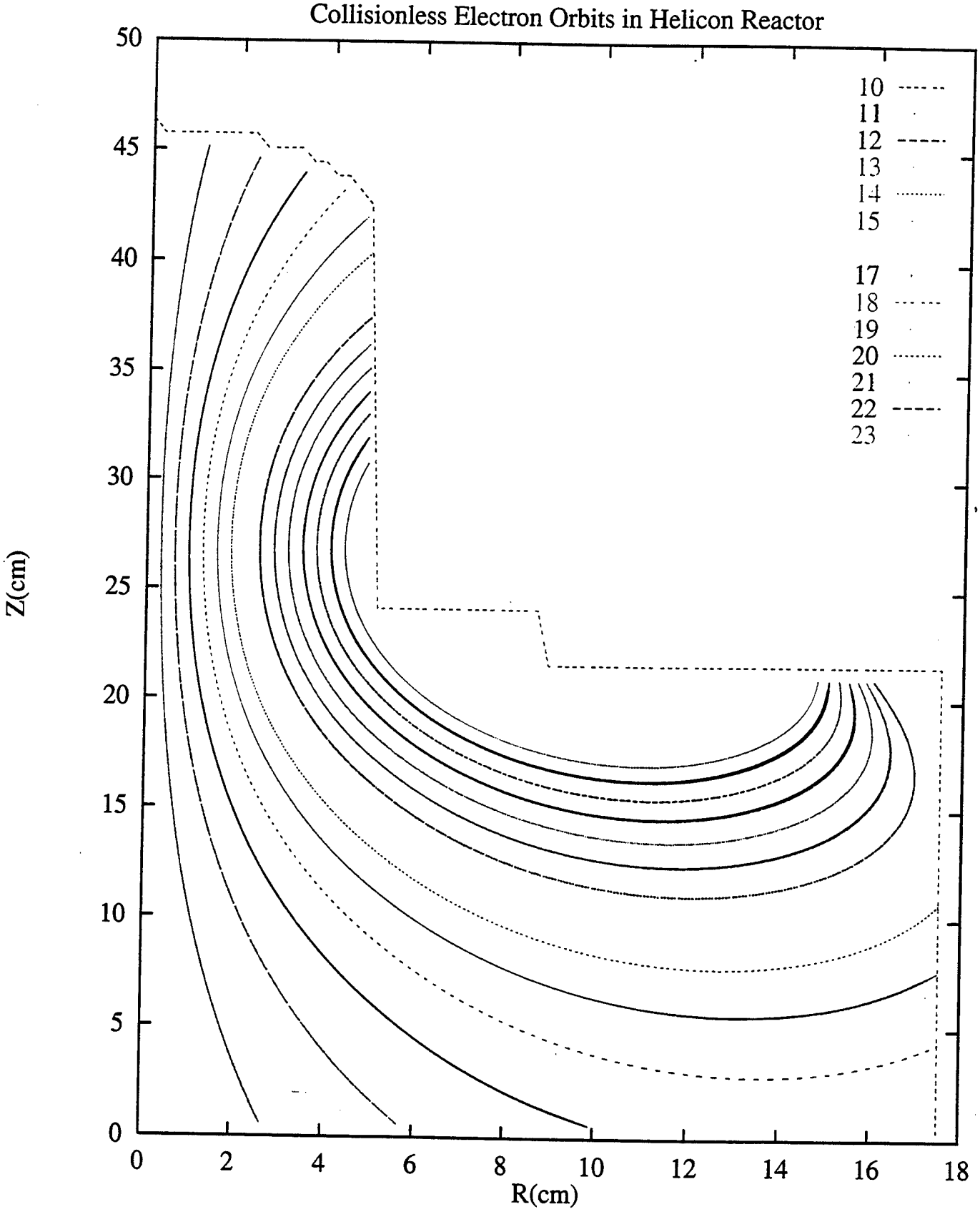


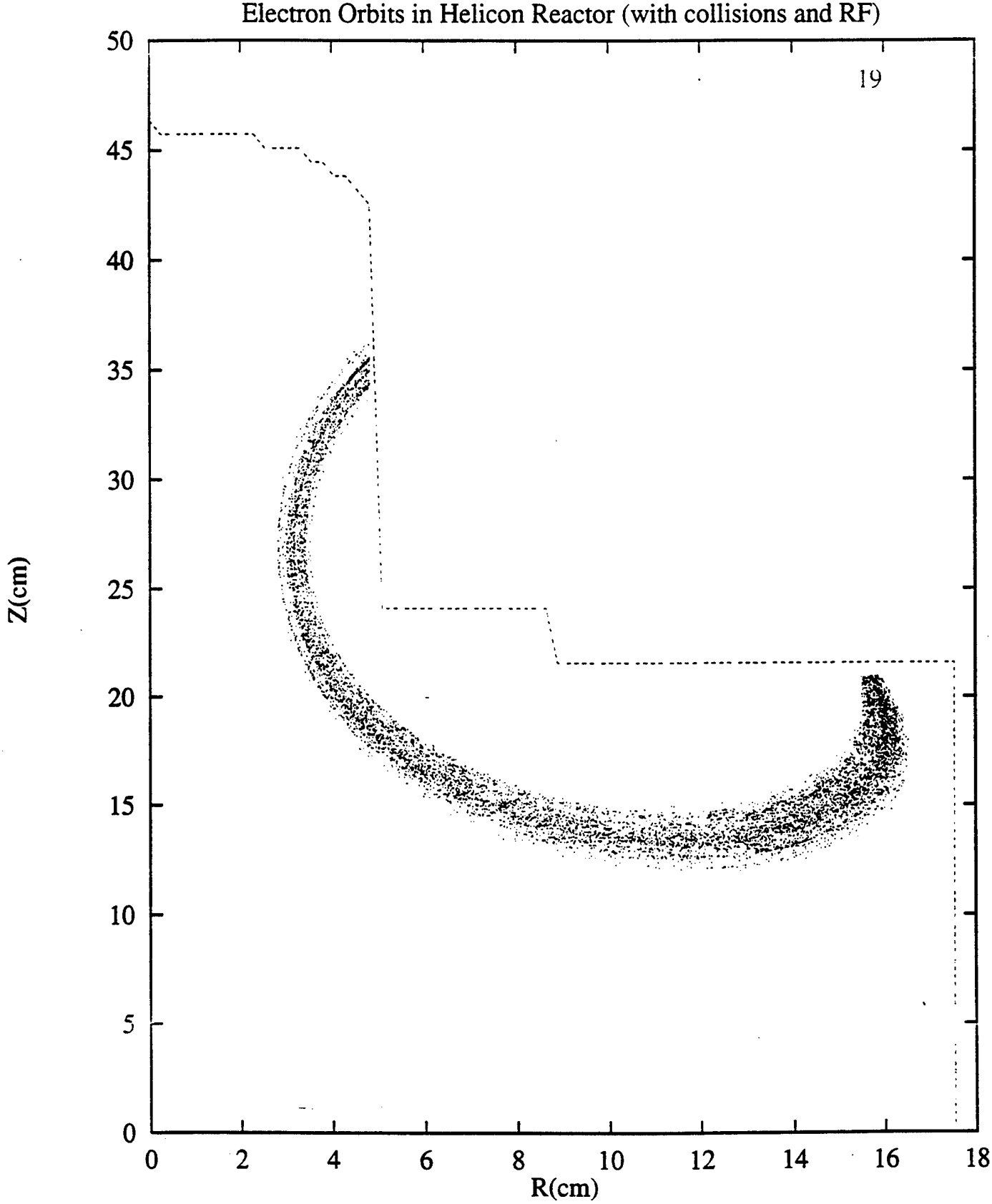
SAIC®
An Employee-Owned Company

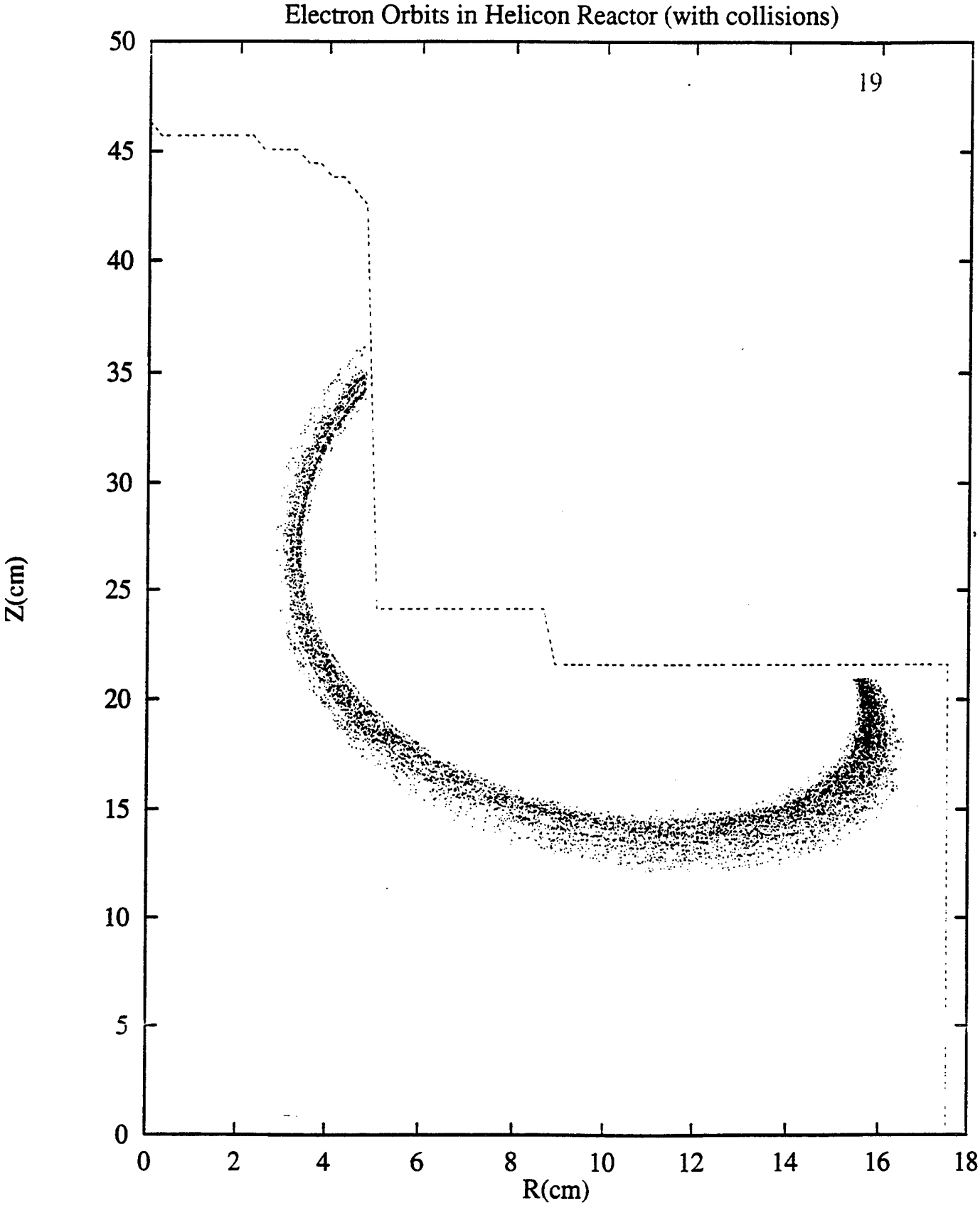
Static B field in Helicon Reactor

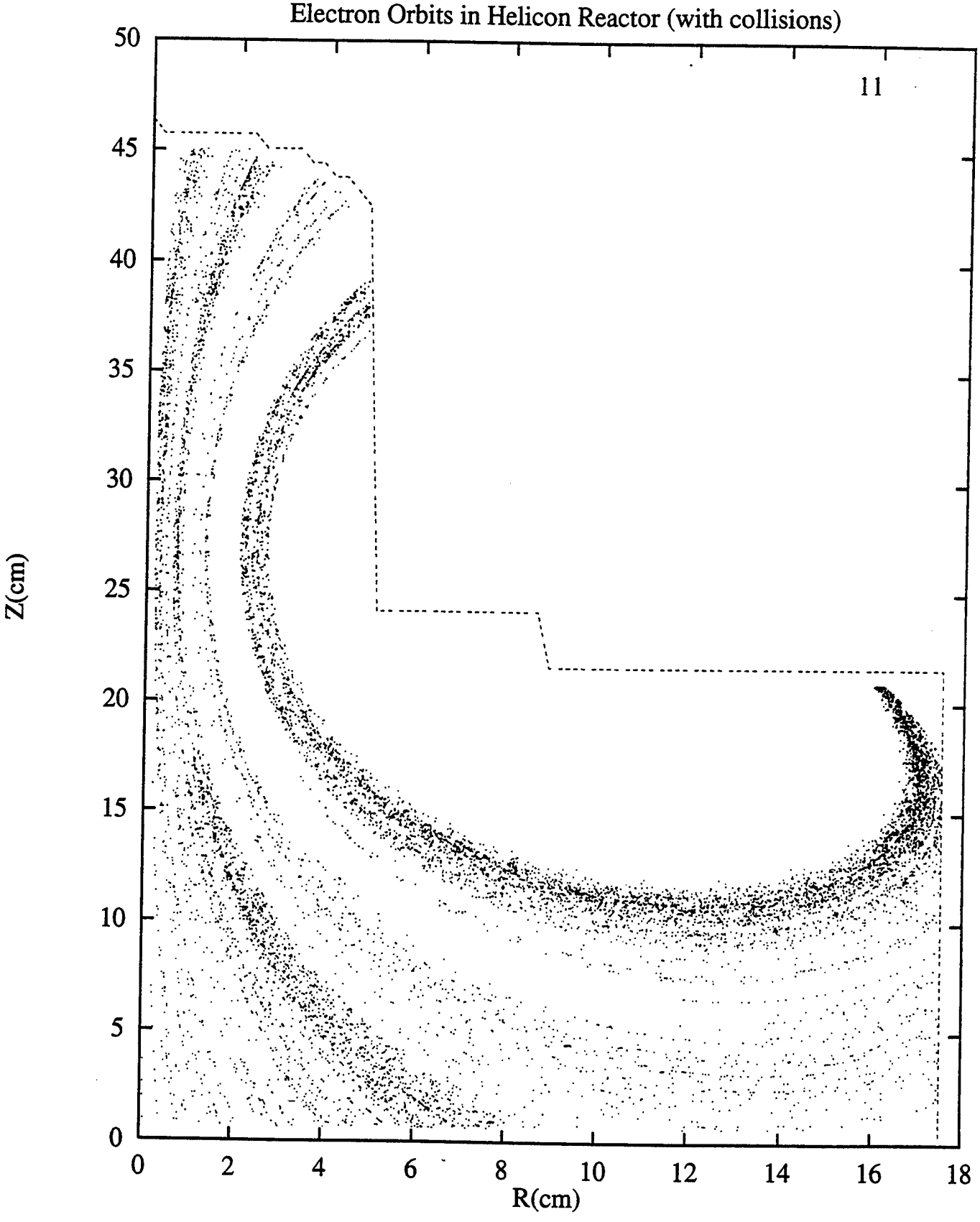


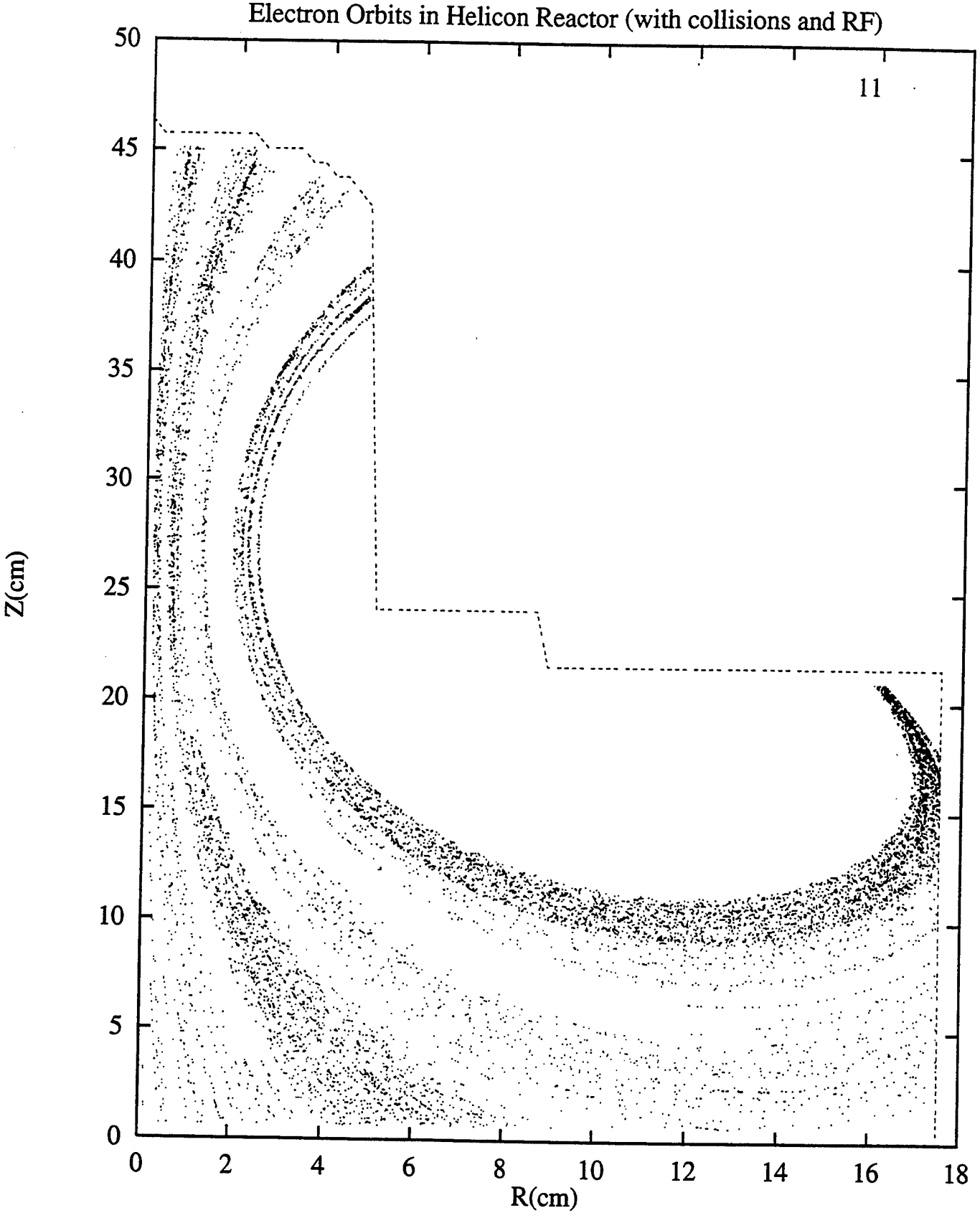
SAIC®
An Employee-Owned Company

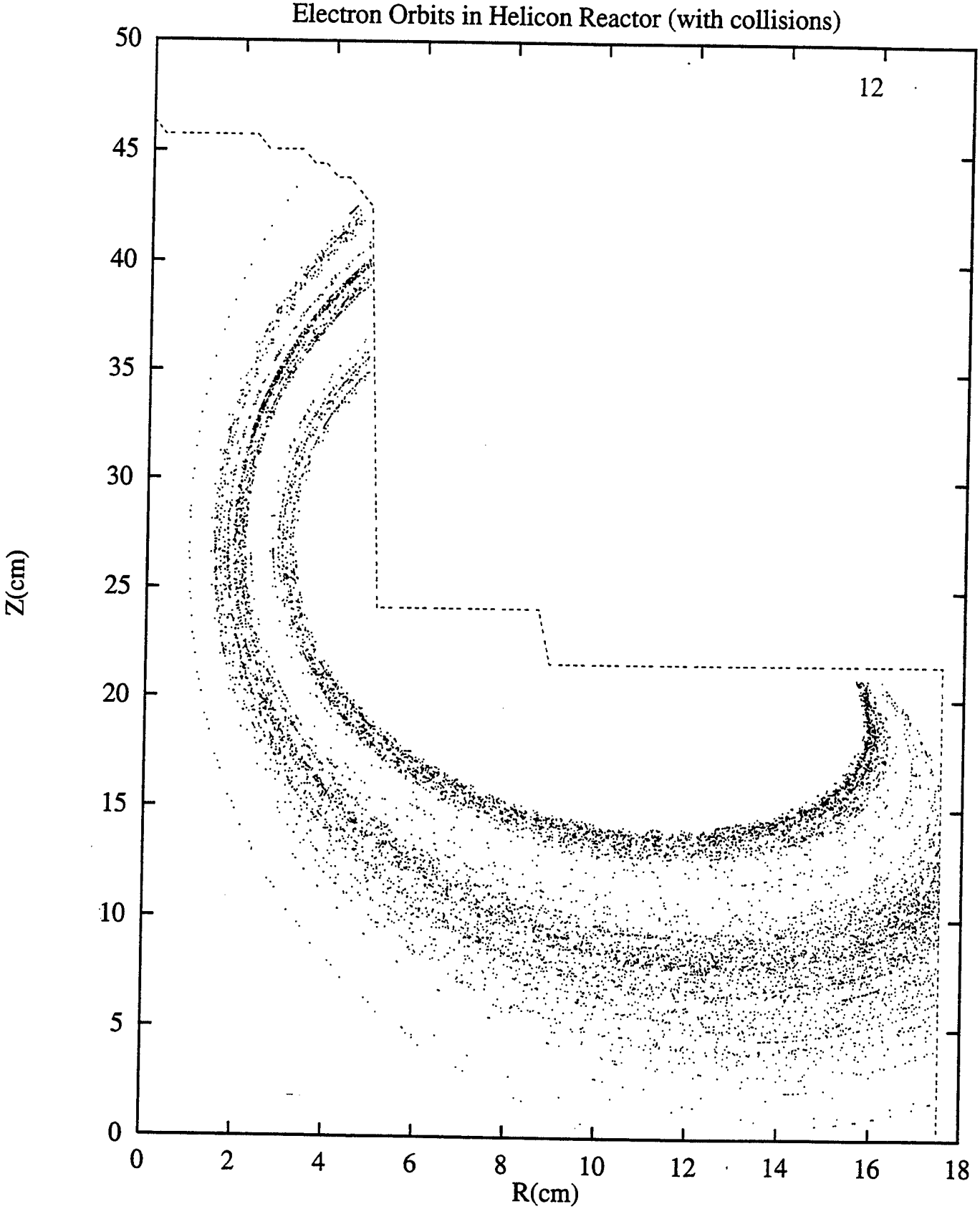

Collisionless Electron Orbits in Helicon Reactor
10
11
12
13
14
15
17
18
19
20
21
22
23
Z(cm)
R(cm)
0
5
10
15
20
25
30
35
40
45
50
0
2
4
6
8
10
12
14
16
18

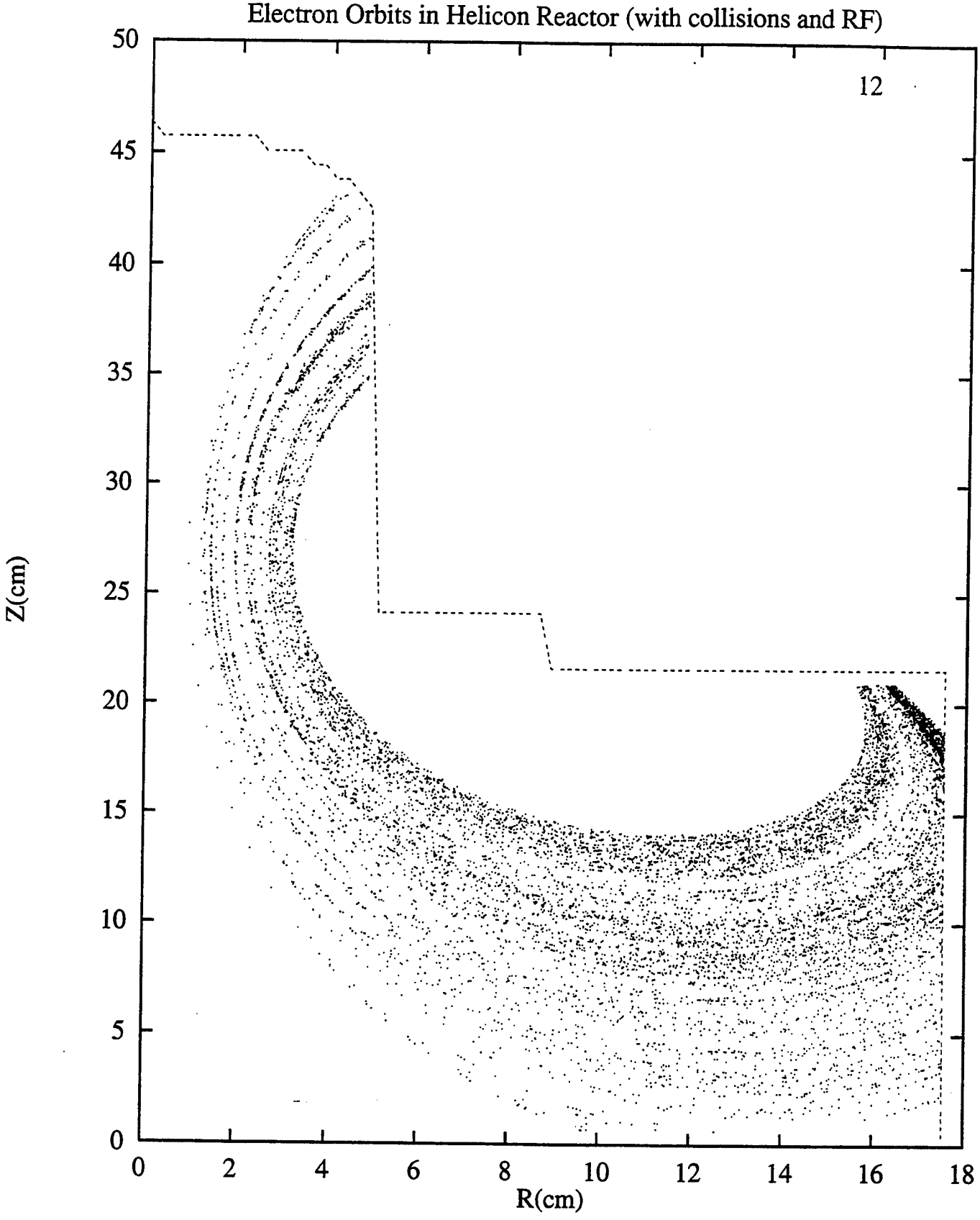

Electron Orbits in Helicon Reactor (with collisions and RF)
Z(cm)
R(cm)
0
5
10
15
20
25
30
35
40
45
50
0
2
4
6
8
10
12
14
16
18


Electron Orbits in Helicon Reactor (with collisions)
19
Z(cm)
R(cm)
50
45
40
35
30
25
20
15
10
5
0
0
2
4
6
8
10
12
14
16
18


Electron Orbits in Helicon Reactor (with collisions)
11
Z(cm)
R(cm)
0
5
10
15
20
25
30
35
40
45
50
0
2
4
6
8
10
12
14
16
18

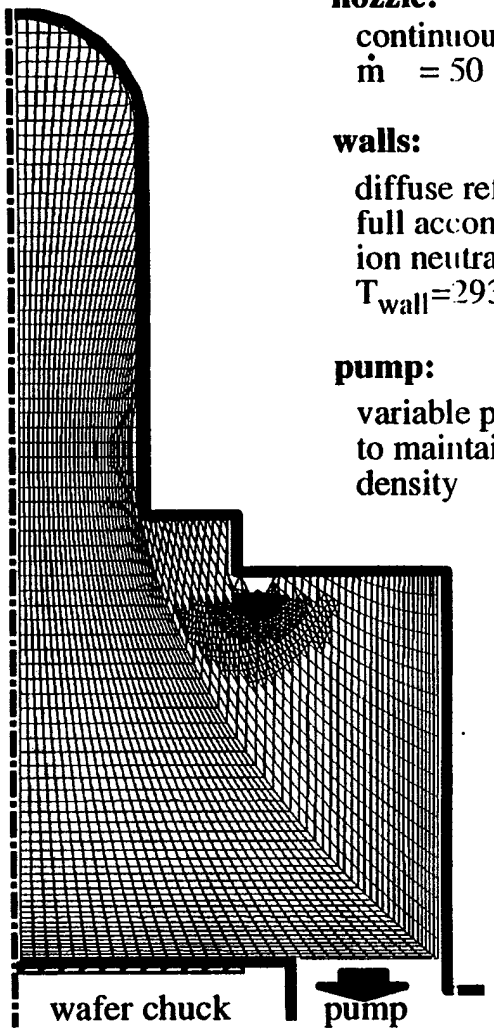

Electron Orbits in Helicon Reactor (with collisions and RF)
11
Z(cm)
R(cm)
0
5
10
15
20
25
30
35
40
45
50
0
2
4
6
8
10
12
14
16
18


Electron Orbits in Helicon Reactor (with collisions)
12
Z(cm)
R(cm)
50
45
40
35
30
25
20
15
10
5
0
0
2
4
6
8
10
12
14
16
18


Electron Orbits in Helicon Reactor (with collisions and RF)
12
Z(cm)
R(cm)
0
5
10
15
20
25
30
35
40
45
50
0
2
4
6
8
10
12
14
16
18

COMPUTATIONAL DETAILS:

- **SAIC calculated electron energy distribution**
- **experimentally measured electron density**
- **DSMC simulation:**
  - **10^6 particles**
  - **IBM SP2: 8 processors**
  - **five particle species: Cl_2, Cl, Cl_2^+, Cl^+, Cl^-**
  - **axisymmetric**

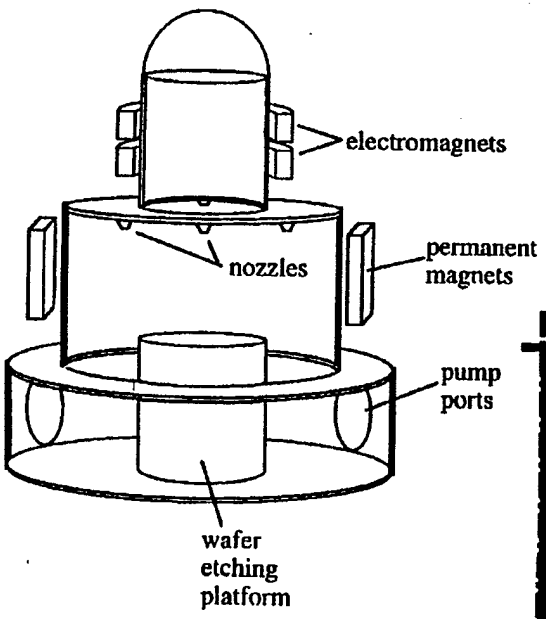


nozzle:
continuous ring
$\dot{m}$ = 50 sccm

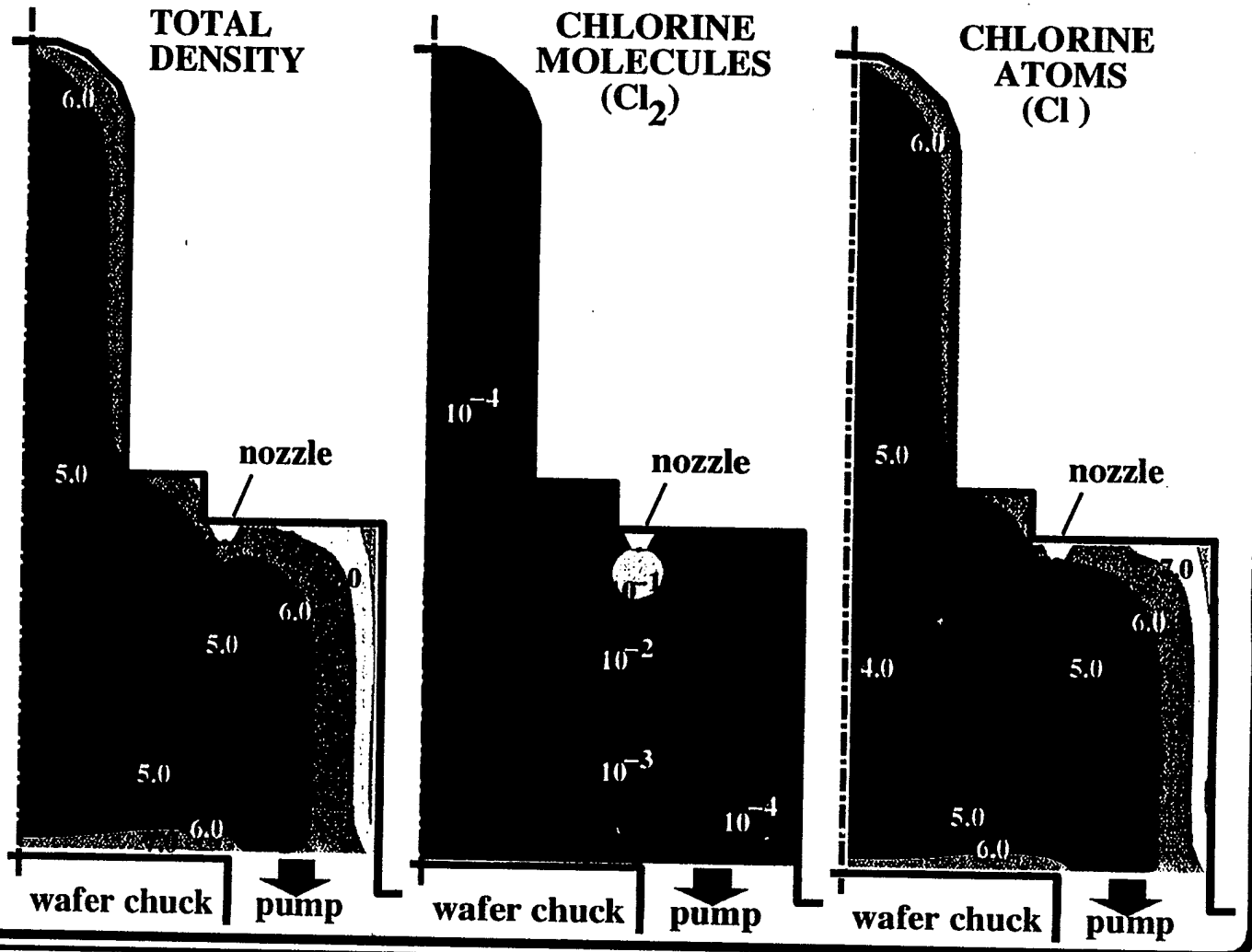
walls:
diffuse reflection
full accomodation
ion neutralization
T_{wall}=293 K

pump:
variable porosity
to maintain target density

PLASMA ETCH REACTOR FLOW FIELD



Density Contours (n x10^{19}#/m^3)



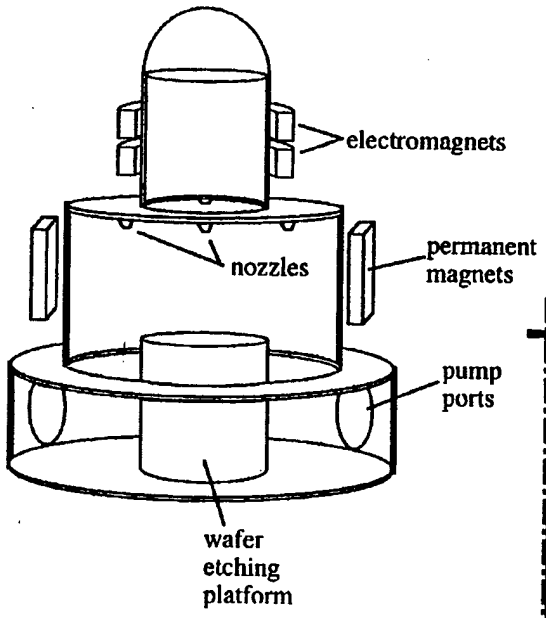
DSMC Axisymmetric Simulation:

- o 10^6particles
- o IBM SP2: 8 processors
- o five particle species: Cl_2, Cl, Cl_2^+, Cl^+, Cl^-
- o experimentally measured electron background
- o calculated electron energy distribution

Gabriel I. Font–Rodriguez

Cornell University

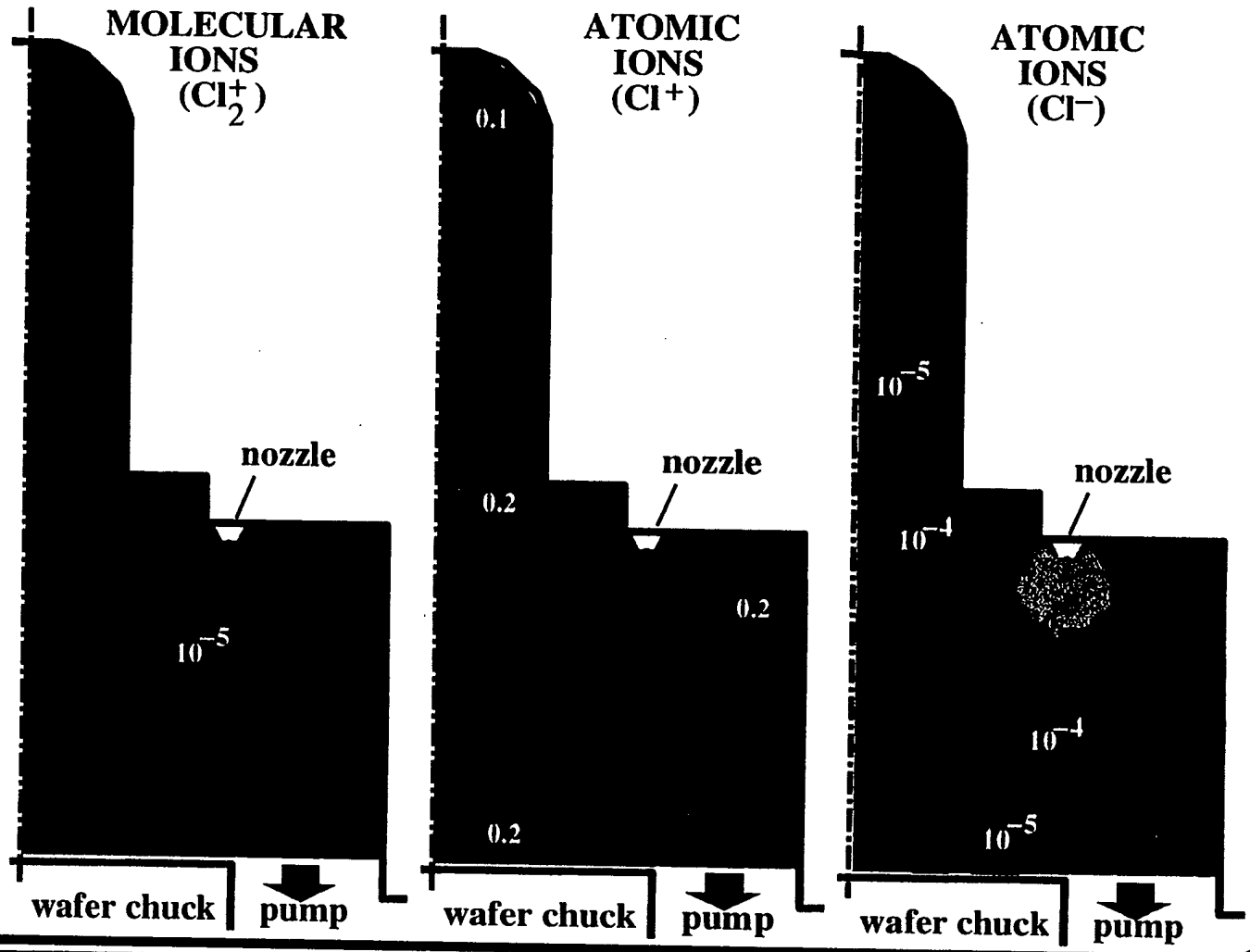
PLASMA ETCH REACTOR FLOW FIELD



DSMC Axisymmetric Simulation:

- o 10^6 particles
- o IBM SP2: 8 processors
- o five particle species: Cl_2, Cl, Cl_2^+, Cl^+, Cl^-
- o experimentally measured electron background
- o calculated electron energy distribution

Density Contours (n x10^{19}#/m^3)



Gabriel I. Font–Rodriguez

Cornell University

GaSonics Reactor Code Status

- Boltzmann Code for Electron Energy Distribution, Local Approximation, Tested for CDE Reactor
- Module for Oxygen-Nitrogen Chemistry Developed and Tested
- Modules for Thermal Conductivity, Diffusion, Viscosity Developed and Tested
- Fluid Code Using Finite Analytic Navier-Stokes (FANS) Algorithm Implemented for GaSonics Reactor - Focus on Plasma Tube Simulation
- Simulation of Plasma Discharge Tube - High Aspect Ratio Simplification

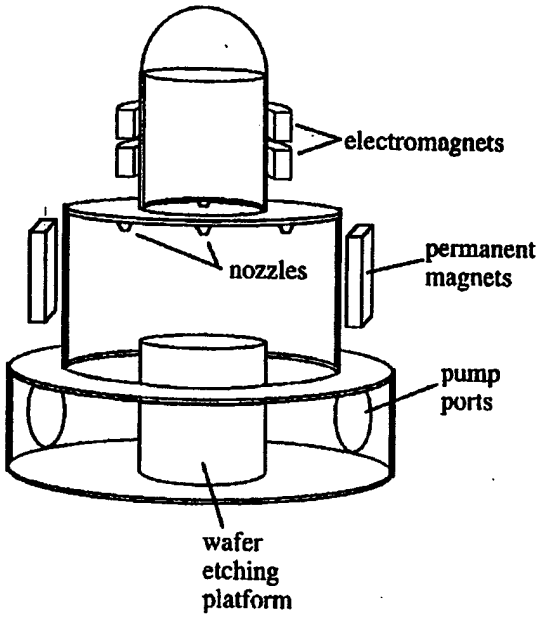
<u>Next Steps</u>

- Full Simulation of Discharge Tube with FD Solver Calculated E Field, Boltzmann Solver, Chemistry, Diffusion, Heat Transport
- Simulate Process Chamber (Neutral Flow and Chemistry)

SAIC®
An Employee-Owned Company

COMPARISON WITH EXPERIMENT

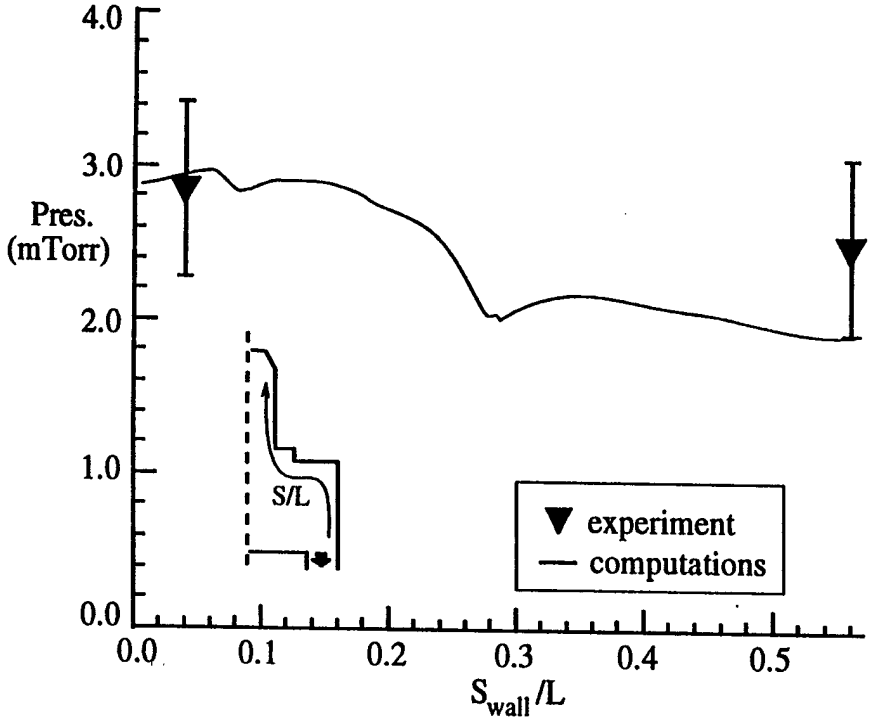
Configuration



DSMC Axisymmetric Simulation:

- o 10^6 particles
- o IBM SP2: 8 processors
- o five particle species: Cl_2, Cl, Cl_2^+, Cl^+, Cl^-
- o experimentally measured electron background
- o calculated electron energy distribution

Wall Pressure



Modeling the Microwave Plasma Chemical Downstream Etch Reactor

Flow in the Discharge Chamber

- Discharge chamber, typically a long dielectric tube
- Gases flow through tube, acted on by microwave source
- Plasma and reactive gases produced; gases heated
- Flow is subsonic (except near orifices) - pressure (varies slowly along tube)

Eliminate sound waves –

Following Collela, replace $\nabla \cdot v = 0$ by

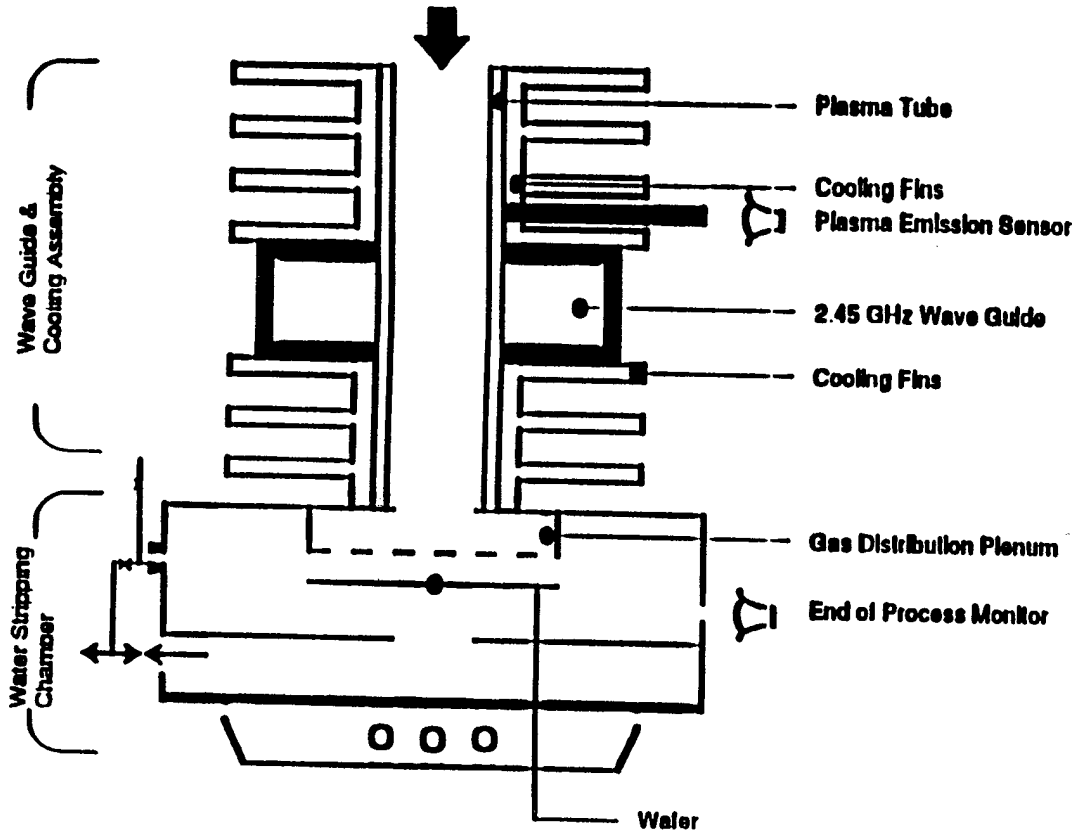
$\nabla \cdot v = S^*(r,z)/c_p$; $S^* = S+\nabla\cdot\lambda\,\nabla T$; S =microwave heat source,

$\lambda =$ thermal conductivity

<u>pressure is constant, but pressure gradients can occur</u>

SAIC®
An Employee-Owned Company

Chemical Downstream Etch Reactor



The chemical downstream etch (CDE) reactor is used for wafers that are sensitive to ion/electron damage. Therefore, it is designed to prevent charged particles from reaching the wafer.

SAIC®
An Employee-Owned Company

Flow in the Discharge Chamber (continued)

- Because pressure is constant, where microwave heating occurs,

  as T increases, ρ decreases

- To conserve mass flux along the tube, v must increase

- In high aspect ratio (tube length/tube radius) limit, complex 3D flow can be simplified

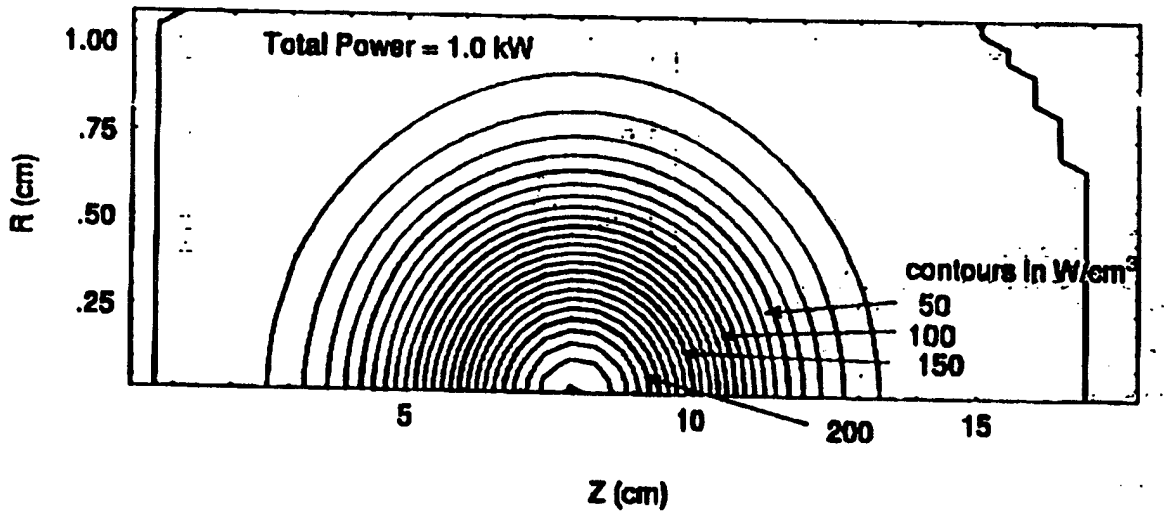
  - provides a means of benchmarking the 3D results

  - fast approximation tool for calculating flow in high aspect ratio reactors

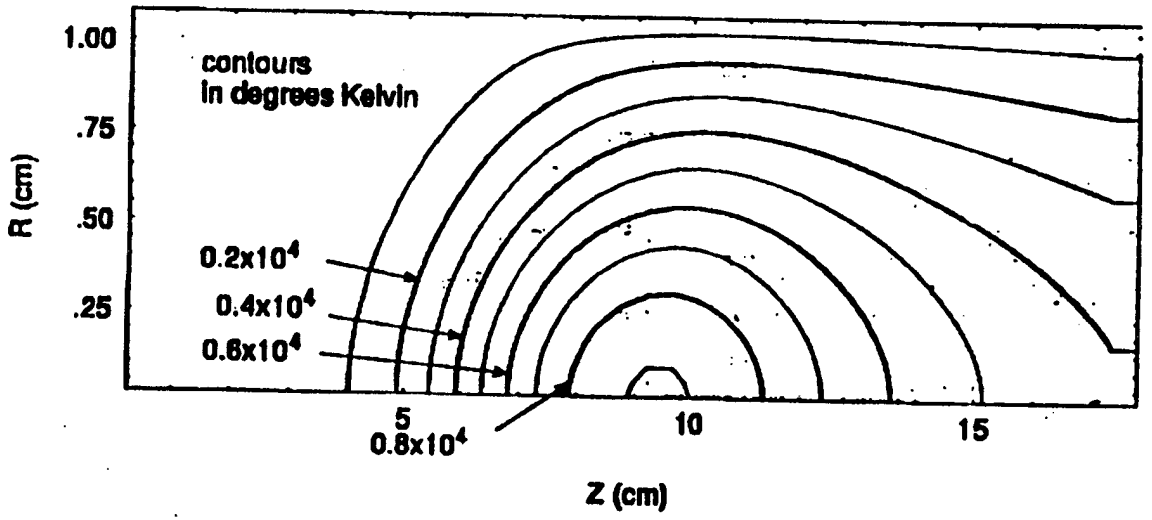
SAIC®
An Employee-Owned Company

On Axis Case

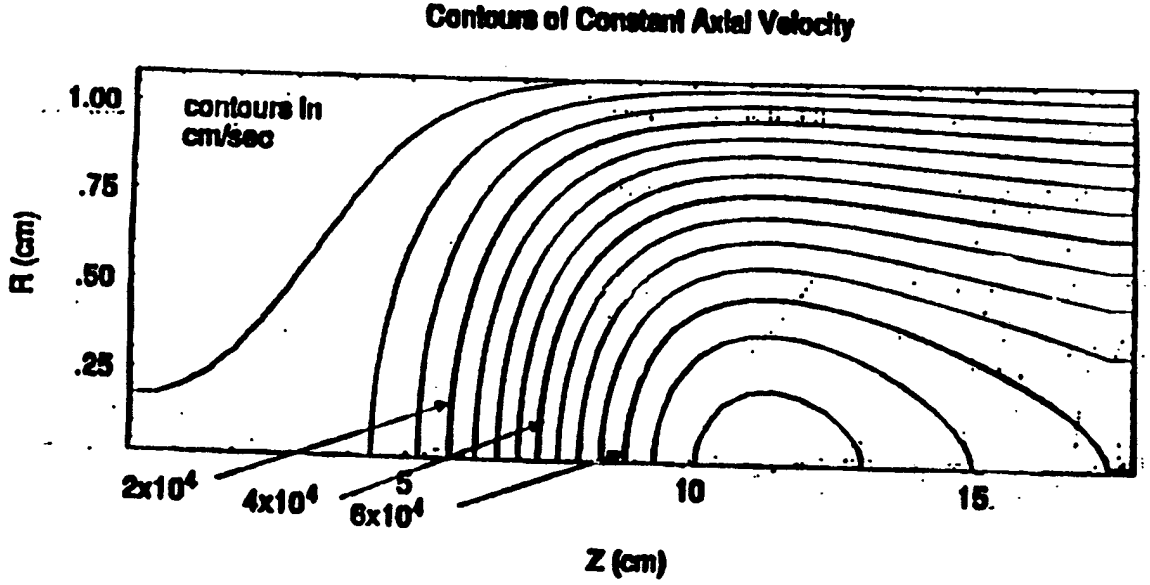
Contours of Constant Heat Source

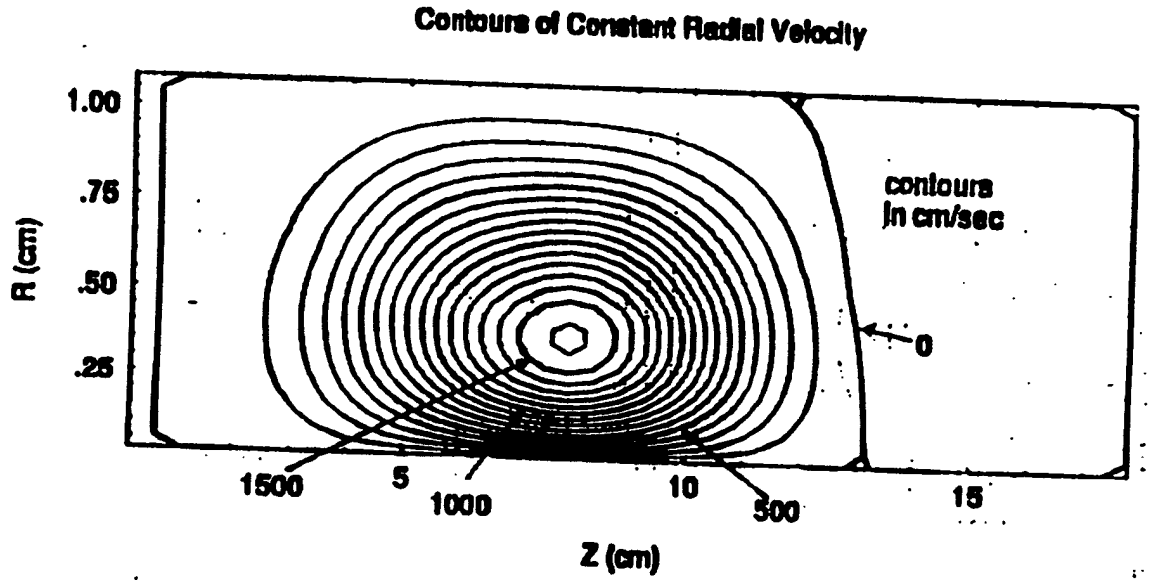


Contours of Constant Temperature

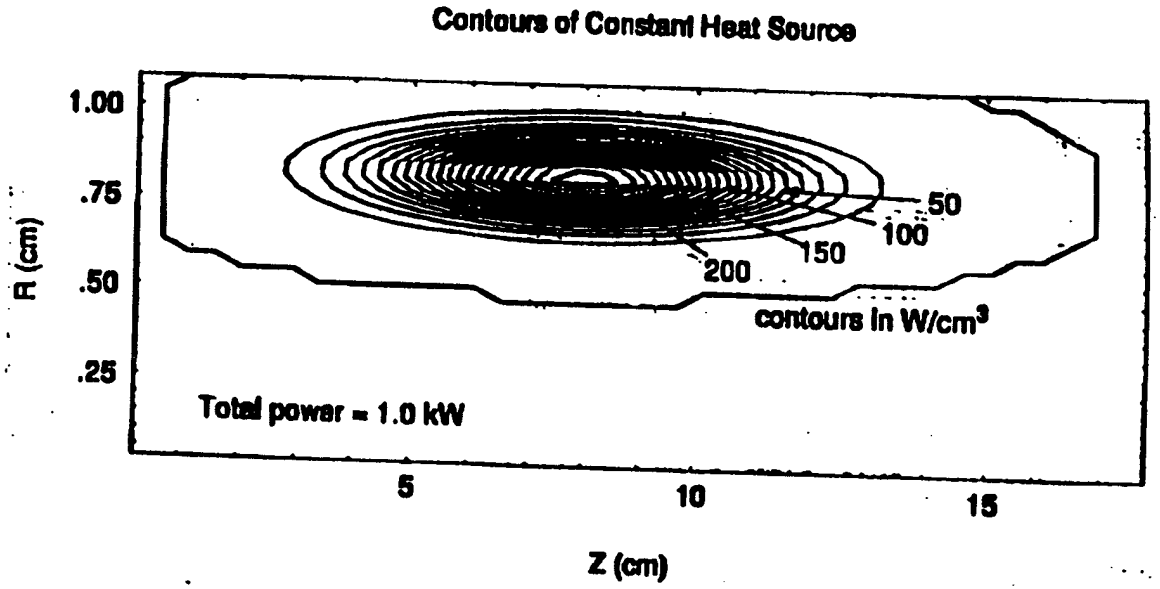


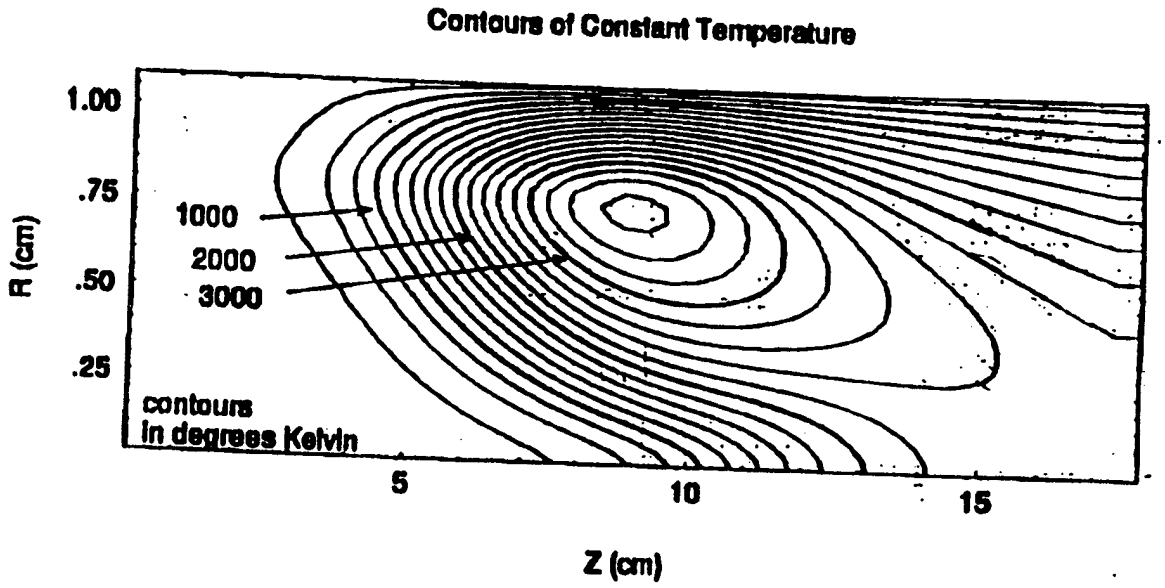
On Axis Case


Contours of Constant Axial Velocity
contours in cm/sec
1.00
.75
.50
.25
R (cm)
2x10⁴
4x10⁴
5
6x10⁴
10
15.
Z (cm)


Contours of Constant Radial Velocity
1.00
.75
.50
.25
R (cm)
contours in cm/sec
0
1500
5
1000
10
500
15
Z (cm)

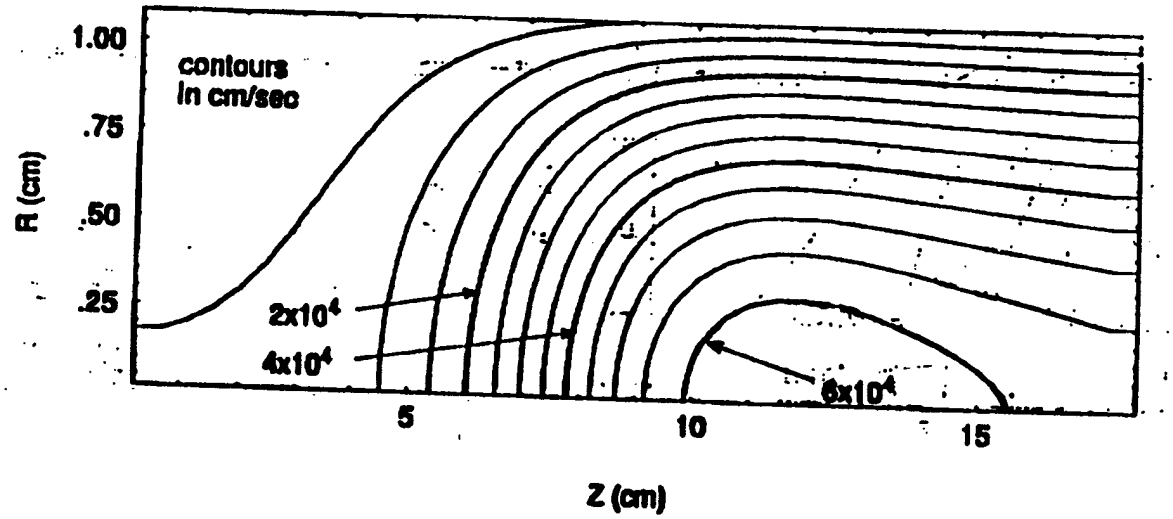
Off Axis Case



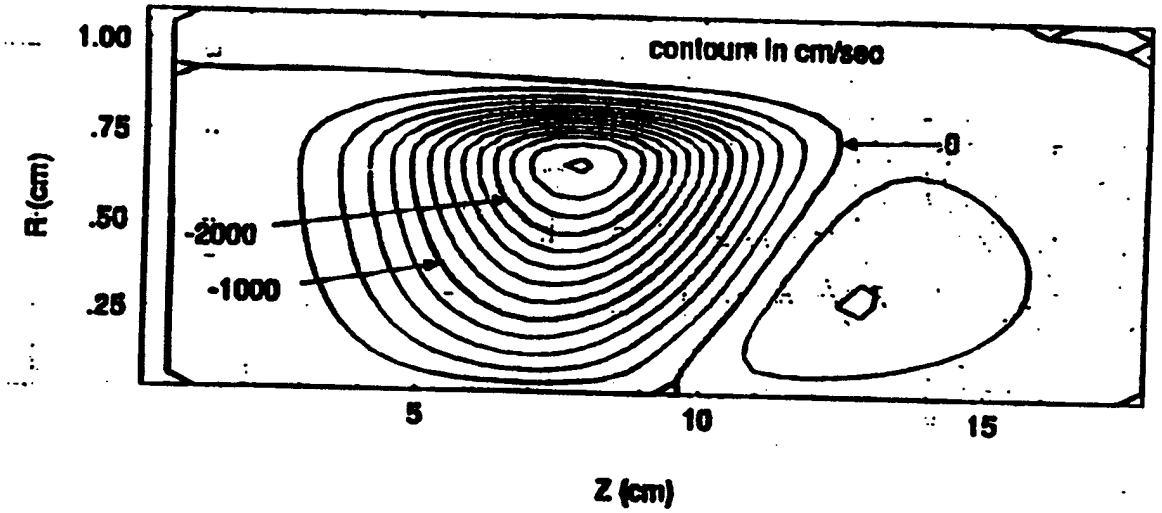


Off Axis Case

Contours of Constant Axial Velocity



Contours of Constant Radial Velocity



Role of N_2 Impurities in O_2 CDE Resist Etching

Experimental measurements by Dr. Oehrlein at SUNY (Albany) on CDE reactor using ASTeX microwave plasma applicator with alumina walls.

Power - 500-1200 W; p - 0.5-1.0 Torr; Flow rate ~ 1100 sccm; Distance from discharge to substrate ~ 30 cm; O_2 - 99.999% purity, add small amounts of N_2 (or H_2)

Results:

Resist Etch Rate: 4-5X enhancement with addition of 1-3% N_2 (or H_2); Enhancement continues to increase with N_2 (but not H_2) up to ~25X.

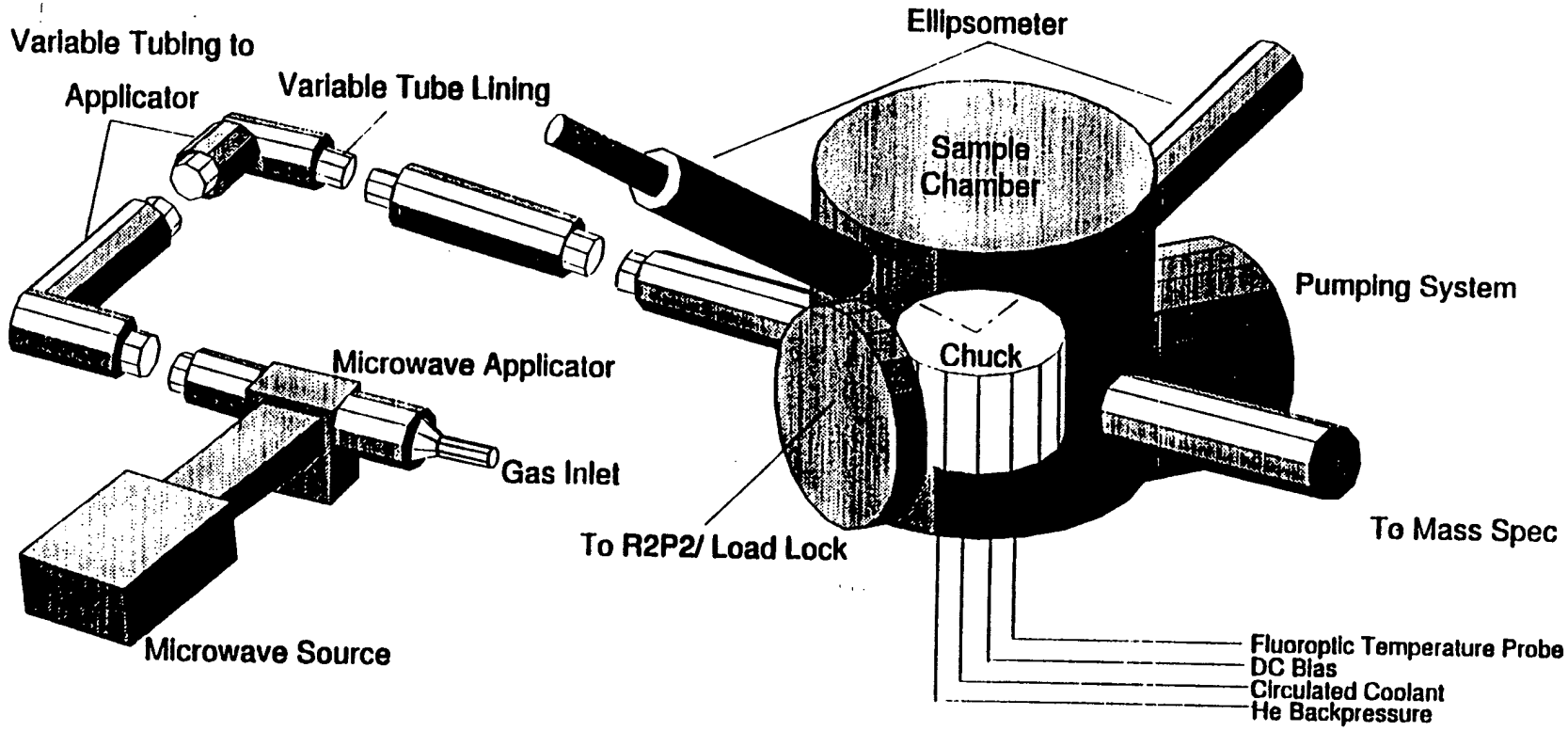
Oxygen Atoms in Discharge Region: Shows corresponding increase in O density up to ~ 3% impurity but decreases above that level.

Excited O Atoms: $O(^1D)$ unaffected by impurity level.

Surface Analysis: Oxidation of resist layer remains constant as N_2 flow rate increases; drops rapidly as H_2 flow increases.

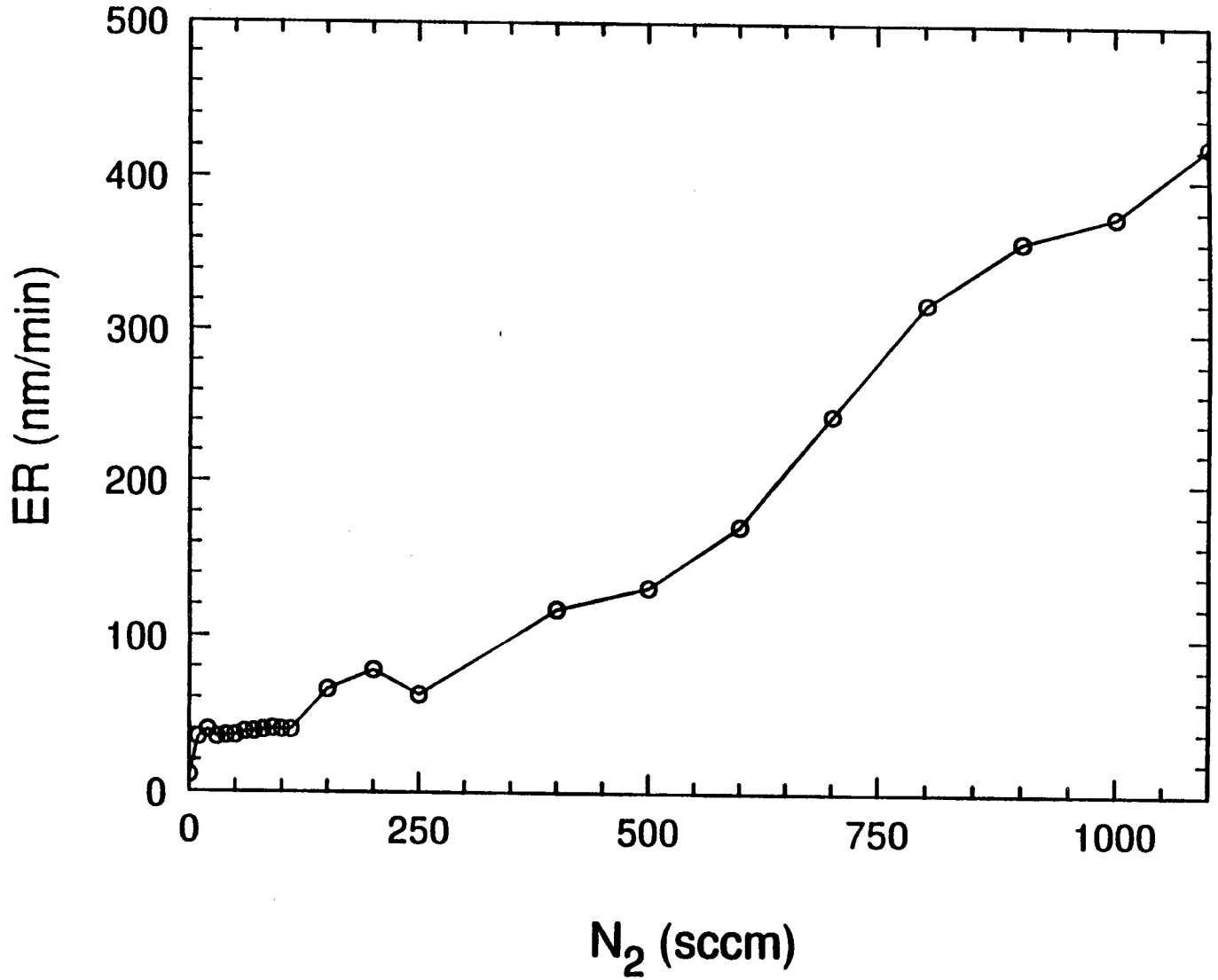
SAIC®
An Employee-Owned Company

Schematic of CDE Tool Used in Oehrlein Experiment



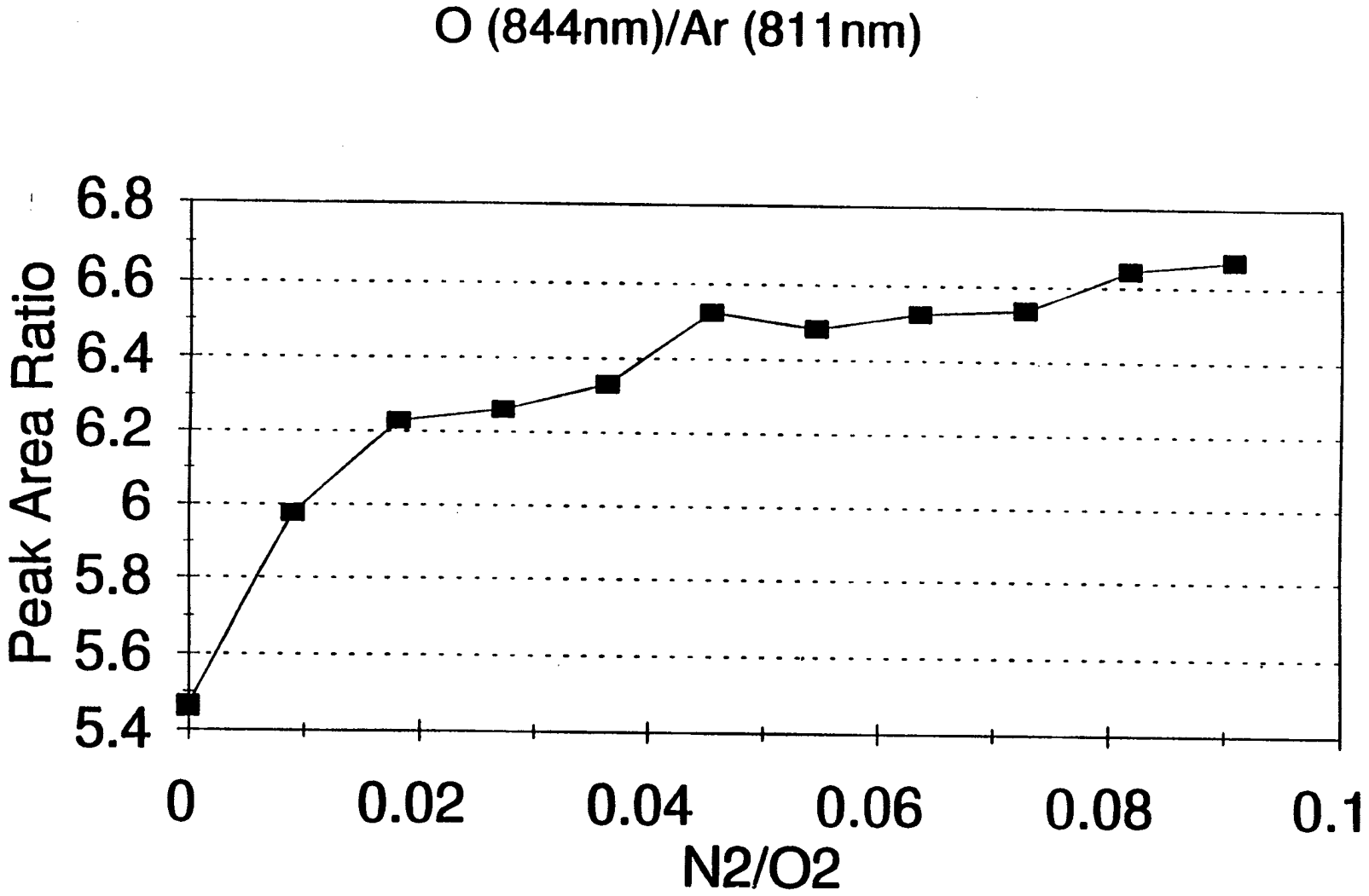
SAIC®
An Employee-Owned Company

Photoresist Etch Rate vs N_2/O_2



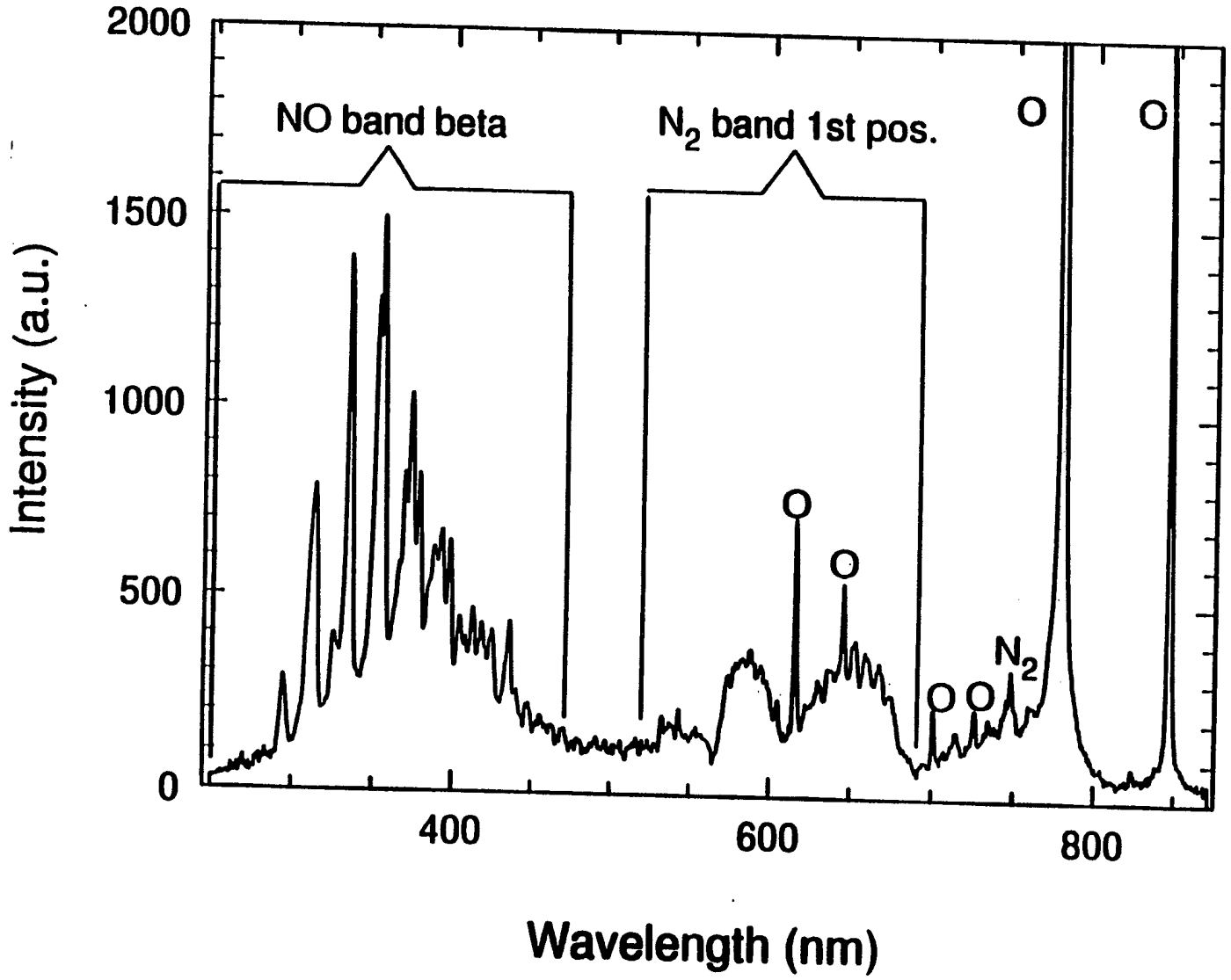
pressure = 1.0 Torr; O_2 Flow = 1100 sccm
Microwave Power = 1 kW; Chuck Temperature = 300 C

Oxygen Atom Concentration in the Discharge Region vs N_2/O_2 from OES

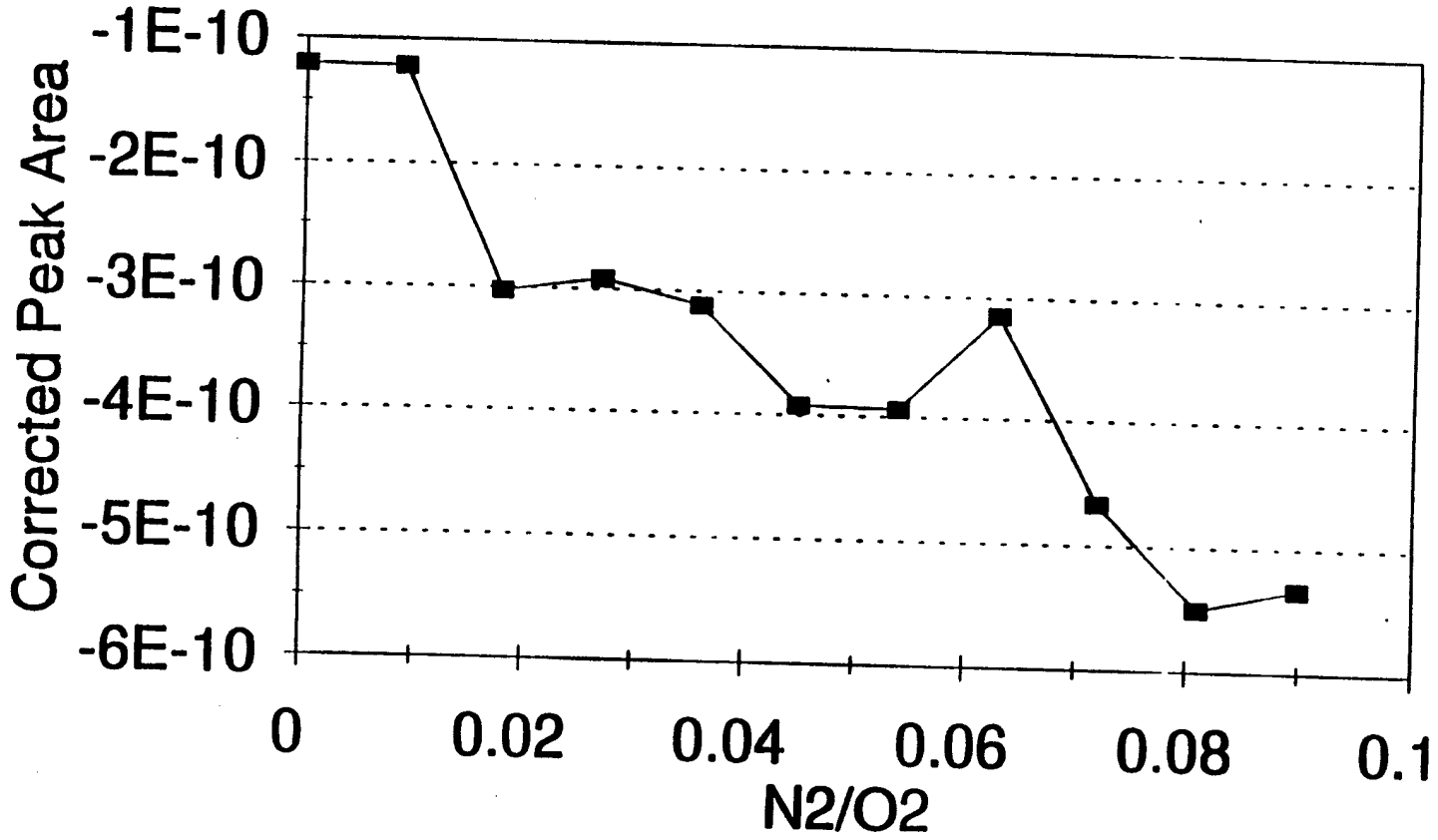


pressure = 1.0 Torr; O_2 Flow = 1100 sccm
Microwave Power = 1 kW

Broad OES Spectrum for an O_2/N_2 Ratio of 0.1

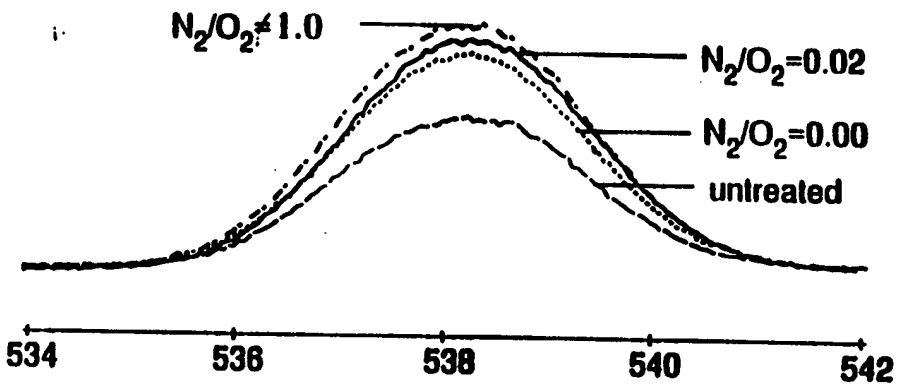


Atomic Oxygen Arrival Rate in Reaction Chamber by Mass Spectrometry

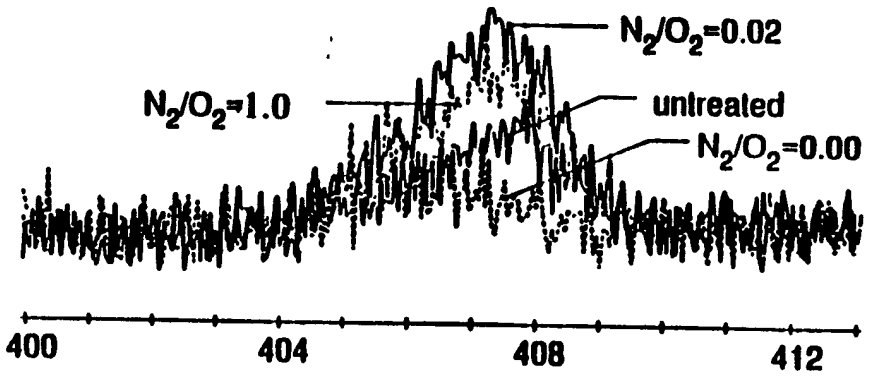


pressure = 1.0 Torr; O_2 Flow = 1100 sccm
Microwave Power = 1 kW

X-ray Photoemission Spectroscopy



oxygen 1s



nitrogen 1s

SAIC®
An Employee-Owned Company

N_2 Impurities: What Have We Learned?

- Based on Experimental Measurements and Chemistry Simulations: Resist Etch Rate Enhancement Must Be a Surface Effect

- Why the Discrepancy Between GaSonics and Oehrlein Results?
  - difference in wall material (quartz vs alumina)
  - difference in path length to substrate

- With Modern Experimental Diagnostic Capabilities the Source of the Discrepancy, ie. the Nature of the Surface Physics, Can Be Determined
  - x-ray photoemission spectroscopy
  - optical emission spectroscopy
  - mass spectrometry
  - ellipsometry

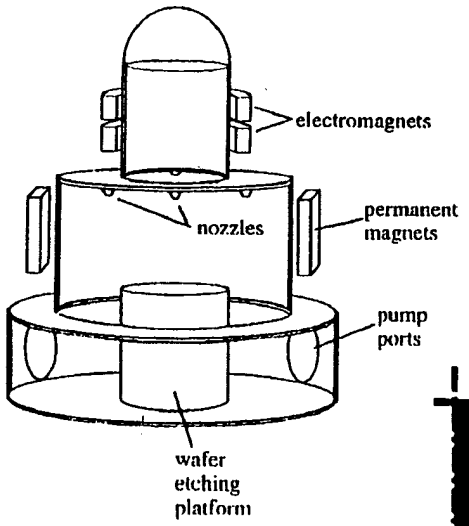
SAIC®
An Employee-Owned Company

Role of Nozzle Placement in PMT Reactor

- Effect on Etch Efficiency
  - fraction of etchant to reach substrate
- Effect on Etch Uniformity
  - uniformity of etchant flux at substrate
  - uniformity of ion flux at substrate

SAIC®
An Employee-Owned Company

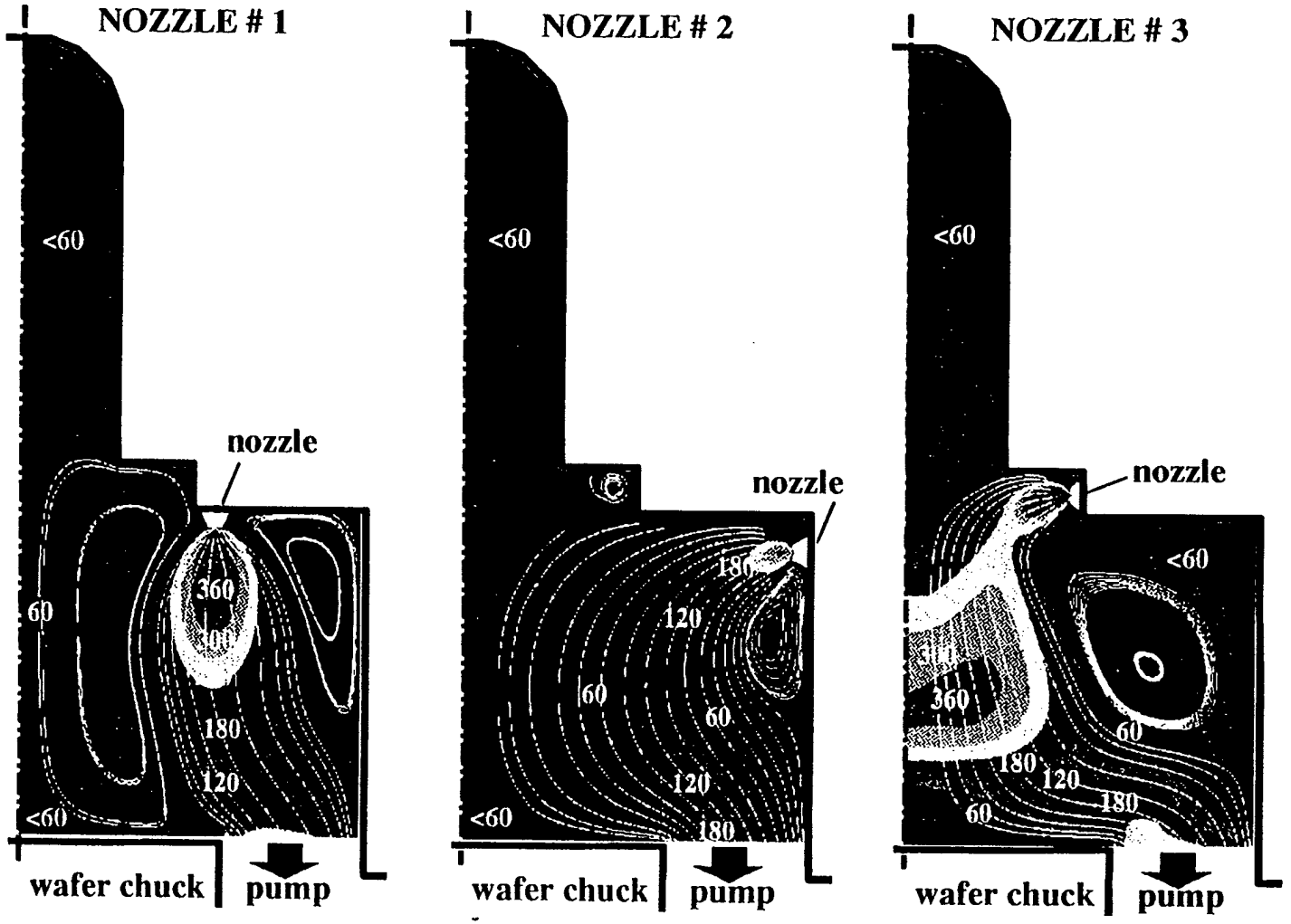
EFFECT OF NOZZLE LOCATION ON THE NEUTRAL FLOW FIELD IN A PLASMA ETCH REACTOR



DSMC Axisymmetric Simulation:

- 10^6 particles
- IBM SP2: 8 processors
- neutral flow only: Cl_2

Velocity Contours and Flow Streamlines (m/s)



Appendix NN

Computer Modeling of Deposition and Etching

Computer Modeling of Deposition and Etching

by E. Hyman and S. Eidelman

In the aerospace, automotive, chemical, electronics, and tooling industries, many operations rely on two processes: *deposition*, precision coating of one material on another; and *etching*, controlled removal of a material from its substrate. These processes can be used to develop improved materials for consumer products and defense needs, if we can maintain precise manufacturing tolerances in uniformity, consistency, coating quality or etch feature, and efficiency.

Unfortunately, the fundamental understanding of the physics and chemistry of these processes has been incomplete, so engineers have counted on experience and intuition in designing new deposition or etching equipment for special applications. Therefore, adapting a process to new-product requirements always entails considerable financial risk.

That risk could be reduced by using computer modeling to examine alternative design concepts and to investigate different ranges for important parameters. Most computer models, however, have not captured enough of the complex physics and chemistry needed to make reliable predictions. Computer running time and memory requirements have forced modelers to treat limited aspects of a process or use idealized geometries, even though engineers know that a process may be extremely sensitive to such details.

In addition, some aspects of physics and chemistry have not been well enough understood to model reliably, although improved diagnostic capabilities show promise of providing needed information. As computer capabilities become greater and product requirements become increasingly stringent beyond the capability of intuition to guide process improvements, numerical models are likely to play a larger role in the future.

Industrial needs for reliability and for meeting product specifications require numerical-simulation tools that can be applied to a specific process. These tools must incorporate the real geometry of every design element that affects the process, and they must capture all the important physical and chemical processes. Inevitably, developing such tools is a large-scale, multiyear, multidisciplinary undertaking that must be accompanied by experimental studies, reactor diagnostics, and model validation. Such an undertaking surpasses the financial interest of individual industrial firms, even the largest ones, because the cost is too high relative to potential payback. On the other hand, developing such tools could provide immense benefits for industry collectively. Consequently, major government funding has supported such work.

The Department of Defense's Advanced Research Projects Agency (ARPA), for example, has supported modeling efforts here at Science Applications International Corporation (SAIC) and at Sematech, Inc., Austin, Texas. In addition, major computer codes developed at the Department of Energy's

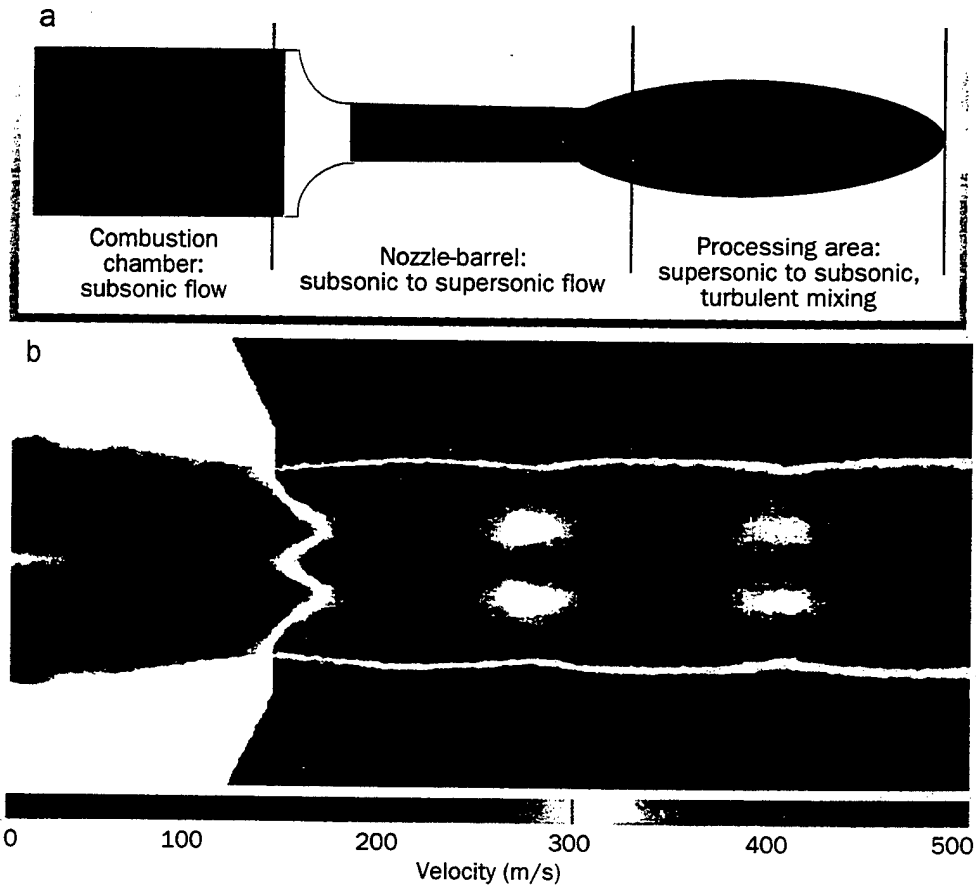


Figure 1. (a) Schematic diagram of a high-velocity oxygen-fuel thermal spray (HVOF TS) system. (b) Modeling results of HVOF TS system gas streams in cross section in the barrel and interacting to form the supersonic jet, which carries particles to the substrate. The color scale indicates velocity variations, from highest (magenta) to quiescent (black) flow.

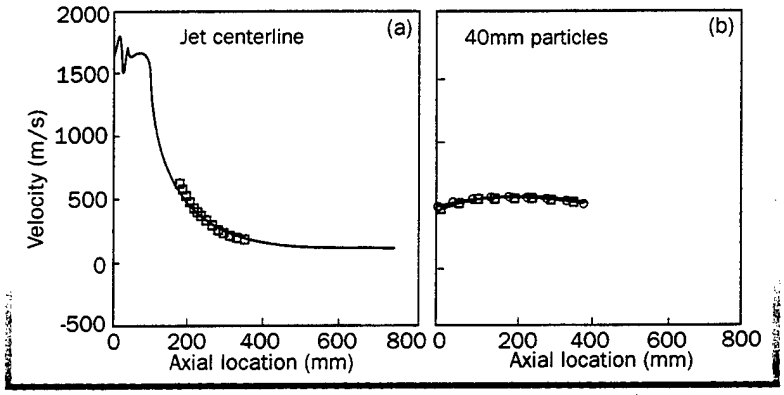


Figure 2. Comparison of experimental □ and simulated ○ gas-velocity distribution in an HVOF TS system.

Lawrence Livermore National Laboratory are being used to model etching reactors through cooperative R&D agreements with AT&T and IBM. Many other academic and industrial modeling projects also depend on support from federal agencies, such as the National Science Foundation (NSF) and the National Institute of Standards and Technology (NIST).

Two industrial-system models that are being developed at SAIC reveal the complexity and potential utility of these efforts.

Thermal spray system

A high-velocity oxygen-fuel thermal spray (HVOF TS) system uses the velocity and temperature of a supersonic jet to accelerate and heat metal or ceramic particles that subsequently coat a substrate. The high rates of deposition produced with this system make it a low-cost and versatile technology. Ceramic, composite, metal, and plastic coatings are being used to improve a material's biological, chemical, electrical, structural, thermal, and tribological properties.

In this system, a set of parameters—including particle size, gas flow, and substrate conditions—determines a coating's quality, consistency, and deposition rate. At present, these relationships are determined experimentally by plating sample after sample while varying the conditions. In other words, developing such a deposition system depends on trial and error. As a result, coating quality has not been well controlled and has been inadequate for many potential applications. Recent computer models, however, will make many new applications feasible and improve process characteristics by improving the deposition efficiency, reducing erosion of the device, lowering its energy consumption, and making the equipment more versatile and scalable.

The HVOF TS system (Figure 1a) consists of a combustion chamber with fuel- and oxidizer-injection systems, a de Laval nozzle, a barrel, particle-injection ports, and a water cooling system. High-pressure, high-temperature gas from oxygen combustion expands into the de Laval nozzle and barrel, which accelerate the gas flow to supersonic velocity. The supersonic jet flow expands into the atmosphere and quickly decays to subsonic velocity. In a TS gun, metallic particles or ceramic powder is injected into the supersonic-gas stream.

The particles are turbulently mixed, heated, accelerated, and propelled to the substrate. The system involves two-phase (gas and particle) flow with turbulence, heat transfer, chemical reactions, and supersonic–subsonic flow transition. The physical properties of the plated surface depend on the physical and chemical conditions of the impinging particles, which in turn depend on parameters including gun design, gas-jet formation, particle size, shape, material, method of injection, and position relative to substrate.

Simulation of this system requires a number of components: (1) detailed gun geometry, (2) Navier–Stokes equations for compressible flow, (3) conservation equations and constitutive relations for gas and particle phases, (4) a turbulence model, (5) a Lagrangian particle-trajectory formulation, and (6) heat and momentum transfer. Fluid velocity fluctuations caused by gas turbulence are simulated using a stochastic model, and the resulting particle motion is calculated with a Lagrangian stochastic–deterministic method. The simulation accurately models several factors: (1) gas expansion and flow in the barrel and in the free jet, (2) particle injection, (3) interaction of the particles with turbulent gas flow inside and outside the barrel, and (4) particle-wall interactions. An example of the resulting gas streams is plotted in Figure 1b.

To validate our model, the steady-state flow regime was simulated for conditions in a de Laval nozzle and gun barrel, including expansion of hot gas into the atmosphere, and compared with experimental results. The model agreed to within 10% of the calculated velocity distribution of gas on the jet centerline with measured data from W. D. Swank and his colleagues at the Idaho National Engineering Laboratory (INEL) (Figure 2a). The axial velocity distribution predicted for 40-mm particles agreed with measurements to within 5%, well within the experimental accuracy (Figure 2b).

We conducted a parametric study, including varying particle size and barrel length, to predict coating regimes that raise energy efficiency, increase coating rate, improve coating quality, reduce plating spot (the size of the jet when it reaches the substrate), and reduce equipment wear over current operating conditions. The simulations indicate that all these improvements are possible without modifying the TS equipment, and the study's optimum plating conditions are being tested at INEL through a cooperative R&D agreement with SAIC. If those experiments confirm our model, it will represent a first case in guiding TS plating conditions entirely through numerical simulations.

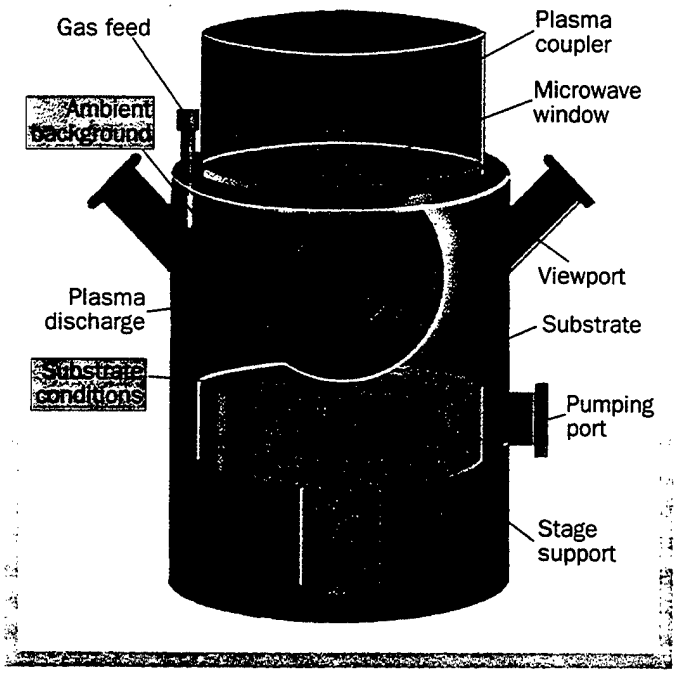


Figure 3. Schematic of a microwave-driven diamond plasma-enhanced chemical vapor deposition reactor.

PECVD

Plasma-enhanced chemical vapor deposition (PECVD) and etching are widely used in industry. The physical and chemical processes that must be modeled in these systems can be daunting. Consider a microwave-driven diamond PECVD reactor (Figure 3). A gas mixture of hydrogen and small amounts of methane and oxygen are fed through a microwave field region. A plasma is formed, and high-energy electrons dissociate some hydrogen molecules and other species. With the proper gas mixture, pressure, power input, substrate temperature, and so on, a diamond coating is deposited on the substrate. It is known empirically that a high flux of hydrogen atoms reaching the substrate is required to deposit high-quality diamond, and one goal of reactor design is producing a plasma that generates this flux uniformly over a large area.

We simulated the performance of an Applied Science and Technology, Inc. (ASTeX) microwave diamond-deposition reactor by using code modules that calculate electromagnetics, electron energy distribution (Boltzmann equation), hydrocarbon chemistry, and fluid and heat transport. The code was used to vary reactor shape and to determine the flux and uniformity of hydrogen atoms reaching the substrate versus power input and pressure. Experimental measurements by ASTeX scientists validated the results of the simulation. This simulation capability facilitated the development of a more efficient reactor with improved uniformity in large-area deposition. This success notwithstanding, the time scales of the deposition process made it difficult to extend the code's applicability without significant numerical developments.

Pressure=40 torr	Time scale (s)
Electric field establishment	10^{-9}
Electron density build-up	10^{-6}
Diffusion time	10^{-3}
System throughput time (for a diffusion reactor)	10^{0}
Thermal dissociation of hydrogen (T=3000K)	5×10^{-2}
Atomic hydrogen recombination (T=1000K, 50% dissociated)	2×10^{-3}
Vibrational relaxation (T=1000K)	3×10^{-5}
Molecular ion recombination (n_e~$10^{12}cm^{-3}$) (eg, $e^-+H_3O^+\rightarrow H_2O+H$)	10^{-5}
Atomic ion recombination (n_e~$10^{12}cm^{-3}$) (eg, $e^-+H^+\rightarrow H_2+h\nu$)	~10^{0}
Neutral chemistry equilibrium	10^{-3}–10^{-6}

Table 1. Time scales for the components of a plasma-enhanced system.

Self-consistently following plasma generation depends on the time scales of the component processes (Table 1). Using straightforward time integration to model the longtime (steady-state) behavior of these processes is clearly impractical. However, the fast time-scale processes, such as electric-field establishment with self-consistently generated plasma and fixed-background gas, can be simulated by solving for microwave modes in the frequency domain. Then the chemistry, diffusion, and convection processes can be integrated over their much longer time scales. Unfortunately, a frequency-domain solver that includes plasma-response terms in Maxwell's equations does not lead to a converging solution using standard techniques. So we developed an approach that combines a generalized minimal-residual (GMRES) algorithm (which provides fast, guaranteed convergence) and features of a conjugate-gradient technique (which requires much less memory than GMRES).

Using this approach, we are developing a reactor model to simulate high-pressure (about 1 torr) plasma-etching reactors, which are used for removing resists with atomic oxygen and for etching silicon with atomic fluorine. We have also begun modeling low-pressure (about 1 mtorr), high plasma density reactors, in which, in addition to the elements required in high-pressure reactors, we must also model magnetic-field effects, ion transport and dynamics, and ambipolarity and sheath effects. In this reactor model, the fluid approximation for the gas breaks down and is replaced by direct-simulation Monte Carlo techniques.

Although computer simulations of industrial deposition and etching processes have shown recent progress and utility, developing comprehensive, realistic, and useful models remains a long-term, multidisciplinary program. Furthermore, model validation is difficult because most industrial reactors do not have the necessary diagnostics. Nevertheless, increasingly stringent product requirements will give computer simulation an increasing role in parameter optimization and design improvements in coming years.

E. Hyman and S. Eidelman are senior staff scientists at Science Applications International Corporation, Mclean, Virginia.

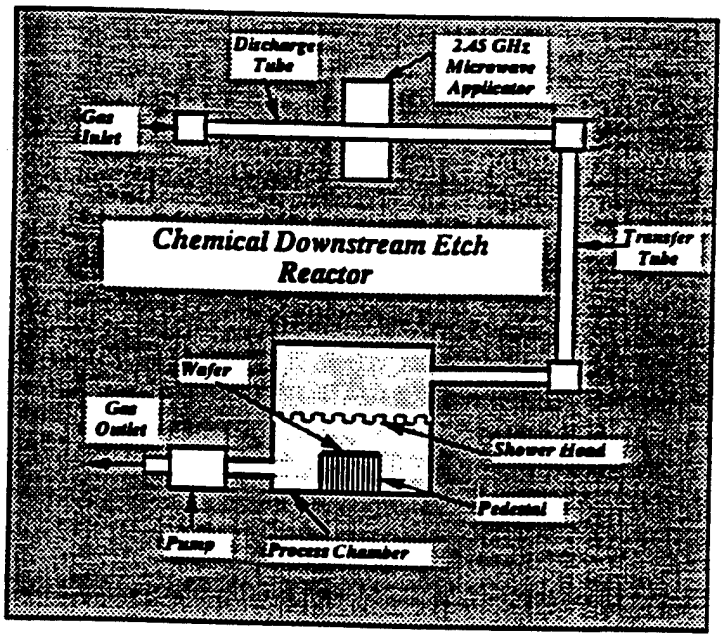
Appendix OO

Modeling Chemical Vapor Deposition and Etching Processes

AIAA 95-0250

Modeling Chemical Vapor Deposition and Etching Processes

E. Hyman, and K. Tsang
Science Applications International Corporation
McLean, VA 22102
and
B. Lane
Plasma Dynamics
Belmont, MA 20178



33rd Aerospace Sciences Meeting and Exhibit
January 9-12, 1995 / Reno, NV

MODELING CHEMICAL VAPOR DEPOSITION AND ETCHING PROCESSES

E. Hyman and K. Tsang
Science Applications International Corporation
McLean, Virginia
and
B. Lane
Plasma Dynamics
Belmont, Massachusetts

Abstract

We describe a modeling program to simulate plasma assisted chemical vapor deposition (CVD) and etching processes generated by a microwave source at 2.45 GHz. The deposition pressure regime we investigate ranges from greater than one Torr to several hundred Torr. The etching regime ranges from a few tenths of a Torr to several Torr. In this pressure regime it is a reasonably good approximation to treat the electrons and ions as essentially local in terms of their energy gain and loss and the resulting neutral chemistry also as local. These phenomena are distinguished from the electromagnetic field development, gas diffusion and convection, and thermal processes each of which is essentially global. As examples of the modeling we present 2D simulations of diamond deposition and detailed one-point models of a CF_4, O_2 etching chemistry.

1. Introduction

Plasma assisted deposition and etching are important industrial processing techniques for coating materials and for etching substrates. Many fundamental elements of the plasma process are not well understood. As more stringent requirements are established for coating uniformity over ever larger substrates, for higher quality of the deposited film, and for increased process reproducibility in the deposition process and as the requirements in the etching process encompass ever smaller feature sizes, better etching selectivity, and stronger anisotropy in the etch it becomes more important to establish which elements in the process are critical to successfully accomplishing these goals. This requires a more complete understanding of the diverse, complex, and highly interactive physical and chemical elements that constitute the plasma assisted reactor process.

In the general plasma assisted etch reactor all of the elements indicated in Figure 1 need to be modeled to generate a comprehensive simulation capability. For example at low pressures (eg., the millitorr regime) the etch process is influenced by the ions reaching the substrate as well as by the etchant, generally a neutral radical. At these pressures the electrons have a mean free path that is comparable to reactor dimensions and the electron heating is nonlocal. In this paper we will consider a specific class of etching reactor that operates in a higher pressure regime, at ~ 1 Torr. One industrial example of this type of reactor is the chemical downstream etch (CDE) reactor. In this reactor the ions do not participate in the process; their presence at the wafer causes unwanted damage to the etched substrate and care is taken in the reactor design to try to exclude them from the process chamber. In addition to this etch reactor we present an example of an important CVD process; we describe and present results of simulations of a 1.2 kW, 2.45 GHz ASTeX microwave diamond deposition reactor operating at a pressure of 40 Torr.

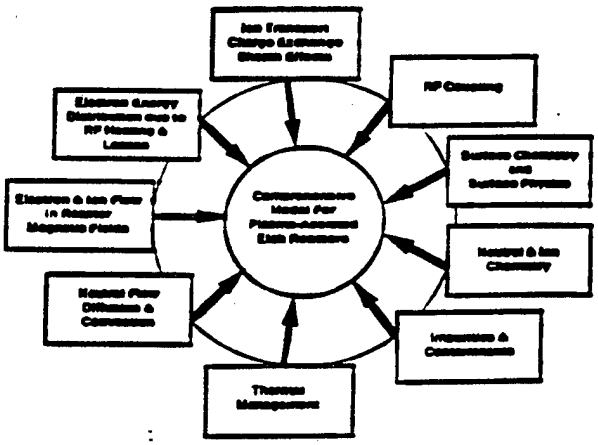


Figure 1. The components of a comprehensive model for the general plasma assisted etch reactor. Here no assumption is made about the locality of electron processes.

The simulation of the complex interactive processes in these plasma reactors requires modeling of:

• Electromagnetics - the electric fields in realistic reactor configurations including the effects of reactor

*Research Scientist, Member AIAA

walls and windows and the self consistently determined plasma generated by the microwave source.

• Plasma Physics - the non-thermal electron energy distribution of the plasma as a function of space and time; needed to calculate the time dependent ionization, dissociation, and temperature rise of the gas.

• Computational Fluid Dynamics - both convective and diffusive regimes in multispecie compressible flow.

• Time Dependent Gas Chemistry - generation of critical radical species reaching the substrate; they can be far out of equilibrium with the gas temperature and will be affected by the presence of other minor species.

• Thermal Control - transfer of heat to the reactor walls by conductivity and diffusion will impact cooling requirements and the development of convective cells.

• Substrate/Wall Heterogeneous Chemistry - the deposition rate on a substrate is finally impacted by the interaction between the substrate surface and the gaseous state radical densities that mediate the deposition or etching process. In addition, the effect of chemistry at reactor walls is often substantial in influencing the radical species that reach the substrate.

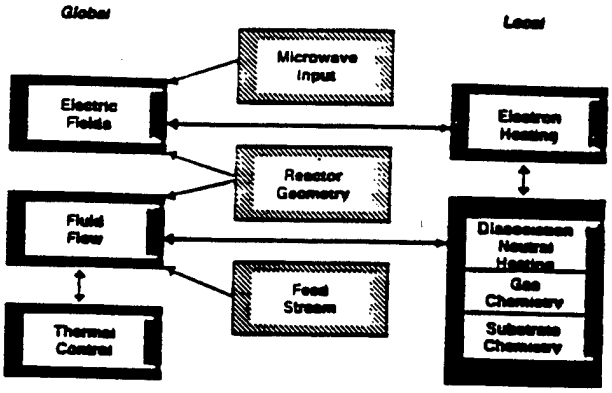


Figure 2. Interaction elements in a plasma assisted CVD or etch reactor in which the gas density is high enough to separate the processes into local and global components. Input elements include the microwave power, reactor configuration, and feed stream. Global elements include the electric field distribution, fluid flow, and thermal transport. Local elements include the electron heating, energy transfer to the gas, gas chemistry, and substrate and wall chemistry.

High Pressure Case

The pressure regime we are modeling is low enough that the plasma electrons do not thermalize. Critical to a reliable simulation is determining the energy distribution of the electrons. It is the high energy tail of the distribution that is most important in determining the ionization, dissociation, and vibrational excitation rates that control the build-up of plasma, the active radical neutral chemistry, and the heating rate of the reactor gas. The pressure regime is, however, high enough that the electron is energized by the microwave field to its steady state value and loses energy to the gas while moving a distance, d, short compared to reactor dimensions (for example, $d = 0.07$ cm for $p = 40$ Torr and $T \sim 2000$ K). This permits a separation of the different interactive physical/chemical processes in the reactor into local and global components. Figure 2 graphically indicates this separation.

As a result of this separation it is convenient to use a one-point Boltzmann equation to determine the electron energy distribution appropriate to a particular microwave generated electric field amplitude, couple this to a very detailed ion and neutral chemistry for that point, and by carrying out one-point simulations for a range of appropriate electric field values develop algorithms that relate the critical electron rates (ionization, dissociation, vibrational excitation) to the electric field at a point. These algorithms can then be used in place of full Boltzmann calculations in multidimensional simulations that incorporate the reactor geometry, fluid flow, and thermal transport. This scheme renders the calculation much more tractable. The one-point calculations are inherently of interest, particularly with the addition of zero order approximations that account for diffusion, wall effects, and heat transport, as they establish with reasonable reliability the gas temperature attained and determine those species that have appreciable densities and those that can be eliminated from the chemistry scheme. Detailed one-point simulations for the ASTeX reactor have been published previously.[1]

In Section 2 we describe the 2D simulations of the diamond deposition process in a microwave CVD reactor. In Section 3 we present results of one-point simulations of a chemical downstream etch reactor with CF_4 and O_2 feed gases. We conclude with a brief summary, Section 4, and acknowledgements for funding support.

2. Diamond Deposition

Figure 3 is a schematic view of a microwave plasma assisted CVD reactor. Typically in current reactors substrates are planar, although one goal of modeling the process is to assist in the design of reactors that can deposit uniformly over conformal

substrates. The input gas is predominantly hydrogen with a small fractional concentration of a carbon containing substance such as methane (CH_4). A small amount of oxygen is often added to improve the process, although detailed understanding of the mechanism whereby this is accomplished is lacking. A general description of diamond deposition rates and the quality of the diamond deposited as a function of the relative fractions of atoms of C, H, and O has been given in a paper by Bachmann, et al.[2] Critical to the efficient deposition of a high quality diamond film (low graphitic component and few defects) is the generation of a high flux of hydrogen atoms to the substrate. The atomic hydrogen apparently serves two functions[3]: (1) it attaches to hydrogen atoms on the surface causing abstraction of molecular hydrogen from the surface, activating a surface site so that a hydrocarbon radical can attach to the surface and (2) is instrumental in healing defects and reducing graphitic deposition, thus improving the quality of the deposition. Which hydrocarbon radical is critical to the growth of diamond on the surface has been controversial. Leading candidates include the methyl radical, CH_3, and acetylene, C_2H_2. It may be that the particular hydrocarbon is less important than whatever is the dominant radical hydrocarbon flux that is reaching the surface.

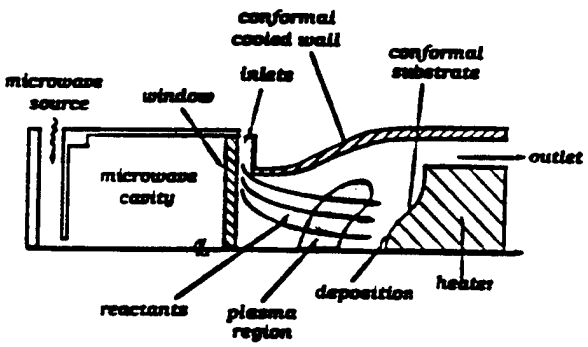


Figure 3. Schematic view of a plasma assisted CVD reactor showing the microwave input region, cavity, and window, the gas inlets and outlets, the location of the substrate, and the plasma region.

At 40 Torr the time scale for different phenomena in the ASTeX microwave reactor varies from ~ 10^{-9} s for the time to establish the microwave electric field to ~ 10^{0} s for the system throughput time of the diffusion reactor. Table 1 illustrates the range of time scales that occur. Reference 1 has presented detailed one-point simulations of this reactor. In this paper we present 2D calculations including the self consistent determination of the plasma generated by the electromagnetic field and diffusion, specifically for the ASTeX reactor configuration. The relevant equations were given previously.[4]

Table 1. Time Scales in Microwave Reactor

Pressure	Time Scale (s)
Electric Field Establishment	10^{-9}
Electron Density Build-up	10^{-6}
Diffusion Times	10^{-3}
System Throughput Time (for a Diffusion Reactor)	10^{0}
Thermal Dissociation of Hydrogen (T = 3000 K)	$5x10^{-2}$
Atomic Hydrogen Recombination (T = 1000 K, 50% Dissociated)	$2x10^{-3}$
Vibrational Relaxation (T = 1000 K)	$3x10^{-5}$
Molecular Ion Recombination (n_e ~ 10^{12} cm^{-3}) (eg, $e^- + H_3O^+ \rightarrow H_2O + H$)	10^{-5}
Atomic Ion Recombination (n_e ~ 10^{12} cm^{-3}) (eg, $e^- + H^+ \rightarrow H + h\nu$)	~ 10^{0}
Neutral Chemistry Equilibrium	10^{-3} - 10^{-6}

Results

When the microwave energy source is turned on, an electric field builds up and generates a plasma. The plasma electrons dissociate hydrogen and heat the gas. The temperature of the gas in the center of the plasma will rise until it reaches about 3000 K. At this temperature thermal dissociation (a cooling process) begins to dominate over electron dissociation (a heating process). The gas temperature stabilizes unless sufficient power is supplied to totally dissociate the hydrogen; in that case the temperature can go higher. Once the temperature within the active region of the gas has become high enough that atomic hydrogen readily converts methane to the methyl radical, for example, electron initiated reactions become of minor importance, except that by exciting the vibrational levels of hydrogen they maintain the elevated gas temperature. In this calculation we assume that all molecules reaching the cold walls reflect off unchanged. At the wafer the different species are divided into three classes based on assumed reactivity with the wall in terms of sticking

coefficients: highly reactive = 0.1, moderately reactive = 0.01, and unreactive = 0.0. A highly reactive species such as H will react with the wafer and come off as H_2.

After a few milliseconds a steady state will be established: many radical species will, however, have densities far in excess of their thermal equilibrium values. The steady state is determined by the interaction among chemical, plasma, and diffusion processes. The following calculation assumed a feed stream of 99% hydrogen and 1% methane.

Figure 4 shows a plot of the magnitude of the steady state electric field in volts/cm in an ASTeX reactor (microwave power = 1.2 kW, pressure = 40 Torr). In this figure the left vertical axis (Z) corresponds to the axis of the reactor. The microwaves are fed in through a dielectric window along the horizontal axis (R) between the reactor axis and ~ 6 cm. The substrate (wafer) is at Z ~ 27 cm and is located between the reactor axis and ~ 4 cm. The output pumping region is near the top right between 6 and 8 cm. The reactor walls are assumed electrically conducting as is the wafer. The walls are maintained at ~ 300 K and the wafer surface at 1200 K. The wafer is thermally conducting. Maxima in the electric field occur at sharp corners and along the reactor axis near the dielectric window. A minimum occurs along the axis between 19 and 20 cm. The next maximum that would occur is altered by the existence of the plasma, centered on axis between 23 and 24 cm (Figure 5), and comprises a relatively flat region in electric field. This is a result of the shielding of the field by the plasma.

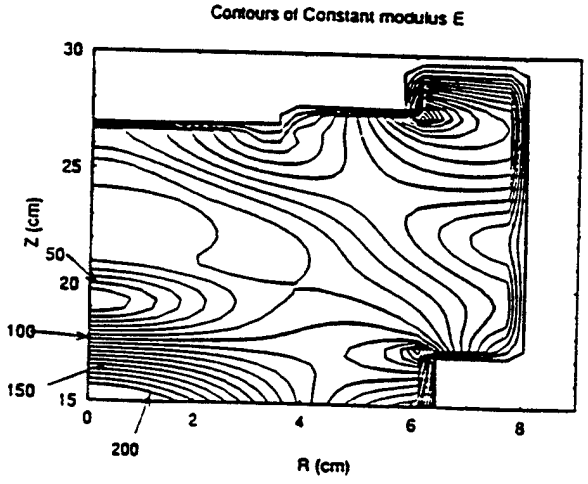


Figure 4. Contours of the amplitude of the electric field, E, in volts/cm. Vertical axis, Z, is the reactor axis. Dielectric microwave window is coincident with the horizontal axis out to ~ R = 6. cm. Substrate surface is at ~ Z = 27. cm out to R = 4. cm. Gas pumping outlet region is at top right from about 6. to 8. cm.

Figure 5 shows contours of electron density in the configuration described for Figure 4. It indicates a maximum in electron density on axis just above 10^{11} cm^{-3}. Figure 6 shows contours of gas temperature (K) for the same configuration as in Figure 4. The temperature maximum occurs near the center of the plasma region and is ~ 3000 K as anticipated from the discussion above. The temperature at the substrate surface is 1200 K and drops off within the substrate, due to its thermal conductivity.

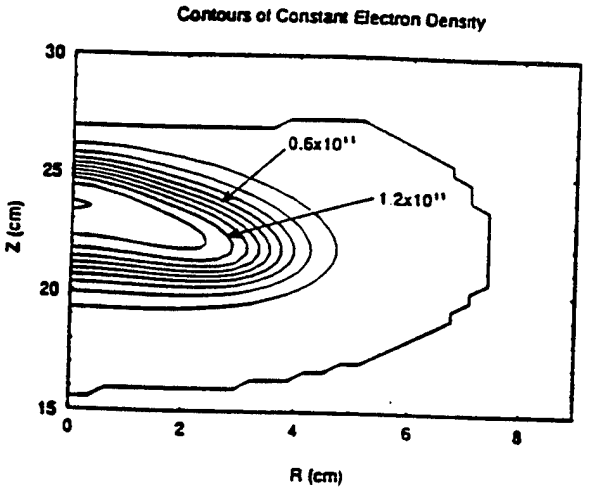


Figure 5. Contours of electron density (cm^{-3}) in the reactor configuration described in Figure 4.

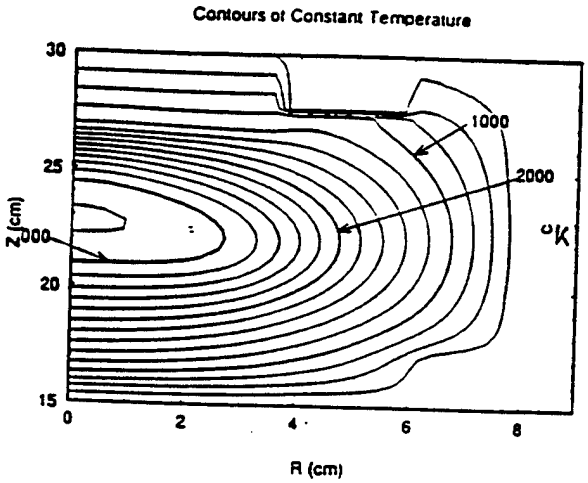


Figure 6. Contours of temperature (K) in the reactor configuration described in Figure 4.

Figure 7 shows contours of H_2 density in cm^{-3}. This is the same reactor as in the configuration described in previous figures but here we are viewing a vertical cut through the reactor axis and have inverted the figure relative to the previous ones. Now the substrate is near the bottom of the figure Z = ~ 3 cm and the microwave window is at the top near Z = 15 cm. The reduction in H_2 density near the center of the reactor is partly a temperature effect; high temperature requires a reduced density to maintain constant pressure. In addition, it is partly due to dissociation of the hydrogen.

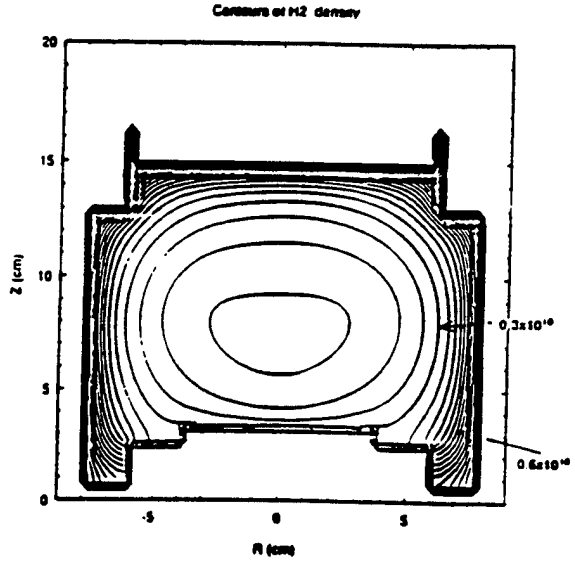


Figure 7. Contours of H_2 density in cm^{-3}. Here the figure has been inverted relative to the previous ones and the reactor axis is the vertical line at R = 0.0 cm. The substrate is at Z ~ 3 cm and from R = 0.0 to R ~ 4. cm. The microwave window is near Z = 15. cm.

Figure 8 shows contours of atomic hydrogen, H. The density is maximum, ~ 10^{17} cm^{-3}, at the center of the plasma as expected. Note that the H density is very uniform over the wafer surface. Figure 9 shows contours of the methyl radical, CH_3, in cm^{-3}. Here the maximum is not at the center of the plasma. This is partly a temperature effect but also reflects the fact that the high temperatures have broken it down somewhat into other molecules such as CH_2 and also converted it to molecules with two C atoms by means of reactions like $CH_3 + CH_3 + M \rightarrow C_2H_6 + M$. The dominant resulting two C atom molecule is acetylene, C_2H_2. Figure 10 shows acetylene contours in cm^{-3}. While the acetylene densities are quite high, they are not maximum in the center due to the temperature effect. Finally, Figure 11 illustrates a relatively minor specie, molecular carbon, C_2. This is maximum in the center because only here are temperatures high enough for it to be produced. The interest in C_2 follows because there is a very strong emission band in the visible spectrum, the Swan band, that serves as an important diagnostic in the diamond deposition process.

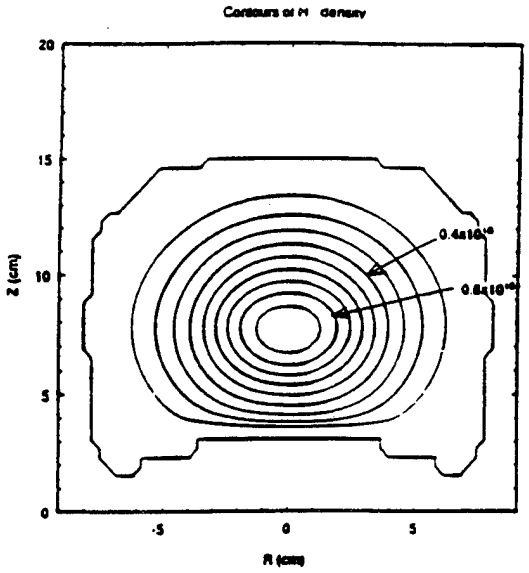


Figure 8. Contours of H density in cm^{-3} with the configuration described in Figure 7.

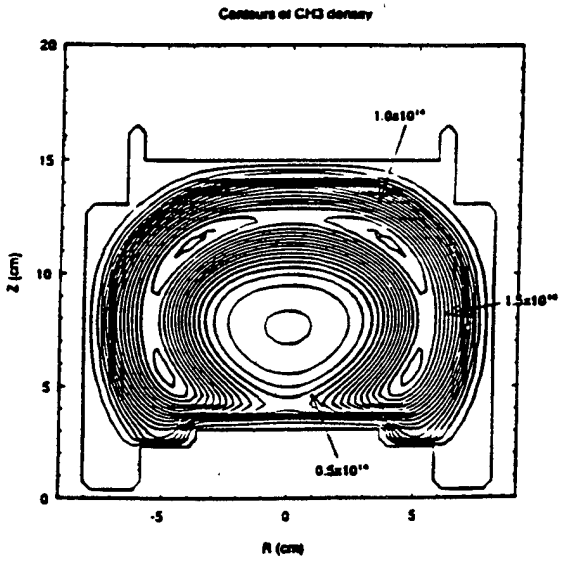


Figure 9. Contours of CH_3 density in cm^{-3} with the configuration described in Figure 7.

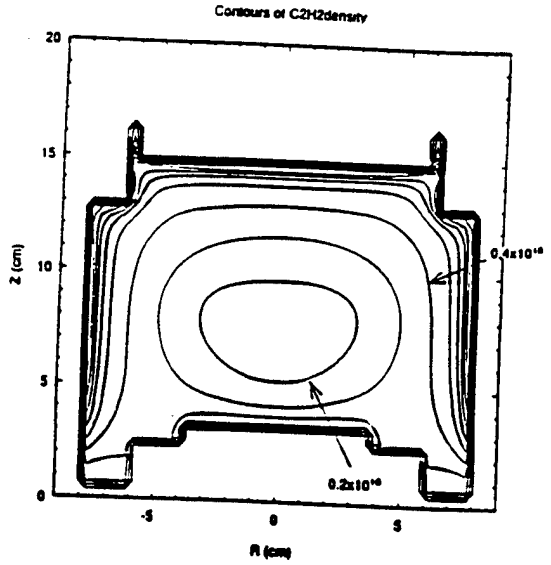


Figure 10. Contours of C_2H_2 density in cm^{-3} with the configuration described in Figure 7.

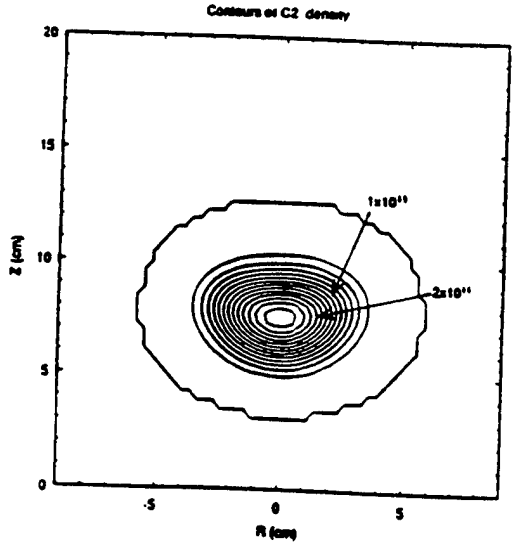


Figure 11. Contours of C_2 density in cm^{-3} with the configuration described in Figure 7.

3. Chemical Downstream Etch Reactor

We present here one-point calculations appropriate to a chemical downstream etch (CDE) reactor using CF_4 and O_2 as the input gases. The source region in the CDE reactor consists of a long thin tube of quartz through which a microwave discharge is initiated. Here the chemistry is less well known than in the hydrocarbon case. A reasonable first step toward understanding the essentials of the CF_4, O_2 chemistry comes as a result of the work of Plumb and Ryan.[5] They performed a number of measurements of critical rates, measured specie concentrations in CF_4/O_2 discharges, and performed numerical modeling to identify the critical reactions and eliminate from the model those that do not make any measurable difference. Starting with a total of 49 reactions they extracted 13 reactions that they claimed are sufficient to simulate all of their experimental results. While this is an extremely valuable study and we have used it tentatively as the basis of our model, caution is in order. First, all results were based on a gas temperature of 300 K; the effect of plasma heating of the gas was not considered. In our modeling, at the power levels in the CDE reactors we model, we predict temperatures at the center of the tube will rise well above 1000 K. This will change some reaction rates and may make some reactions that had a negligible effect at 300 K important. In addition, independent of the temperature effect, the electron density and the electron energy distribution from a microwave source may be sufficiently different from that of the 13.56 MHz rf source used by Plumb and Ryan as to induce different electron dissociation rates, heating rates, etc. and thereby change the chemistry to the extent that other reactions need to be considered.

In an independent experimental study by Kiss and Sawin[6] of a pure CF_4 discharge, using modulation of the power source to uncover dominant reaction rates, they confirm much of the work of Plumb and Ryan but propose the existence of a long lived excited state of CF_3 to explain their experimental results. They also include the specie, C_2F_6, which is formed in a three body reaction involving two CF_3's. This reaction was omitted by Plumb and Ryan from their simplified model and we also have omitted it from our model. Finally, the Kiss and Sawin study allowed them also to establish sticking coefficients for F, CF_2, and CF on the walls of their tube (anodized aluminum) and they concluded that the wall chemistry dominated over the gas chemistry.

The relative importance of chemical recombination and ambipolar diffusion in the reactor will be influenced by the details of the ion chemistry, in particular, by the mass of the dominant ions. The ion chemistry module associated with CF_4/O_2 chemistry includes the positive ions: CF_3^+, O_2^+, F^+, and O^+. It also tracks the following negative ions: CF_3^-, O_2^-, O^-, and F^-. Electron processes produce CF_3^- (dissociative ionization), the positive oxygen ions (ionization), and the negative ions (attachment and dissociative attachment). The major chemical deionization processes include dissociative recombination via electrons and mutual recombination of positive and negative ions. A fast ion-neutral rearrangement reaction: $O^+ + CF_4 \rightarrow CF_3^+ + O + F$ is responsible for converting the major ion from O^+ to CF_3^+.

To provide an approximate estimate of the effect of thermal conductivity in the one-point model we take the distance from the center of the tube (the single calculation point) to the quartz wall in the source region and incorporate the affect of thermal conductivity through the quartz wall, fixing the outside wall temperature at 300 K. Because of the dimensions of the tube most of the heat is lost radially through the wall. A very simple model for wall effects has been incorporated in the model that assumes all species other than atomic fluorine, F, reflect off the wall. F is assumed to interact with the quartz and come off as the volatile product SiF_4.

Results

We present results of a one-point calculation for a point on the tube axis at the center of the plasma discharge region. The assumed conditions are: power = 700 W $\rightarrow$ ~ 3 W/cm^3, pressure = 0.75 Torr, $[CF_4]_0/[O_2]_0 = 1$. The simulation was carried out over 2 ms.

Figure 12 plots the average electron energy and the neutral temperature on axis, and the temperature of the inside wall of the quartz tube, as a function of time from turn-on of the microwave source to 2 ms. The spiky behavior of the electron energy is not real but numerical and can be ignored. By ~ 0.5 ms the electron energy reaches a steady value corresponding to ~ 1 ev. The inside wall temperature is just a little above 300 K. The gas temperature on axis approaches a steady state value near 1400 K.

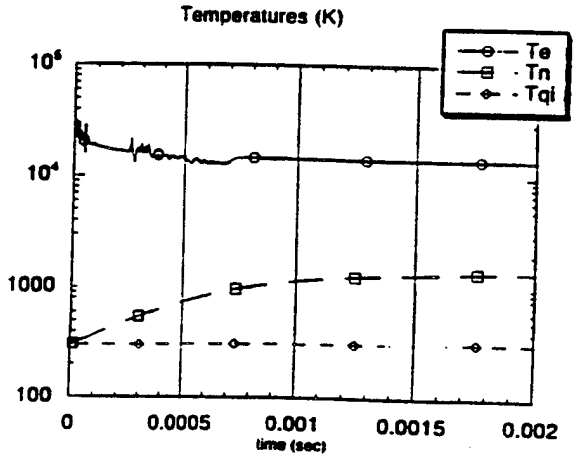


Figure 12. Average electron energy (T_e) in units of Kelvin, gas temperature (T_n), and inside temperature of the quartz tube (T_{qi}) as a function of time from turn on of the microwave source to 2 ms.

Figure 13 shows the time dependent generation of the neutral carbon-fluorine species and Figure 14 the carbon-oxygen species as well as the build-up of SiF_4 on axis. Because pressure stays constant, as the gas temperature rises the number densities are reduced, independent of any effects of chemical reactions. In 2 ms the CF_4 has been reduced by about a factor of 5 and the O_2 by about an order of magnitude. The major species generated by the chemistry are atomic oxygen (O) and atomic fluorine (F). Next in importance are carbon dioxide (CO_2) and COF_2. The SiF_4 from the quartz wall is building up but is still a relatively minor specie. Figure 15 exhibits the development of the electron density and the carbon-fluorine ions and Figure 16 the oxygen ions. The dominant ion is clearly CF_3^+; O_2^+ makes a small contribution but the other species are generally negligible.

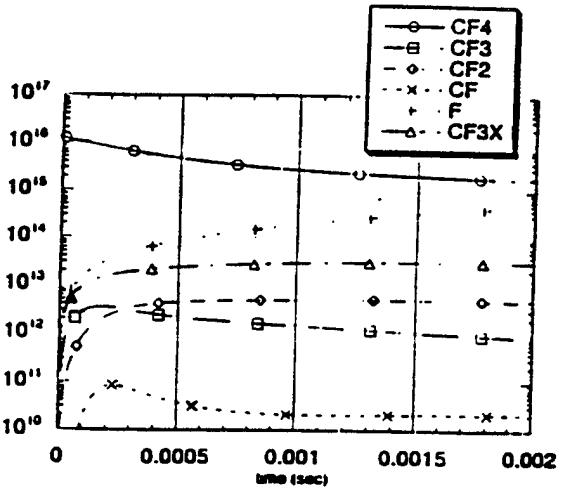


Figure 13. Time dependent generation of carbon-fluorine species from microwave turn on to 2 ms.

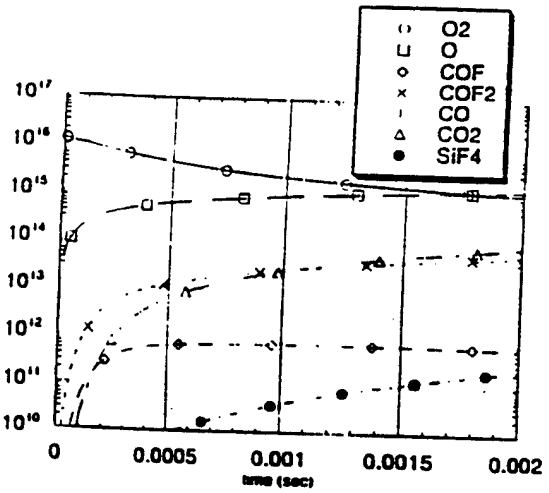


Figure 14. Time dependent generation of carbon-fluorine-oxygen species and SiF_4 from microwave turn on to 2 ms.

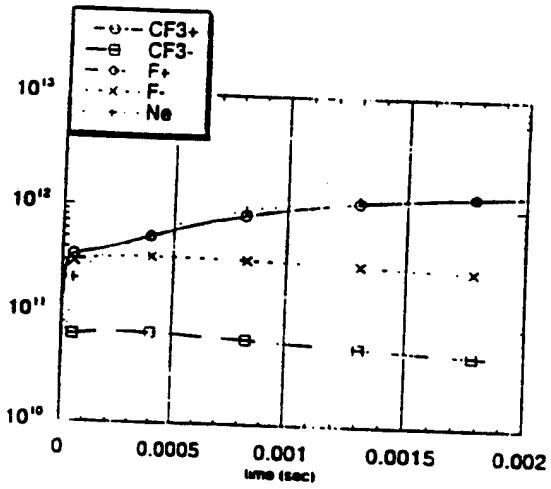


Figure 15. Time dependent generation of the electron density and non-oxygen containing ions from microwave turn on to 2 ms.

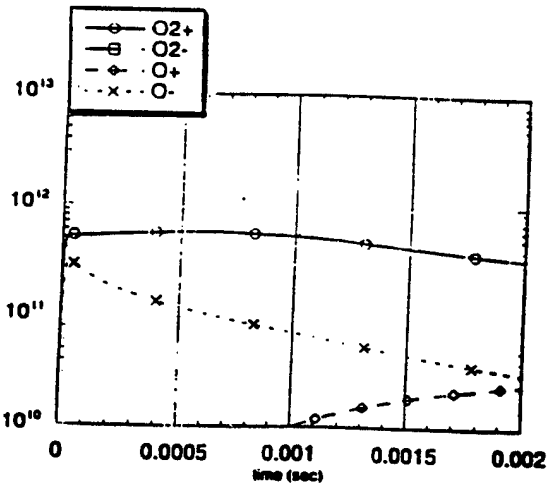


Figure 16. Time dependent generation of the oxygen containing ions from microwave turn on to 2 ms.

To summarize the results of this calculation, starting with equal amounts of CF_4 and O_2 at a pressure of 40 Torr and a microwave power density of $3W/cm^3$, after 2 ms approximately 25% of the neutral gas at the center of the plasma was atomic oxygen and 10% was atomic fluorine. The gas temperature at that point had risen to about 1400 K. The dominant ion was CF_3^+.

4. Summary

Plasma assisted deposition and etch processes comprise an important industrial tool. Historically, the design of reactors and the choice of working parameters has been made based on the experience of engineers and on their intuition as to the effect a given change will have on the process. Lack of detailed knowledge of the complex interactive physical and chemical processes and the ability to simulate those processes inhibits the development of improvements, particularly as process requirements become more stringent. We have exhibited a numerical simulation capability that includes detailed modeling of the chemistry, fluid flow, electromagnetic, and plasma effects, appropriate where electron heating occurs on spatial scales small compared to reactor dimensions. As examples of the capability we showed results of (1) 2D calculations of an ASTeX microwave chemical vapor deposition reactor for diamond deposition and (2) one-point calculations of a chemical downstream etch reactor using CF_4 and O_2 as the feed gases.

Acknowledgments

The work on diamond CVD was partially performed under subcontract to Applied Science and Technology and funded by DARPA under contract DAAH01-92-C-R109 as well as the SAIC Independent Research and Development Program. The work on etch chemistry has been funded by SEMATECH under contract 74054993.

References

1. Hyman, E., K. Tsang, and B. Lane, "One-point Numerical Modeling of Microwave Plasma Chemical Vapor Deposition Diamond Deposition Reactors," J. Vac. Sci. Technol. A **12**, 1474 (1994).
2. Bachmann, P. K., D. Leers, and H. Lydtin, "Towards a General Concept of Diamond Chemical Vapor Deposition," *Diamond and Related Materials I*, Elsevier Science Publishers B.V., 1 (1991).
3. Butler, J. E. and R. L. Woodin, "Thin Film Diamond Growth Mechanisms," Phil. Trans. R. Soc. Lond. A **342**, 209 (1993).
4. Hyman, E., K. Tsang, A. Drobot, and B. Lane, "Modeling Diamond Deposition in a Plasma Assisted CVD Reactor," AIAA 92-3016, AIAA 23rd Plasmadynamics & Lasers Conference, July 1992, Nashville, TN.
5. Plumb, I. C. and K. R. Ryan, "A Model of the Chemical Processes Occurring in CF_4/O_2 Discharges Used in Plasma Etching," Plasma Chemistry and Plasma Processing 6, 205 (1986).
6. Kiss, L. D. B. and H. H. Sawin, "Evaluation of CF_4 Plasma Chemistry by Power Modulation," Plasma Chemistry and Plasma Processing **12**, 523 (1992).

Appendix PP

Role of Impurities in O_2 Chemical Dry Etching

Role of Impurities in O_2 Chemical Dry Etching

G. S. Oehrlein, B. E. E. Kastenmeier, and P. M. Matsuo
Department of Physics
State University of New York, Albany
1400 Washington Avenue
Albany, N.Y. 12222

Abstract

This report summarizes results of an investigation of the role of N_2 and H_2 impurities on etching behavior of a resist thin film located downstream from the primary discharge region. Investigations of the effect of N_2 and H_2 in O_2 on (i) downstream resist etch rates, (ii) oxygen atom density in the discharge region, (iii) reactive intermediates in the reaction chamber as measured by mass spectrometry, and (iv) surface modifications of the resist are reported. Two distinct mechanisms of the enhancement of the resist etch rate upon adding either H_2 or N_2 were identified.

TABLE OF CONTENTS

1 Introduction

In stripping of organic films using downstream plasma systems, gases are frequently added to research grade O_2 (99.999% purity) to improve the gasification of the organic materials the etch rate of organic material placed in the downstream chamber is low. This report describes the progress which was made in September through November 1995 by the group at the State University of New York, Albany, under a contract with SAIC on the influence of N_2 and H_2 addition to 99.999% purity O_2 on the efficacy of downstream resist stripping.

2 Experimental Apparatus and Procedures

An ultrahigh vacuum-compatible microwave-based chemical dry etching apparatus was used in this work. It has been described previously[1] and is depicted in Fig. 1. Briefly, a microwave plasma is produced in an ASTeX microwave plasma applicator that employs an alumina coupling dielectric to minimize discharge-induced erosion of the wall material. The microwave power (2.54 GHz) was varied between 500 W and 1200 W for this work. The sample is located in a cylindrical process chamber (inner diameter 27 cm, height 30 cm). A flow of typically 1100 sccm 99.999% purity O_2 was employed and various amounts of N_2 or H_2 were added to study the effect of these impurities on downstream stripping performance. The system is pumped using a turbomolecular pump backed by a roots blower and a vane pump. For the experiments described here, the pressure was between 0.5 and 1 Torr. A throttle valve was used to control the pumping speed and maintain this pressure in the chamber. For the current studies the microwave applicator is connected directly to the dry etching chamber and the species produced in the microwave discharge have to travel horizontally a distance of 30 cm to reach the sample.

A number of 200 mm wafers covered with resist films were provided by SEMATECH. For the etching and surface analysis experiments, 1-inch by 1-inch squares were cut from the 200 mm wafers. These pieces were mounted on 125 mm diam. silicon carrier wafers using a thermal glue. The wafers were mounted on a heated platen at a temperature of 300°C to maximize the stripping rate of the resist layers. For these experiments the thermal contact between the platen and the sample was determined by the pressure of the process gas. Experiments were also conducted with an electrostatic chuck with a backside helium pressure of 5 Torr to improve temperature control of the sample. For these studies the temperature of the wafer was in the range 30°C to 50°C. Qualitatively similar enhancements of the stripping rates as a result of adding N_2 or H_2 to O_2 than seen at the higher substrate temperatures were observed, but the absolute rates were several orders of magnitude lower. The low temperature results were of interest since they allowed to gain insight in the surface processes at the gas-resist interface by "freezing-in" the surface modifications produced by the exposure of the resist surface to the excited O_2/N_2 and O_2/H_2 gas mixtures.

In-situ ellipsometry, optical emission spectroscopy (OES) and mass spectrometry were used for real-time plasma diagnostics. The etching chamber is also connected via a UHV central wafer handler to both a load-lock and a multi-technique surface analysis chamber. The UHV wafer handler is used to load/unload samples to/from the CDE chamber while rigorously maintaining the cleanliness of the chamber. Processed specimens were transported under vacuum to the multi-technique surface analysis system for x-ray photoemission spectroscopy measurements.

3 Experimental Results and Discussion

3.1 Resist Etching

The effect of adding H_2 to O_2 and N_2 to O_2 on the downstream stripping rates of resist has been summarized in Figs. 2 and 3 for O_2/H_2 and Figs. 4 and 5 for O_2/N_2, respectively. Different microwave power and pressure combinations were used. The significant observations are:

- Adding about 1% of H_2 to 99.999% pure O_2 increases the stripping rate of resist by a factor of 5. If more H_2 is added, the etch rate increases only slightly, reaches a maximum at about 2% to 3% H_2, and subsequently decreases again (at 1 Torr pressure, Fig. 2), or no longer changes (0.6 Torr pressure, Fig. 3).
- A similar rapid increase (approximately 4x etch rate enhancement) and subsequently plateau for low N_2 flows is seen if N_2 is added to O_2 (Fig. 4 (a), and expanded in Fig. 4 (b)). If the amount of N_2 is increased beyond 15%, a steady increase of the etch rate is seen in this case (Fig. 4(a)). This is shown more clearly in Fig. 5, where the N_2 flow is increased to a value of 2x the O_2 flow. The etch rate does not saturate in this case.

These results show that for low impurity concentrations (up to 3% of the O_2 flow) H_2 and N_2 apparently enhance the resist etch rate in a similar fashion by a factor of 4x to 5x. At higher flows H_2 no longer enhances the stripping rate, but the etch rate actually starts to decrease. On the other hand, for N_2 a very significant increase is observed at hgih levels of N_2, which is much greater (about a factor 25x in Fig. 5) than the etch rate enhancement at the 2% N_2 level.

3.2 Optical Emission Spectroscopy of Oxygen Atom Abundance

Actinometry was used to obtain information on relative changes in oxygen abundance in the discharge region upon injecting either H_2 or N_2. The 844.625 nm, 844.636 nm and 844.676 nm triplet was used for this work, since it correlates well with the atomic oxygen abundance. These results are summarized in Figs. 6 (a) and (b), 7 and 8. We note the following:

- Adding H_2 at the 1 to 3 % level to O_2 increases the O atom intensity. Increases of H_2

beyond 3% no longer lead to a significant increase of the O atom intensity in the discharge region.

- A similar effect on O atom emission is seen if N_2 is added at levels from 1 to 10%. This is the region where a 4-fold increase of the etch rate is seen.
- Adding N_2 at levels exceeding 50% of the O_2 flow leads to a continuous decrease of the O atom emission.

These data shows that the initial rapid increase of the resist etch rate upon adding H_2 or N_2 at levels of 1 to 3% can be explained by the increased production of atomic oxygen. On the other hand, the much greater etch rate increase seen at high levels of N_2 addition cannot be explained by increased oxygen atom production, since the oxygen atom concentration actually decreases at that point.

It has been stated that oxygen atoms in the singlett D state are chemically much more reactive than oxygen atoms in the triplett P state. We used a high-resolution OES system with a 750 mm monochromator to determine the ratio of oxygen atoms in the singlett D state to those in the triplett P state. These results are shown in Figs. 9 and 10 for O_2/H_2 and O_2/N_2, respectively. No significant change in the relative population of these states is indicated by these data.

3.3 Comparison of Discharge OES and Air Afterglow

Figure 11 shows a broad optical emission spectrum from the discharge region of a H_2/O_2 ratio of 0.1. A corresponding spectrum for an N_2/O_2 ratio of 0.1 is shown in Fig. 12. For the H_2/O_2 gas mixture emission from O, O_2^+, H, and OH is observed. For the N_2/O_2 gas mixture emission from O, N_2 and NO is seen. In this case a faint afterglow is seen in the reaction chamber. The afterglow spectrum is shown in Fig. 13 and extends from about 400 nm to 850 nm. The air afterglow has been studied by Kaufman[2] and is due to the reaction of NO with atomic O to form NO_2. The integrated intensity of the afterglow as a function of added N_2 is shown in Fig. 14.

3.4 Mass Spectrometry

Mass spectrometry was applied in the reaction chamber to obtain information on relative changes in arrival rates of important chemical species with gas composition in the discharge region. In Fig. 15 (a) and (b) we show the data for O_2/H_2 gas mixtures. Figure 15 (a) shows an apparent decrease of the oxygen atom arrival rate in the downstream chamber with increasing H_2 flow. This is surprising, considering that (i) the OES measurements on the discharge region show an increase in the production of O atoms, and (ii) the etch rate of the resist films increases. Since O atoms in the mass spectrometer are produced from a variety of different parent molecules, e.g. O_2 in addition to O, this result may just indicate a depletion of one of these, rather than an

actual depletion of oxygen atoms. Figure 15 (b) shows that the production of OH increases linearly with the flow of added H_2.

Figures 16 (a), (b), (c) and (d) show results for O atoms, NO, NO_2, and N_2O, respectively, as a function of N_2 added to O_2. Again we find that the O atom signal just decreases with added N_2 (the O_2 flow was held constant at 600 sccm). The NO signal increases and then decreases. The qualitative shape of the NO curve resembles very closely the air afterglow signal. The signal from NO_2 initially increase rapidly with added N_2, and then shows a continuous decrease. On the other hand, N_2O increases in a roughly linear fashion over the whole range of N_2 flows studied.

Figures 17 (a), (b), (c) and (d) show the corresponding results for O atoms, NO, NO_2, and N_2O, respectively, at low flows of N_2. The O_2 flow was 1100 sccm in this case. The observed behavior is consistent with that seen in Fig. 16.

The dependence of the air afterglow intensity on N_2 flow is qualitatively very similar to the NO mass spectrometry signal. Since the air afterglow intensity is proportional to the concentrations of O atoms and NO, a comparison of Fig. 14 and Figs. 16 (b) and 17 (b) should allow us to obtain the relative variation of the O atom concentration in the downstream chamber. The signal obtained from this comparison is plotted in Fig. 18 as a function of N_2 flow. It increases roughly linearly with N_2 flow. A comparison of the O atom data of Figs. 7, 8, 16(a), 17(a) and 18 shows considerable differences which need to be explained. The actinometric method should provide reliable results in the discharge region. For the downstream regime titration of oxygen atoms is the method of choice and needs to be applied in future work to obtain more reliable data.

3.5 Surface Analysis Results

The influence of gas composition on the modifications of resist surfaces was examined using x-ray photoemission spectroscopy after vacuum-transfer. Figures 19 (a) and (b) show the carbon 1s and oxygen 1s of resist surfaces after treatment in O_2/H_2 downstream plasmas. A control resist surface without treatment is shown for comparison. Because of the insulating character of the resist layers, the resist surfaces charge up positively as a result of the electron emission process which increases the apparent binding energy of photoelectrons. No corrections were made for this artifact. The carbon 1s spectrum shows three components, which can be assigned to C-H, C-C type bonding, C-O, C=O, and CO_2 bonding (in order of increasing binding energy). We see the most highly oxidized resist surface for pure O_2 discharges. As the amount of H_2 is increased, the oxidation of the carbon atoms in the surface region decreases and the C-H carbon intensity increases. The same trends can be seen in the oxygen 1s data of Fig. 19(b).

Results for O_2/N_2 discharges are shown in Figs. 20 (a) through (c) for carbon 1s, oxygen 1s and nitrogen 1s, respectively. The N_2/O_2 flow ratios are 0, 0.02 and 1.0. In this case

the carbon 1s signal shows significant oxidation for all N_2/O_2 flow ratios. A nearly constant level is exhibited by the O 1s signal in Fig. 20 (b). Figure 20 (c) shows a small amount of nitrogen incorporation in the case of nitrogen containing gas mixtures.

We conclude that O_2/H_2 and O_2/N_2 gas mixtures behave very differently with regard to the surface oxidation of the resist layer: For O_2/H_2 gas mixtures the amount of surface oxygen decreases rapidly as small levels of H_2 are added, whereas for O_2/N_2 the amount of surface oxygen is basically constant up to very high flows of N_2.

4 Conclusions

The dependence of resist etch rate, oxygen atom emission, and mass spectrometric signal of important gas fragments has been studied as a function of gas composition for O_2/H_2 and O_2/N_2. We find a 4x to 5x enhancement of the resist etch rate upon addition of 1% to 3% of H_2 or N_2. At the same time the density of O atoms in the discharge region increases. This effect is apparently the same for H_2 and N_2. Increasing the H_2 concentration further does not result in further increases of the resist etch rate. This can be explained by the surface analysis results, which show a rapid reduction of oxygen on the resist surface as the amount of H_2 in the O_2/H_2 gas mixture is increased. The results are very different for O_2/N_2. Very large enhancements (25x) of the resist etch rate at high flows of N_2 are observed. These enhancements may be related to the formation of NO, NO_2, and N_2O. The surface characterization shows in this case a nearly constant level of surface oxidation up to a N_2/O_2 flow ratio of 1, and a small level of nitrogen incorporation.

References:

1. J. J. Beulens, B. E. E. Kastenmeier, P. J. Matsuo, and G. S. Oehrlein, Appl. Phys. Lett. **66,** 2634 (1995).

2. F. Kaufman, Proc. R. Soc. **247,** 123 (1958).

FIGURE CAPTIONS

Figure 1: Schematic of the CDE tool used in this work.

Figure 2: Photoresist etch rate vs H_2/O_2 at a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm. The different curves represent microwave powers of 500W and 1000W.

Figure 3: Photoresist etch rate vs H_2/O_2 at a constant oxygen flow of 1100 sccm. The different curves represent chamber pressures of 600 and 1000 mTorr. The chuck temperature was 300°C and the plasma was fed with 1000W microwave power.

Figure 4a: Photoresist etch rate vs N_2/O_2 for a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm up to a N_2/O_2 ratio of 1. The chuck temperature was 300°C and the plasma was fed with 1000W microwave power.

Figure 4b: Photoresist etch rate vs N_2/O_2 for a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm up to a N_2/O_2 ratio of 0.1. A blow up of the low N_2 region of figure 4.

Figure 5: Photoresist etch rate vs N_2/O_2 for a chamber pressure of 1.0 Torr and an O_2 flow of 600 sccm up to a N_2/O_2 ratio of 2. The chuck temperature was 300°C and the plasma was fed with 1000W microwave power.

Figure 6a: Atomic oxygen concentration in the discharge vs H_2/O_2 as determined by Ar actinometry for a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm. The plasma was fed with 1000W microwave power.

Figure 6b: Atomic oxygen concentration in the discharge vs H_2/O_2 as determined by Ar actinometry for a chamber pressure of 1.0 Torr and an O_2 flow of 600 sccm. The plasma was fed with 1000W microwave power.

Figure 7: Atomic oxygen concentration in the discharge vs N_2/O_2 as determined by Ar actinometry for a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm. The plasma was fed with 1000W microwave power.

Figure 8: Atomic oxygen concentration in the discharge vs N_2/O_2 as determined by Ar actinometry for a chamber pressure of 1.0 Torr and an O_2 flow of 600 sccm. The plasma was fed with 1000W microwave power.

Figure 9: Ratio of the O $3p'^1D$ to the O $3p^3P$ states in the discharge as a function of H_2/O_2 for a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm. The plasma was fed with 1000W microwave power.

Figure 10: Ratio of the O $3p'^1D$ to the O $4p^3P$ states in the discharge as a function of

N_2/O_2 for a chamber pressure of 1.0 Torr and an O_2 flow of 600 sccm. The plasma was fed with 1000W microwave power.

Figure 11: Broad OES spectrum for a H_2/O_2 ratio of 0.1.

Figure 12: Broad OES spectrum for a N_2/O_2 ratio of 0.1.

Figure 13: Air afterglow spectrum for an N_2/O_2 ratio of 0.1.

Figure 14: Air afterglow intensity vs N_2/O_2 at a chamber pressure of 1.0 Torr. The different curves represent O_2 flows of 600 and 1100 sccm.

Figure 15a&b: Mass spectrometry plasma on-plasma off determined (a) O and (b) OH concentrations vs H_2/O_2 at a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm.

Figure 16a-d: Mass spectrometry plasma on-plasma off determined (a) O, (b) NO, (c) NO_2 and (d) N_2O concentrations vs N_2/O_2 at a chamber pressure of 1.0 Torr and an O_2 flow of 600 sccm up to a N_2/O_2 ratio of 2.

Figure 17a-d: Mass spectrometry plasma on-plasma off determined (a) O, (b) NO, (c) NO_2 and (d) N_2O concentrations vs N_2/O_2 at a chamber pressure of 1.0 Torr and an O_2 flow of 1100 sccm up to a N_2/O_2 ratio of 0.1.

Figure 18: Atomic oxygen concentration as determined by comparing the air afterglow intensity to the NO and NO_2 mass spec signals as a function of admixed nitrogen.

Figure 19a&b: XPS surface analysis vs H_2 processing for (a) C(1s) and (b) O(1s). Panel (a) represents an electron emission angle of 15°, while panel (b) represents an electron emission angle of 90°.

Figure 20a-d: XPS surface analysis vs N_2 processing for (a) C(1s) (b) O(1s) and (c) N(1s). Panel (a) represents an electron emission angle of 15°, while panel (b) represents an electron emission angle of 90°.

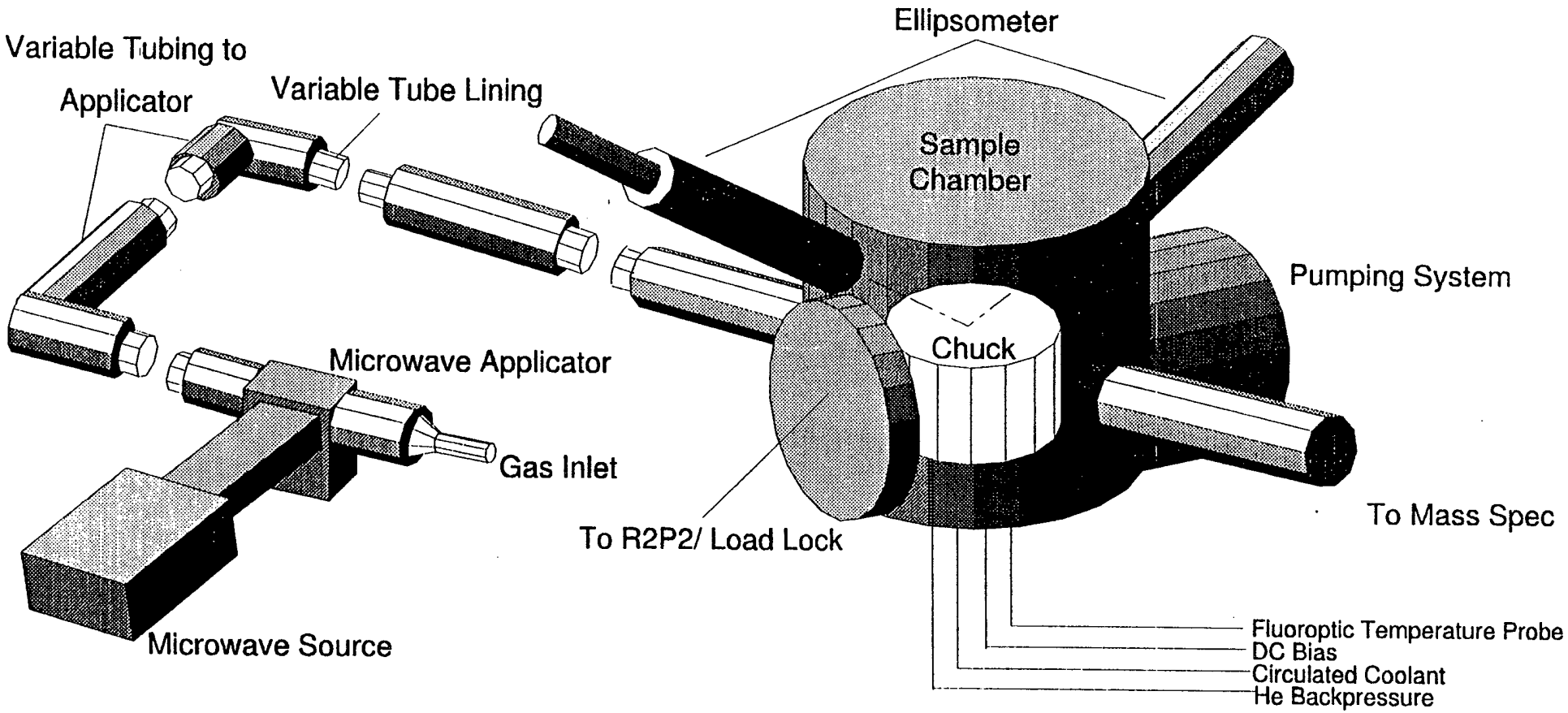

Variable Tubing to
Applicator
Variable Tube Lining
Ellipsometer
Sample
Chamber
Pumping System
Chuck
Microwave Applicator
Gas Inlet
To R2P2/ Load Lock
To Mass Spec
Microwave Source
Fluoroptic Temperature Probe
DC Bias
Circulated Coolant
He Backpressure

Fig. 1

Figure 2

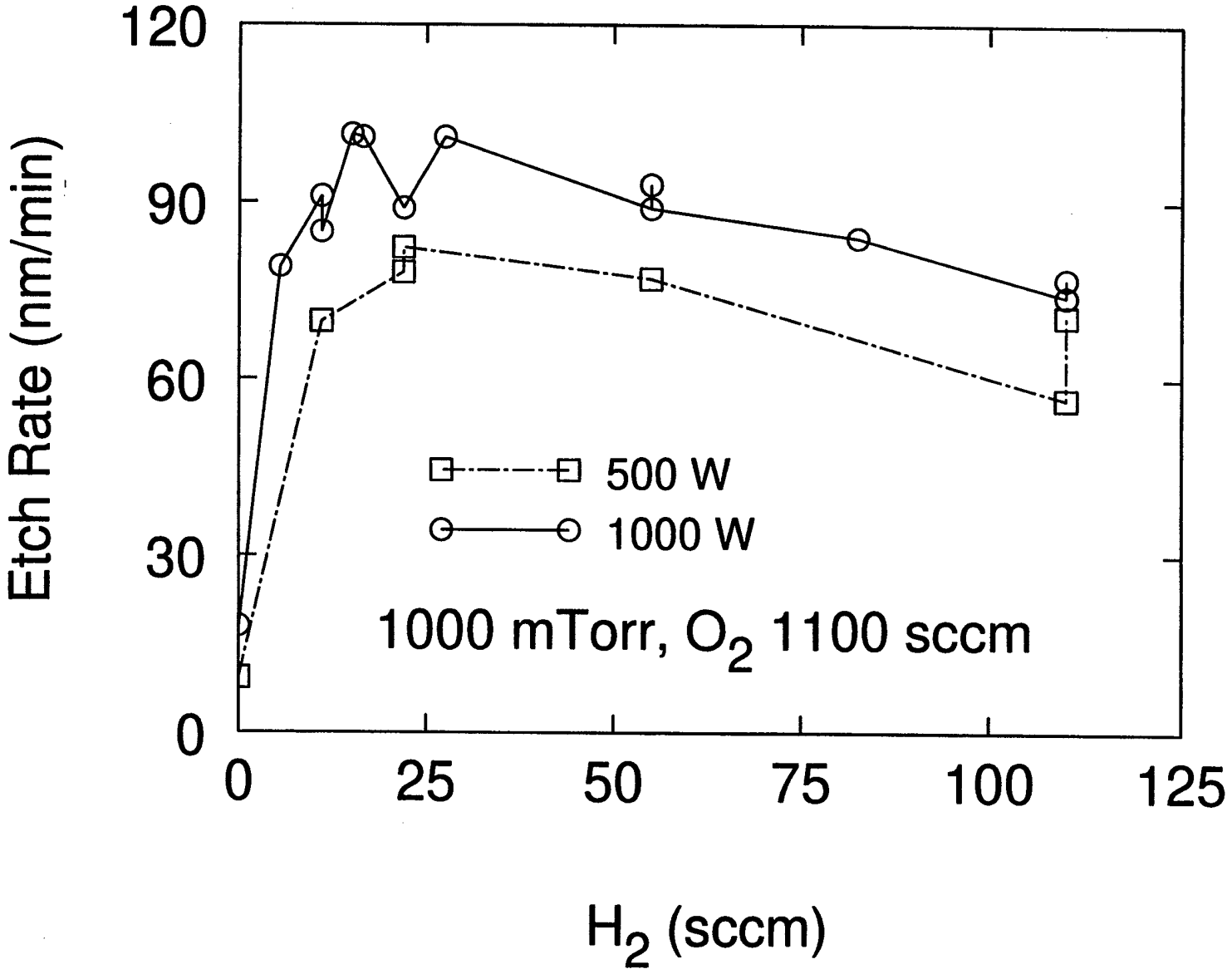

120
90
60
30
0
Etch Rate (nm/min)
0
25
50
75
100
125
H2 (sccm)
500 W
1000 W
1000 mTorr, O2 1100 sccm

Figure 3

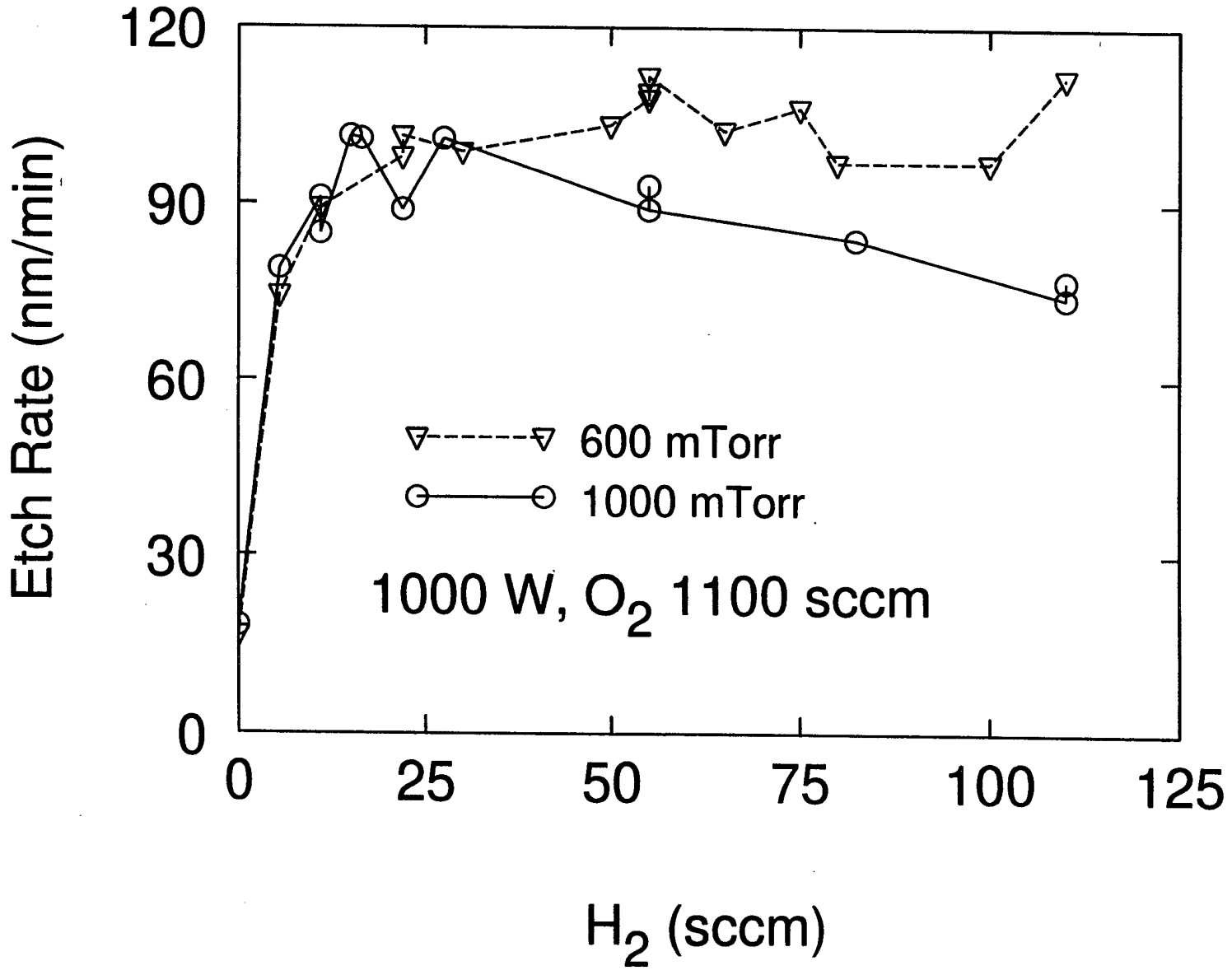

120
90
60
30
0
Etch Rate (nm/min)
0
25
50
75
100
125
H2 (sccm)
600 mTorr
1000 mTorr
1000 W, O2 1100 sccm

Figure 4a

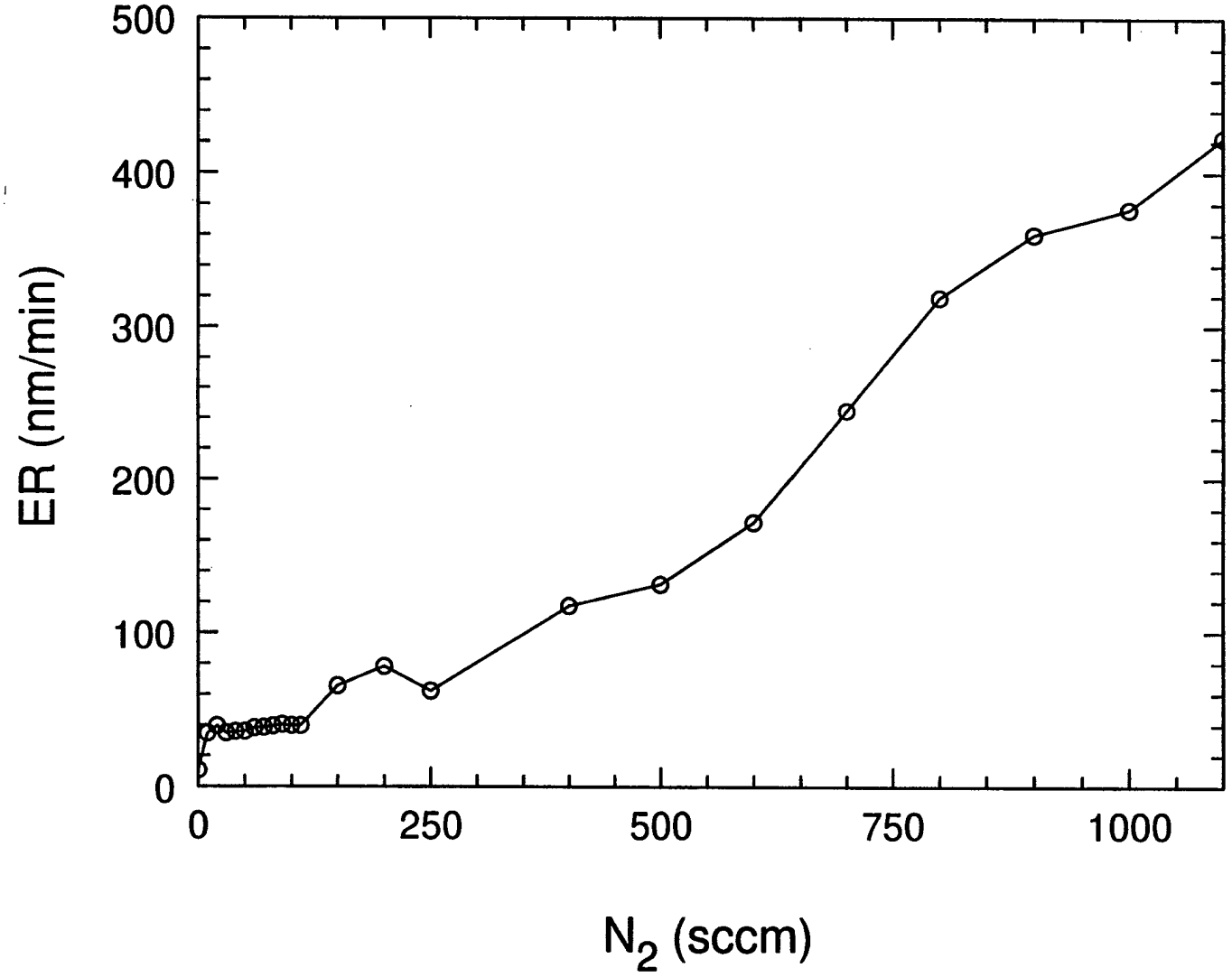


Figure 4b

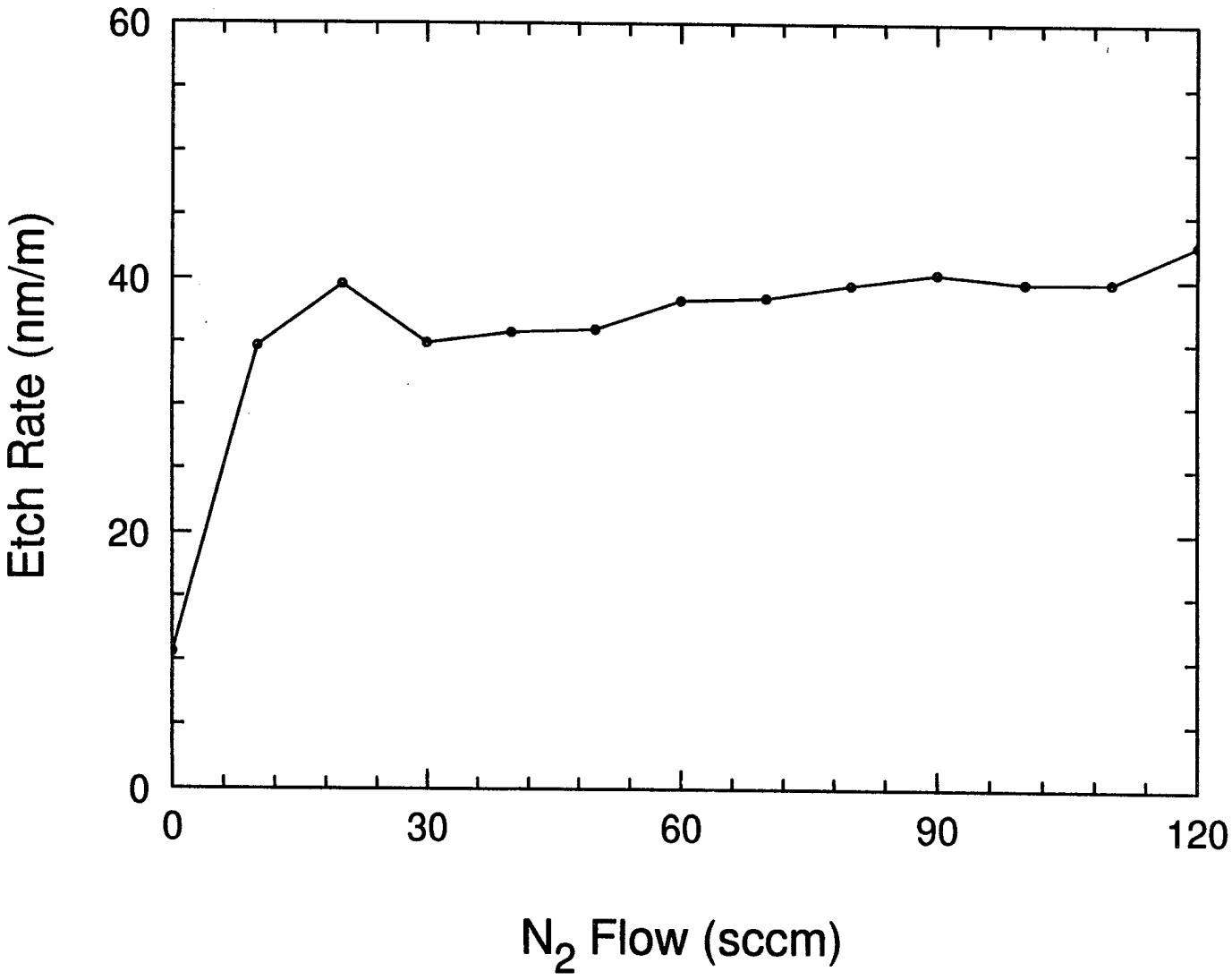


Figure 5

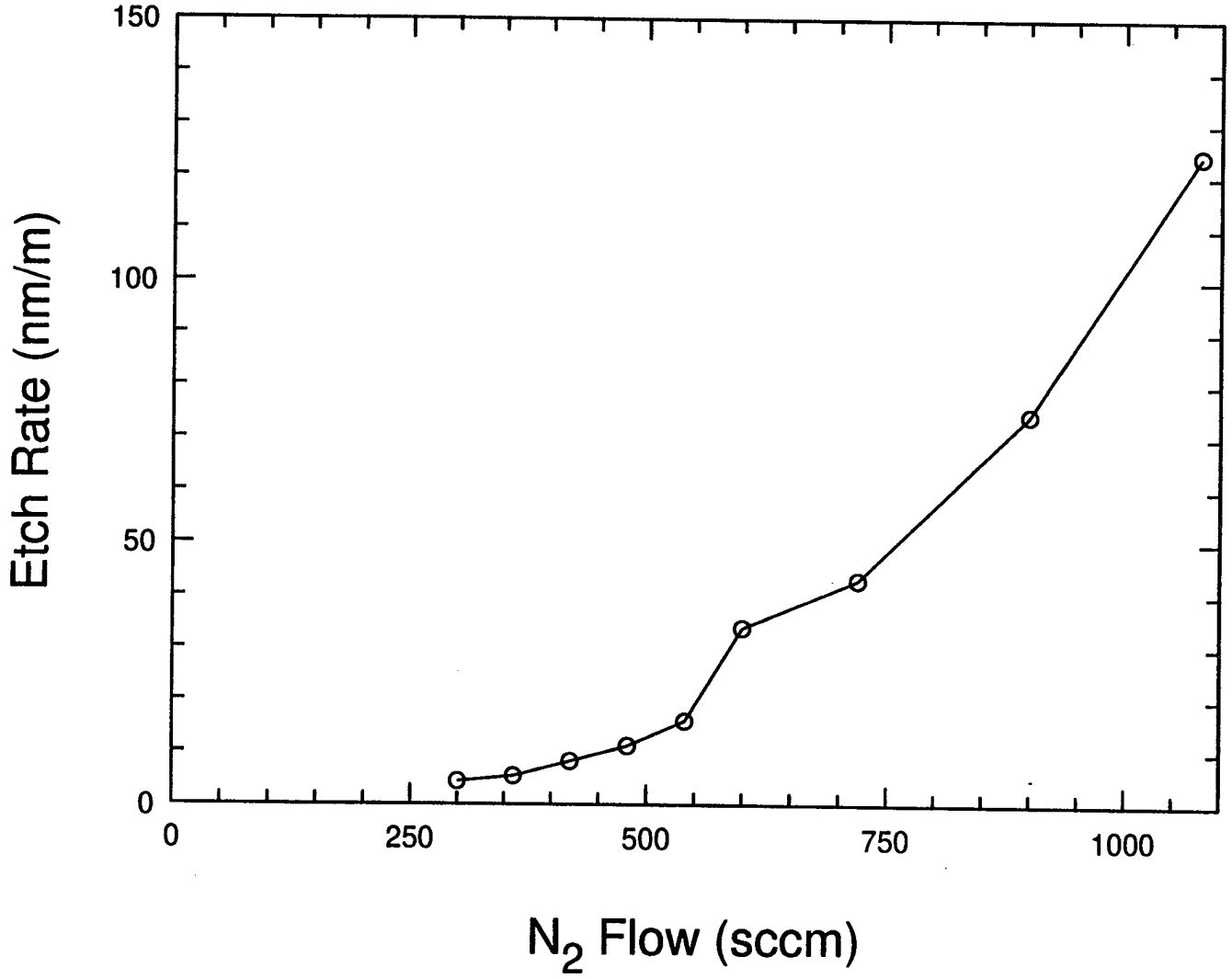


Figure 6a

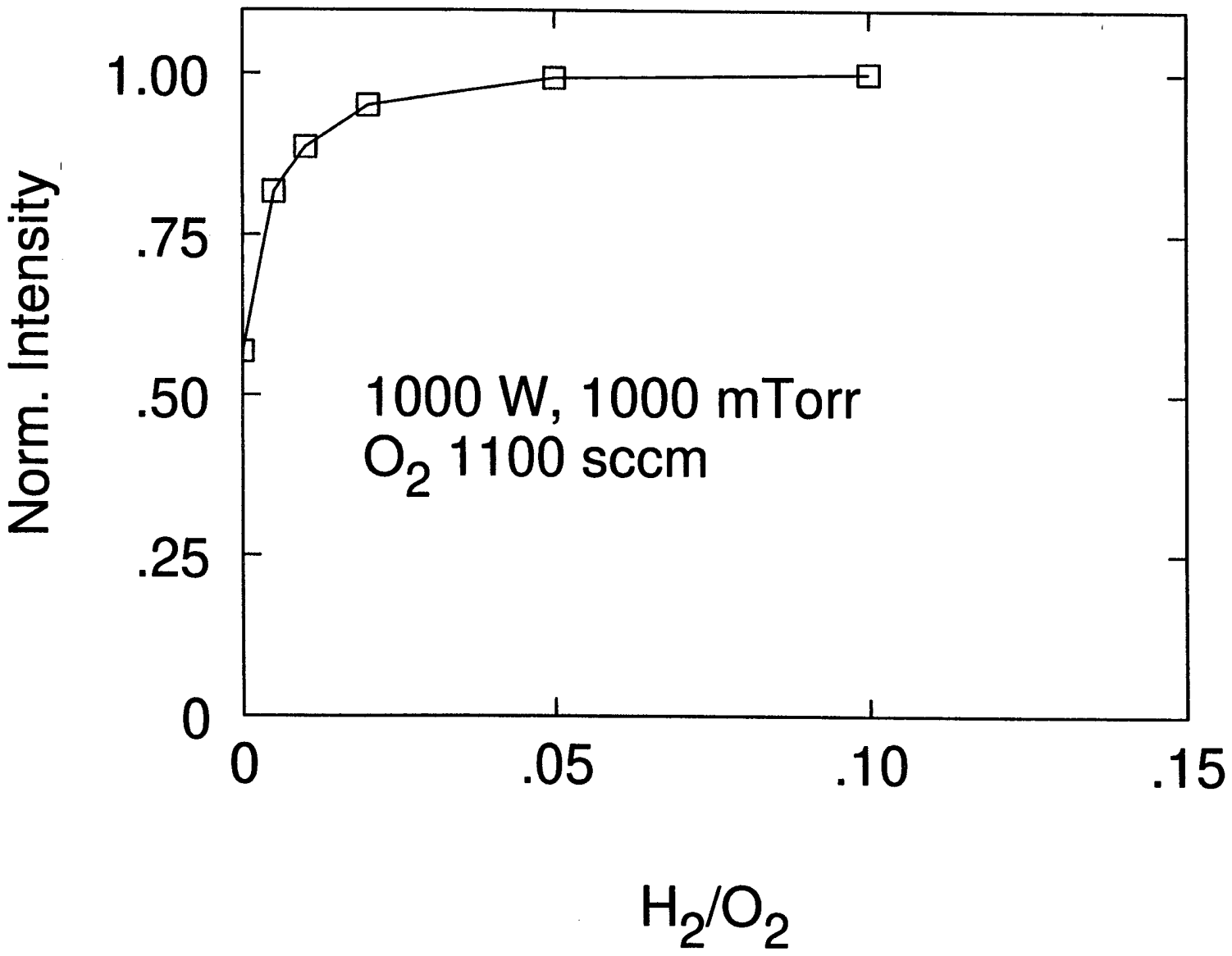

1.00
.75
.50
.25
0
Norm. Intensity
1000 W, 1000 mTorr
O_2 1100 sccm
0
.05
.10
.15
H_2/O_2

Figure 6b

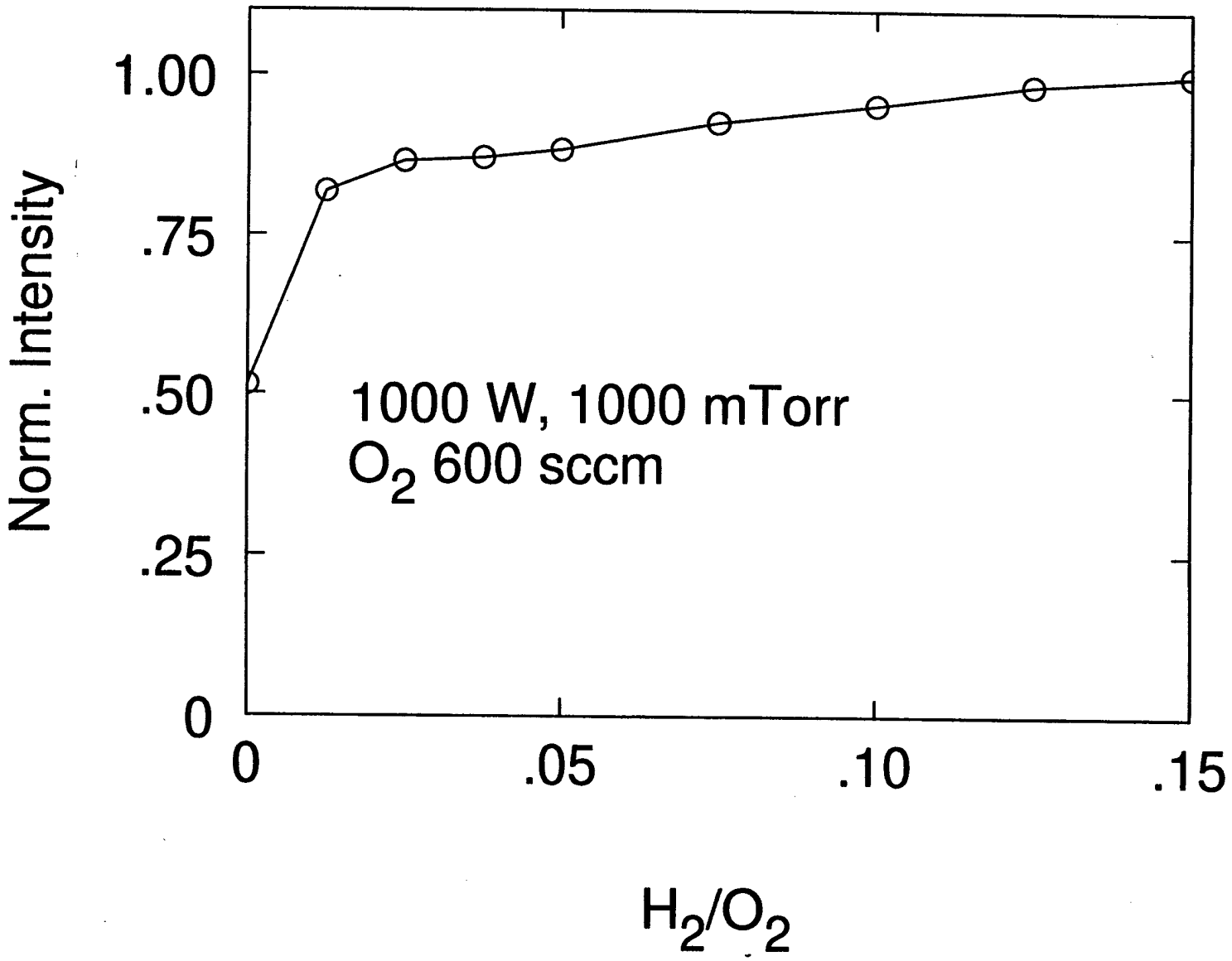

1.00
.75
.50
.25
0
Norm. Intensity
1000 W, 1000 mTorr
O_2 600 sccm
0
.05
.10
.15
H_2/O_2

Figure 7

O (844nm)/Ar (811nm)

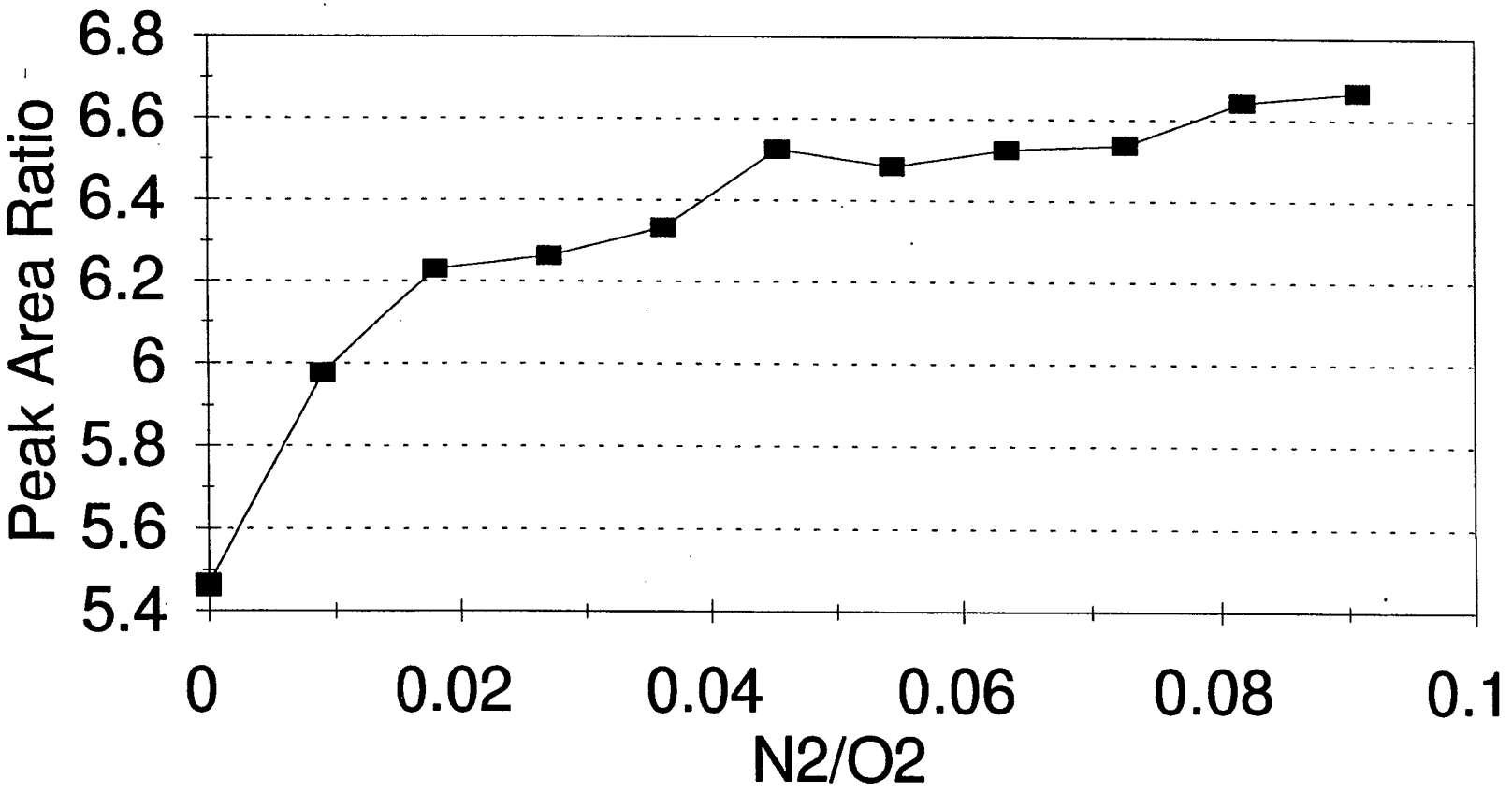


Figure 8

O (844nm)/Ar (811nm)

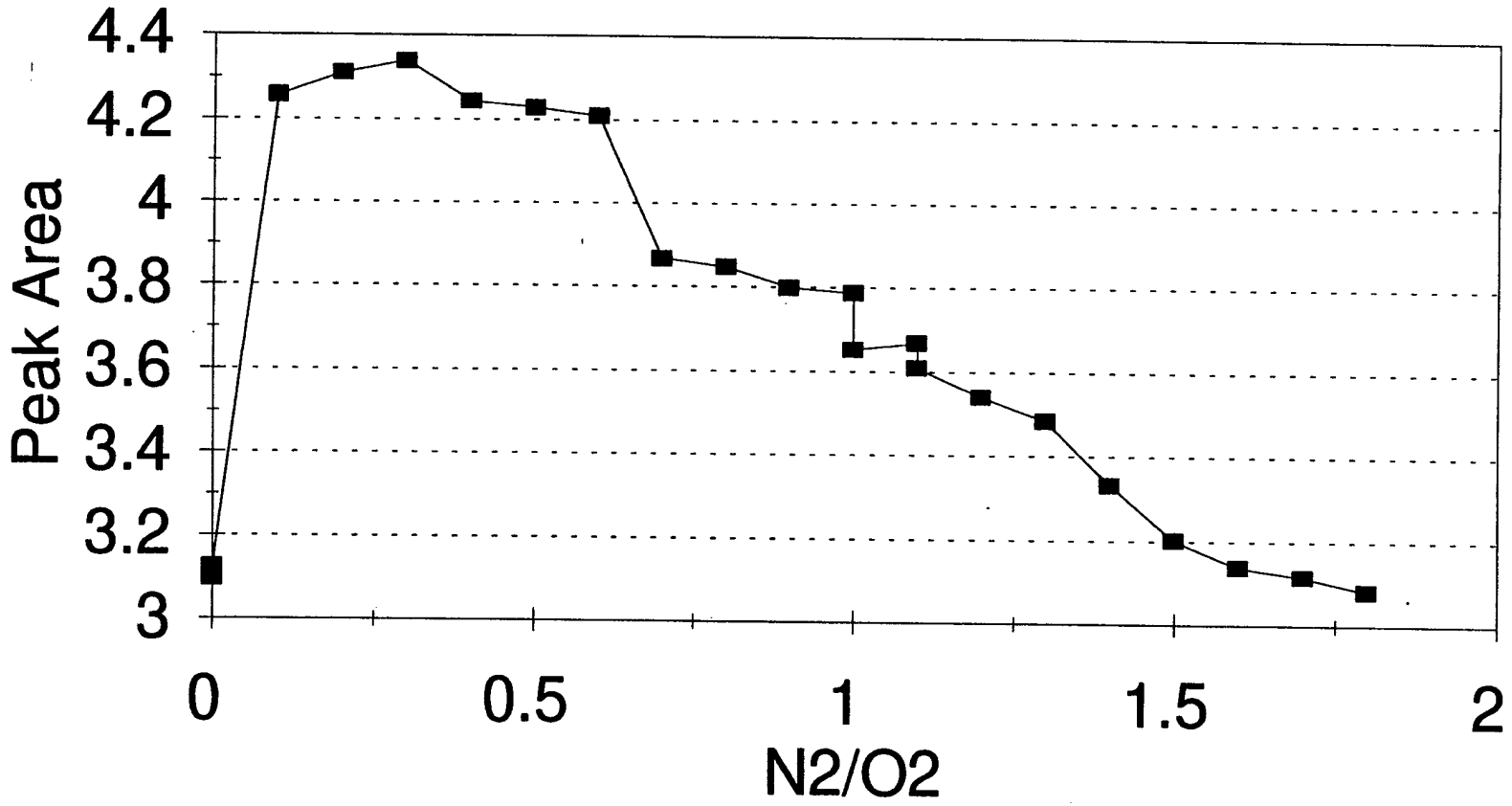


Figure 9

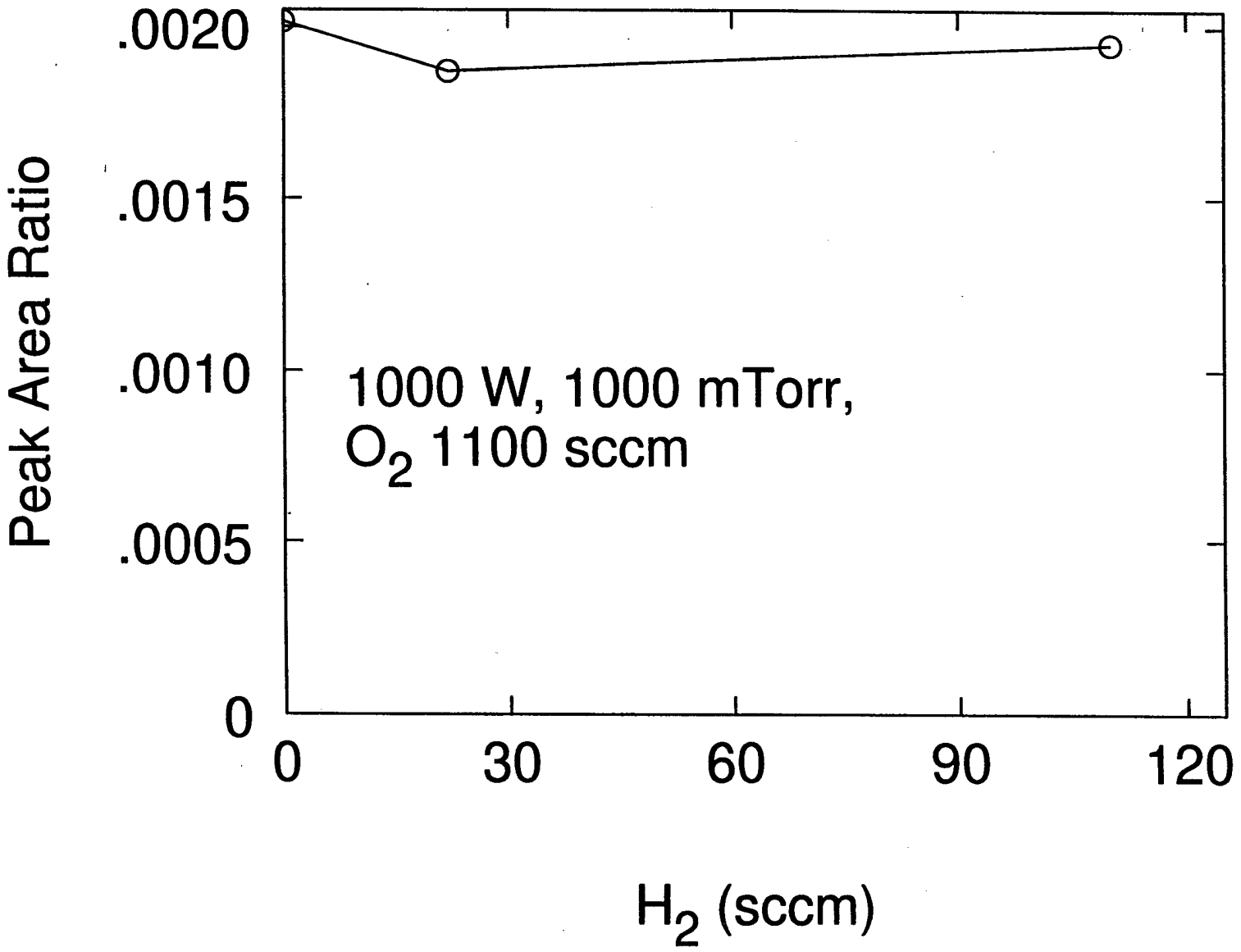

Peak Area Ratio
.0020
.0015
.0010
.0005
0
1000 W, 1000 mTorr,
O2 1100 sccm
0
30
60
90
120
H2 (sccm)

Figure 10

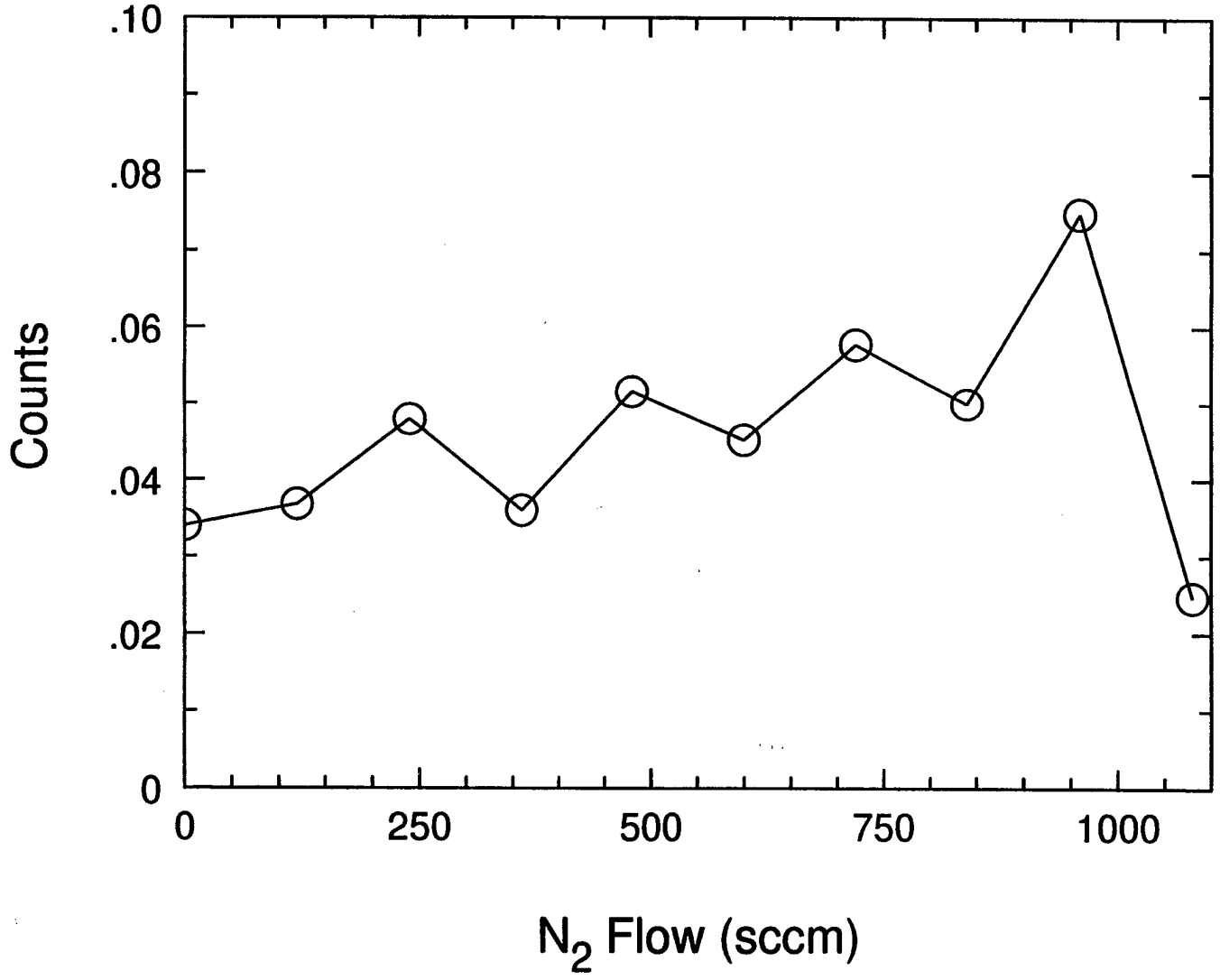


Figure 11

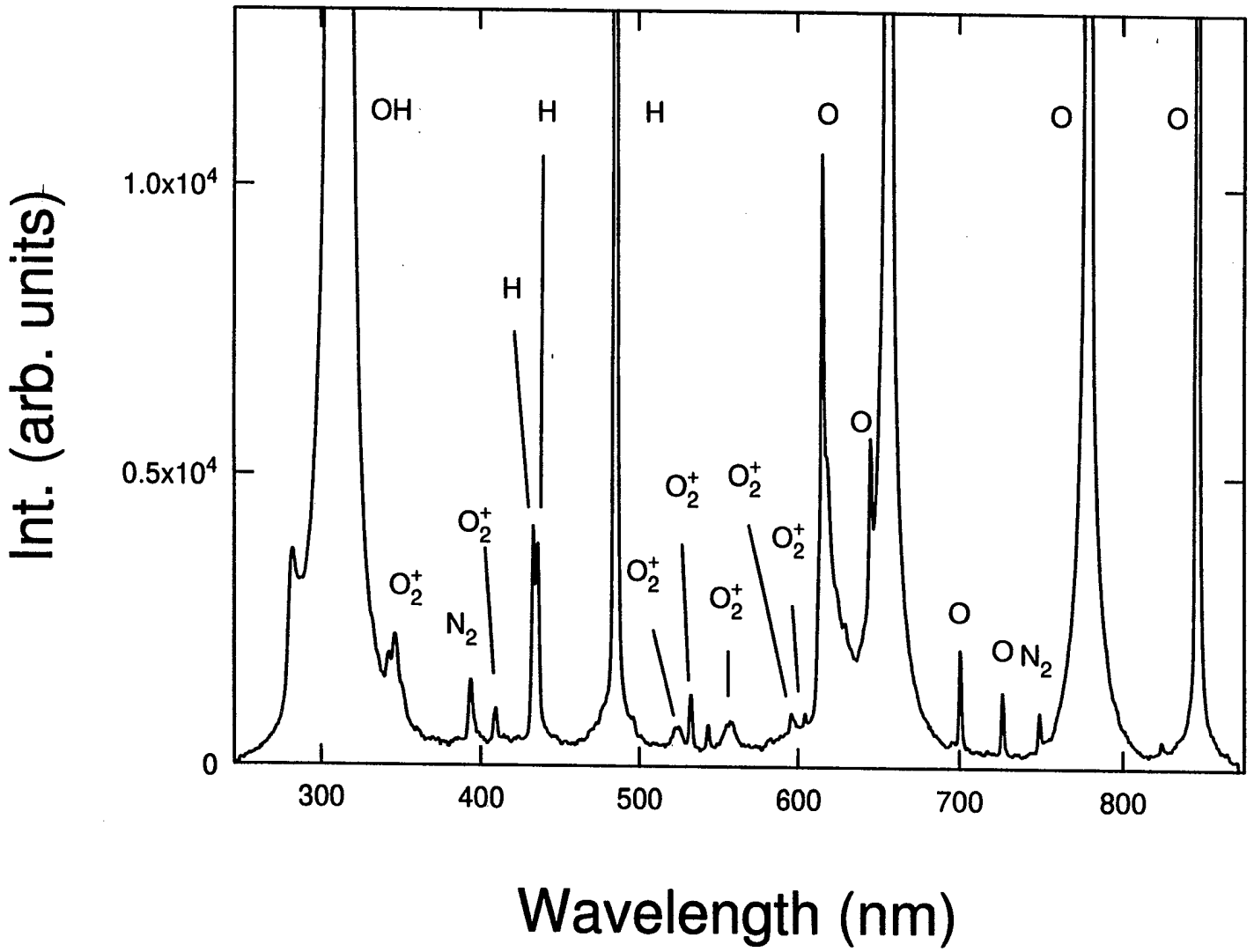

Int. (arb. units)
1.0x10^4
0.5x10^4
0
300
400
500
600
700
800
Wavelength (nm)
OH
H
H
H
O
O
O
O
O
O
O
O
O
O
O
O
N2
N2

Fig 12

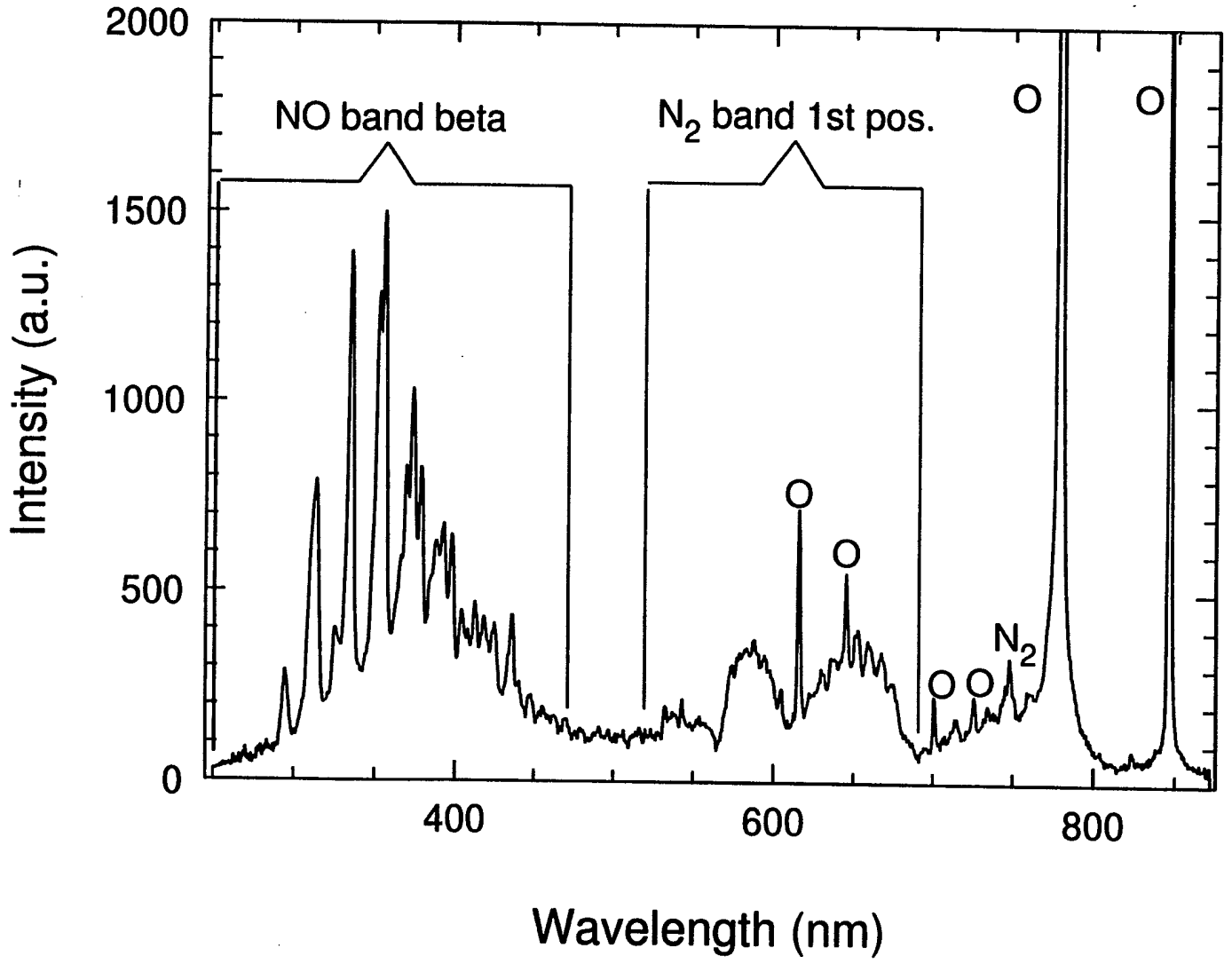

2000
1500
1000
500
0
Intensity (a.u.)
NO band beta
N2 band 1st pos.
O
O
O
O
O
O
N2
400
600
800
Wavelength (nm)

Figure 13

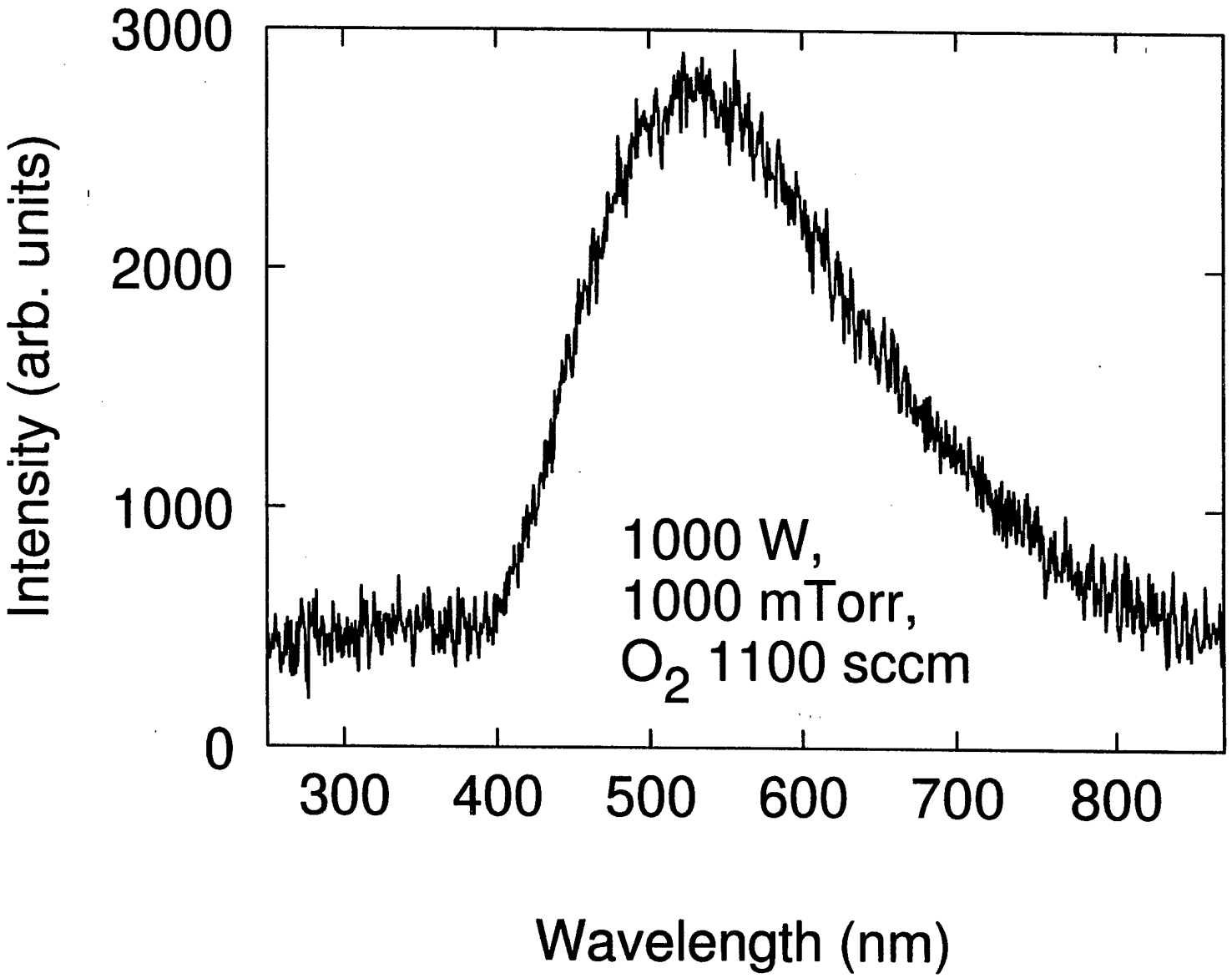

3000
2000
1000
0
Intensity (arb. units)
1000 W,
1000 mTorr,
O_2 1100 sccm
300
400
500
600
700
800
Wavelength (nm)

Figure 14

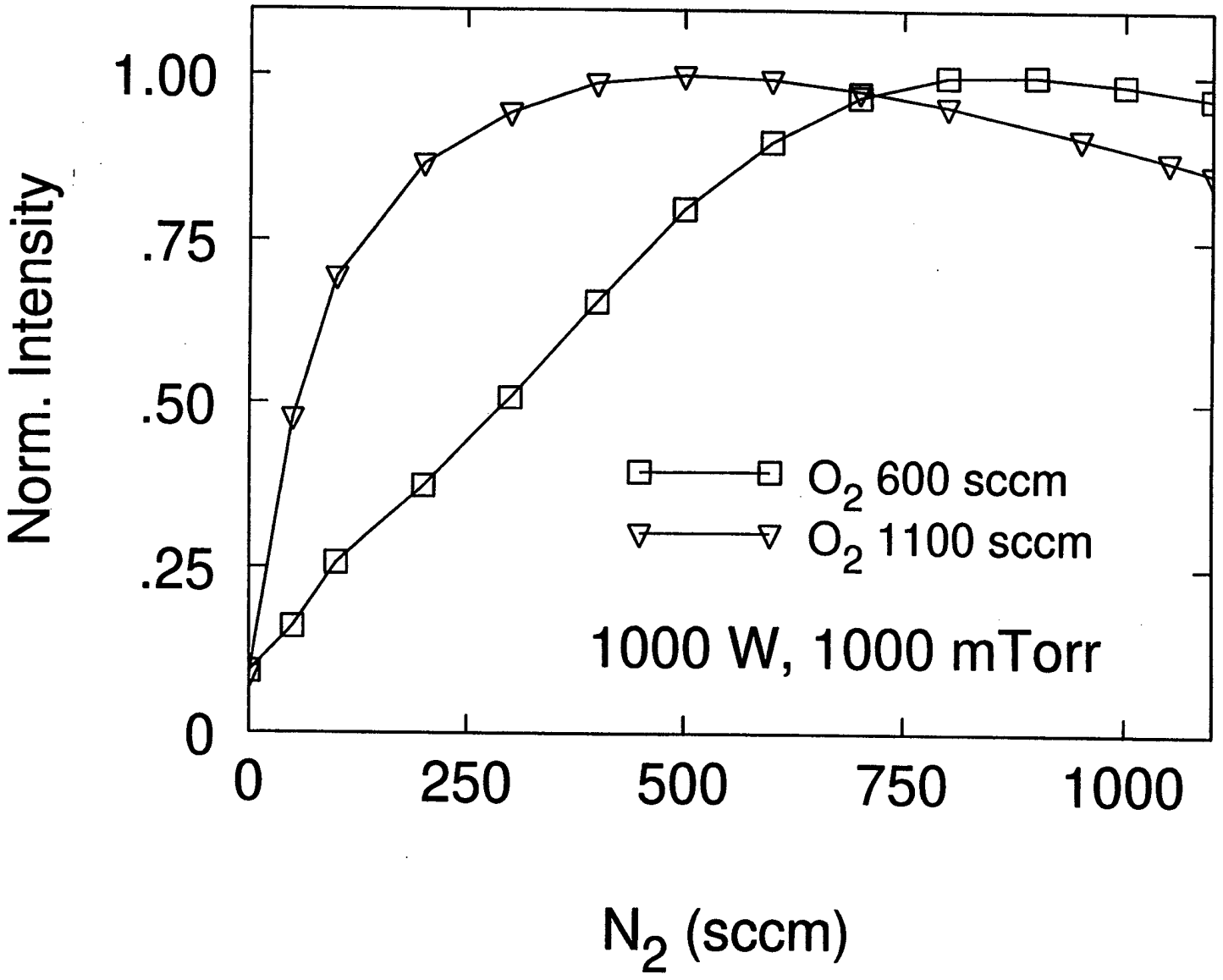

1.00
.75
.50
.25
0
Norm. Intensity
0
250
500
750
1000
N_2 (sccm)
O_2 600 sccm
O_2 1100 sccm
1000 W, 1000 mTorr

Figure 15a

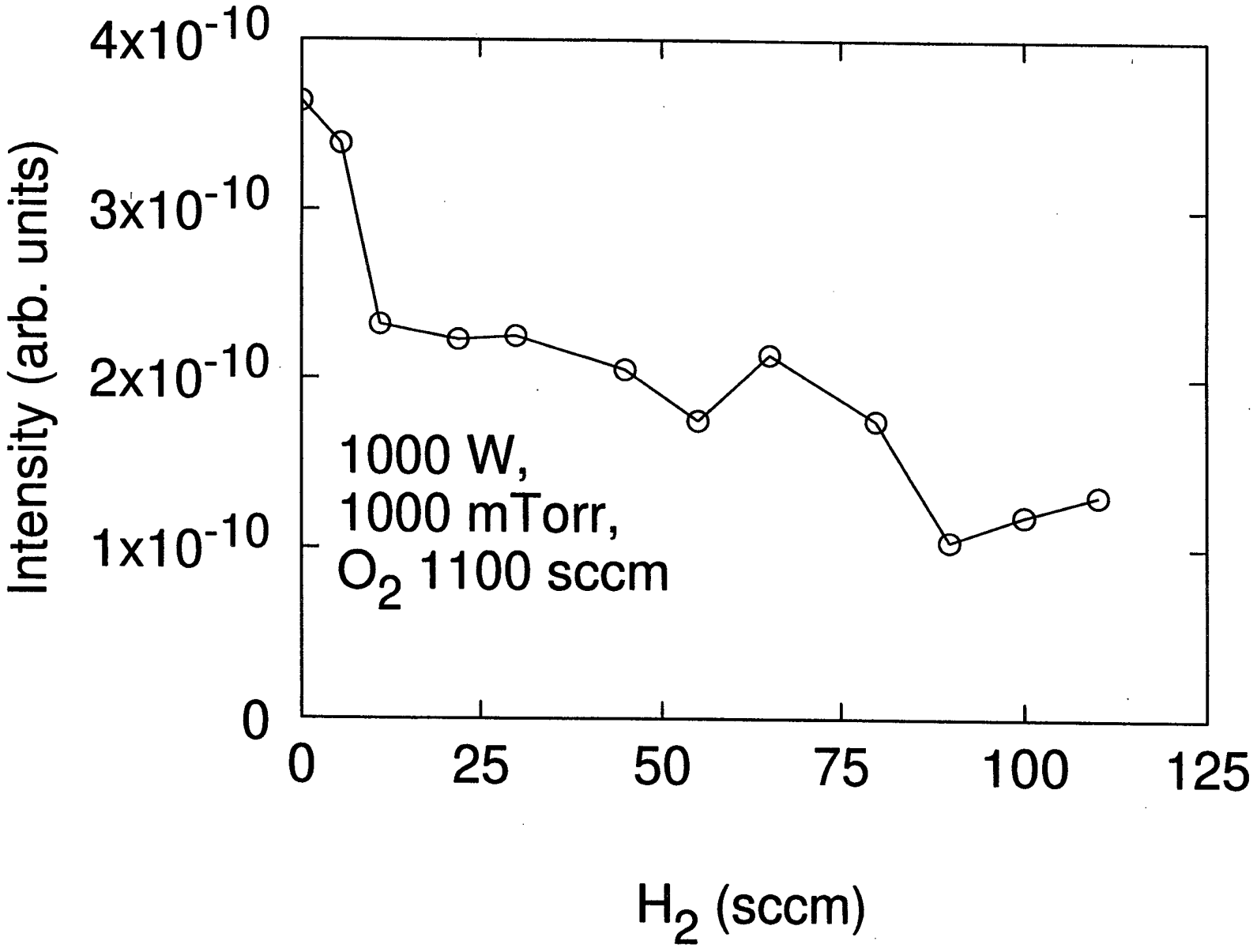

4x10-10
3x10-10
2x10-10
1x10-10
0
Intensity (arb. units)
1000 W,
1000 mTorr,
O2 1100 sccm
0
25
50
75
100
125
H2 (sccm)

Figure 15b

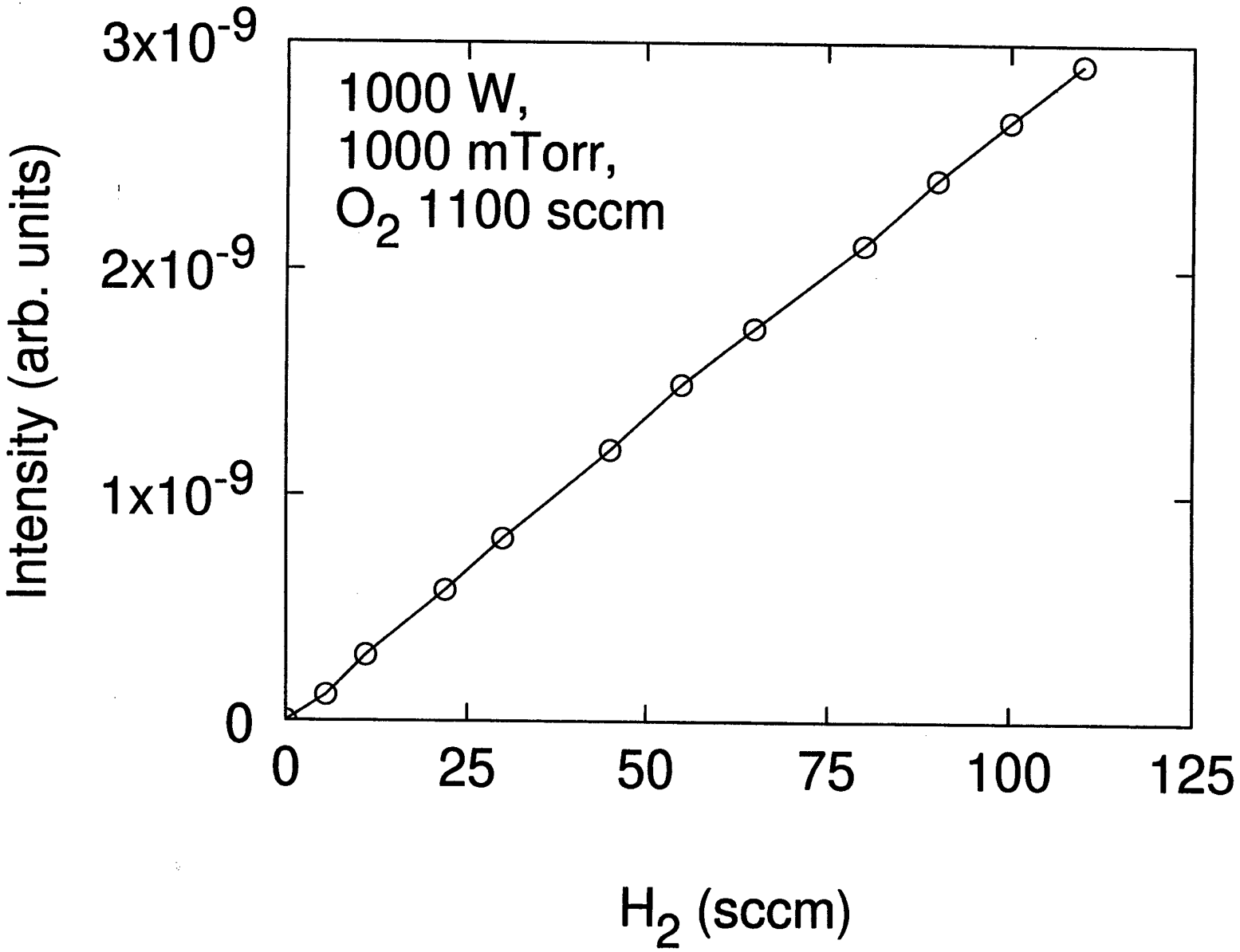

3x10-9
2x10-9
1x10-9
0
Intensity (arb. units)
1000 W,
1000 mTorr,
O2 1100 sccm
0
25
50
75
100
125
H2 (sccm)

Figure 16a

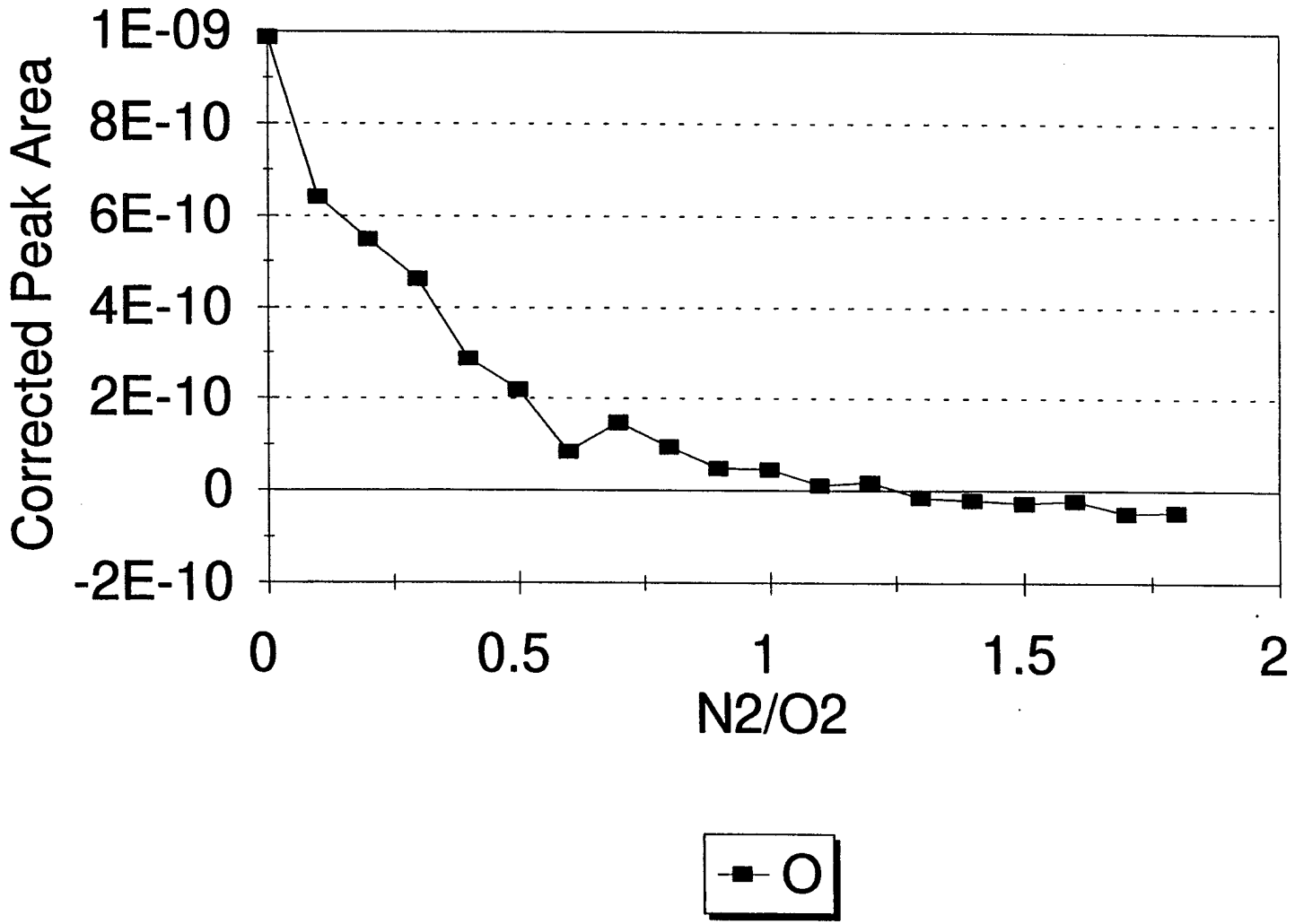

Corrected Peak Area
1E-09
8E-10
6E-10
4E-10
2E-10
0
-2E-10
0
0.5
1
1.5
2
N2/O2
O

Figure 16b

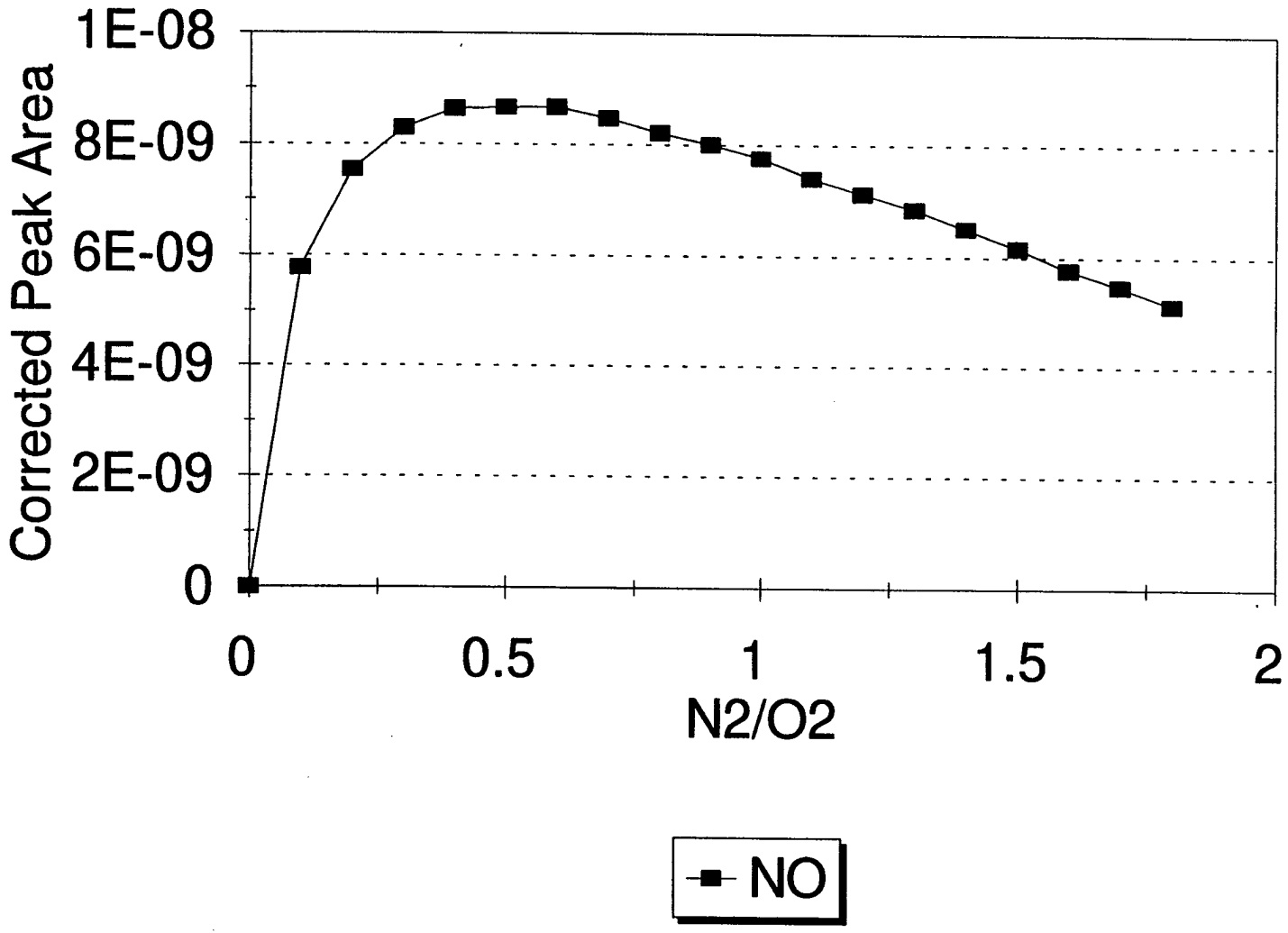

Corrected Peak Area
1E-08
8E-09
6E-09
4E-09
2E-09
0
0
0.5
1
1.5
2
N2/O2
NO

Figure 16c

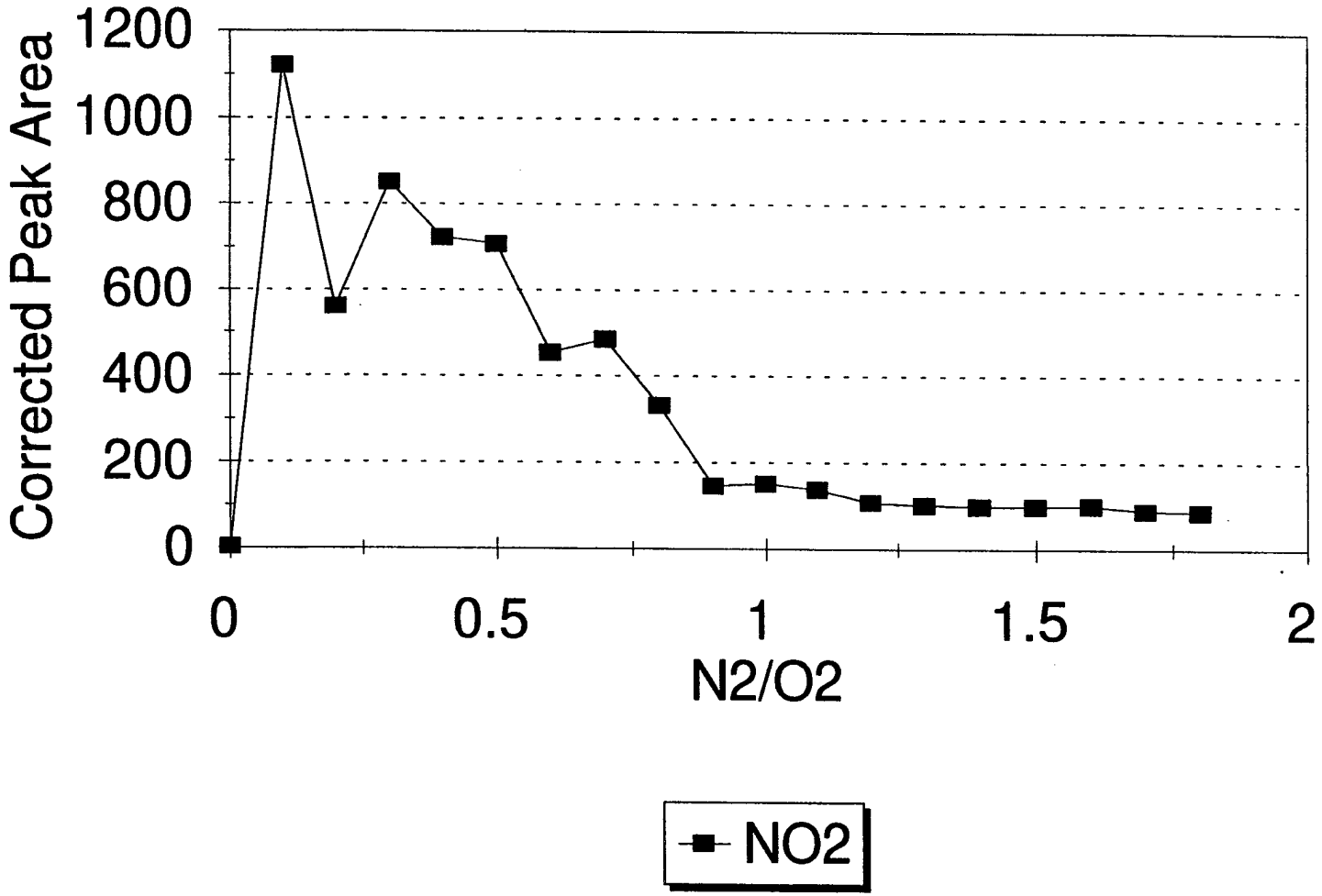

Corrected Peak Area
1200
1000
800
600
400
200
0
0
0.5
1
1.5
2
N2/O2
NO2

Figure 16d

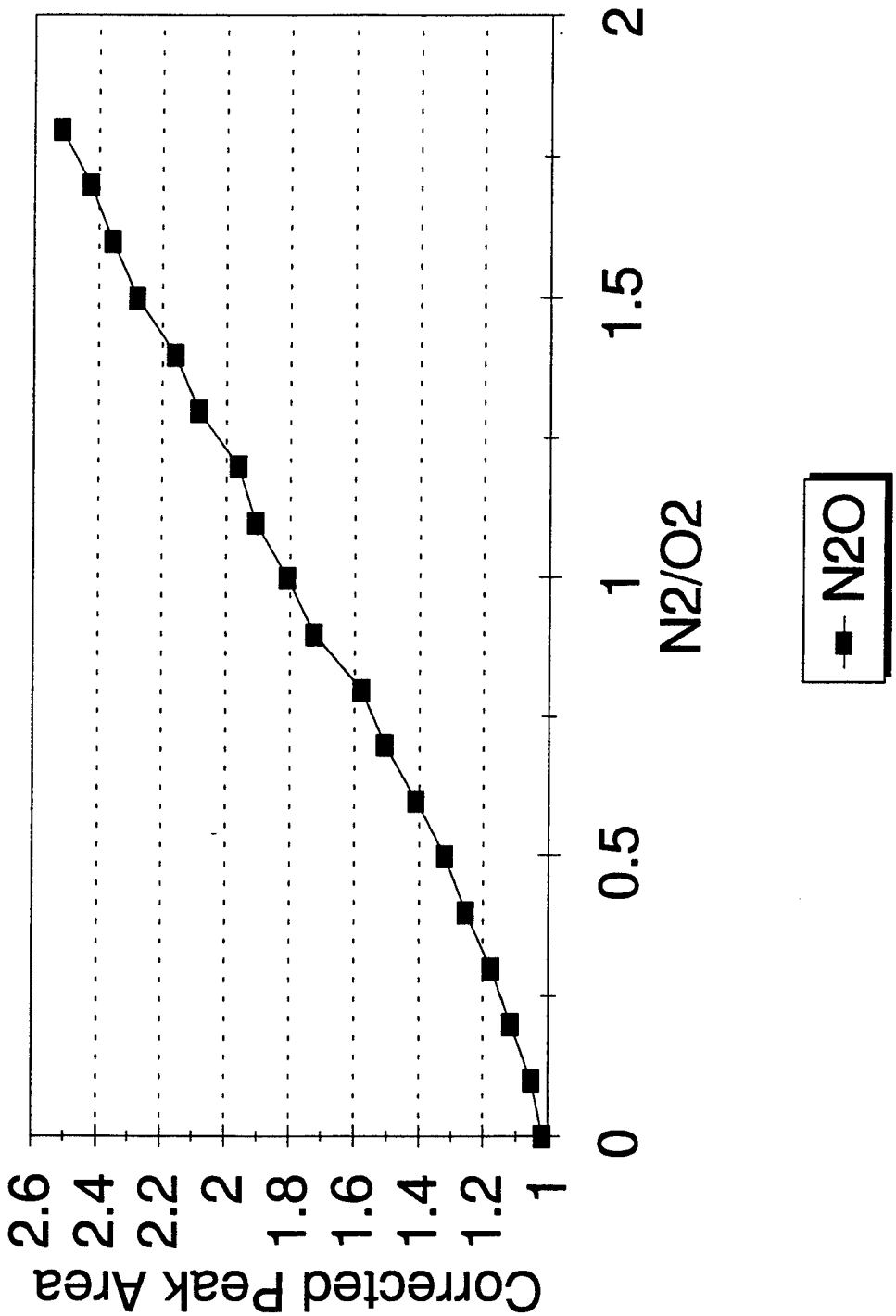


Figure 17a

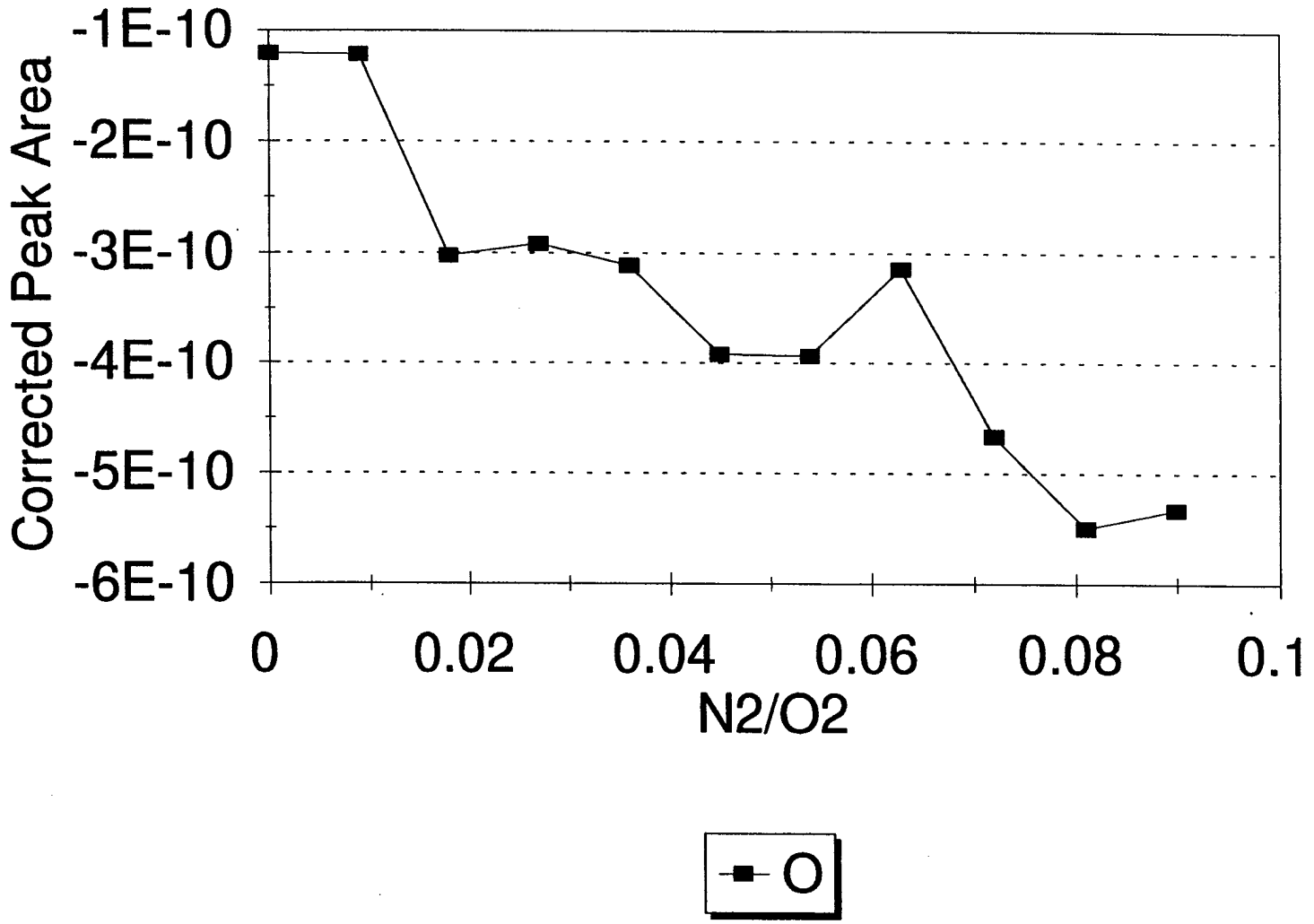

Corrected Peak Area
-1E-10
-2E-10
-3E-10
-4E-10
-5E-10
-6E-10
0
0.02
0.04
0.06
0.08
0.1
N2/O2
O

Figure 17b

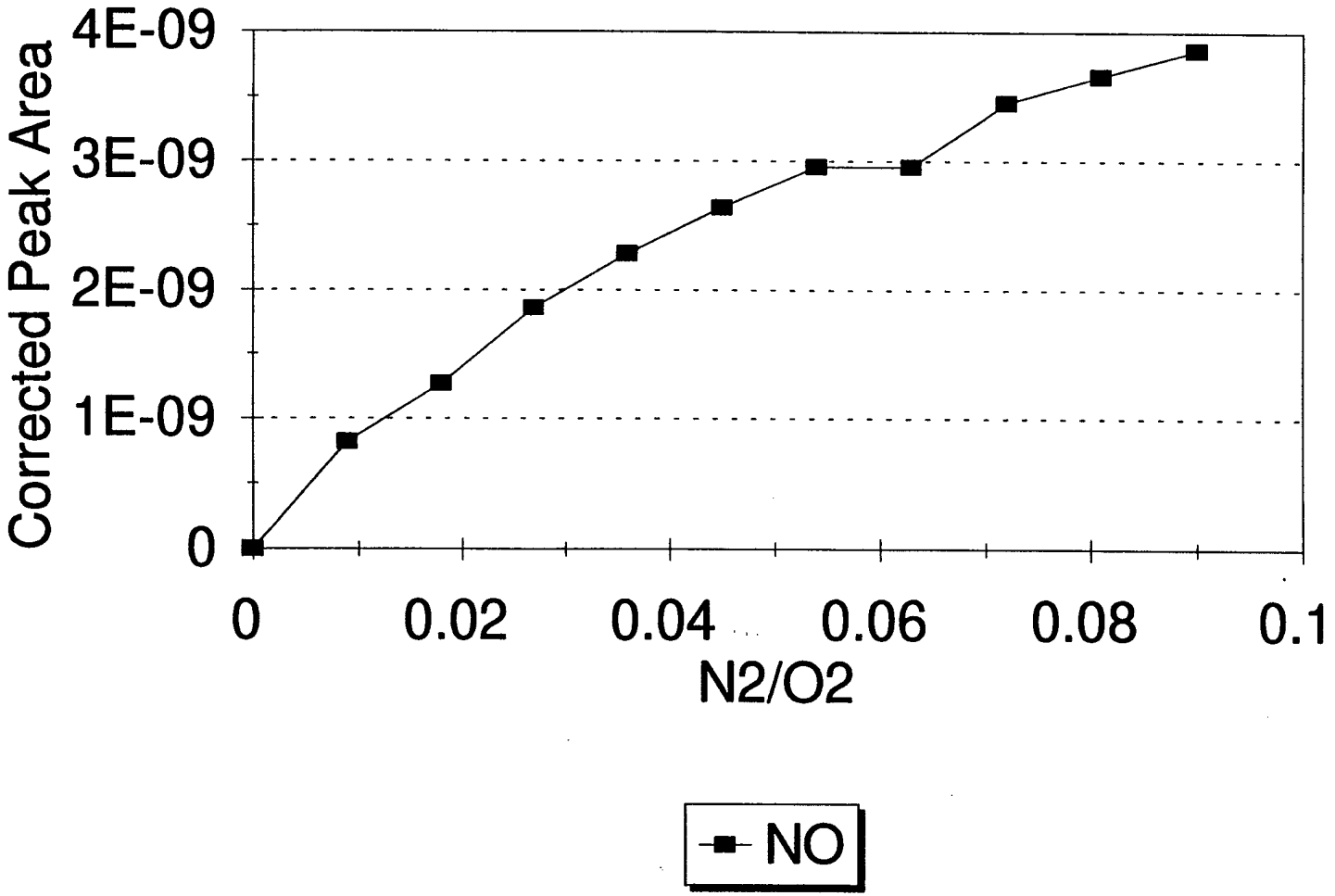

Corrected Peak Area
4E-09
3E-09
2E-09
1E-09
0
0
0.02
0.04
0.06
0.08
0.1
N2/O2
NO

Figure 17c

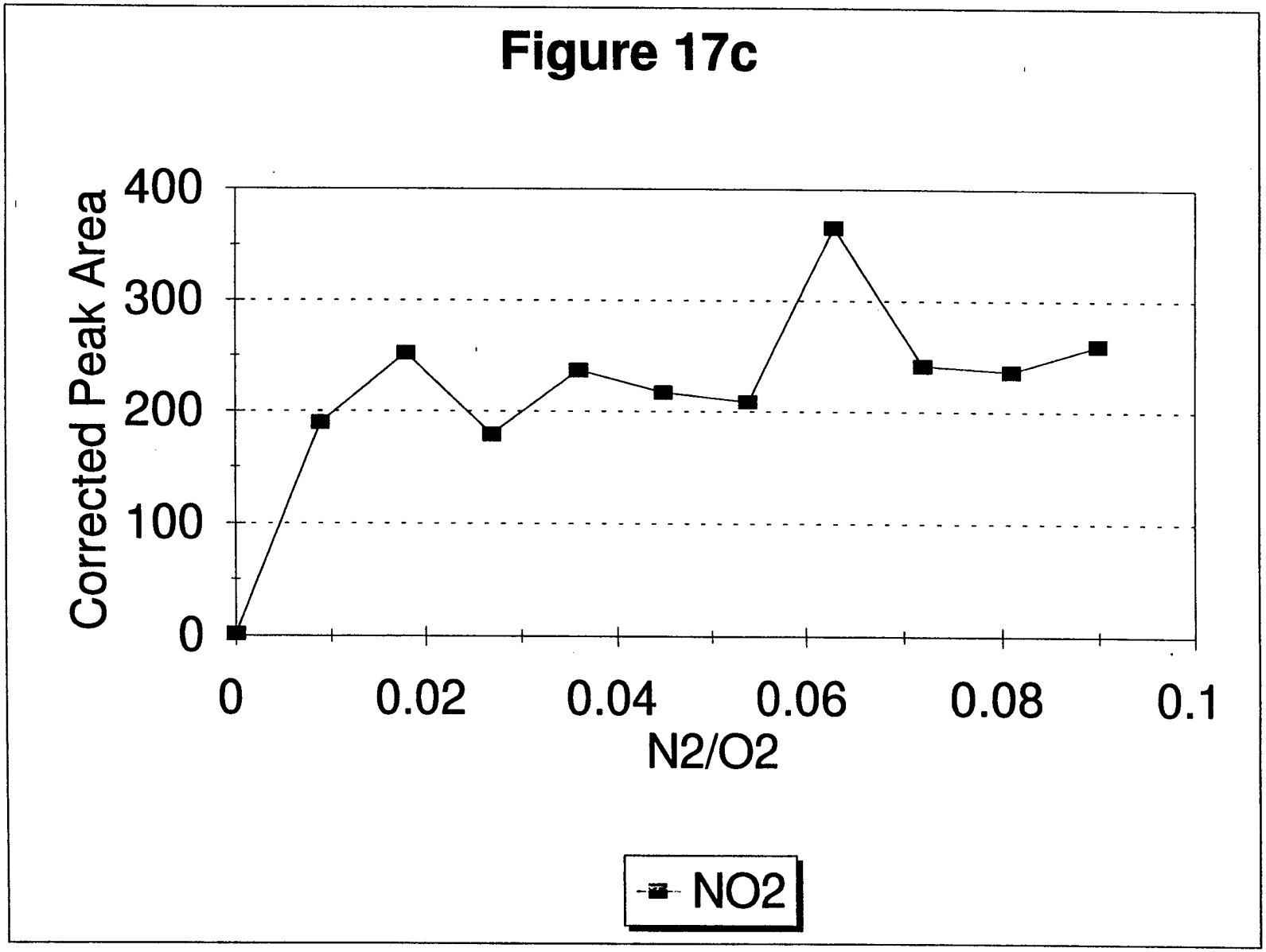

Corrected Peak Area
400
300
200
100
0
0
0.02
0.04
0.06
0.08
0.1
N2/O2
NO2

Figure 17d

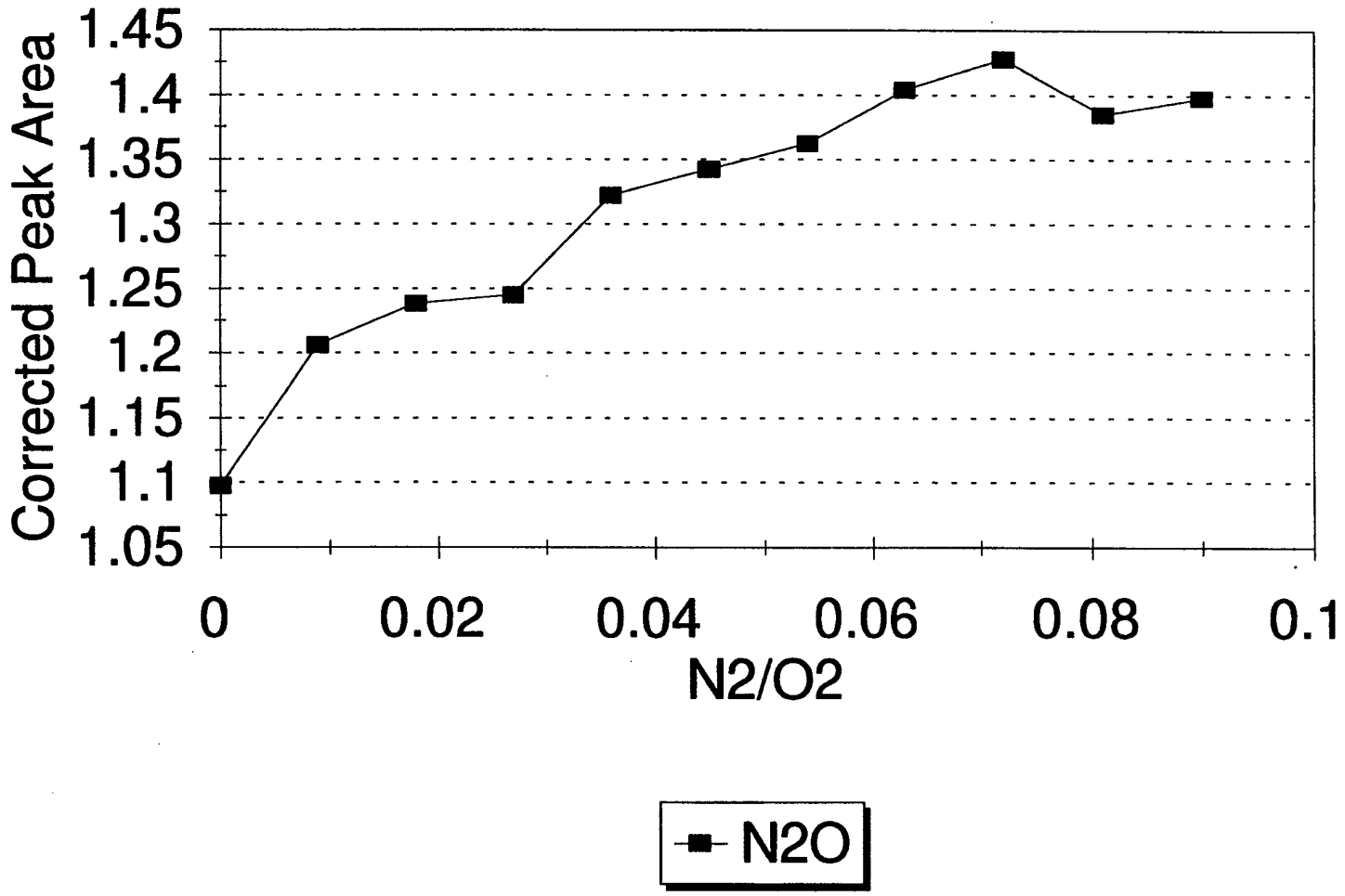

Corrected Peak Area
1.45
1.4
1.35
1.3
1.25
1.2
1.15
1.1
1.05
0
0.02
0.04
0.06
0.08
0.1
N2/O2
N2O

Figure 18

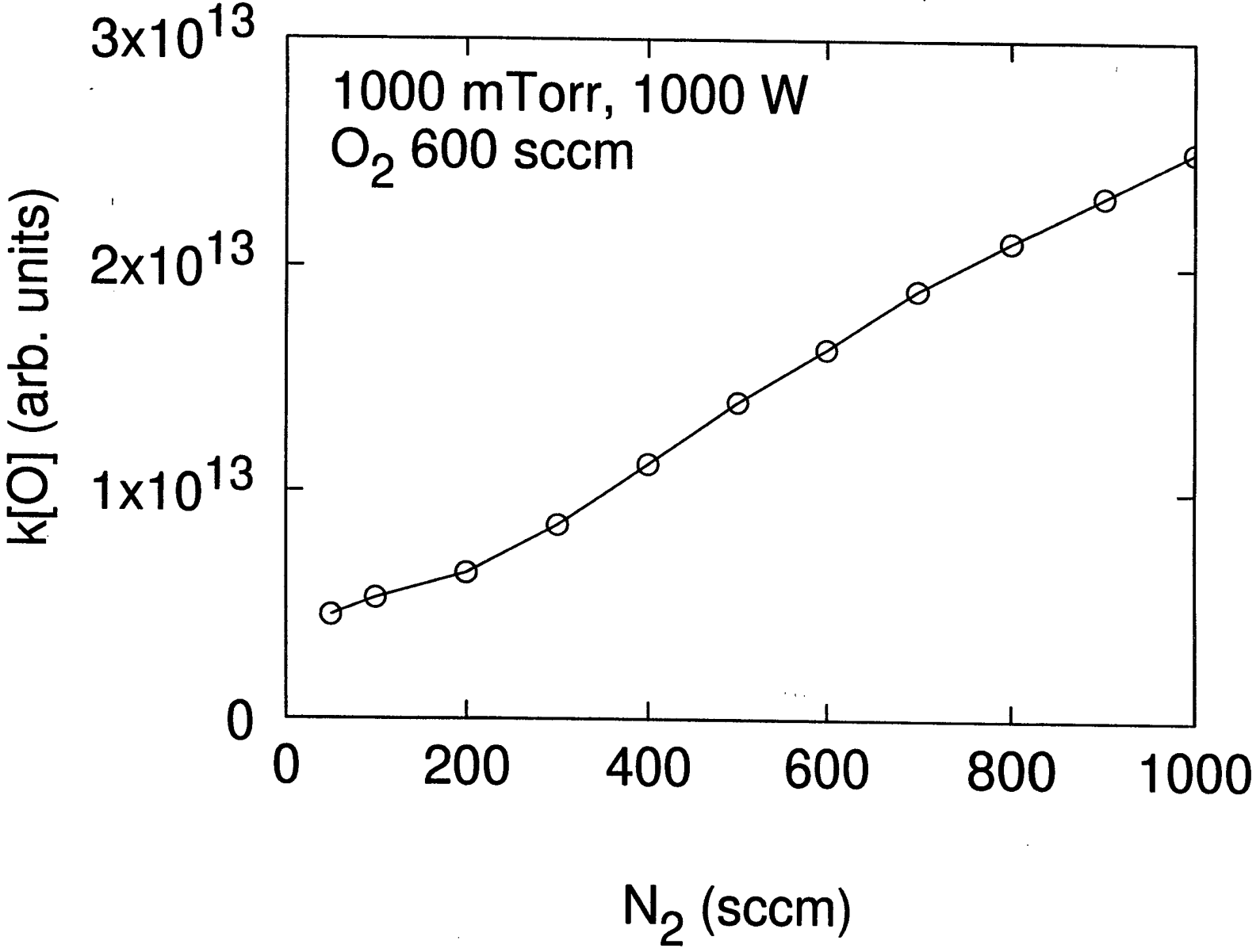

3x10^13
2x10^13
1x10^13
0
k[O] (arb. units)
1000 mTorr, 1000 W
O2 600 sccm
0
200
400
600
800
1000
N2 (sccm)

Figure 19a

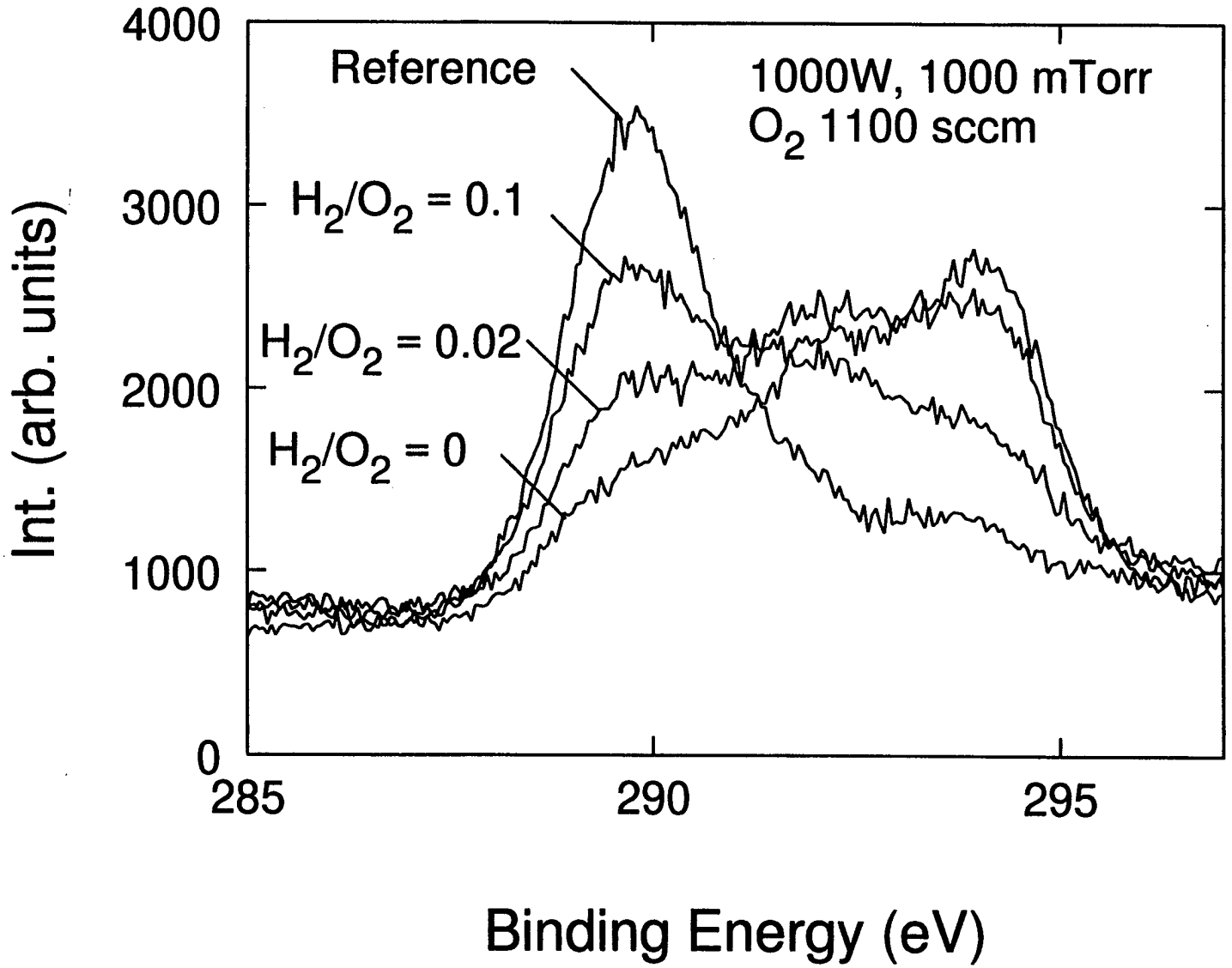

4000
3000
2000
1000
0
Int. (arb. units)
Reference
1000W, 1000 mTorr
O_2 1100 sccm
H_2/O_2 = 0.1
H_2/O_2 = 0.02
H_2/O_2 = 0
285
290
295
Binding Energy (eV)

Figure 19b

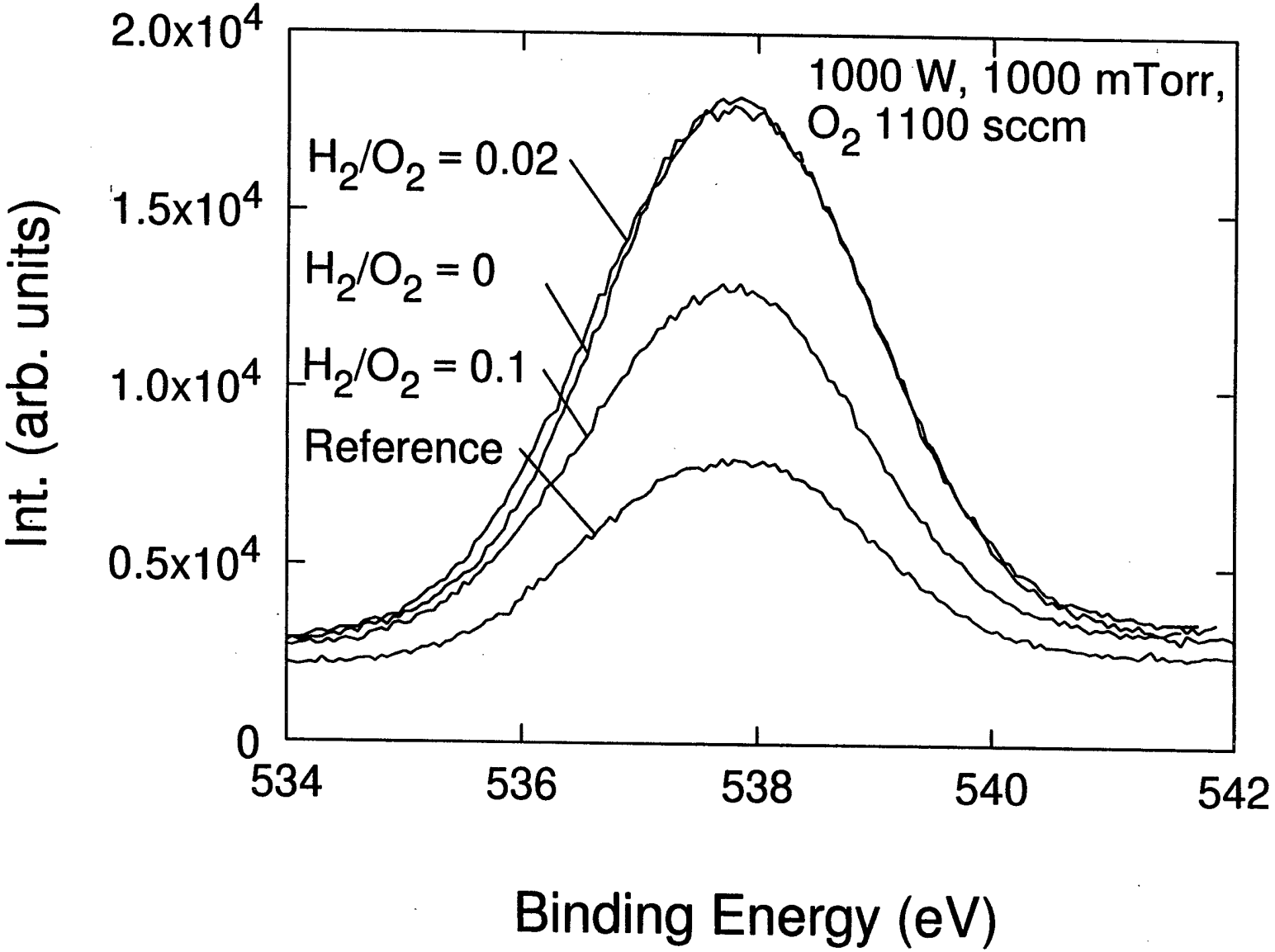
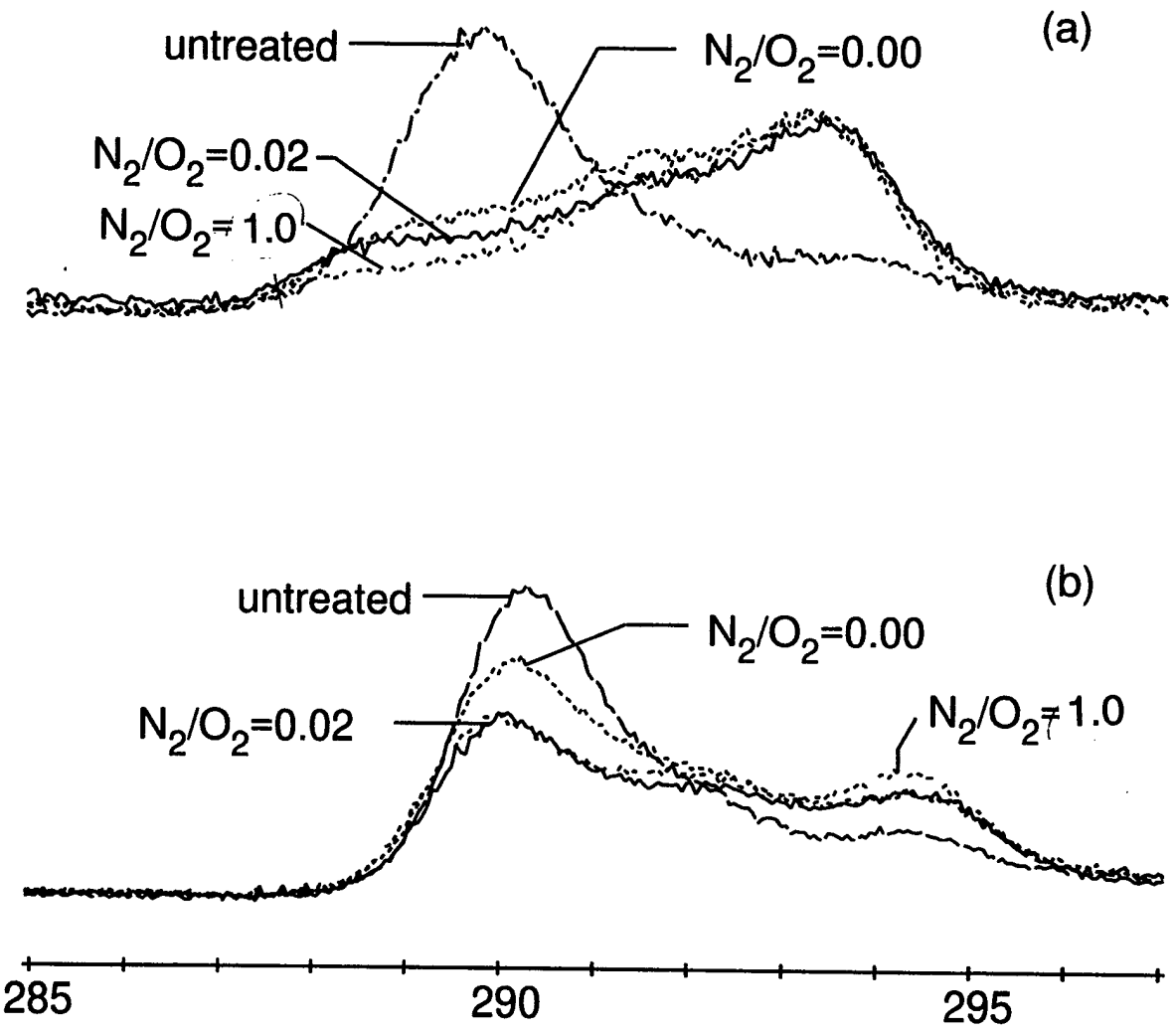

2.0x10^4
1.5x10^4
1.0x10^4
0.5x10^4
0
Int. (arb. units)
534
536
538
540
542
Binding Energy (eV)
1000 W, 1000 mTorr,
O2 1100 sccm
H2/O2 = 0.02
H2/O2 = 0
H2/O2 = 0.1
Reference

Figure 20a



Binding Energy (eV)

Figure 20b

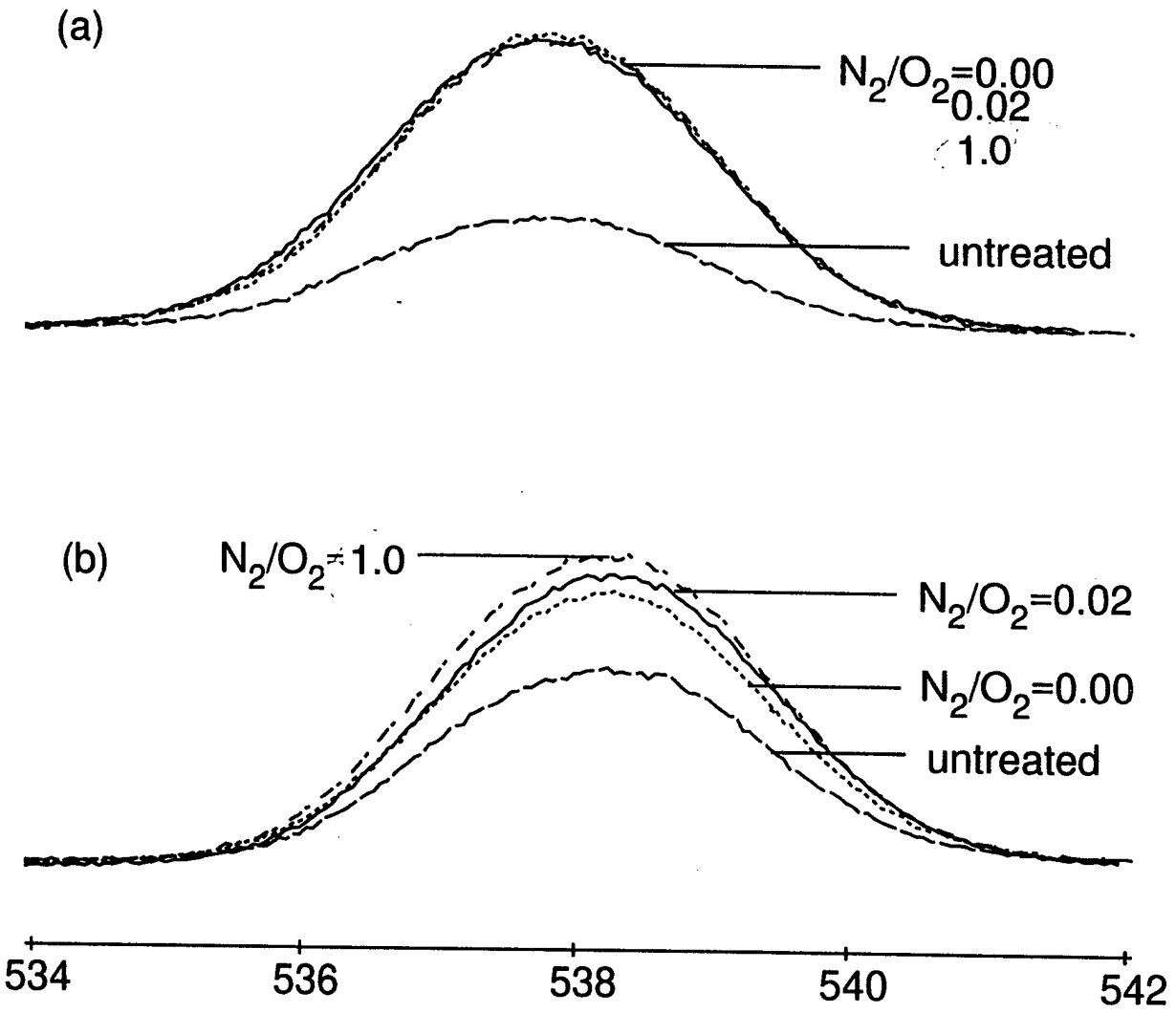
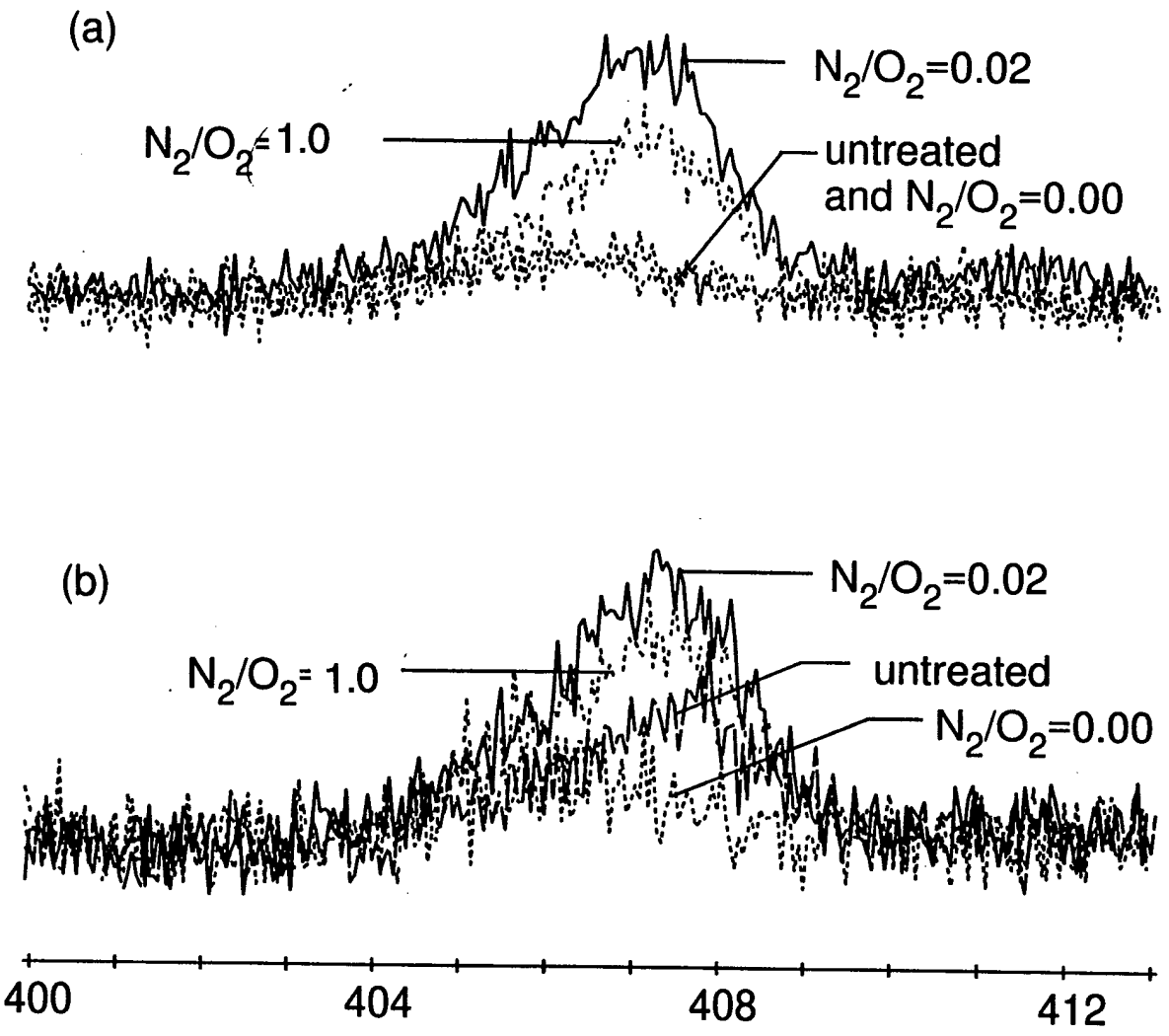

(a)
N_2/O_2=0.00
0.02
1.0
untreated
(b)
N_2/O_2=1.0
N_2/O_2=0.02
N_2/O_2=0.00
untreated
534
536
538
540
542
Binding Energy (eV)

Figure 20c


(a)
N_2/O_2=1.0
N_2/O_2=0.02
untreated
and N_2/O_2=0.00
(b)
N_2/O_2= 1.0
N_2/O_2=0.02
untreated
N_2/O_2=0.00
400
404
408
412
Binding Energy (eV)